ОБРОБЛЕННЯ СИГНАЛІВ І ЗОБРАЖЕНЬ ТА РОЗПІЗНАВАННЯ ОБРАЗІВ

Signal/Image Processing and Pattern Recognition

ПРАЦІ * PROCEEDINGS

П'ЯТА ВСЕУКРАЇНСЬКА МІЖНАРОДНА КОНФЕРЕНЦІЯ
УкрОБРАЗ'2000 • 27листопада - 1 грудня 2000 року

THE FIFTH ALL-UKRAINIAN INTERNATIONAL CONFERENCE
UkrOBRAZ'2000 • November 27 - December 1, 2000

Українська асоціація з оброблення інформації та розпізнавання образів
Ukrainian Association for Information Processing and Pattern Recognition

Національна академія наук України та Міністерство освіти і науки України -
Міжнародний науково-навчальний центр інформаційних технологій та систем
National Academy of Sciences of Ukraine & Education and Science Ministry of Ukraine -
International Research and Training Center for Information Technologies and Systems

Національна академія наук України - Інститут кібернетики
National Academy of Sciences of Ukraine - Institute of Cybernetics

Академія інженерних наук України - Відділення інформатизації та автоматизації
Academy of Engineering Sciences - Department of Informatization and Automatization

Українська інженерна спілка - Секція інформатики
Ukrainian Engineering Association - Section of Informatics

ОБРОБЛЕННЯ СИГНАЛІВ І ЗОБРАЖЕНЬ ТА РОЗПІЗНАВАННЯ ОБРАЗІВ

Signal/Image Processing and Pattern Recognition

ПРАЦІ ✵ PROCEEDINGS

П'ЯТА ВСЕУКРАЇНСЬКА МІЖНАРОДНА КОНФЕРЕНЦІЯ
27 листопада —1 грудня 2000 року, Україна, Київ

THE FIFTH ALL-UKRAINIAN INTERNATIONAL CONFERENCE
November 27— December 1, 2000 Kyjiv, Ukraine

УкрОБРАЗ'2000 ✵ UkrOBRAZ'2000

Київ 2000

Відповідальний упорядник і редактор

Вінцюк Тарас **— професор, доктор технічних наук, Голова Національного комітету Української асоціації з оброблення інформації та розпізнавання образів, академік Академії інженерних наук України, дійсний член НТШ, віце-президент Української інженерної спілки**

Видавнича колегія:

Гаврилюк Ольга **— канд. фіз.-мат. наук**

Гриценко Володимир **— канд. техн. наук, член-кореспондент Академії технологічних наук України**

Краковський Володимир **— канд. техн. наук**

Куляс Анатолій **— канд. техн. наук**

Сажок Микола **— мол. наук. співробітник**

Chief Editor

Vintsiuk Taras **— Professor, Dr. of Eng. Sc., President of Ukrainian Association for Information Processing and Pattern Recognition, Academician of the Academy of Engineering Sciences of Ukraine, Active Member of Shevchenko Scientific Society, Vice-President of Ukrainian Engineering Association**

Editorial Board:

Havryliuk Ol'ha **— Cand. of Phys.-Math. Sc.**

Hrytsenko Volodymyr **— Cand. of Eng. Sc., Correspondent Member of the Academy of Technological Sciences of Ukraine**

Krakovs'kyj Volodymyr **— Cand. of Eng. Sc.**

Kulias Anatolij **— Cand. of Eng. Sc.**

Sazhok Mykola **— junior researcher**

Статті публікуються без змін, зі збереженням авторського оригіналу

The papers are published without any change in the text that have been presented by authors

ISBN 966-02-1800-1

Видання Української асоціації з оброблення інформації та розпізнавання образів:
40, проспект Академіка Глушкова, Київ 03680, Україна

Наклад 200 примірників

ЗМІСТ
CONTENTS

I. Методологічні проблеми розпізнавання образів
Methodological Problems of Pattern Recognition

Тарас Вінцюк. Образний комп'ютер: концепція, методологія, підходи 9
Taras K. Vintsiuk. Pattern Computer: Concepts, Methodology, Approaches

Oleh J.Tretiak, Maria Gabrani. Multidimensional Nonlinear Registration: a Stochastic Formulation 17
Олег Третяк, Марія Ґабрані. Багатовимірна нелінійна реєстрація: стохастичне формулювання

Carlos Eduardo Thomaz, Raul Queiroz Feitosa, Alvaro Veiga. Separate-Group Covariance Estimation with Insufficient Data for Object Recognition 21
Карлос Е. Томаз, Рауль К. Фейтоса, Альваро Вейга. Оцінка коваріації при неповних даних для розпізнавання об'єктів

Савчинський Б.Д. Порівняльний аналіз алгоритмів стереозору в рамках байєсівської теорії прийняття рішень . 25
Savchyns'kyj B.D. Comparative Analysis of Computational Stereo Algorithms from the Point of View of Bayessian Approach

Zoltan Rusak, Imre Horvath, Gyorgy Kuczogi, Joris S. M. Vergeest . Techniques for Generating Shape Instances from Domain Distributed Vague Models 29
Золтан Русак, Імре Горват, Ґйоргі Кучоґі, Жоріс С.М.Вергеест. Методи генерування зразків форм за нечіткими моделями розподілених доменів

Ljudmila Sarycheva. Invariant Topological Characteristics of a Geometrical Object 33
Людмила Саричева. Інваріантні топологічні характеристики геометричних об'єктів

А.С. Краснопоясовський, С.О. Заговора. Розпізнавання електронограм за методом функціонально-статистичних випробувань 37
A.S. Krasnopojasovs'kyj, S.O. Zahovora. Electron-Diffraction Recognition Based on Functional-Statistic Testing Method

V.P. Mashtalir, E.P. Putyatin. Conflictfree Parallel Memory Access Models for Image Processing 41
В.П. Машталір, Є.П. Путятин. Моделі безконфліктного паралельного доступу до пам'яті для оброблення зображень

I.B. Sirodzha, B. Bar, I.I. Sirodzha. Pattern Recognition by Method of Probable Algorithm Quantums of Knowleges (PAQK-Method) 45
І.Б. Сіроджа, Б. Бар, І.І. Сіроджа. Розпізнавання образів методом імовірнісних алгоритмічних квантів знань (ІАКЗ-метод)

II. Теоретичні проблеми оброблення та розпізнавання сигналів і зображень
Theoretical Problems of Signal Processing and Recognition

Віктор Михайлишин, Ігор Яворський. Когерентні оцінки імовірнісних характеристик ритмічних процесів із невідомим періодом 51
Viktor Mykhajlyshyn, Ihor Javors'kyj. Coherent Estimates of Probabilistic Characteristics of Rhythmical Processes with Unknown Period

Nikolay N.Ponomarenko, Karen Egiazarian, Vladimir V.Lukin, Jaakko T. Astola. Speeding-up the Fractal Compression with Clustering 55
Микола Пономаренко, Карен Егіазарян, Володимир Лукін, Йаакко Т. Астола. Поліпшення фрактальної компресії методом кластеризації

Sergey K. Abramov, Vladimir V. Lukin, Jaakko T. Astola. Myriad Filter Properties and Parameter Selection 59
Сергій Абрамов, Володимир Лукін, Йаакко Т. Астола. Властивості та вибір параметрів множинних фільтрів

Р.В. Лотоцький. Методи стискання даних за допомогою перетворення сплесків 63
R.V. Lotots'kyj. Using Wavelet Transform for Data Compression

O.V. Babak, A.S. Gasanov, T.A. Babayev, J.V. Backlan. Common Approach to Solution of Some Problems of Discriminant and Claster Analysis 67
О.В. Бабак, А.С. Гасанов, Т.А. Бабаєв, І.В. Баклан. Єдиний підхід до розв'язання деяких задач дискримінантного та кластерного аналізу

III. Навчання та самонавчання розпізнаванню образів
Training and Selftraining in Pattern Recognition

Васильєв В.І., Ланге Т.І., Кобець Н.М. Принцип зовнішнього доповнення у методі групового врахування аргументів (МГВА) та в методі граничних спрощень (МГС) 71
Vasyl'jev V.I., Lange T.I., Kobets' N.M. Principle of External Complement in Group Arguments Considering (GAC) and Maximum Simplifications (MS) Methods

© UkrObraz'2000

Васильєв В.І., Струк С.П. Ітеративна процедура, яка використовує принцип дуальності в задачах навчання розпізнаванню образів 75
Vasyl'jev V.I., Struk S.P. The Iterative Procedure, which Uses a Duality Principle in Pattern Recognition Training Tasks

Ю.Ю. Милейко, Б.В. Рубльов. Побудова еліпса мінімальної площі на площині та еліпсоїда мінімального об'єму в *m*-вимірному просторі для скінченого набору точок 79
Ju.Ju. Mylejko, B.V. Rubliov. Minimal Area Elypsis Construction on Plane and Minimal Volume Ellipsoid Construction in *m*-Dimensional Space for Finite Set of Points

М.Ф. Кириченко, Роман Куц, М.П. Лепеха. Множини належності в задачах класифікації сигналів 83
M.F. Kyrychenko, Roman Kuc, M.P.Lepekha. Relationship Sets in Signal Classification Problems

Роман Базилевич . Конструктивні алгоритми розбиття складних систем на основі оптимального згортання 87
Roman Bazylevych. Complex Systems Constructive Partitioning Algorithms on the Basis of the Optimal Circuit Reduction Method

Роман Базилевич. Ітераційні алоритми оптимізації розбиття складних систем на основі оптимального згортання 91
Roman Bazylevych. Complex Systems Iterative Partitioning Algorithms on the Basis of the Optimal Circuit Reduction Method

IV. Автоматичне розпізнавання та синтез усномовних сигналів
Automatic Recognition and Synthesis of Speech Signals

Taras K. Vintsiuk. Generalised Automatic Phonetic Transcribing of Speech Signals 95
Тарас Вінцюк. Узагальнене автоматичне фонетичне транскрибування усномовного сигналу

Potapova R.K., Sobakin A.N. Development of the Approach to Recognition of Speech with a Support on Visual Images 99
Потапова Р. К., Собакін А.М. Розроблення підходу до розпізнавання усної мови на основі зорових образів

Dionysios Politis, Alexander Dolia, Panagiotis Linardis. Identifying Transient Patterns in Idiomatic Greek and Slavonic Music 103
Діонісіос Політіс, Олександр Доля, Панагіотіс Лінардіс. Ідентифікування перехідних образів в ідіоматичній грецькій і слов'янській музиці

В.М. Магера, І.І. Горбань, С.В. Лєвий. Автоматизація криміналістичних досліджень мовленнєвих сигналів 107
V.M. Mahera, I.I. Horban', S.V. Lievyj. Voice Calls Criminalistic Research Automatization

Микола Сажок. Генерування правил розставляння наголосів у багатомовному аспекті 111
Mykola Sazhok. Generating Stress Rules in Multilingual Aspect

V. Технічна діагностика об'єктів і машин за їх сигналами та полями
Automatic Diagnostic of Objects and Machines on its Signals and Fields

Ігор Ісаєв, Георгій Трохим, Ігор Яворський. Статистичний аналіз ритмічних сигналів для діагностики технічного стану рейок 113
Ihor Isajev, Heorhij Trokhym, Ihor Javors'kyj. Statistical Analysis of Rhytmical Signals for Rail Technical State Diagnostics

Андрій Гвоздак, Олександр Ахметшин. Контроль якості сталевих труб в реальному часі з використанням нейронної мережі, яка прогнозує сигнал помилки наявності дефектів 117
Andrij Hvozdak, Oleksander Akhmetshyn . Real-Time Steel Pipe Quality Testing with Using a Neural Network that Forecasts Defect Presence Error Signal

Петро Драбич, Олег Драбич. Пошук та обстеження захованих об'єктів за їх сигналами 121
Petro Drabych, Oleh Drabych. Search and Investigation of Hidden Objects by Their Signals

W.Kongrattanaprasert, S.Arunrungrusmi, B.Pungsiri, K.Chamnongthai, M.Okuda. Nondestructive Maturity Determination of Durian by Force Vibration 125
В. Конграттанапрасерт, С. Арунрунгрусмі, Б. Пунгсірі, К. Чамнонгтай, М. Окуда. Неруйнівний контроль зрілості кокосового горіха за допомогою вібрацій

Ibarra-Pico F., Garcia-Crespi F., Cuenca-Asensi S.A., Morales-Benavente J.L., Lorenzo-Quintanilla J.J. A DSP Implementation of an AOM and its Application to Defects Detection in Textile Material 129
Ібарра-Піко Ф., Гарсіа-Креспі Ф., Куенка-Асенсі С. А., Моралес-Бенавенте Й.Л., Лоренцо-Квінтанілла Й. Й. Втілення AOM на DSP і його використання до виявлення дефектів у текстильному матеріалі

В.В. Кошовий, Є.В. Кривін, В.П. Курій, О.Є. Левицький, І.М.Романишин. До проблеми технічної діагностики товстолистових виробів методом ультразвукової томографії 133
V.V. Koshovyj, Je.V. Kryvin, V.P. Kurij, O.Je. Levyts'kyj, I.M. Romanyshyn. Thechnical Diagnostics of the Thick-Sheet Products by the Computerized Ultrasonic Tomography Method

В.Ф. Миргород, С.М. Ковешников, А.Г. Кисель. Формування А-ехограм у вигляді розтину акустичних когерентних зображень 137
V.F. Myrhorod, S.M. Koveshnykov, A.H. Kysel'. A-Echogram Formation as a Dissection of Acoustic Coherent Images

В.Д. Павленко, О.О. Фомін, В. О. Шибаєв. Інтегрована інструментальна система для розв'язання задач діагностичного контролю 141
V.D. Pavlenko, O.O. Fomin, V.O. Shybajev. Integrated Instrumental System for Decision Problems of Diagnostics Control

Anagnostopoulos C., Hasan S., Kayafas E., Knaak U., Loumos V., Nassar S., Saqer Abdel-Rahim, Stassinopoulos G., Vergados D. Requirements and Specification for a High-Speed Computing Unit for Flaw Detection in Fabrics 145
Анагностопулос К., Гасан С., Кайафас Е., Кнаак У., Лумос В., Нассар С. , Сакер А.-Р., Стассінопулос Г., Вергадос Д. Вимоги та специфікація швидкого обчислювача для виявлення дефектів у тканинах

В.Г. Рябцев, Н.Л. Кириленко, М. Х. Кхан. Метод вибору тестів діагностування мікросхем пам'яті при нечітких апріорних даних 149
V.H. Riabtsev, N.L. Kyrylenko, M.H. Khan. Method of Diagnostics Tests Selection for Memory Microcircuits under Fuzzy Prior Data

VI. Оброблення та розпізнавання зображень і сцен
Image and Scene Processing and Recognition

Akhmetshin A.M., Shutin D.V. Non-Linear Model of Qualitative Classification on Base of Wave-Packet Decomposition in Space of Volterra Kernels 153
Ахметшин О.М., Шутін Д.В. Нелінійна модель якісної класифікації на основі методу хвильового пакетного перетворення у просторі ознак ядер Вольтерра

Володимир Калмиков, Дмитро Калмиков, Тетяна Романенко . Структурний метод опису і виділення відрізків прямих ліній у контурі 157
Volodymyr Kalmykov, Dmytro Kalmykov, Tetiana Romanenko. The Strucrural Method: the Contour Direct Line Pieces to be Detected and Allocated

A.M. Akhmetshin, A.E. Fedorenko. Improvement of Low-Contrast Medical Image Analysis Sensitiveness and Exactness Using Markov Random Field Multidimentional Segmentation Method 161
О.М. Ахметшин, А.Є. Федоренко. Підвищення чутливості і точності аналізу низькоконтрастних медичних зображень на основі методу багатовимірної сегментації марківських полів

Ch.Feist, V.Kiiko, R.Knofel, G.Stanke. Attributed Relational Graphs Based Description and Retrieval of Shoeprints in the Image Database System SHARS 165
Х. Файст, В. Кийко, Р. Кнофель, Ґ. Штанке. Застосування атрибутних графів для опису та пошуку зображень відбитків взуття

Левко О.В., Нюнькін К.М. Використання "власних облич" для розпізнавання людських облич 169
Levko O.V., Niun'kin K. M. "Proper Faces" Usage for Human Face Recognition

A.M. Akhmetshin, D.A. Mikhalev. Classification of Texture Images with the Usage of Artificial Neural Nets and Higher-Order Statistics 173
О.М. Ахметшин, Д.О. Михальов. Класифікація текстурних зображень за допомогою штучних нейронних мереж та статистик вищих порядків

М.І. Шлезінгер, Б. Флах, І.В. Онищенко. Оптимізація кусково-лінійної деформації знімків у задачах машинного стереозору 177
M.I. Schlesinger, B. Flach, I.V. Onyshchenko. Piecewise-Linear Photographs Deformation Optimization in Machine Stereovision Problems

Д.І. Рябоконь. Калібровка зображень для порядкових методів розпізнавання рельєфу 179
D.I. Riabokon'. Image Calibration for Line-by-Line Relief Recognition Methods

А.С. Краснопоясовський, А.М. Скаковська. Автофокусування електронного мікроскопа за зображенням 183
A.S. Krasnopojasovs'kyj, A.M. Skakovs'ka. Electronic Microscope Auto-Focusing on the Image

VII. Архітектура процесорів оброблення та розпізнавання сигналів і зображень
Processor and System Architecture for Signal/ImageProcessing and Recognition

Y.V. Baybak, N.O. Shweiki, Y.F. Kutaev, A.A. Gertsiy, L.I. Timchenko. Multistage Approach to Image Processing 187
Ю.В. Байбак, Н.О. Швейкі, Ю.Ф. Кутаєв, А.О. Герцій, Л.І. Тимченко. Підхід до багаторівневого оброблення зображень

Тетяна Мартинюк, Антоніна Буда, Андрій Кожем'яко, Тетяна Васильєва, Віра Козлова. Математична модель нейрона на принципах паралельної порогової обробки інформації 191
Tetiana Martyniuk, Antonina Buda, Andrij Kozhemiako, Tetiana Vasilyeva, Vira Kozlova. Mathematical Model of Neuron on Principles of Parallel Threshold Information Processing

I.M. Gvozdeva, A.S. Paevskiy. Increasing the Resolution of Space Filters 193
І.М. Гвоздєва, А.С. Паєвський. Підвищення розрізняльної здатності просторових фільтрів

Михайло Семенюк. Засоби пам'яті послідовної та паралельної дії на оптоелектронних пристроях 197
Mykhajlo Semeniuk. Successive and Parallel Action Memory on Optoelectronic Devices

VIII. Фільтрація, аналіз, відновлення та синтез сигналів і зображень
Filtering, Analysis, Reconstruction and Synthesis of Signals and Images

Pertti T. Koivisto, Vladimir V. Lukin, Vladimir P. Melnik, Oleg V. Tsymbal. Removal of Impulse Bursts in Radar Images 201

Перті Т.Койвісто, Володимир В.Лукін, Володимир П.Мельник, Олег В.Цимбал. Усунення імпульсних викидів у радарних зображеннях

Олексій Коваль, Богдан Русин. Вибір вейвлету та аналіз впливу степені регулярності біортогональних фільтрів, побудованих на основі теорії полюсів, для стиску зображень 205

Oleksij Koval', Bohdan Rusyn. Image Compression Wavelet Choice and Biorthogonal Regularity Degree Influence Analysis for Filters Constructed on Poles Theory

Роман Воробель, Ігор Журавель, Наталія Опир. Підвищення візуальної якості затемнених деталей на зображенні 209

Roman Vorobel', Ihor Zhuravel', Natalia Opyr. Visual Quality Increasing for Image Darken Components

A.M. Akhmetshin, D.A. Fenoga. Increasing of Low Contrast Medical Multispectral Image Segmentation Sensitivity on Base Regularized Method of the Gram-Schmidt Orthogonalization 213

О.М. Ахметшин, Д.А. Фенога. Підвищення чутливості сегментації низькоконтрасних медичних багатоспектральних зображень на основі регуляризованого методу ортогоналізації Грама-Шмідта

О.П. Бухало, Б.П. Клим, Г.В. Микитин, Є.П. Почапський, Р.Ф. Федорів. Сигнал акустичної емісії матеріалів і конструкцій – вимірювальні перетворення 217

O.P. Bukhalo, B.P. Klym, H.V. Mykytyn, Je.P. Pochaps'kyj, R.F. Fedoriv. Acoustic Emission Signal of Materials: Constructions-Measuring Transformations

Антоніна Буда, Тетяна Мартинюк, Генадій Лисенко, Ольга Король, Сергій Буда. Математичні моделі аналізу контурів плоских симетричних зображень 221

Antonina Buda, Tetiana Martyniuk, Henadij Lysenko, Olha Korol', Serhij Buda. Mathematical Model for Analysis of Flat Symmetry Image Contours

N.V. Ruzhentsev, Yu.M. Gerasimov, A.V. Antonov. The Particularities of Radio Heat Images of Cloudy Fields as a Priori Information for Remote Sensing Algorithms 223

Н.В. Руженцев, Ю.М. Герасімов, О.В. Антонов. Особливості теплових радіозображень хмарних полів як апріорна інформація для алгоритмів дистанційного зондування

Тарасов В.О., Кропачов Д.О., Коржов В.І. Практична реалізація алгоритмів первинного оброблення сигналів у системах пасивної звукової локації 227

Tarasov V.O., Kropachov D. O., Korzhov V.I. Signal Preprocessing Practical Realization for Passive Sound Ranging Systems

A.M. Akhmetshin, O.P. Kirichenko. Increasing Sensitivity and Space Resolving Power of Low Contrast Radiological Image Analysis as a Task Inverse Analytical Continuation of Virtual Pseudocoherent Wave Field 231

О.М. Ахметшин, О.П. Кириченко. Підвищення чутливості і розрізняльної здатності аналізу низькоконтрасних радіологічних зображень як задача оберненого аналітичного продовження віртуального псевдокогерентного хвильового поля

D.V. Dovnar, K.G. Predko, O.N. Lysikova. The Combined Systems of Supervision 235

Д.В. Довнар, К.Г. Предко, О.М. Лисікова. Комбіновані системи відновлення зображень

Корольов В.Ю. Реставрація образів у частотній області методом локально-адаптивної умовної деконволюції .237

Koroliov V.Ju. Image Restoration in Spatial Frequency Domain by Using the Local-Adaptive Conditional Deconvolution Method

Г.М. Востров, М.В. Полякова, В.В. Любченко. Розпізнавання образів з самонавчанням за допомогою детермінованої фрактальної моделі 241

G.M. Vostrov, M.V. Poliakova, V.V. Liubchenko. Pattern Recognition with Self-Learning Based on Deterministic Fractal Theory

В.В. Кошовий, А.Б. Лозинський, Б.А. Лозинський. Застосування томографічного підходу в задачах відновлення радіозображень космічних радіоджерел 245

V.V. Koshovyj, A.B. Lozyns'kyj, B.A. Lozyns'kyj. Using Tomographic Technique for Restoring the Cosmic Sources Radio Images

І.О. Процько, В.А. Радомський. Узагальнений підхід швидкого трансформування класу Фур'є на основі згорток 249

I.O. Prots'ko, V.A. Radoms'kyj. Generalized Fast Fourier Transform Based on Convolutions

Микола Будник. Двовимірне перетворення Фур'є з комплексною фазою 253

Mykola Budnyk. Two-Dimensional Fourier Transform with Complex Phase

В.П. Савчук, І.М. Глушко. Фільтрація ближнього поля системи взаємодіючих випромінювачів 257

V.P. Savchuk, I.M. Hlushko. Interactive Emitters System Near-Field Filtration

С.Г. Антощук, В.М. Крилов, В.О. Давидов. Попереднє оброблення сигналів та зображень 261

S.H. Antoshchuk, V.M. Krylov, V.O. Davydov. Signal and Image Preprocessing

IX. Нові інформаційні технології та системи, які базуються на обробленні сигналів і зображень
Advanced Information Technologies Based on Signal/Image Processing and Pattern Recognition

Radu Pentiuc,Stefan-Gheorghe Pentiuc, Adrian Graur, Mariana Milici. Integrated System for Mechanic Band Conveyors with Hybrid Motor, TV Camera and Robot 265
Раду Пентюк, Стефан-Ґеорґ Пентюк, Адріан Ґраур, Маріана Мілічі. Інтегрована система керування механічними стрічковими конвейерами, що використовує робота, TV-камеру і гібридний мотор

Gyorgy Kuczogi, Zoltan Rusak, Imre Horvath. Towards a Natural User Interface for Comprehensive Support of Conceptual Shape Design 271
Ґйорґі Кучоґі, Імре Хорват, Золтан Русак. На шляху до природнього інтерфейсу для інтелектуальної підтримки розроблення систем

Ростислав Косаревич, Ігор Ісаєв, Богдан Русин. Визначення періоду папілярного відбитка на зображенні на основі побудови функціоналів математичного сподівання 275
Rostyslav Kosarevych, Ihor Isajev, Bohdan Rusyn. Fingerprint Image Period Determination Based on Functionals Mathematical Expectation

Д.М. Піза, С.В. Морщавка. Алгоритм класифікації рослинних об'єктів з можливістю швидкої адаптації 279
D.M. Piza, S.V. Morshchavka. Vegetative Objects Classification Algorithm with an Opportunity of Fast Adaptation

Боюн В.П. Довгань В.В. Інтелектуалізація відеосистем колового огляду 283
Bojun V. P., Dovhan' V. V. Intellectualization of Round-Observation Videosystems

М.Ф. Козак, В.В. Ніколенко. Завадостійкі штрихові коди 287
M.F. Kozak, V.V. Nikolenko. The Noise-Resistant Bar Codes

A.Benkhalil, S.Ipson, W.Booth. New Generation Video Surveillance System Based on Image Processing and FPGA Processor towards Autonomous System 291
А. Бенхаліл, С. Іпсон, В. Боот. Нове покоління автономних систем відеонагляду на основі оброблення зображень

X. Моделювання та дослідницькі системи
Modelling and Experimental Systems

Tarasov V. A. Simulation of the Radiation Creep of a Reactor Fuel 297
Тарасов В.О. Моделювання поширення радіації від пального реактора

Воронова О.С. Про один метод адаптивної апроксимації неперервних сигналів апроксимуючими імпульсними спектрами 301
Voronova O. S. On One Method of Adaptive Continuous Signal Approximation by Impulse Spectra

V.N. Borovytsky, V.V. Fesenko. Fidelity Analysis of Imaging Systems 305
В.М. Боровицький, В.В. Фесенко. Аналіз вірності систем відтворення зображень

Сергій Юхимчук. Алгоритм і пристрій оцінки стійкості нелінійних нестаціонарних систем, які допускають лінеаризацію 309
Serhij Jukhymchuk. Nonlinear Non-Stationary System Stability Evalution Algorithm and Device

Н.С. Кравець. Інтерактивна імітація та аналіз розкрашеної сітки Петри з використанням алгебри предикатних операцій 313
N.S. Kravets'. Painted Petri Net Interactive Imitation and Analysis by Predicate Operations Algebra

Володимир Овсяк, Василь Бритковський. Математичні моделі структур текстово-графічних даних редактора формул алгоритмів-секвенцій 317
Volodymyr Ovsiak, Vasyl' Brytkovs'kyj. Mathematical Models of Structures for Algorithms-Secventhens Formulas Data Text Graphic Editor

АЛФАВІТНИЙ ПОКАЖЧИК. AUTHOR INDEX 319

Образний комп'ютер: концепції, методологія, підходи

Тарас Вінцюк

Міжнародний науково-навчальний центр ЮНЕСКО інформаційних технологій та систем
40 проспект Академіка Глушкова, Київ 03680, Україна
Тел.: +380 44 266-4356 Факс: +380 44 266-1570
vintsiuk@uasoiro.freenet.kiev.ua

ABSTRACT

Taras K. Vintsiuk. Pattern Computer: Concepts, Methodology, Approaches. Problems of designing, creation, developing, production and dissemination for principally new so called pattern computers and respective intellectual information technologies and systems are discussed. Research strategy and technical policy for Ukraine on 2000-2010 years period as well as on nearest 2000-2003 years period are considered too.

Pattern computer (PatCom) is such intellectual cybernetic system that provides a functional simulation of intelligent, mainly subconscious activity of all living and a human being particularly. In PatCom this activity is related to image, sound and other patterns perception, scene analysis, action and movement planning, generalisation of observations, discovering of regularities, prediction, decision making etc. Pattern computer operates with patterns and other complex notions. It actualises both pattern and logical reasoning.

Pattern computer is a parallel system. It has several information perception channels (acoustic, visual, scential) that is multimodal perception, pattern operation system, improved human-machine interface. Unlike usual computer, which is based on a rapid arithmetic-logical processor and a large RAM space, PatCom is grounded on external world models including physical, geometry, acustical, language, linguistic, semantical, canonical forms etc models.

Two ways how to create a pattern computer are debated. The first approach is based on so-called generative model of pattern recognition and understanding. A multimodal information synthesis is used as a feedback in the pattern analysis and understanding. The second way to create PatCom is so-called multilevel multidecision model. As example, ways how to create a dictation and spoken translation machine are discussed.

ВСТУП

Сьогодні науково-технічний прогрес визначається значними проривами екстенсивного характеру в галузі розроблення та впровадження інформаційних систем, засобів телекомунікації та комп'ютеризації. Досягнені фантастичні швидкодії та обсяги пам'яті у комп'ютерів, глобалізується Інтернет, надається широкий спектр інформаційних послуг обчислювального та пошукового характеру.

Спостерігаються й чіткі негативні тенденції, пов'язані з світовою монополізацією розроблення та виробництва комп'ютерів та засобів комунікації, зі штучним стримуванням розвитку передових інформаційних технологій та систем, в яких має нагальну потребу людство.

А людство потребує створення принципово нових комп'ютерів, які здатні сприймати та розуміти звуки, зображення, людську мову, рукописні тексти, креслення, просторові та звукові сцени, інші образи, описувати та озвучувати зображення, перекладати з однієї мови на іншу тощо.

Розроблення таких комп'ютерів, які виконують не тільки обчислення, але й моделюють образне сприйняття світу та образне прийняття рішень відносять до проривних напрямів у науково-технологічному поступі.

Кабінет Міністрів України своєю Постановою від 8.11.2000 № 1652 схвалив Державну науково-технічну програму "Образний комп'ютер" та встановив два періоди її розвитку: 2000-2003 та 2004-2010 роки.

НАЗВА ТА МЕТА ПРОГРАМИ

"Розроблення нових, так званих образних комп'ютерів (ОК)". Скорочена назва програми — "Образний комп'ютер".

Програма має на меті створення принципово нових інформаційних технологій та систем — образних комп'ютерів, входження України в міжнародний науково-технологічний розподіл праці, гарантування ринків збуту вітчизняної наукомісткої продукції.

ЩО ТАКЕ ОБРАЗНИЙ КОМП'ЮТЕР?

Образний комп'ютер — це така кібернетична система, в якій виконується функційне моделювання інтелектуальної, головно підсвідомої, діяльності людини та всього живого, що пов'язана зі сприйняттям зорових, слухових та інших образів, аналізом сцен та складних ситуацій, плануванням дій та рухів, узагальненням спостережень, встановленням закономірностей, прогнозуванням, прийняттям рішень тощо. ОК оперує образами та іншими складними поняттями, реалізує як образне, так і логічне мислення.

Образний комп'ютер є паралельною мультимодальною системою, яка має у своєму складі декілька каналів сприйняття інформації (слухової, зорової, текстової, смакової, нюхової тощо), образну операційну систему, моделі зовнішнього світу (в тому числі акустичну, оптичну, геометричну, лінгвістичну, семантичну, канонічних форм тощо), розвинений інтерфейс з людиною, засоби взаємодії з існуючими комп'ютерними та телекомунікаційними мережами.

Образна операційна система "синхронізує" оброблення інформації, що надходить різними каналами її сприйняття, та, оперуючи моделями зовнішнього світу, виконує комплексну семантичну інтерпретацію всієї отриманої інформації.

© UkrObraz'2000

На відміну від звичайного комп'ютера, основу якого складають арифметико-логічний процесор та оперативна пам'ять, у центрі образного комп'ютера знаходяться моделі зовнішнього світу з усіма його об'єктами, явищами та їх проявами.

Розкриємо поняття образного комп'ютера переліком деяких базових його властивостей, які розглядатимемо як вимоги до ОК.

Сприймає та аналізує звукові сигнали, визначає характер звуків, їх джерело, напрямок та відстань до джерела, автоматично розпізнає тип джерела.

Сприймає людську мову, автоматично розпізнає та розуміє її, озвучує довільні тексти. Автоматично редагує та друкує тексти під диктування.

Сприймає та розуміє надруковані та написані від руки тексти. Вміє перекладати з однієї мови на іншу.

Робить усний переклад з однієї мови на іншу.

Аналізує та розуміє зображення (фото, написи, креслення, графіки тощо). Аналізує та розуміє сцени, обстежує об'єкти, вимірює відстань до об'єктів, вказує їх місцерозташування.

Орієнтується в часі та просторі. Аналізує причини та наслідки.

Верифікує та ідентифікує людину за її голосом.

Сприймає запахи та тактильну інформацію.

Стежить за здоров'ям користувача.

Здатен комплексно сприймати інформацію про зовнішній світ. Здатен до узагальнення спостережень, встановлення закономірностей, прогнозування, цілеспрямованих дій та рухів.

Дає зорову інтерпретацію усних повідомлень, описує текстами та усною мовою ті зображення, які спостерігає.

Сприймає, інтерпретує та описує аудіовідеоінформацію.

Виконує автоматичний пошук аудіовідео- та текстової інформації за заданою темою. Автоматично реферує тексти та аудіовідеоінформацію.

КОНЦЕПЦІЯ ПРОГРАМИ

Ряд ОК. Створюється ряд моделей образних комп'ютерів – від супер-ОК до персонального ОК масового використання.

Єдність ближніх і дальніх цілей. Створення образних комп'ютерів у повному обсязі потребує не менш ніж 10-річного періоду науково-технологічного поступу. Перші прототипи ОК можуть бути розроблені за дво(чотири)річний період. Окремі компоненти образних комп'ютерів можуть бути підготовлені для впровадження в найближчий рік-два. Ці компоненти вбудовуватимуться в існуючі ЕОМ, суттєво покращуючи споживчі властивості останніх.

Створення прототипів ОК кожного поточного року та коригування науково-технічної програми є необхідною умовою робіт.

Основні напрями робіт. Розроблення образного комп'ютера виконується за такими напрямами:

1. Методологія, концепція та архітектура ОК.

2. Образні операційні системи та системи програмування ОК.

3. Автоматичне розпізнавання, розуміння та синтез мовних сигналів для усного спілкування з ЕОМ природною мовою, створення диктувальних машин та машин усного перекладу. ОК "Усна Мова".

4. Сприйняття, автоматичне розпізнавання, розуміння та генерація зображень (просторових сцен, плоских зображень, текстової та креслярської інформації, зокрема друкованої та рукописної). ОК "Око".

5. Автоматичне розуміння, інтерпретація та генерація текстової інформації та її переклад з однієї мови на іншу. ОК "Перекладач".

6. Зорова інтерпретація усномовних повідомлень, усномовна інтерпретація зображень, комплексне сприйняття та генерація аудіовідеозображень. ОК "Інтерпретатор".

7. Самопрограмування цілеспрямованих дій комп'ютерних систем в реальному світі. ОК "Робот".

8. Автоматичне реферування та пошук текстової та аудіовідеоінформації за заданою темою. ОК "Автореферент".

Дослідження за пп. 1-2 є фундаментальними, за пп. 3-8 — фундаментальними та науково-дослідними і дослідно-конструкторськими.

Перелік пілотних проектів 2000-2001 років:

1. Пакет програмних засобів для озвучення текстів українською та іншими мовами в ПК.

2. Портативний текстовий словник-перекладач з української та на українську з озвученням українською мовою.

3. Усний словник-перекладач (англійсько-український та українсько-англійський) для персонального комп'ютера.

4. Програмно-апаратний комплекс для оброблення та змістовної інтерпретації креслярсько-графічних зображень (схем, карт, планів місцевості, креслень споруд тощо).

5. Система телевізійного сприйняття навколишнього світу.

6. Засоби усномовного введення-виведення даних для ЕОМ з технологіями мультимедія.

7. Голосові замки.

8. Комп'ютерна фоноскопія.

9. Самопрограмування цілеспрямованих рухів маніпуляційних роботів для обминання перепон (наприклад, для збирання уламків радіоактивниго палива).

10. Тривимірний дисплей з контролером для нього.

Першочергові завдання і проекти програми зведені у перелік, що наводиться нижче.

1. Методологія, концепція та архітектура ОК

1.1. Методологія, концепція та архітектура образних комп'ютерів

1.2. Розроблення моделей та формальних мов для опису зовнішнього світу та подій у ньому

1.3. Автоматизація процесів формування понять про об'єкти, явища та закономірності за зразками, подібностями та аналогіями

1.4. Навчання, самонавчання та самоорганізація моделей

2. Образні операційні системи та системи програмування ОК

2.1. Ієрархічна структуризація знань і даних про зовнішній світ через мультимодальне сприйняття та організація управління пошуком інформації й прийняттям рішень

2.2. Принципи організації та управління в образних операційних системах

2.3. Системи програмування образних комп'ютерів

3. Автоматичне розпізнавання, розуміння та синтез мовних сигналів для усного спілкування з ЕОМ природною мовою, в тому числі створення диктувальних машин та машин усного перекладу. ОК "Усна Мова"

3.1. Набір засобів для озвучення текстів українською та іншими мовами в ПК

3.2. Портативний текстовий словник-перекладач з української та на українську з озвученням українською мовою

3.3. Усний словник-перекладач (англійсько-український та українсько-англійський) для персонального комп'ютера

3.4. Засоби усного діалогу для персонального комп'ютера

3.5. Диктувальна машина (машина для автоматичного друкування та редагування текстів під послівне диктування)

3.6. Диктувальна машина під зв'язну мову

3.7. Машина усного перекладу для англійської та української мов при послівному вимовлянні

3.8. Машина усного перекладу для англійської та української мов при злитому вимовлянні

3.9. Засоби стискання, компресованої передачі та відтворення усномовної інформації та їх інтеграція з Інтернет для організації комп'ютерної телефонії

3.10. Автоматична ідентифікація та верифікація особи за її голосом, голосові замки

3.11. Комп'ютерна фоноскопія

3.12. Генерація усної мови з моделюванням індивідуальних особливостей голосу людини

4. Сприйняття, автоматичне розпізнавання, розуміння та генерація зображень (просторових сцен, плоских зображень, текстової та креслярської інформації, зокрема друкованої та рукописної). ОК "Око"

4.1. Оброблення та змістовна інтерпретація креслярсько-графічних зображень (схем, карт, планів місцевості, креслень споруд тощо)

4.2. Програмно-апаратні засоби сприйняття рукописної інформації

4.3. Автоматичне сприйняття текстової інформації з книжок, газет, друкованих рукописів, її редагування та переформатування

4.4. Сприйняття, аналіз та розуміння просторових сцен

4.5. Автоматична ідентифікація та верифікація підписів

4.6. Генерація зображень та просторових сцен

4.7. Тривимірний дисплей з засобами його підтримки для персонального комп'ютера

5. Автоматичне розуміння, інтерпретація та генерація текстової інформації та її переклад з однієї мови на іншу. ОК "Перекладач"

5.1. Автоматичне розуміння та інтерпретація текстової інформації

5.2. Автоматичний синтез текстової інформації за заданим смислом

5.3. Автоматизація перекладу текстової інформації з однієї мови на іншу — лексично-граматичний переклад

5.4. Автоматичний переклад текстової інформації з однієї мови на іншу, що ґрунтується на розумінні текстів — семантичний переклад

5.5. Створення базового набору комп'ютерних мультимедійних словників української мови

5.6. Створення комп'ютерних версій українсько-іншомовних словників

6. Зорова інтерпретація усномовних повідомлень, усномовна інтерпретація зображень, комплексне сприйняття та генерація аудіовідеозображень. ОК "Інтерпретатор"

6.1. Зорова інтерпретація текстової інформації – перетворювач "текст—зображення"

6.2. Смислова, текстова та усномовна інтерпретація зображень – перетворювач "зображення—смисл—текст—усна мова"

6.3. Комплексне сприйняття та опис аудіовідеоінформації – перетворювач "аудіовідео—смисл-текст"

6.4. Комплексна генерація аудіовідеоінформації за заданим "смисл-текстом"

7. Самопрограмування цілеспрямованих дій комп'ютерних систем в реальному світі. ОК "Робот".

7.1. Система телевізійного сприйняття навколишнього світу

7.2. Самопрограмування цілеспрямованих рухів маніпуляційних роботів для обминання перепон (наприклад, для збирання уламків радіоактивниго палива).

8. Автоматичне реферування та пошук текстової та аудіовідеоінформації за заданою темою. ОК "Автореферент"

8.1. Автоматичне реферування текстів

8.2. Пошук текстової інформації за темою реферату

8.3. Автоматичне реферування фонограм

8.4. Автоматичний пошук фонограм на задану тему

8.5. Автоматичне реферування та пошук аудіовідеоінформації

МЕТОДОЛОГІЯ СТВОРЕННЯ ОК

Першу в континентальній Європі електронну обчислювальну машину, що мала назву "МЭСМ", було створено в Києві ще на початку 50-х років. Уже тоді її творці усвідомлювали, що машину треба навчити не тільки робити обчислення, а й мислити, сприймати й розуміти людську мову, а також звуки та зображення.

З того часу комп'ютери досягли гігантської швидкодії та обсягів пам'яті, значних успіхів у виконанні обчислень, пошуку інформації, формальної логіки. Але дуже скромні результати мають місце в моделюванні підсвідомої діяльності людини та всього живого. По правді кажучи, комп'ютери все ще залишаються "сліпими" та "глухими" в широкому розумінні цих слів.

Знаменитий Білл Гейтс, керівник фірми Microsoft, вважає, що майбутнє за інтелектуальними інформаційними технологіями та системами, які сприймають і розуміють людську мову, друковані та рукописні

тексти, обробляють і розуміють різноманітні зображення, креслення тощо, сприймають просторові та звукові сцени. Розробкою цих технологій Гейтс збирається заволодіти світовим інформаційним простором та отримати надприбутки.

З часу створення першої ЕОМ "МЭСМ" в Україні сформувалась та утвердилась наукова школа світового рівня з проблем автоматичного розпізнавання сигналів і зображень та створення інтелектуальних інформаційних технологій і систем на їх основі.

Складність проблеми створення ОК зумовлена багатьма чинниками. Чи не найголовнішим з-посеред них є той, що розумова діяльність людини, підсвідомі дії всього живого, що пов'язані з розпізнаванням образів, є "за сімома замками", тобто прихованими від нас Природою. Часто цю діяльність та дії називають підсвідомими або інтуїтивними, а, отже, такими, що погано піддаються формалізації, дослідженням точними науками.

Українські науковці пішли шляхом функційного моделювання підсвідомої діяльності людини та всього живого, грунтуючись на загальновизнаних принципах оброблення інформації в живій природі, техніці та суспільстві, таких як: аналіз через синтез, індуктивне та дедуктивне виведення, генерація та направлений перебір варіантів, зворотній зв'язок, використання апріорної інформації, адаптація, навчання та самонавчання, розпізнавання образів, побудова моделей зовнішнього світу та розумової діяльності. Проблема полягає в тому, як конструктивно об'єднати та реалізувати ці принципи, щоб створити інтелектуальні інформаційні технології та системи.

Такими конструктивними моделями стали названі нами генеративні моделі розпізнавання сигналів і зображень. Відомо, що кожне вимовляння одного й того ж слова чи написання однієї й тієї ж літери навіть однією й тією ж людиною, скажімо, з інтервалом в одну секунду часу, завжди передаються різними ("двічі в одну й ту ж воду ввійти неможливо"), але чимось схожими сигналами або зображеннями.

Отож, універсальним алгоритмом розпізнавання міг би бути такий. Спочатку запам'ятати всі можливі сигнали чи зображення, які передають один і той самий образ (клас, слово, букву тощо), а потім при автоматичному розпізнаванні порівнювати пред'явлений для оброблення сигнал чи зображення з усіма раніше запам'ятованими сигналами чи зображеннями й віднести розпізнаваний сигнал (зображення) до того образу (класу), з чиїм раніше запам'ятованим сигналом (зображенням) він збігся. На жаль, такий алгоритм розпізнавання хоч і є "сильним", але він не є конструктивним: не знайдеться такий звичайний комп'ютер, ні тепер, ні в майбутньому, який був би в змозі запам'ятати всі можливі, розмаїті сигнали (зображення) та виконувати в реальному часі необхідні порівняння.

Труднощі, що виникають при реалізації цього універсального алгоритму розпізнавання можна подолати, якщо спробувати навчитись запам'ятовувати, структуризувати всі можливі сигнали (зображення) якимсь економним способом (адже вони пов'язані сильними залежностями, бо є схожими!), а порівняння сигналів (зображень) виконувати направленим перебором варіантів.

Відомо, що за допомогою диференційних рівнянь або рівнянь з частинними похідними можна генерувати величезну кількість близьких, "схожих" розв'язків. Для економного ж опису сигналів і зображень з подальшим їх розпізнаванням були розроблені ієрархічно організовані одно- та двовимірні стохастичні породжувальні граматики (генеративні моделі), які використані в якості зворотного зв'язку для направленого перебору та порівняння сигналів і зображень, що стало ефективною конструктивною реалізацією ідеї аналізу сигналів (зображень) через їх синтез.

Українськими науковцями розроблені принципово нові, інтелектуальні, системи, теорія навчання та самонавчання розпізнаванню образів, принципи розумної поведінки кібернетичних пристроїв, принципи планування цілеспрямованої поведінки роботів у складних умовах, теорія і методи автоматичного розпізнавання, розуміння та синтезу мовних сигналів, математична теорія розпізнавання зображень та теорія комп'ютерного стереозору, методи розпізнавання об'ємних тіл та просторових сцен і побудови моделі навколишнього середовища, теорія кластеризації спостережень та індуктивного виведення, принципи самоорганізації нейронних мереж, принципи побудови людино-машинних інформаційних технологій оброблення зображень.

На базі отриманих фундаментальних результатів побудовані та впроваджені інтелектуальні інформаційні технології та системи для оброблення інформації в таких ділянках: розпізнавання, аналіз, синтез та розуміння усної мови; розпізнавання зображень, зокрема креслень, картографічних, текстурних, стерео-скопічних зображень; аерофотограмметричний аналіз земної поверхні для відновлення рельєфу місцевості за стереопарами аерофотознімків; аналіз та розуміння сцен для керування автономними роботами з автоматичним плануванням дій та прийняттям рішень; адаптивні системи прогнозування складних процесів в природі та суспільстві.

В Україні були створені перші системи автоматичного розпізнавання окремо вимовлюваних слів, системи типу "око-рука", системи автоматичного розпізнавання, розуміння та синтезу усної мови та системи усного діалогу, системи автоматичного розпізнавання креслярсько-графічних зображень, системи технічного зору для контролю якості виробів та керування виконавчими механізмами, перші в світі експериментальна система автоматичного розпізнавання зв'язного мовлення та багатомовна система усного діалогу для ЮНЕСКО, що оперувала сімома мовами та виконувала усний переклад з однієї мови на іншу. За даними японської фірми WACOM система оброблення креслярсько-гра-

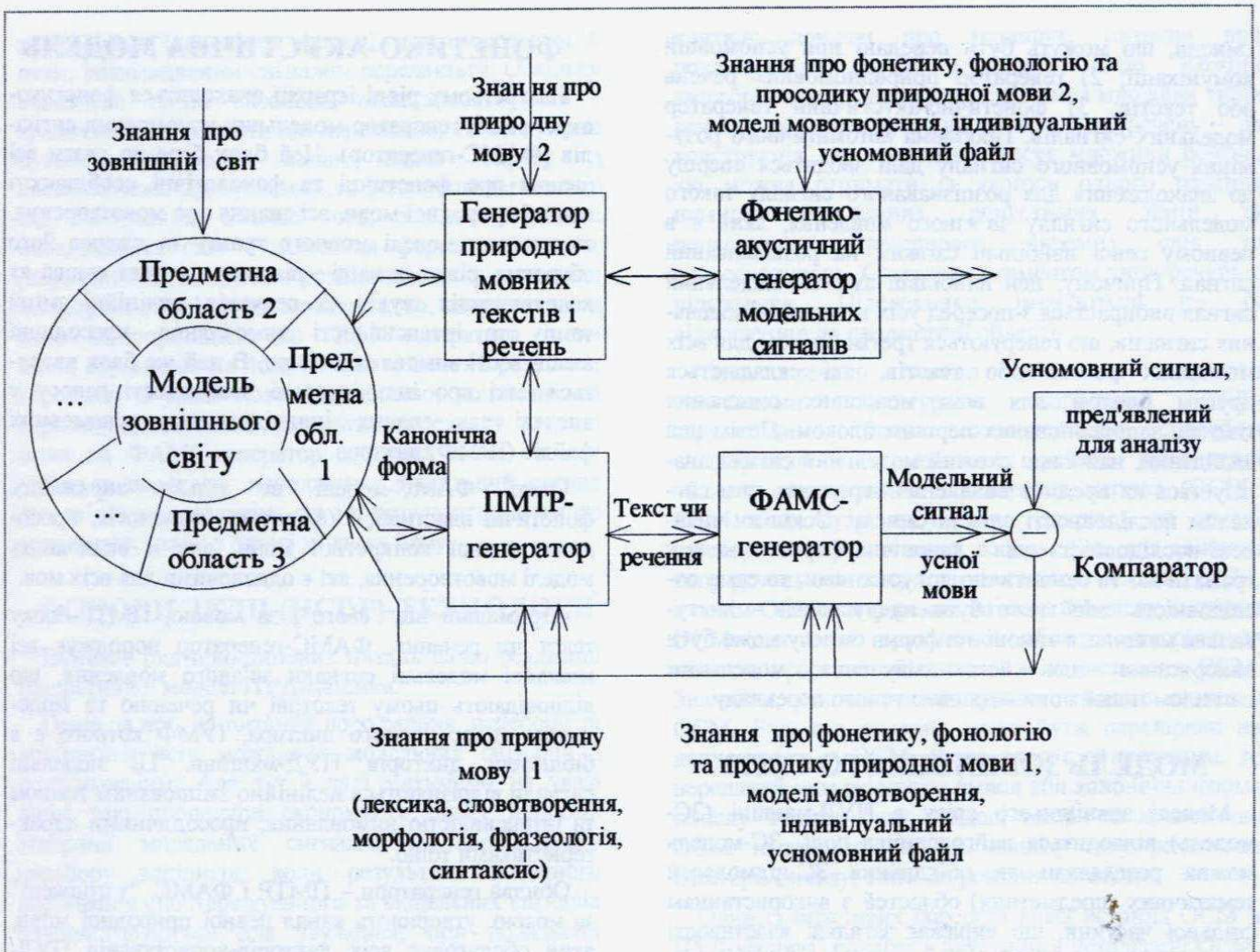


Рис. 1. Структура диктувальної машини та машини для усного перекладу.

фічних документів є найефективнішою з-посеред систем такого ж призначення. Розроблено програмне забезпечення першої вітчизняної, конкурентноздатної на зовнішньому ринку, цифрової фотограмметричної станції "Дельта".

В Україні сформувалася визнана у світі наукова школа, що обіймає широке коло наукових та прикладних проблем створення інтелектуальних інформаційних технологій та систем. Ці роботи є загальновизнаним у світі внеском у становлення та розвиток автоматичного розпізнавання образів (сигналів і зображень) як розділу науки. Українські роботи широко цитуються, визнані піонерними та такими, що мають значний вплив на світовий рівень досліджень у цій галузі і зміцнюють авторитет української науки.

Науковці України беруть участь у виконанні міжнародних контрактів з автоматичного розпізнавання сигналів і зображень, запрошуються до читання лекцій в західних університетах. На жаль, частина вчених виїхала на наукову роботу за кордон.

До сьогодні розробки з автоматичного розпізнавання, розуміння та синтезу образів велись розрізнено: кожна група науковців зосереджувалась на окремих вузьких проблемах. Тепер же на часі є об'єднання науковців навколо розроблення методології, концепції та архітектури ОК, навколо розроблення образних операційних систем та систем програмування ОК, окремих підсистем ОК, створення прототипів ОК.

ПІДХОДИ ДО СТВОРЕННЯ ОК

Розглянемо два підходи до створення образного комп'ютера на прикладі Диктувальної машини та машини Усного Перекладу (ПУД-машини). А потім покажемо, як моделі ПУД-машини можуть бути доповнені каналами сприйняття іншої, неусномовної, інформації та перетворені в прототипи ОК.

Перший підхід ґрунтується на так званій генеративній моделі автоматичного розпізнавання, розуміння та синтезу мовних сигналів.

Другий підхід використовує багатозначну багаторівневу модель.

ГЕНЕРАТИВНА МОДЕЛЬ ПУД-МАШИНИ

Узагальнена структура ПУД-машини, що ґрунтується на генеративній моделі, подана на Рис. 1.

Трьома найголовнішими блоками ПУД-машини є: 1) модель зовнішнього світу, яка описує всі можливі

смисли, що можуть бути передані при усномовній комунікації; 2) генератор природномовних речень або текстів; 3) фонетично-акустичний генератор модельних сигналів. Проблема автоматичного розуміння усномовного сигналу далі зводиться спершу до знаходження для розпізнаваного сигналу такого модельного сигналу зв'язного мовлення, який є в певному сенсі найбільш схожим на розпізнаваний сигнал. Причому, цей найбільш схожий модельний сигнал вибирається з-посеред усіх можливих модельних сигналів, що генеруються третім блоком для всіх можливих речень або текстів, які складаються другим блоком для всіх можливих смислових гіпотез, запропонованих першим блоком. Потім цей знайдений, найбільш схожий модельний сигнал аналізується на предмет виявлення переданих цим сигналом послідовності слів та смислу. Оскільки виявлені послідовність слів і канонічна форма смислу є граматично та семантично допустимими, то сама послідовність слів може бути надрукована – диктувальна машина, а канонічна форма смислу може бути використана для його озвучення модельним сигналом іншої мови – машина усного перекладу.

МОДЕЛЬ ЗОВНІШНЬОГО СВІТУ

Моделі зовнішнього світу в ПУД-машині (ЗС-модель) відводиться найголовніша роль. ЗС-модель можна розглядати як об'єднання ЗС-підмоделей тематичних (предметних) областей з використанням спільної частини, що виражає загальні властивості зовнішнього світу. ЗС-модель є спільною для всіх природних та штучних мов і, власне, мало від них залежить. Описується ЗС-модель спеціальною математичною мовою, наприклад мовою канонічних форм. Опускаючи тут подробиці задання ЗС-моделі, підкреслимо лише, що ЗС-модель генерує канонічні форми смислів, які можуть передаватись в процесі усномовної комунікації. Очевидно, що для кожної предметної області може бути вказана скінченна множина канонічних форм, можливих при діалозі. Наприклад, якщо йдеться про усномовний калькулятор на чотири арифметичні дії, то в цьому випадку допустимі канонічні форми мають просту структуру: операція та перший і другий операнди.

МОДЕЛЬ ПРИРОДНОЇ МОВИ

Наступною, другою в ієрархії частиною ПУД-машини є генератор природномовних речень або текстів (ПМТР-генератор). Цей блок грунтується на знаннях про лексику, словотворення, морфологію, фразеологію та синтаксис кожної конкретної природної мови. Отримуючи від ЗС-моделі канонічну форму передаваного смислу, ПМТР-генератор породжує всі допустимі для даної мови речення і тексти, котрі виражають один і той же смисл, що визначений принятою канонічною формою. Можна сказати, що ПМТР-генератор є семантико-граматичною текстовою моделлю конкретної природної мови.

ФОНЕТИКО-АКУСТИЧНА МОДЕЛЬ

На третьому рівні ієрархії знаходиться фонетико-акустичний генератор модельних усномовних сигналів (ФАМС-генератор). Цей блок бере до уваги всі знання про фонетичні та фонологічні особливості певної природної мови, всі знання про мовотворення. включаючи моделі мовного тракту та джерел його збурення, рівно ж наші уявлення про такі явища як коартикуляція звуків, їх редукція, нелінійні зміни темпу та інтенсивності вимовляння, просодичні властивості вимовляння тощо. В цей же блок вводяться дані про індивідуальні особливості голосу у вигляді так званих індивідуальних усномовних файлів (ІУМФ) диктора.

Отже, ФАМС-моделі не тільки виражають фонетичні властивості (фонетика, фонологія, просодика) певної конкретної мови, але й включають моделі мовотворення, які є однаковими для всіх мов.

Отримавши від "свого", за мовою, ПМТР-блоку текст чи речення, ФАМС-генератор породжує всі можливі модельні сигнали зв'язного мовлення, що відповідають цьому текстові чи реченню та моделюють голос певного диктора, ІУМФ котрого є в бібліотеці дикторів ПУД-машини. Ці модельні сигнали відрізняються нелінійно змінюваним темпом та інтенсивністю вимовляння, просодичними характеристиками тощо.

Обидва генератори – ПМТР і ФАМС, "узгоджені" за мовою, утворюють канал певної природної мови, який обслуговує всіх дикторів-користувачів ПУД-машини.

ІДЕЯ ВИРІШЕННЯ ПРОБЛЕМИ

Далі в рамках генеративної моделі автоматичного розпізнавання, розуміння та синтезу мовлення проблема автоматичного усного перекладу з однієї мови на іншу або проблема автоматичного редагування та друкування під диктування формулюється та вирішується так.

Одночасно з подачею пред'явленого до аналізу усномовного сигналу за допомогою генеративної моделі породжуються всі можливі модельні сигнали, які порівнюються з розпізнаваним сигналом. Ці модельні сигнали, при всьому їх розмаїтті, відповідають певним текстам та реченням, які, в свою чергу, як це випливає з опису ПУД-машини, виражають певний смисл, що заданий канонічною формою. Перебираючи всі можливі сигнали зв'язного мовлення, а, отже, всі допустимі тексти чи речення та всі можливі канонічні форми передаваних смислів, знаходимо такий модельний сигнал, який з-посеред усіх можливих модельних сигналів є найбільш схожим, у певному розумінні, на розпізнаваний сигнал.

Далі, аналізуючи цей найбільш схожий модельний сигнал, визначаємо, якій послідовності слів, тобто якому реченню або тексту, він відповідає, а заодно,

який смисл (канонічна форма) цим модельним і, отже, розпізнаваним сигналом передається. Оскільки визначена таким способом послідовність слів є граматично та семантично коректною, то вона може бути надрукована, якщо маємо справу з автоматичним друкуванням та редагуванням текстів під диктування. Більше того, оскільки знайдені послідовність слів і канонічна форма виражають результат автоматичного розпізнавання та смислової інтерпретації зв'язного мовлення, то, звертаючись з цією інтерпретацією до ПМТР-генератора іншої мови, отримаємо на його виході текстовий результат перекладу з усної мови. Аналогічно, подавши цей отриманий текстовий результат перекладу з усної мови на ФАМС-генератор для тієї мови, на яку робиться переклад, породжуємо модельний сигнал для зв'язного мовлення, що і є результатом повного усного перекладу з однієї мови на іншу.

ОСНОВИ ІКДП (HCDP)-ТЕХНОЛОГІЇ

Виникає ряд принципових питань щодо реалізації генеративної моделі ПУД-машини.

Перш за все, це питання породження, перебору та порівняння всіх можливих модельних сигналів з розпізнаваним. Але тут і далі матимемо на увазі лише такі процедури економного опису процесів генерації модельних сигналів та направленого перебору варіантів, коли результати поточного порівняння спостережуваного та модельних сигналів використовуються для звуження, при подальшій генерації, підмножини модельних сигналів, відповідно текстів і канонічних форм, що претендують у відповідь розпізнавання, так, щоб гарантувати не втратити оптимальне рішення. Саме ці згадані дві суперечливі вимоги: економна специфікація породжувальних процедур та направлений перебір варіантів, – задовольняються ІКДП(HCDP)–технологією автоматичного розпізнавання, розуміння та синтезу усномовних сигналів [1–6]. Перша вимога задовольняється використанням ієрархічно (І (Н – *Hierarchy*)) структурованих автоматних породжувальних граматик, які синтезують (композиція (К (С – *Composition*))) модельні сигнали, а друга – заснована на динамічному програмуванні (ДП (DP – *Dynamic Programming*)).

На Рис. 1 зображена не тільки ієрархія ПМТР- і ФАМС-генераторів, але й ідея направленого перебору варіантів – зворотніми зв'язками, що йдуть від канального компаратора до всіх рівнів (блоків) ПУД-машини.

ПРИКЛАД РЕАЛІЗАЦІЇ

Розіб'ємо всі речення предметної області за передаваним смислом на категорії або типи смислу. Так, для інформаційно-довідникової служби аеропорту можна виділити такі категорії смислу: питання, які стосуються прибуття літаків; питання стосовно відльоту літаків; довідки про наявність квитків; довідки про маршрут; питання про розташування служб аеропорту тощо. Кожній категорії смислу відповідає скінченна множина типів речень, які передають цей смисл. Тип речень – це конструкція, що економно описує множини речень, які можна отримати із якогось одного речення шляхом незалежних допустимих замін та допустимих перестановок окремих слів та словосполучень. Основним елементом типу речень є підсловник. Підсловники іменуються по їх відношенню до предметної області.

Типи речень можна економно задавати списковими структурами, наприклад, *LISP*-структурами. Іншим еквівалентним способом специфікації всіх можливих речень предметної області є орієнтована семантична мережа (ОСМ). ОСМ може бути побудована, зокрема, на основі категорій смислу та типів речень. Так, для кожного типу речень спочатку будується своя часткова ОСМ, а потім ОСМ для окремих категорій смислу та повна ОСМ для всієї предметної області утворюється шляхом простого об'єднання всіх часткових ОСМ. Звичайно, важливо мінімізувати кількість станів в ОСМ. Будь-яке речення може бути перевірене на допустимість в ОСМ. Якщо воно є допустимим, то переданий цим реченням смисл або канонічна форма смислу можуть бути найдені також за допомогою ОСМ, наприклад, шляхом аналізу типу речень та категорії смислу, яким це речення належить.

Один із можливих способів використання ОСМ в ПУД-машині полягає в акустичній деталізації ОСМ. Для цього кожне слово-місце в ОСМ заміщується його графом (як в *HCDP*- чи *HMM*-технології). Тоді цей граф генеруватиме всі можливі сигнали цього слова в контексті злитого мовлення, а сама ОСМ стає усномовною, тобто такою, яка генерує модельні усномовні сигнали для вибраної предметної області. Тепер вирішення проблеми ПУД-машини для цієї предметної області полягає у знаходженні найбільш схожого модельного сигналу та у його аналізі на предмет виявлення послідовності слів та канонічної форми смислу, переданого пред'явленим усномовним сигналом [1-3].

ГЕНЕРАТИВНА МОДЕЛЬ ОК

Доповнивши генеративну модель ПУД-машини (див. Рис. 1), зокрема, приєднавши до моделі зовнішнього світу канал генерації зображень та просторових сцен, що складається з композитора зображень і сцен та генератора модельних зображень і сцен, отримаємо модель образного комп'ютера, який не тільки є диктувальною машиною та машиною усного перекладу, але й сприймає зображення та аналізує просторові сцени, дає їм усномовну інтерпретацію та робить їх текстовий опис, будує зображення за усномовним повідомленням.

БАГАТОРІВНЕВА БАГАТОЗНАЧНА МОДЕЛЬ

Другий підхід до створення ПУД-машини – це багаторівнева багатозначна модель. Вона також є ієрархічно організованою. Тут використовуються багатозначні рішення на всіх рівнях оброблення усномовного сигналу. Наприклад, на першому, нижньому, рівні знаходять $N1>>1$ найкращих розбивок усномовного сигналу на сегменти, що відповідають фонемам. Далі, на другому рівні, ця багатозначна сегментація в умовах, наприклад, вільного порядку слідування фонем, відображається в $N2>>1$ найкращих послідовностей фонем. Отже, на другому рівні ієрархії маємо справу з багатозначним фонемним розпізнавачем. На третьому рівні ієрархії розташований розпізнавач слів, який в умовах вільного порядку слідування слів знаходить $N3>>1$ найкращих послідовностей слів. Це – так званий узагальнений розпізнавач послідовностей слів. І так далі. На наступних рівнях ієрархії виконується багатозначна синтаксична та семантична фільтрація послідовностей слів. Нарешті, на найвищому рівні ієрархії знаходять $Nr>>1$ найкращих результатів смислової інтерпретації усномовного сигналу та приймають остаточне рішення.

В багаторівневій багатозначній моделі зворотний зв'язок явно не показаний. В ній замість повного перебору варіантів використовується "обережна" стратегія прийняття багатозначних рішень, яка має на меті не "загубити" і врешті-решт віднайти найкраще рішення.

Структура багаторівневої багатозначної ПУД-машини показана на Рис. 2.

Стосовно багаторівневого багатозначного оброблення зображень і сцен — то діємо аналогічно. Наприклад, на першому, нижньому, рівні знаходимо $N1>>1$ найкращих розбивок зображення на сегменти, що відповідають дугам, відрізкам, дрібним елементам. Далі, на другому рівні, ця багатозначна сегментація відображається в $N2>>1$ найкращих композицій з більш крупних елементів (літер, окремих деталей, дрібних предметів, елементів фону тощо). На третьому рівні ієрархії знаходить $N3>>1$ найкращих сегментацій зображення, що віповідають можливим композиціям з більш крупних елементів (надписів, предметів, людей). На вищих рівнях ієрархії виконується опис та смисловий аналіз зображення і сцени в цілому.

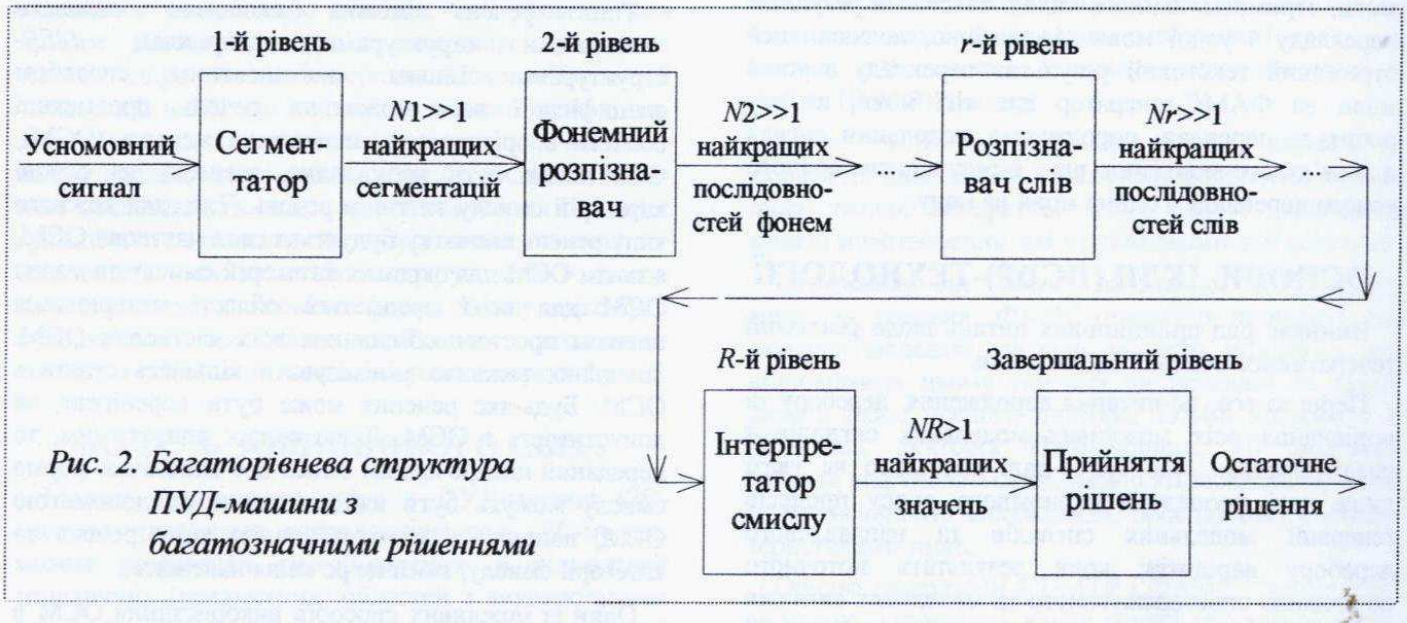


Рис. 2. Багаторівнева структура ПУД-машини з багатозначними рішеннями

ПРИКІНЦЕВІ ПОЛОЖЕННЯ

Щоб взяти участь в розробках за Державною науково-технічною програмою "Образний комп'ютер", необхідно виграти тендерні змагання (конкурс запитів та пропозицій). Ці змагання проводитиме Міжнародний науково-навчальний центр ЮНЕСКО інформаційних технологій та систем, на який покладено функції органу управління програмою.

ЛІТЕРАТУРА

[1] T.K. Vintsiuk, "Element-Wise Recognition of Continuous Speech Composed of Words from Specified Vocabulary", *Kibernetika*, 1971, No. 2, pp 133-143.

[2] T.K. Vintsiuk, "Speech Recognition and Understanding", *Kibernetika*, 1982, No 5, pp 101-111.

[3] Т.К. Винцюк. "Анализ, распознавание и интерпретация речевых сигналов", Киев: "Наукова думка", 1987, 264 с.

[4] T.K. Vintsiuk, "HCDP-Technique for Automatic Analysis, Recognition and Understanding of Speech Signals", *Proc. First Intern. Conf. on Information Technology for Image Analysis and Pattern Recognition*, L'viv, 1990, Vol 1, pp 108 - 112.

[5] Taras K.Vintsiuk, "Two Approaches to Create a Dictation/ Translation Machine", *Proc. Second Intern. Workshop "Speech and Computer"*, Cluj-Napoca, 1997, pp 1-6.

[6] Тарас Вінцюк, "Образний комп'ютер", *Зб. наук. праць "Сучасні проблеми в комп'ютерних науках"*, Вид-во Нац. ун-ту "Львівська політехніка", Львів, 2000, сс 5-14

Multidimensional Nonlinear Registration: a Stochastic Formulation

Oleh J. Tretiak and Maria Gabrani
Imaging and Computer Vision Center, Drexel University, Philadelphia PA 19104, USA
tretiak@coe.drexel.edu (215) 895 2214

ABSTRACT

Many image processing applications require the comparison of images or multidimensional data sets. While rigid body transformations are in common use, applications in medicine and biology where images from different individuals are compared require nonlinear transformations. A number of investigators have explored nonlinear or 'rubber sheet' transformations, which have great expressive power. We develop a random field model for such transformations. Each transformation is a sample function from a random process. Maximum likelihood estimation of the transformation has a direct relation to radial interpolation theory. The stochastic process formulation provides a model that facilitates prediction of alignment accuracy. The theory is illustrated with application to three-dimensional alignment of brain images that arise in neuroscience research. The proposed transformations lead to algorithms that provide better accuracy.

1. INTRODUCTION

In an alignment task we are given two "images" $f_1(\mathbf{x})$, $f_2(\mathbf{x})$ defined over d-dimensional space. We are required to find a space transformation $T(\mathbf{x})$ and to compute $f_1'(\mathbf{x}) = f_1(T(\mathbf{x}))$ in which homologous points in $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ are aligned. In biomedical applications where volumetric studies ($d = 3$) from different individuals are aligned rigid-body transformations do not provide sufficient accuracy. Most of the work to date has been empirical: algorithms are proposed, and their efficacy is evaluated through empirical trials. In this research we propose a random process model for the deformations that are required to align biological objects. We feel that a suitable model can suggest better alignment algorithms, and predict the accuracy with which such alignment is possible. Our model has been motivated by a new class of nonlinear transformations, called elastic transformations [1]. The formulation is based on multidimensional spline theory developed by Duchon [2] and further developed by Meinguet [3]. Elastic transformations have a closed-form interpolation formula. The formulation follows the elastic deformation methodology introduced by Bajcsy et al. [4, 5], and further studied by Miller and co-workers [6], and Kass and his colleagues [7]. The above authors use correlation of gray scale data to drive their registration. Our research is based on the use of geometric features extracted from the images, and is motivated by the interpolation theory introduced to this field by Bookstein [8]. The effectiveness of an interpolation technique is measured by the accuracy with which it predicts the values of the function at ordinates not used to compute the function. In the case of brain mapping, we measure how well inner structures of an atlas or reference image are mapped to the corresponding locations in a test image. To measure this accuracy we align a brain and its known second order polynomial transformation.

The paper is organized as follows. In the next section, we describe a model for the proposed transformations, and relate it to our elastic transformation algorithm. In the third section, we report results of alignment trials. In the last section, we discuss the relation between the results of these trials and the proposed model.

2. RANDOM TRANSFORMATIONS

We assume that the transformation between the two data sets has the following form:

$$T(\mathbf{x}) = \mathbf{x} + \mathbf{u}(\mathbf{x}) \tag{1}$$

where $\mathbf{u}(\mathbf{x}) = \nabla\phi(\mathbf{x})$. This definition is motivated by elasticity theory. The function $\phi(\mathbf{x})$ is a sample function of a Gaussian random process with density function given by

$$f(\phi(\bullet)) = k\exp(-\int |[w*(\phi-\mu)](\mathbf{x})|^2\, d\mathbf{x}) \tag{2}$$

where $w*$ represents the application of an operator described below (for example, $w*$ may be equal to Δ, the Laplacian operator), and $\mu(\mathbf{x})$ is the mean of the random process. At present, we assume that $\mu(\mathbf{x}) = 0$. We assume that we are given two point sets S_1 and S_2 in the two images that are to be matched, so that T must satisfy

$$\mathbf{x} \in S_1 \Rightarrow T(\mathbf{x}) \in S_2 . \tag{3}$$

For computational purposes, we approximate the above integral in the following way: we choose a set of N points $\mathbf{A} = (\mathbf{a}_1, \cdots \mathbf{a}_N)$, $\mathbf{a}_i \in S_1$ and look for points $\mathbf{B} = (\mathbf{b}_1, \cdots \mathbf{b}_N)$, $\mathbf{b}_i \in S_2$. Function ϕ must now satisfy the interpolation constraints

$$\mathbf{a}_i + (\nabla\phi)(\mathbf{a}_i) = \mathbf{b}_i, \quad i = 1, \cdots, N . \tag{4}$$

and a function that satisfies (4) will be denoted as $\phi(\mathbf{x}; \mathbf{A}, \mathbf{B})$, and the likelihood of this function is $f(\phi(\mathbf{x}; \mathbf{A}, \mathbf{B}))$. We now find ϕ and $\mathbf{B}$ by the principle of maximum likelihood. For notational convenience, we

© UkrObraz'2000

choose to work with $Q = -\ln(f)$, so that we need to solve the formal problem

$$\min_{\phi,\mathbf{B}} Q(\phi(\bullet;\mathbf{A},\mathbf{B})) = \min_{\phi,\mathbf{B}} \int |[w * \phi](\mathbf{x};\mathbf{A},\mathbf{B})|^2 \, d\mathbf{x}. \quad (5)$$

We seek a minimum over ϕ in the space of temperate distributions.

To examine some of the implications of this formulation, we review some results presented in [1]. We write the integral in the Fourier domain,

$$Q(\phi(\bullet;\mathbf{A},\mathbf{B})) = \int |\hat{w}(\xi)\hat{\phi}(\xi;\mathbf{A},\mathbf{B})|^2 \, d\xi, \quad (6)$$

where $\hat{w}$, $\hat{\phi}$ are, respectively, the Fourier transforms of w and ϕ. We assume that w is radially symmetric, so that $\hat{w}(\xi) = g(\|\xi\|)$, and that for $\rho = \|\xi\| > 0$ $g(\rho) > 0$,

$$\lim_{\rho\to 0} g(\rho) = \rho^m, \quad \lim_{\rho\to\infty} g(\rho) = \rho^k. \quad (7)$$

The functions $\phi(\bullet)$ for which the Q is zero (the kernel of Q) is the linear space of polynomials of degree less than m. A minimum over ϕ of (6) subject to the interpolation constraints (5) is achieved (in the space of temperate distributions) when $k > d/2 + 1$ and when the point set $\mathbf{A}$ contains a unisolvent subset for polynomials of degree $m-1$. The solution can be constructed as follows: let

$$p_j(\mathbf{x}), \; j = 1,\cdots,M, \; M = C(d+m-1,d) \quad (8)$$

be a set of linearly independent polynomials of degree $m-1$ in d variables. The function $\mathbf{u}(\mathbf{x})$ is given by

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^{N} \mathbf{H}(\mathbf{x}-\mathbf{a}_i)\bullet\boldsymbol{\gamma}_i + \sum_{j=1}^{M} d_j \nabla p_j(\mathbf{x}), \quad (9)$$

where $\mathbf{H}(\bullet)$ is the tensor-valued function

$$H_{kl}(\mathbf{x}) = \frac{\partial^2 E(\mathbf{x})}{\partial x_k \partial x_l}, \quad (10)$$

$\boldsymbol{\gamma}_i$ are d-dimensional vectors, and $E(\mathbf{x})$ satisfies the equation

$$\hat{E}(\xi) = \frac{1}{\hat{w}^2(\xi)}. \quad (11)$$

Coefficients $\boldsymbol{\gamma}_i$, d_j are found from the interpolation constraints (4)

$$\mathbf{b}_k = \mathbf{a}_k + \sum_{i=1}^{N} \mathbf{H}(\mathbf{a}_k - \mathbf{a}_i)\bullet\boldsymbol{\gamma}_i + \sum_{j=1}^{M} d_j \nabla p_j(\mathbf{a}_k), \quad (12)$$

$k = 1,\cdots,N$ and the *orthogonality* constraints

$$\sum_{i=1}^{N} \nabla p_j(\mathbf{a}_i)\bullet\boldsymbol{\gamma}_i = 0, \; j = 1,\cdots,M. \quad (13)$$

To summarize, the solution of the minimization of Q over ϕ reduces to the solution of a set of linear equations. Define

$$Q(\mathbf{A},\mathbf{B}) = \min_{\phi} Q(\phi(\bullet;\mathbf{A},\mathbf{B})). \quad (14)$$

$Q(\mathbf{A}, \mathbf{B})$ is a quadratic function in the variables $(\mathbf{b}_i)$. The maximum likelihood solution to the estimation problem is found by locating

$$\min_{\mathbf{B}\in S_2} Q(\mathbf{A},\mathbf{B}).$$

We now return to the probabilistic formulation of the problem. From a formal perspective, we would like to ask whether the problem is well-posed, i. e. under what conditions does the solution of (5) exist? The structure of the density function (2) implies that the random process $(\phi - \mu)(\mathbf{x})$ is stationary, and its power spectrum is

$$S(\xi) = |\hat{w}(\xi)|^{-2}. \quad (15)$$

This spectrum is, typically, singular. For example, when $d = 1$ and $w = D$ (the differential operator) the process generating functions ϕ is a random walk, and the variance of ϕ is infinite. However, certain linear combinations of the random variables have finite variance. Specifically, if we define the random variable

$$\mathbf{z} = \sum_{k=1}^{P} c_k \mathbf{u}(\mathbf{a}_k), \quad (16)$$

and if the coefficients c_k satisfy the orthogonality constraints

$$\sum_{k=1}^{P} c_k p_j(\mathbf{a}_k) = 0, \; j = 1,\cdots,M, \quad (17)$$

then the variance of $\mathbf{z}$ is finite, and is furthermore given by the equation

$$V(\mathbf{z}) = \sum_{k=1}^{P}\sum_{l=1}^{P} c_k c_l E(\mathbf{a}_k - \mathbf{a}_l), \; \hat{E} = S. \quad (17)$$

We see that this framework is well adapted for interpolation problems.

EXPERIMENTS

We examine the performance of the proposed transformation method in the alignment of rat brains. Our research deals with the development of tools for neuroscience studies. Inbred laboratory rats are a standard model for neurochemistry research, since the low genetic variability allow easy experimental designs in which treatment effects, such as the response of the brain to neuroactive drugs, can be separated from intra-individual variability. The large amount of published data for this model also make it attractive for further experimental studies. Autoradiographic measurement of metabolic activity, neurotransmitter distributions, and protein synthesis with in-situ hybridization of mRNA produce voluminous data in the form of sectional images, and it is desirable to transform these images into a standard coordinate system for inter-animal comparisons.

We have developed a technique that aligns rat brains by using rigid-body [9] and affine [10] transformations based on distance between brain surfaces. The surface of the brain is a reliable feature which can be extracted from the images produced by most labeling methods. The accuracy of the technique is tested by comparing how well internal structures are aligned. We found [10] that the rigid-body and affine transformations produced a median error of about 168 μm in the alignment of brain surfaces, and an error of

about voxels 88 μm for various internal structures. In these experiments, the locations of these internal structures were not known to the alignment algorithms. ALIGN, a MATLAB implementation of this algorithms, is available thorough the web (http://coe.drexel.edu/ICVC). Though these results are encouraging, they were obtained with very precise techniques for preparing the rat brains. In usual practice more variability is encountered.

To test the effectiveness of surface-guided alignment we used a brain dataset from the above study, and distorted it with a second-order polynomial transformation. The ALIGN program was used to register the two brains with rigid body and affine transformations. The accuracy of these methods were compared with the results of an elastic transformation. In all cases, the surfaces of the brains provided the data for alignment. Note that the rigid-body and affine transformations are not interpolatory, they can not achieve perfect alignment of the brain surface. We measured the accuracy with which all algorithms registered the brain surface, an internal structure, as well as the accuracy with which the control points ($\mathbf{a}_i$ in the above discussion) were aligned with their known images. The distance between sections in the data set was 100 μm, and the thickness of each section is approximately 40 μm.

The outer surface of the rat brain is shown in Fig. 1 (a), the transformed brain surface is in Fig. 1 (b), and the two datasets are superimposed in Figure 1(c). The outcomes of the rigid body, the affine and the elastic transformations are given in Figure 1 (d), (e) and (f), respectively. In all cases $N = 40$ control points were used. It is evident that the best alignment is produced by the elastic transformation.

Numerical evidence of the performance is given in Table 1, where we provide the median Euclidean distance error in pixels between the target image and the outcomes of the transformations. In addition to the distance for the outer surface, we compute the median distances for the following internal structures: the third and fourth ventricles (V34), middle cerebellar peduncle (MCP), lateral ventricle (LV), midline of the cerebrum (MC), seventh nerve (N7), trigeminal nerve (TG5), corpus callosum (CC), cerebellar tracks (CT), and superior colliculus (SC).

Table 1. Median distance error (pixels) between the target image and the outcomes of alignment algorithms. One pixel = 42 μm.

	Original	RB	AFF	ET
Outer Surface	30.23	19.5	3.4	2.2
V34	30.63	14.6	1.1	1
MCP	24.12	17.4	2.8	2.4
LV	26.9	14.0	1.6	1
MC	28.55	15.3	1.8	1.4
N7	29.18	16.8	4.2	2.4
TG5	29.92	20.0	4.9	2.4
CC	28.98	16.6	1.5	1
CT	32.52	10.7	3.8	1
SC	34.35	18.9	1.7	1

Knowing the exact correspondence of the control points we are able to examine the accuracy with which the algorithm locates them. Table 2 gives the energy between the initial point configuration and the final control points of the alignment algorithms. The known correspondence points give an energy Q of 0.0309. The estimation algorithm gives control points with energy 0.0357, which is very close to the minimum. The rigid-body and affine transformations produce control points with higher energy. However, the main criterion of performance is not the energy but the outcome registration quality of the algorithms. In Table 3 we give the median and mean distance error of the outcome of the elastic transformations considering the known correspondence and that found by the estimation algorithm. Notice that the difference is very small. The median errors are almost all the same, and the mean errors differ mainly in the second decimal point. The above results establish the use of the proposed surface based interpolation technique.

Table 2. Energy defined from the distance between the points in the original image and the final control points of the alignment

Control Points	Energy
Known Correspondence	0.0309
Elastic Estimation	0.0357
Affine	0.0829
Rigid Body	1.0454

Table 3. Distance error (pixels) between the target image and the outcomes of elastic transformations with known correspondence and estimating control points locations.

	Known		Estimated	
Error	Median	Mean	Median	Mean
Outer Surface	2.2	1.6	2.2	1.6
V34	1	0.8	1	0.9
MCP	2.4	2.5	2.4	2.6
LV	1	0.8	1	0.8
MC	1.4	1.7	1.4	1.8
N7	2.2	2.6	2.4	2.7
TG5	2.2	2.2	2.4	2.3

4. DISCUSSION

We have developed a multidimensional alignment algorithm that is based on multidimensional interpolation theory [2, 3], and that can furthermore perform an effective transformation when only the surfaces of the objects are known (no known point correspondences). In this paper we show that the algorithm has a natural relationship to a random process formulation of multidimensional alignment. In the future we propose to further explore the implications of the random process model. For example, both the alignment algorithm and the random process make essential use of the interpolating function $E(\cdot)$ defined in (11) which depends on the smoothness operator w. In the interpolation theory this function is ad-hoc, while the random process theory indicates that this is the Fourier transform of the power spectrum. This suggests that the

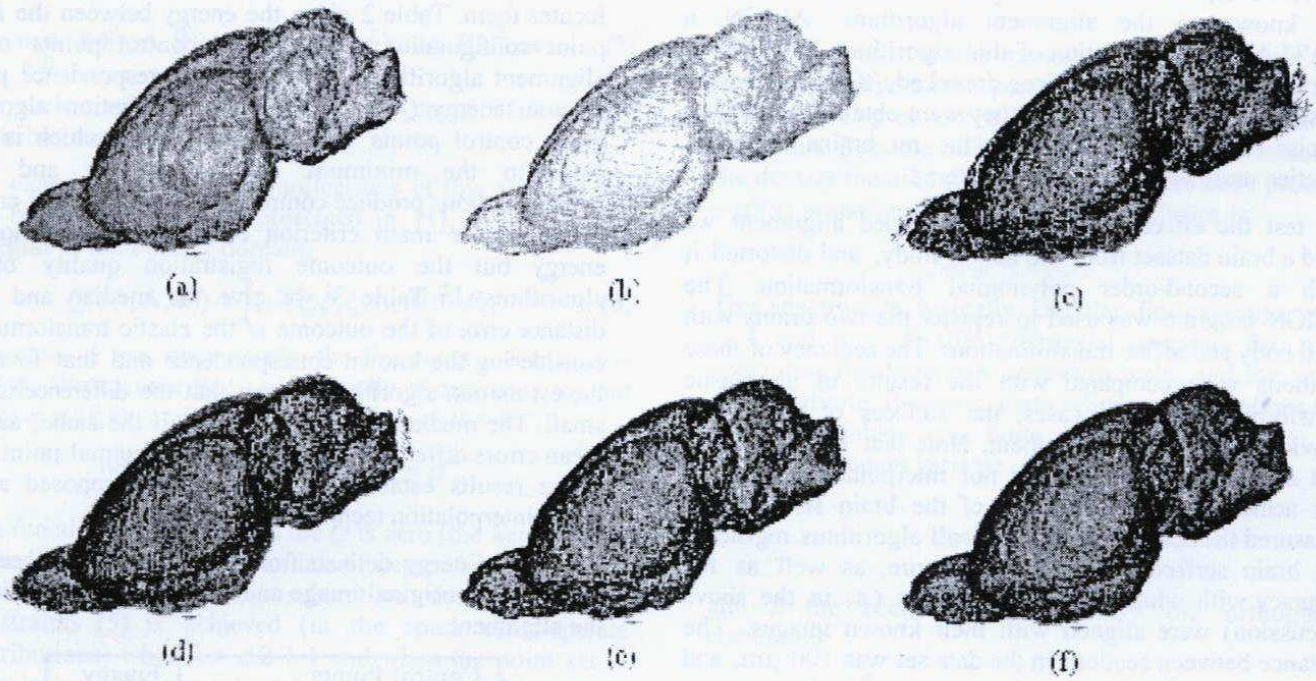


Figure 1. (a) Surface of a rat brain, (b) the same dataset, transformed with a quadratic polynomial, (c) for comparison, the two brains superimposed, (d) rigid body alignment, (e) affine, and (f) elastic transformation.

appropriate w can be found by performing experiments which measure the fluctuations in the interpolating transformation (found by some other method), and by deducing the form of E from these measurements. Another feature suggested by the random process model is inclusion of the random process mean $\mu(\bullet)$ in the likelihood function (2). The inclusion of this function may lead to improvements in the alignment algorithm. Finally, the random process model can lead to analytical predictions about the magnitude of alignment errors. This may be of considerable value in making optimal choices of the control points $\mathbf{a}_i$. We expect that the accuracy of alignment will depend on both $S(\bullet)$, the random process spectrum, and on **A**, the control point set. Choosing a larger control point set will improve accuracy of prediction of internal structure location until the limit of natural variability is reached. The proposed theory may allow us to identify this point of diminishing returns.

Acknowledgments. This work is supported by NIH Grant P41-RR01638. We would like to express our thanks to Dr. Jonathan Nissanov, Dr. Cengizhan Ozturk, and Ms. Louise Bertrand for their graceful assistance.

REFERENCES

[1] Gabrani, M.; and Tretiak, O. J., "Surface Based Matching using Elastic Transformations," *Pattern Recognition*, in press.

[2] Duchon, J., "Interpolation des fonctions de deux variables suivant le principle de la flexion des plaques minces," *RAIRO Analyse Numérique*, vol. 10, pp. 5-12, 1976.

[3] Meinguet, J., "Multivariate Interpolation at Arbitrary Points Made Simple," *Journal of Applied Mathematics (ZAMP)*, vol. 30, 1979.

[4] Broit, C., "Optimal registration of deformed images," Doctoral dissertation, Univ. Pennsylvania, 1981.

[5] Bajcsy, R.; and Kovacivc, S. "Multiresolution Elastic Matching." *Computer Vision, Graphics, and Image Processing*, vol. 46, pp. 1-21, 1989.

[6] Miller, M.; Christensen, G.; Amit, Y.; and Grenander, U., "Mathematical textbook of deformable neuroanatomies," *Proc. Natl. Acad. Sci. USA*, vol. 90, pp. 11944-11948, December 1993.

[7] Kass, M.; Witkin, A.; and Terzopoulos, D., "Snakes: Active contour models," *Int. J. Comput. Vision*, vol. 1, pp. 321-331, 1987.

[8] Bookstein, F.L., "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations," *IEEE trans. on PAMI*, vol. 11, no. 6, June 1989.

[9] Kozinska, D.; Tretiak, O. J.; Nissanov, J.; and Ozturk, C., "Multidimensional Alignment Using the Euclidean Distance Transform," Graphical Models and Image Processing, vol. 59, November, pp. 373-387, 1997.

[10] Ozturk, C., "Rat Brain Variability, Utility of its Surface for Guiding Alignment and Development of A Structured Light Based Brain Surface Scanner," Doctoral dissertation, Drexel Univ., 1997.

Separate-Group Covariance Estimation With Insufficient Data for Object Recognition

*Carlos Eduardo Thomaz[1], Raul Queiroz Feitosa[2], Álvaro Veiga[3]

[1,2,3]Catholic University of Rio de Janeiro
Department of Electrical Engineering
r. Marquês de São Vicente 225,22453-900, Rio de Janeiro, Brazil

[2]Department of Computer Engineering
University of Rio de Janeiro
r. São Francisco Xavier, 524, 20559-900, Rio de Janeiro, Brazil

{cethomaz,raul,alvf}@ele.puc-rio.br

Abstract. Many similarity measures used for classification involve the inverse of the group covariance matrices. However, the number of observations available in the training set for each group is, in many cases, significantly inferior to the dimension of the feature space, what implies that the sample covariance matrix is singular. A common solution to this problem is to assume the same covariance matrix for all groups using the pooled covariance matrix computed from the whole training set. This paper investigates an alternative estimate for the group covariance matrices, called Mixed Covariance, given by a linear combination of the sample group and pooled covariance matrices. This estimate has the same rank of the pooled covariance matrix without assuming equal covariance for all groups. Experiments were carried out to evaluate the performance associated with the proposed estimate in two automatic recognition applications: face and facial expression. The average recognition rates obtained by using the mixed covariance were higher than the usual sample group and pooled covariance estimates.

1. Introduction

Many similarity measures used for classification involve the inverse of the group covariance matrices. Since in practical cases these matrices are not known, estimates must be computed based on the patterns available in a training set. The usual choice for the estimate of the covariance matrices is the sample group covariance. However, the number of training examples for each group is, in many cases, significantly less than the dimension of the feature space. This implies that the sample covariance matrix will be singular.

A common solution to this problem is to assume that all populations have the same covariance matrix and to use the pooled covariance estimate computed from the whole training set. The resulting matrix will have the same rank as the data matrix of the training set.

This paper investigates a new estimate for the group covariance matrices, called mixed covariance, given by a linear combination of the sample group and the pooled covariance estimates. It has the property of having the same rank as the pooled estimate, while allowing a different estimate for each group, what may imply in a better modeling of the population involved in the problem.

In order to evaluate the proposed approach two pattern recognition applications were considered: automatic face recognition and automatic facial expression recognition. The evaluation used different image database for each application and two dimensionality reduction techniques were applied: Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). The mixed covariance estimate attained the best performance in the experiments for both applications.

2. Maximum Probability Classifiers

The basic problem in the decision-theoretic methods for pattern recognition consists of finding a set of g discriminant functions $d_1(\mathbf{x})$, $d_2(\mathbf{x})$, ..., $d_g(\mathbf{x})$, (g is the number of groups) with the property that if the n-dimensional pattern vector $\mathbf{x}$ belongs to the group π_i ($1 \le i \le g$) then $d_i(\mathbf{x}) \ge d_j(\mathbf{x})$, for all $i \ne j$, $1 \le i \le g$.

The Bayes classifier designed to maximize the total probability of correct classification, where equal prior probabilities for all groups are assumed, corresponds to a set of discriminant functions equal to the respective probability density functions, that is, $d_i(\mathbf{x}) = f_i(\mathbf{x})$ for all classes.

Many proposed pattern recognition systems assume that the population of all groups can be properly modeled by a multivariate normal distribution [13]. Its density function can be expressed by:

$$d_i(\mathbf{x}) = f_i(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\Sigma_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1}(\mathbf{x}-\mu_i)\right], \quad (1)$$

for $1 \le i \le g$, where μ_i and Σ_i are respectively the mean and the covariance of group π_i. Since those values are seldom available, estimates must be provided. This works

* C.E. Thomaz (cet@doc.ic.ac.uk) is currently starting a Ph.D. in Computer Science (Visual Information Processing) at Imperial College, London.

focus on the usual sample estimate for the mean and on three estimates for the covariance, as described below.

2.1. Sample Group Covariance Matrix

The most commonly used estimate for Σ_i is the Sample Group Covariance matrix defined by:

$$\mathbf{S}_i = \frac{1}{k_i - 1} \sum_i (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i)^T (\mathbf{x}_{ij} - \bar{\mathbf{x}}_i), \text{ for } 1 \le i \le g \quad (2)$$

where $\mathbf{x}_{ij}$ are the training examples of group π_i, k_i is the number of them, and $\bar{\mathbf{x}}_i$ is the corresponding sample mean. By using this estimate, equation (1) takes the form:

$$d_{sample\,i}(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\mathbf{S}_i|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_i)^T \mathbf{S}_i^{-1}(\mathbf{x} - \bar{\mathbf{x}}_i)\right], \quad (3)$$

The matrix $\mathbf{S}_i$ will be singular if k_i is less than the dimension of the feature space.

2.2. Pooled Covariance Matrix

One way to get around this problem is to assume that all groups have equal covariance matrices, and to use as its estimate the weighted average of each sample group covariance matrix, given by

$$\mathbf{S}_{pooled} = \frac{(k_1 - 1)\mathbf{S}_1 + (k_2 - 1)\mathbf{S}_2 + \cdots + (k_g - 1)\mathbf{S}_g}{k_1 + k_2 + \cdots + k_g - g}. \quad (4)$$

By introducing the equal covariance assumption in equation (1), and after some simplifications [13], the following set of discriminant functions can be derived:

$$d_{pooled\,i}(\mathbf{x}) = (\mathbf{x} - \bar{\mathbf{x}}_i)^T \mathbf{S}_{pooled}^{-1}(\mathbf{x} - \bar{\mathbf{x}}_i), \text{ for } 1 \le i \le g \quad (5)$$

Since more observations are taken to calculate $\mathbf{S}_{pooled}$, it will pottentially have a higher rank than $\mathbf{S}_i$ and will be eventually full rank. Although the pooled estimate does provide a solution for the algebric problem arising from the insufficient number of training patterns in each group, assuming equal covariance for all groups may bring about distortions in the modeling of the recognition problem.

2.3. Mixed Sample Covariance Matrix

The Mixed Covariance Matrix is a tradeoff between $\mathbf{S}_{pooled}$ and $\mathbf{S}_i$. It is given by

$$\mathbf{S}_{mix\,i} = a\mathbf{S}_{pooled} + (1 - a)\mathbf{S}_i, \text{ where } 0 < a < 1. \quad (6)$$

Figure 1 gives the geometric interpretation of the proposed Mixed Sample Covariance Matrix. The ellipsoids correspond to the contour of the constant density for three groups. The dashed gray lines represent the different sample group covariance estimates, while the pooled estimate is represented in dotted gray lines. The proposed mixed sample estimates assume that the ellipsoid corresponding to the true covariance is placed somewhere between both ellipsoids, as shown by the solid black lines.

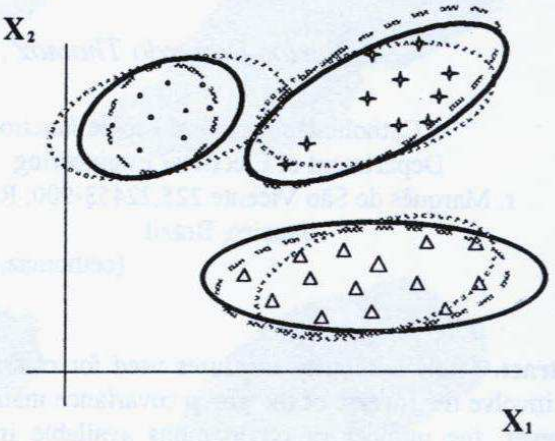


Figure 1: Geometric interpretation of $\mathbf{S}_i$, $\mathbf{S}_{pooled}$ and $\mathbf{S}_{mix}$.

Each $\mathbf{S}_{mix\,i}$ matrix has the important property of admitting an inverse if $\mathbf{S}_{pooled}$ does. Let the dimension of the feature space n be such that $\mathbf{S}_{pooled}$ is full rank (invertible) and $\mathbf{S}_i$ is not full rank (non invertible). Thus $\mathbf{S}_{pooled}$ and $\mathbf{S}_i$ are respectively positive definite and positive semi-definite matrices. Since $0<a<1$, a and $(1-a)$ are both positive numbers. Therefore, $a\,\mathbf{S}_{pooled}$ and $(1-a)\mathbf{S}_i$ are still positive definite and positive semi-definite matrices. For a matrix **A** positive definite and a matrix **B** positive semi-definite the next inequality is valid [5]

$$\det(\mathbf{A} + \mathbf{B}) \ge \det(\mathbf{A}) \quad (7)$$

Hence,

$$\det(\mathbf{S}_{mix\,i}) = \det[a\mathbf{S}_{pooled} + (1 - a)\mathbf{S}_i] \ge a\det(\mathbf{S}_{pooled}) > 0 \quad (8)$$

Since $a > 0$, this implies that $\mathbf{S}_{mix}$ will be non singular. By using $\mathbf{S}_{mix}$ in the place of the group covariance matrix, equation (1) takes the form:

$$d_{mix\,i}(\mathbf{x}) = \frac{1}{(2\pi)^{p/2}|\mathbf{S}_{mix\,i}|^{1/2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \bar{\mathbf{x}}_i)^T \mathbf{S}_{mix\,i}^{-1}(\mathbf{x} - \bar{\mathbf{x}}_i)\right]. \quad (9)$$

3. Projection spaces in recognition

One of the most successful approaches to the problem of creating a low dimensional image representation is based on Principal Component Analysis (PCA). It was firstly proposed by Sirovich and Kirby [7] for representing face images. After the classical work of Turk and Pentland [8], many extensions were proposed [1,2,10,11,12,15]. Swets and Weng called the selected eigenfaces the Most Expressive Features (MEF) [3], since they give the minimum mean square reconstruction error and describe the major variations in the set of samples.

While the goal of PCA is to minimize the reconstruction error, the Linear Discriminant Analysis (LDA) provides a procedure to determine a set of axes whose projections of different groups have the maximum separa-

tion. It can be shown [3,13] that the discriminants axes are given by the eigenvectors of $\mathbf{C} = \mathbf{W}^{-1}\mathbf{B}$, where $\mathbf{W}$ is the sample within groups covariance matrix, and $\mathbf{B}$ is the sample between groups covariance matrix.

In many applications, like these considered in this work, the number of patterns available in the sample (K) is less than the number of features and $\mathbf{W}$ is singular. As Swets and Weng [3] observed, this problem can be circumvent by first applying the PCA to the entire sample and then choosing p, the number of Most Expressive Features, such that $p+g \leq K$. The LDA procedure is then applied on the sample projected on the MEF subspace, where the matrix $\mathbf{W}$ will be full rank. The resulting g-1 axes define the basis of a new subspace, called the Most Discriminant Features space (MDF) [3].

4. Experiments

With the evaluation purpose six distinct recognition systems were built. Each system is characterized by one out of two basis (MEF and MDF) and one out of three estimates for the covariance matrices ($\mathbf{S}_i$, $\mathbf{S}_{pooled}$,$\mathbf{S}_{mix\ i}$). In all systems the maximum probability classifier was applied, each time using one of the three covariance estimates. The linear combination factor a corresponding to the mixed sample covariance matrix proposed assumed the following values: 0.1, 0.3, 0.5, 0.7 and 0.9.

4.1. Database

The experiments to evaluate the classification schemes for the face recognition problem make use of the ORL Face Database [4,14], containing ten images for each of 40 individuals, a total of 400 images.

The facial expression database was provided by the Tohoku University [6,9]. It is composed of 193 images of expressions posed by nine Japanese females. Each person posed three or four examples of each six fundamental facial expressions: angry, disgusting, fear, happy, sad and surprised, as define in [9]. The database has at least 29 images for each fundamental facial expression. For implementation convenience all images were first resized to 64x64 pixels.

4.2. Training and Testing Sets

The face recognition classifiers were implemented using for each individual 5 images to train and 5 images to test. For the MEF's computation, the PCA eigenvectors corresponding to the top 70 eigenvalues were kept. The Experiments have shown that the use of more than 70 components brings no performance improvement.

Since the number of MDF's is limited by the number of groups and there are 40 people to recognize, the MDF features took values in the range [1, 39]. The MDF's were computed based on the first 50 MEF's. No improvement was observed in the experiments by using more than 50 MEF's in the computation of the MDF's.

For the facial expression recognition problem, 29 images of each fundamental facial expression were used. The training image set included a total of 120 images, consisting of 20 images of each facial expression, and the testing set contains the remaining 54 expression images. Our experiments on the MEF subspace used till 65 principal components and the computation of the MDF's was based on the first 55 MEF's. The number of MDF's was limited by 5, since there are 6 fundamental facial expression to recognize.

5. Results

The main results of the experiments are summarized in figures 2, 3 and 4. Figure 2 shows the average recognition rate for each covariance matrix estimate as a function of the number of MEF components. The six curves represent the performance of the $\mathbf{S}_{pooled}$ estimate against the five $\mathbf{S}_{mix}$ estimates corresponding to a linear combination factor a equals to 0.1, 0.3, 0.5, 0.7 and 0.9. Since only 5 images of each individual were used to form the training set, the results relative to the sample group covariance estimate were limited to 4 MEF components and, therefore, does not appear on the graph of figure 2. It shows that the $\mathbf{S}_{mix}$ produced a better performance than the $\mathbf{S}_{pooled}$ estimate, for all number of MEF components and for all values of the linear combination factor a considered in these experiments. The best recognition rate for the $\mathbf{S}_{mix}$ estimate was 96.88% with 40 MEF components and a equals to 0.7. The performance is similar to the best results reported in previous works [4,14] which used the same database.

Figure 3 shows the average recognition rate for each covariance estimate as a function of the number of MDF components. Again, because of the training size of each group, the sample group covariance results are not shown. No evident superiority of any estimate is shown in figure 3. Depending on the value of the linear combination factor a and on the number of components, $\mathbf{S}_{mix}$ may have a better or worse performance than $\mathbf{S}_{pooled}$. The best recognition rate – 96.52% – for all the MDF components considered is reached by the $\mathbf{S}_{mix}$ estimate with the linear factor a equals to 0.7 for 39 MDF components.

One of the results of the facial expression recognition is summarized in figure 4. In this case, seven curves representing the performance of the $\mathbf{S}_{group}$, $\mathbf{S}_{pooled}$, and the five $\mathbf{S}_{mix}$ estimates (for a = 0.1;0.3;0.5;0.7;0.9) are presented. From figure 4 it can be observed that for more than 20 MEF components, in which case the $\mathbf{S}_{group}$ becomes singular, the $\mathbf{S}_{mix}$ estimate reaches better recognition rates than the $\mathbf{S}_{pooled}$ estimate for all values of a considered in the experiment. The best recognition rate of these experiments – 85.19% – was obtained by the $\mathbf{S}_{mix}$ estimate with a linear combination factor a equals to 0.3 and for 65 MEF components. This performance is similar to the best results reported so far [6,9] for the same database.

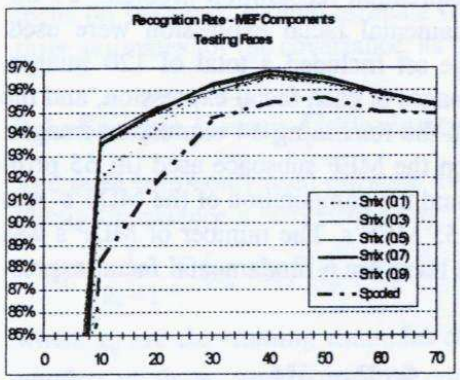


Figure 2: Face Recognition

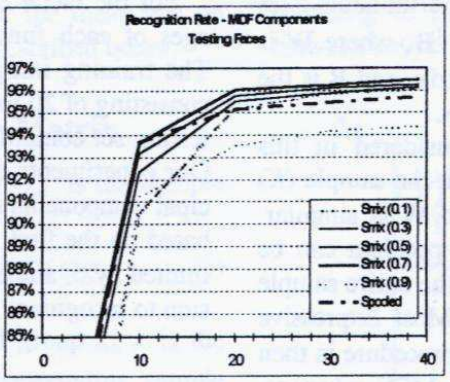


Figure 3: Face Recognition

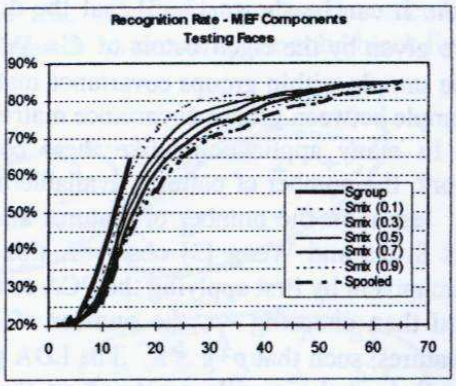


Figure 4: Facial Expression Recognition

There is no clear superiority of any covariance estimator for the facial expression recognition as a function of the number of MDF components. Therefore, these results are not shown.

6. Conclusion

This work proposed a new estimate for the covariance matrix for object recognition applications, called mixed covariance matrix. The new estimate has the same rank of the data matrix and is, therefore, invertible even in the cases where the usual sample group estimate does not admit an inverse due to an insufficient number of patterns for each group in the training set. Behind this advantage, the proposed estimate does not assume equal covariance matrices for all groups. This allows a better representation for the population of each group.

Extensive experiments were carried out to evaluate this approach on two recognition tasks: face recognition and facial expression recognition. A maximum probability classifier was built using the proposed estimate, the usual sample group and pooled estimates. In both tasks the best recognition performance was reached by the mixed group covariance estimate, especially when the MEF projection was used.

Acknowledgments

The first author was partially supported by the Brazilian Government Agencies CNPq and CAPES.

References

1. A. Pentland, B. Moghaddam, and T. Strainer, "View-based and modular eigenspaces for face recognition", CVPR, June, 1994.
2. A. Pentland, et al., "Experiments with Eigenfaces", International Joint Conference on Artificial Intelligence, Chamberry, France, August ,1993.
3. D. L. Swets and J. Weng, "Using Discriminant Eigenfeatures for Image Retrieva," IEEE Ttrans. on PAMI, Vol. 18, pp. 831-836, Aug. 1996.
4. F. Samaria and A. Harter, "Parametrisation of a stochastic model for human face identification", Proc. 2nd IEEE workshop on Applications of Computer Vision, 1994.
5. J.R. Magnus and H. Neudecker, "Matrix Differencial Calculus with Applications in Statistics and Econometrics", Wiley and Sons Limited, pg. 21-21, 1995.
6. M.J. Lyons, J.Budynek, and S.Akamatsu, "Automatic Classification of Single Facial Images", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 21, no. 12, pp. 1357-1362, December 1999.
7. M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve procedure for the characterization of human faces", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 12, No. 1, Jan. 1990.
8. M. Turk and A. Pentland, "Eigenfaces for Recognition, "Journal of Cognitive Neuroscience, Vol. 3, pp. 72-85, 1991.
9. P. Ekman and W.V.Friesen, "Pictures of Facial Affect", Human Interaction Laboratory, Univ. of California Medical Center, San Francisco, 1976.
10. P. J. B. Hancock, A. M. Burton and V. Bruce, "Face processing: human perception and principal components analysis", Memory and Cognition, 1996.
11. P. J. B. Hancock, V. Bruce and A. M. Burton, "Testing Principal Component Representations for Faces", Proc. of 4th Neural Computation and Psychology Workshop, 1997.
12. P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs FisherFace: Recognition Using Class Specific Linear Projection", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 711-720, July 1997.
13. R.A. Johnson R.A. and D.W. Wichern, "Applied Multivariate Statistical Analysis", by Prentice-Hall, Inc., 3d. edition, 1992.
14. S.Lawrence, C. L. Giles, A.C. Tsoi and A. D. Back, "Face Recognition: A Convolutional Neural-Network Approach", IEEE Trans. Neural Networks, vol. 8, no. 1, pp. 98-113, January 1997.
15. W. Zhao, R. Chellappa and A. Krishnaswamy, "Discriminant Analysis of Principal Components for Face Recognition," Proc. 2nd International Conference on Automatic Face and Gesture Recognition, pp. 336-341, Japan, April, 1998.

ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ СТЕРЕОЗОРУ В РАМКАХ БАЙЄСІВСЬКОЇ ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ

Савчинський Б.Д.
Міжнародний науково-навчальний центр ЮНЕСКО/МПІ інформаційних технологій та систем.
пр.Академіка Глушкова 40, Київ 03680 т.266-62-08 bogdan@image.kiev.ua

Анотація. Розглядається байєсівський підхід до задачі стереозору. Існуючі її розв'язки аналізуються з точки зору байєсівських вирішуючих правил. Проводиться порівняльний аналіз алгоритмів стереозору в рамках байєсівського підходу.

1 ВСТУП

Проблема машинного стереозору має свою історію, що триває вже кілька десятиріч. Дослідження цієї проблеми були започатковані видатними результатами, що зводили задачі стереозору до вирішення певних задач динамічного програмування (див.,напр. [1], [2]). В рамках цього підходу стало можливим знаходити найімовірніший рельєф місцевості, що відповідає заданій стереопарі знімків. В цьому напрямі були вирішені численні прикладні проблеми. Завдяки одержаним успіхам такий підхід став свого роду аксіомою, і задачі, що не вдавалося вирішити у рамках цього підходу, вважалися за практично невирішувальні.

Дана робота відходить від встановленого стереотипу знаходження **найімовірнішого** рельєфу, що конче є лиш окремим випадком байєсівської теорії прийняття рішень. Погляд на задачу з точки зору цієї, більш загальної теорії, дозволив знайти нові, прийнятніші формулювання задачі стереозору і, відповідно, побудувати нові алгоритми її вирішення.

Робота складається з чотирьох розділів. Другий розділ присвячений постановці і розв'язку задачі стереозору, як задачі, що формулюється в рамках байєсівського підходу. Тут запропоновано різні функції втрат та алгоритми розв'язку задач, що їм відповідають. Третій розділ має на меті демонстрацію прикладних результатів, які були отримані при застосуванні алгоритмів, наведених у другому розділі. Нарешті, в четвертому розділі формулюються висновки.

2 ПОСТАНОВКА ТА РОЗВ'ЯЗОК ЗАДАЧІ СТЕРЕОЗОРУ ПРИ РІЗНИХ ФУНКЦІЯХ ВТРАТ

Задача стереозору полягає у відновленні просторової конфігурації об'єкта за парою його фотознімків. Вважатимемо, що об'єкт знімається таким чином, що для кожної його точки справедливо: або її зображення присутнє на обох знімках, або на жодному із знімків його нема. Зображення точки об'єкта на лівому та правому знімках називатимемо відповідно лівою та правою проекціями цієї точки. Не втрачаючи загальності вважатимемо, що зображення оцифровані і відкалібровані таким чином, що рядки з номерами k на лівому та правому зображеннях є, відповідно, лівою та правою проекціями деякої лінії на вихідному об'єкті для будь-якого k. Отже, інформацію про просторову конфігурацію цієї лінії несуть лише дані два рядки. Тобто, ми можемо обмежитись розглядом задачі стереозору лише для деякої пари рядків.

Назвемо **полем зору** множину:

$$T = \{i : 0 \leq i \leq I\}.$$

Елементами поля зору є пікселі, або **точки**. **Зображенням** називатимемо функцію:

$$x : T \to R.$$

Множину всіх зображень позначатимемо через R^T. Значення зображення $x(t)$ в точці t називається **кольором піксела**. Нехай x і y – два зображення. Введемо функцію $\delta : R \times R \to R$, яку називатимемо **різницею кольорів пікселів**. Окремими її випадками є $\delta(x(t_1), y(t_2)) = \mid x(t_1) - y(t_2) \mid$ та $\delta(x(t_1), y(t_2)) = \mid x(t_1) - y(t_2) \mid^2$.

Також розглянемо клас Φ монотонних функцій вигляду $\phi : T \to T$, тобто:

$$\Phi = \{\phi : t_1 > t_2 \Rightarrow \phi(t_1) \geq \phi(t_2),\ t_1, t_2 \in T\}. \quad (1)$$

© UkrObraz'2000

Дана властивість означає, що *"точкам, які знаходяться правіше на лівому зображенні, повинні відповідати лише точки, які знаходяться правіше і на правому зображенні"*.

При даному фіксованому ϕ точки t та $\phi(t)$ називатимемо **відповідними**. Для зручності введемо клас $\mathcal{P}$ функцій вигляду $\rho : T \to T$, який визначимо як:

$$\mathcal{P} = \{\rho : \exists \phi \in \Phi \ \forall t \in T \ \rho(t) = t - \phi(t)\}.$$

Функції з класу $\mathcal{P}$ називатимемо **функціями паралакса**, а значення $\rho(t)$, $\rho \in \mathcal{P}$ - **паралаксами**. Очевидним є те, що задачі знаходження $\phi \in \Phi$ та $\rho \in \mathcal{P}$ є еквівалентними. Але паралакси мають прозорий геометричний зміст: висота об'єкта є монотонною фунцією паралакса.

Твердження 2.1 *Нехай $\rho \in \mathcal{P}$. Тоді множина значень, які може приймати функція ρ в точці t, $1 \le t \le I$ при фіксованих значеннях $\rho(0), \rho(1), ..., \rho(t-1)$ залежить лише від $\rho(t-1)$ і не залежить від $\rho(0), \rho(1), ..., \rho(t-2)$.*

▷ Доведення випливає із справедливості даного твердження для функцій з класу Φ, що, в свою чергу, є наслідком того, що $\phi \in \Phi \Leftrightarrow \phi(t-1) \le \phi(t) \ \forall t : 1 \le t \le I$. ◁

Позначимо через $N_i(t)$, $0 \le i \le (I-1)$ множину тих значень, які може приймати функція ρ в точці $i+1$ при фіксованому значенні $\rho(i) = t$:

$$N_i(t) = \{r : \exists \rho \in \mathcal{P} \ \rho(i) = t \ \& \ \rho(i+1) = r\}.$$

Для формулювання задачі в рамках байєсівської теорії (див.,напр. [3],[4],[5]) нам необхідно визначити **множину ознак об'єкта**, **множину станів**, їх сумісний розподіл ймовірностей та **функцію втрат**.

Множина ознак об'єкта

$$X = \{(x_l, x_r) : x_l, x_r \in R^T\}$$

- це пара зображень. Множина станів співпадає з множиною функцій паралакса $\mathcal{P} \ni \rho$. Вважатимемо, що апріорний розподіл $p(\rho)$ всіх можливих функцій паралакса є рівномірним, а умовний розподіл ознак об'єкта при заданому його стані визначається як:

$$p((x_l, x_r)/\rho) = A \cdot \prod_{t \in T} \exp[-\delta(x_l(t), x_r(\phi(t)))] =$$

$$= A \cdot \prod_{i=0}^{I} \exp[-\delta(x_l(i), x_r(i - \rho(i)))] =$$

$$= A \cdot \exp[-\sum_{i=0}^{I} \delta(x_l(i), x_r(i - \rho(i)))],$$

де A – нормувальний множник.

При таких припущеннях апостеріорний розподіл функцій паралакса для заданої пари зображень матиме вигляд:

$$p(\rho/(x_l, x_r)) = p((x_l, x_r), \rho)/p((x_l, x_r)) =$$
$$= p((x_l, x_r)/\rho)p(\rho)/p((x_l, x_r)) = B \cdot p((x_l, x_r)/\rho) =$$
$$= C \cdot \exp[-\sum_{i=0}^{I} \delta(x_l(i), x_r(i - \rho(i)))],$$

де C та B – нормувальні множники.

При заданій функції втрат $W : \mathcal{P} \times \mathcal{P} \to R$ задача полягає в знаходженні такої функції паралакса ρ^*, яка б мінімізувала умовне математичне сподівання втрат:

$$\rho^* = \arg\min_{q \in \mathcal{P}} \sum_{\rho \in \mathcal{P}} p(\rho/(x_l, x_r)) W(q, \rho). \qquad (2)$$

Розглянемо кілька різних функцій втрат:

$$W(q, \rho) = \begin{cases} 0, & \text{якщо } \rho = q \\ 1, & \text{якщо } \rho \neq q \end{cases}, \qquad (3)$$

$$W(q, \rho) = \sum_{\substack{t \in T \\ \rho(i) \neq q(i)}} 1, \qquad (4)$$

$$W(q, \rho) = \sum_{t \in T} (\rho(t) - q(t))^2, \qquad (5)$$

Почнемо з (3). Згідно з (2)

$$\rho^* = \arg\min_{q \in \mathcal{P}} \sum_{\substack{\rho \in \mathcal{P} \\ \rho \neq q}} p(\rho/(x_l, x_r)) = \qquad (6)$$

$$= \arg\min_{q \in \mathcal{P}} (1 - p(q/(x_l, x_r))) =$$

$$= \arg\max_{q \in \mathcal{P}} p(q/(x_l, x_r)) = \qquad (7)$$

$$= \arg\max_{q \in \mathcal{P}} \exp[-\sum_{i=0}^{I} \delta(x_l(i), x_r(i - q(i)))] =$$

$$= \arg\min_{q \in \mathcal{P}} \sum_{i=0}^{I} \delta(x_l(i), x_r(i - q(i))). \qquad (8)$$

Отже, функції (3) відповідає оцінка максимальної правдоподібності (див. (7)) функції ρ. З іншого боку, **формула (8) виражає постановку задачі стереозору при традиційному, небайєсівському підході**. Але чи доречною є в поставленій задачі функція втрат вигляду (3) ? При такій функції залишаються неврахованими кілька факторів. По-перше, дана функція втрат не враховує **в скількох точках** відрізняються різні функції паралакса між собою. Неформально кажучи, ті функції паралакса, які відрізняються лише в одній точці, мусять штрафуватись менше, аніж ті, що відрізняються в десятьох точках. Цей факт враховується функцією втрат (4). Та, більше того,

оскільки паралакси монотонно пов'язані з висотою, то грає роль ще й **величина різниці** між функціями паралакса в тих точках, в яких вони відрізняються. Цей факт може бути врахований за допомогою функції втрат (5).

Але повернімось до розв'язку задачі (8). Введемо позначення:

$$F_k(\rho(0),\rho(1),...,\rho(k)) = \sum_{i=0}^{k} \delta(x_l(i), x_r(i-\rho(i))).$$

Між цими величинами існує рекурентне співвідношення:

$$F_k(\rho(0),\rho(1),...,\rho(k)) = \\ = F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) + \\ +\delta(x_l(k), x_r(i-\rho(k))), 1 \le k \le I, \quad (9)$$

а, отже:

$$\min_{\substack{\rho(0),\rho(1),...,\rho(k) \\ \rho\in\mathcal{P}}} F_k(\rho(0),\rho(1),...,\rho(k)) = \\ = \min_{\substack{\rho(0),\rho(1),...,\rho(k-1) \\ \rho\in\mathcal{P}}} F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) + \\ + \min_{\rho(k)\in N_{k-1}(\rho(k-1))} \delta(x_l(k), x_r(i-\rho(k))), 1 \le k \le I, \quad (10)$$

$$\min_{\rho(0)\in T} F_0(\rho(0)) = \min_{\rho(0)\in T} \delta(x_l(0), x_r(0-\rho(0))). \quad (11)$$

Оскільки:

$$\rho^* = \arg\min_{\rho\in\mathcal{P}} F_I(\rho(0),...,\rho(I)),$$

то алгоритм розв'язку задачі (8) полягає в послідовному застосуванні формул (10) та (11) для $k = I; I-1; ...; 0$.

Перейдемо до функцій втрат, які задаються виразами (4) та (5). Їх можна представити у вигляді:

$$W(q,\rho) = \sum_{i=0}^{I} \omega(q(i),\rho(i)), \ q,\rho\in\mathcal{P}. \quad (12)$$

Функції втрат вигляду (4) відповідає

$$\omega(k,r) = \begin{cases} 0, & \text{якщо } r = k \\ 1, & \text{якщо } r \ne k. \end{cases}$$

А функції втрат вигляду (5) –

$$\omega(k,r) = (k-r)^2.$$

Тому розв'яжемо задачу (2) саме для такого, більш загального вигляду функцій втрат (12) :

$$\rho^* = \arg\min_{q\in\mathcal{P}} \sum_{\rho\in\mathcal{P}} p(\rho/(x_l,x_r)) \sum_{i=0}^{I} \omega(q(i),\rho(i)) = \\ = \arg\min_{q\in\mathcal{P}} \sum_{i=0}^{I} \sum_{\rho(i)\in T} p(\rho(i)/(x_l,x_r))\omega(q(i),\rho(i)), \quad (13)$$

де $p(\rho(i)/(x_l,x_r)) = \sum_{\rho\in\mathcal{P}} p(\rho/(x_l,x_r),\rho(i))$ – сумарна ймовірність всіх функцій паралакса ρ при заданому значенні $\rho(i)$ в точці i. Припустимо, що величини $p(\rho(i)/(x_l,x_r))$ нам відомі. Тоді задача (13) перетворюється в задачу (8) заміною величин $\delta(x_l(i), x_r(i-q(i)))$ на $\sum_{\rho(i)\in T} p(\rho(i)/(x_l,x_r))\omega(q(i),\rho(i))$, а, отже, і розв'язується за допомогою співвідношень, що аналогічні співвідношенням (10) та (11).

Знаходження величин $p(\rho(i)/(x_l,x_r))$ також проводиться за тією ж схемою (10),(11). Але замість операції max необхідно підставити операцію $\sum$, а замість $\sum$ – операцію $\prod$. Запишемо необхідні перетворення:

$$p(\rho(i) = t/(x_l,x_r)) = \sum_{\rho\in\mathcal{P}} p(\rho/(x_l,x_r), \rho(i) = t) = \\ = C \sum_{\rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}} \prod_{j=0}^{I} exp[-\delta(x_l(j), x_r(j-\rho(j)))] = \\ = C \cdot \sum_{\rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}} \prod_{j=0}^{i} exp[-\delta(x_l(j), x_r(j-\rho(j)))] \times \\ \times \prod_{j=i}^{I} exp[-\delta(x_l(j), x_r(j-\rho(j)))] = \\ = \left(\sum_{\substack{(\rho(0),...,\rho(i-1)) \\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} \prod_{j=0}^{i} \exp[-\delta(x_l(j), x_r(j-\rho(j)))] \right) \times \\ \times C \left(\sum_{\substack{(\rho(i+1),...,\rho(I)) \\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} \prod_{j=i}^{I} \exp[-\delta(x_l(j), x_r(j-\rho(j)))] \right). \quad (14)$$

Обчислимо перший співмножник у виразі (14). Другий обчислюється аналогічним чином. Процедура обчислення має вигляд:

$$F_k(\rho(0),\rho(1),...,\rho(k)) \stackrel{def}{=} \\ \stackrel{def}{=} \prod_{j=0}^{k} \exp[-\delta(x_l(j), x_r(j-\rho(j)))]; \quad (15)$$

$$\Rightarrow F_k(\rho(0),\rho(1),...,\rho(k)) = \\ = F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) \times \\ \times \exp[-\delta(x_l(k), x_r(j-\rho(k)))]; \quad (16)$$

$$\Rightarrow \sum_{\substack{(\rho(0),\rho(1),...,\rho(k)) \\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} F_k(\rho(0),\rho(1),...,\rho(k)) = \\ = \sum_{\substack{(\rho(0),\rho(1),...,\rho(k-1)) \\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) \times \\ \times \sum_{\rho(k)\in N_{k-1}(\rho(k-1))} \exp[-\delta(x_l(k), x_r(k-\rho(k)))], \quad (17)$$

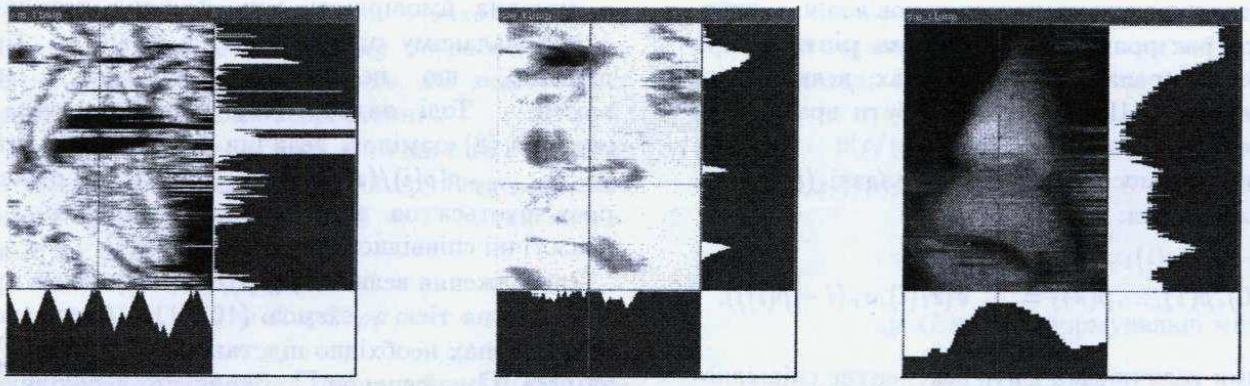


Рис. 1 Результати розв'язання задачі стереозору для трьох описаних випадків. Малюнок зліва відповідає пошуку найімовірнішого рельєфу (див.(8)), по центру - розв'язку байєсівської задачі з "точковим" штрафом (4). І, нарешті, малюнок справа отриманий з розв'язку байєсівської задачі з "квадратичним" штрафом (5).

$$F_0(\rho(0)) = \sum_{\rho(0)\in T} \exp[-\delta(x_l(0), x_r(-\rho(0)))]. \quad (18)$$

Таким чином, алгоритм пошуку першого співмножника у виразі (14) полягає в послідовному застосуванні формул (17) та (18) для $k = i; i-1; ...; 0$.

3 ДЕМОНСТРАЦІЯ ПРИКЛАДНИХ РЕЗУЛЬТАТІВ

В якості вихідних даних (стереопари) було взято відкалібровану стереопару обличчя людини.

На рис.1 зображено результат відновлення рельєфу для трьох різних функцій втрат.

Результати представлені у вигляді так званого "висотного" зображення, тобто більша яскравість точки відповідає її більшій висоті. Справа і знизу від зображення показані "розрізи": вертикальний та горизонтальний. Найлівіше зображення відповідає функції вигляду (3), зображення по центру – (4) і найправіше – (5).

Різниця в результатах є очевидною. Слід все ж відзначити, що при традиційному, небайєсівському підході (при якому використовуються критерії, аналогічні (8)), можна отримати результати, що можуть бути порівняні з результатами, отриманими в рамках байєсівського підходу з функцією втрат (5) . Це досягається головним чином за рахунок ускладнення моделі (тобто структури класу Ф) та використання різних евристичних прийомів.

4 ВИСНОВКИ

В даній роботі продемонстровано відновлення рельєфу на основі байєсівського підходу для трьох різних функцій втрат. При цьому найкращі результати було отримано при інтуїтивно найкращій з них – функції втрат (5).

Метод найбільшої правдоподібності не повинен бути догмою. Байєсівський підхід включає його як частковий випадок і тому є більш потужним інструментом при прийнятті рішень.

Література

[1] Гимельфарб Г.Л. "Симметричный подход к задаче автоматических стереоскопических измерений в фотограммметрии". *Кибернетика.*- 1979,(2).-с.73-82.

[2] Гимельфарб Г.Л. "Симметризованное би- и тринокулярное стереозрение: взаимосвязи между теоретическими основами и эвристическими решениями". *Теоретические и прикладные вопросы распознавания изображений.*- Киев, 1995.-с.4-25.

[3] Schlesinger M.I., V.Hlavač. "Deset prednasek z teorie statistickeho a strukturniho rozpoznavani". Vydavatelstvi CVUT, Praha, 1999.

[4] Ковалевский В.А. "Методы оптимальных решений в распознавании изображений".–М. :Наука, 1976.–328с.

[5] Дуда Р., Харт П. "Распознавание образов и анализ сцен". – Москва: Мир, 1976.

Techniques for generating shape instances from domain distributed vague models

Zoltán Rusák
Imre Horváth
György Kuczogi
Joris S.M. Vergeest
Faculty of Design, Engineering and Production
Delft University of Technology
Jaffalaan 9, 2628 BX, Delft,
The Netherlands
TELEPHONE: +31 15 278 3788
FAX: +31 15 278 1839
E-MAIL: z.rusak@io.tudelft.nl

ABSTRACT

We are developing and implementing a novel method, called vague shape modeling, which is able to model geometric uncertainty by describing distribution domain of the shapes. In order to obtain shape instances, we have implemented specific instance generation operators that extract a set of boundary particles from the distribution domain based on predefined selection functions. In this paper, we present three instance generation methods: (a) monolithic instancing by a single selection function, which is applied to a single particle cloud (b) monolithic instancing by multiple selection functions, which applies different selection functions on different parts of a single particle cloud, and (c) constrained instancing by multiple selection functions, which uses various selection functions on each particle cloud contained by a particle system. In the first case, the application of single selection function naturally assures the required order of discrete continuity. In the second case, a fuzzy technique is applied on the selection functions to achieve the required continuity condition. Finally, for the third case, we developed a technique, called constrained instancing, which uses correction functions in a transition strip specified in between the selected instances.

1. INTRODUCTION

Vague shape modeling has been proposed to support computer internal representation of a cluster or family of shapes without conventional geometric parameterization. This approach enables us to represent the distribution domain of vague shapes between a maximal and a minimal set of discrete entities, called particles. A particle is a weakly defined 3D point specified by its reference vector, and metric occurrence. The reference vector defines the position of the particle in a vector space, and the metric occurrence, which is a finite vector space, represents the geometric uncertainty of this position. A finite set of particles, called particle cloud, forms a discrete, vague shape. The particle system, which is combined from a finite set of particle cloud, is the highest ordered modeling entity in our geometric modeling method. In the space of the metric occurrence, so-called supporting vectors are identified that stretch a minimal and maximal boundary for the vague particle cloud/system. The directions, in which the significance distributions of from a design point of view is described, are represented by the distribution trajectories of the particles. Since each particle is contained by only one vector space, the vague shape modeling technique shows a significant difference with fuzzy modeling, which is discussed in specific papers [1, 2].

We have developed a specific shape generation method, called instancing, to extract parametric shapes from domain distribution. This technique can overcome a shortcoming of the conventional parameterization techniques, since it does not require a strict system of parameters and provides knowledge intensive support of manipulation of geometric entities. Instancing applies knowledge-driven selection functions to specify the position of particles along the distribution trajectories based on characteristic shape attributes. Application of a single selection function on a single cloud, called *monolithic instancing* by single selection function, can naturally assure the required order of discrete continuity defined in [3]. We interpret the discrete continuity as the continuity of the best matching continuos curve in the considered finite region.

In order to be able to apply multiple functions on a single particle cloud, we have developed a method, called *compound instancing*, which uses fuzzy technique in specification of multiple selection functions to maintain the required discrete continuity condition. The third method, called *constrained instancing*, applies multiple selection functions on a particle system, which can result in disconnected components of instance shapes. To obtain the required continuity: (a) a transition strip is specified between the shapes, (b) extra particles are generated in the gaps, and (c) the particles of the transition strip are manipulated. The further parts of the paper discuss the newly introduced instance generation methods in detail.

© UkrObraz'2000

2. METHODS FOR INSTANCE GENERATION

2.1. Monolithic instance generation using single selection functions

In the simplest case, an instance shape can be generated in the distribution domain of the archetype by one selection function. The mathematical form of the selection function is derived based on shape formation rules of a particular application (e.g. aesthetics, ergonomics, functionality). The selection function considers the boundaries of the distribution domain and specifies the set of particles along the distribution trajectories. Figure 1 illustrates the variables used in the instance generation.

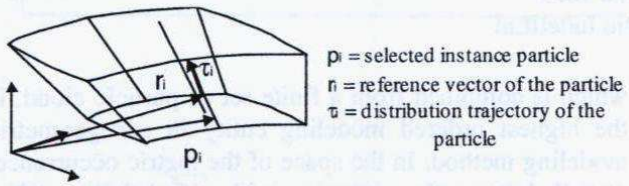


Figure 1: Variables in instance selection

To be able to calculate the selection functions, we parameterize the distribution trajectories of each particle in an interval $\alpha[-1; 1]$. In this manner, all the selection function returns with α, and the position $\mathbf{p}_i$ of the particle is computed as follows:

$$\mathbf{p}_i = \mathbf{r}_i + \alpha \cdot \boldsymbol{\tau}_i \tag{1}$$

The condition is that the instance function has to be continuos in third order at least. It means, sufficiently smooth functions can only be considered in this case. However, in certain cases, lower level of discrete continuity can also be accepted.

2.2. Monolithic instance generation using multiple selection functions

Monolithic instance generation by multiple selection functions is a more sophisticated method, which considers region specifications of a particle cloud. It is able to apply various selection functions in the different regions. This technique, called compound instancing, interfaces the results of the individual instancing functions by applying corrective functions and achieves lower order discrete continuity between the regions this way.

To specify regions R_i on a particle cloud Π, we implemented a method, by which the user can interactively manipulate the particles of a region with user-defined, free-from virtual tools. The particles are assigned to a specific region by moving the tool in the modeling space, and checking the reference vectors of the particles, whether the tool contains them. Afterwards, particles are identified that are not contained by any regions, and assigned to the closest region.

The compound instance generation applies compensated selection functions in the regions of a particle cloud.

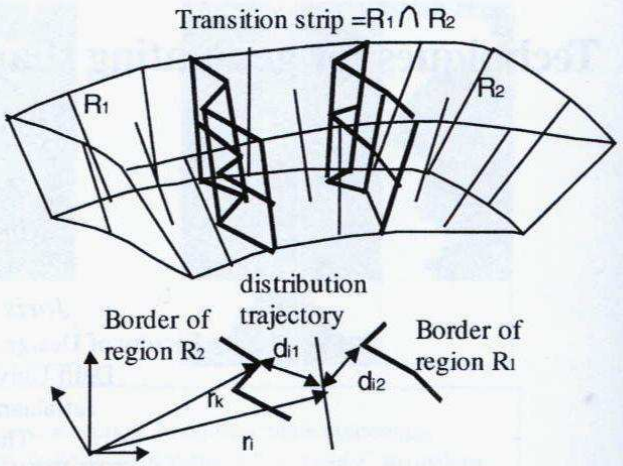


Figure 2: Specification of wieghts of the selection functions

First, the algorithm specifies a transition strip from the overlap regions described by the user. It searches for the borders of the overlap regions, and displaces them to the distance of m-th neighbor particles from each other, keeping the overlap condition. The value of m is selcted to be 8 for C^1 and 32 for C^2. Then for all particles π_i in the transition strip, distances d_{ik} are computed from the borders of the transition strip. Since we assume linearly changing fuzzy membership functions in the transition strip, the proportion of the member values w_{ik} of R_i can be computed based on d_{ik} as follows:

$$w_{ik} = \frac{d_{ik}}{\sum_{k=1}^{N} d_{ik}} \tag{2}$$

Figure 2 shows the example of specification of the member values in the transition strip. Afterwards, the position $\mathbf{p}_i$ of the instance particles can be calculated as follows:

$$\mathbf{p}_i = \mathbf{r}_i + \boldsymbol{\tau}_i \cdot \frac{\sum_{k=1}^{N} w_{ik}\alpha_k}{\sum_{k=1}^{N} w_{ik}} \tag{3}$$

The result of this method can be further elaborated either by increasing the size of the transition strip or by changing the linear membership function.

2.3. Constrained instance generation on a particle system

The constrained instance generation method applies various selection functions on different particle clouds of a particle system. It can result in a set of disconnected surface patches. Therefore, the concerned instancing functions have to be appended by additional correction functions. The correction functions modify the position of the particles in the transition strip, which is defined based on the required continuity conditions. Figure 3 illustrates a pair of disconnected surface patches and the transition strip, whose geometric smoothness needs improvement.

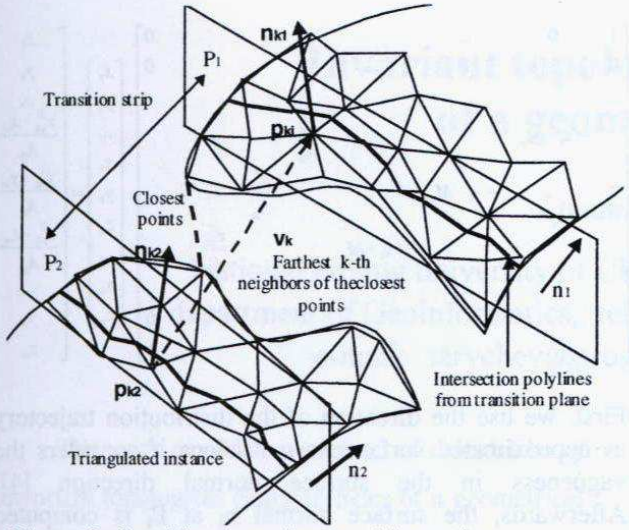


Figure 3: Specification of a transition strip

To generate a smooth transition between disconnected or intersecting shape instances, we propose a method that we call constrained instancing. Figure 4 presents the process of constrained instancing. In the first step, selection functions are applied to the particle clouds that extract shape instances from the domain distribution. Next, the relationship between the particle clouds is investigated based on intersection detection between the triangulated models of the instances. When the shapes are intersecting, the intersecting polylines are calculated, and the extra particles are removed from the model. Afterwards, the starting particle of a transition strip is determined. When there is no intersection, the closest

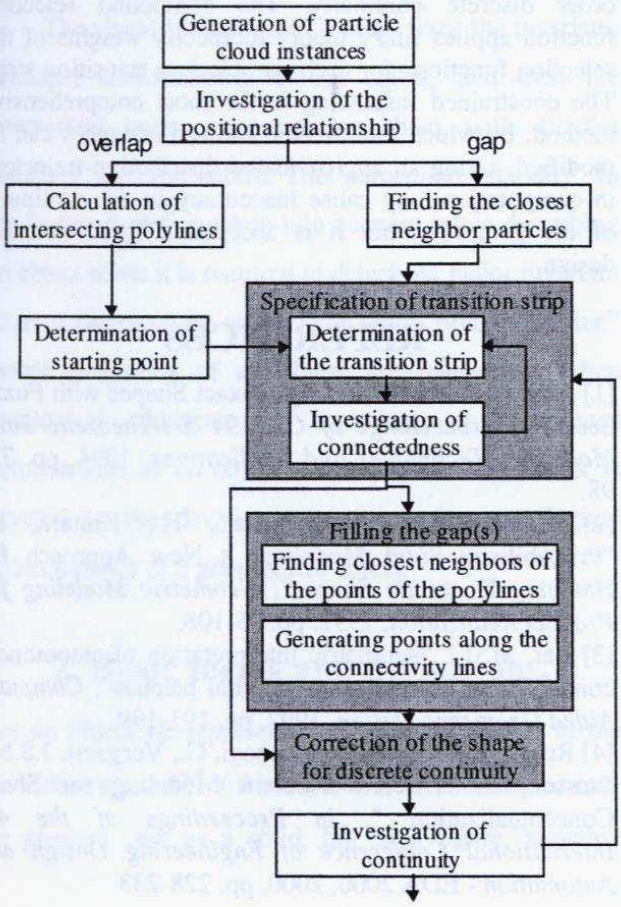


Figure 4: Constrained Instance generation

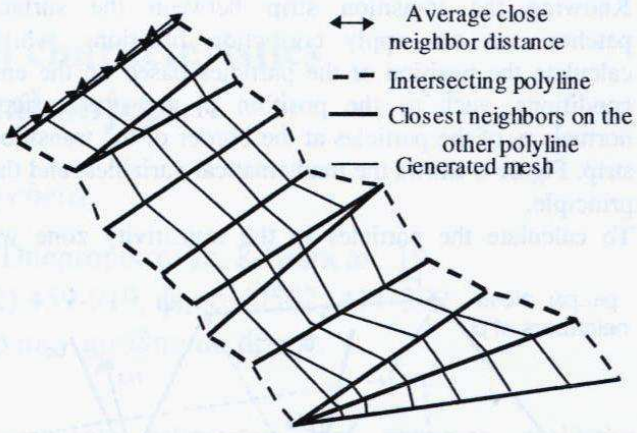


Figure 5: Particle generation in the transition strip

neighbor particles are determined as starting particles.

To specify the transition strip, two planes P_1, and P_2 are determined, as the border of the transition strips, and the intersection polylines between the planes and the instances are calculated. Figure 3 illustrates the specification of a transition strip. The transition strip of two instances is described by planes P_1, and P_2 placed into the k-th neighbor of the starting particle(s), where the value of k is chosen to be 4 for C^1 and 16 for C^2. From all the k-d neighbors on the two instances, the farthest ones are selected as the border particles of the transition strip. The normal vectors of P_1, and P_2 are defined by the positions ($\mathbf{p}_{k1}$, $\mathbf{p}_{k2}$) of farthest k-th neighbors of the starting particle(s) as follows: $\mathbf{n}_1 = \mathbf{n}_{k1} \times \mathbf{v}_{kp}$ and $\mathbf{n}_2 = \mathbf{n}_{k2} \times \mathbf{v}_{kp}$, where $\mathbf{v}_{kp}$ is the perpendicular component of the $\mathbf{v}_k = \mathbf{p}_{k1} - \mathbf{p}_{k2}$, and $\mathbf{n}_{k1}$ and $\mathbf{n}_{k2}$ are the normal vectors at $\mathbf{p}_{k1}$ and $\mathbf{p}_{k2}$ respectively. The transition strip is investigated based on the type of the intersecting polylines. The border of the transition strip is extended, until the intersection results in one closed or one open polyline.

Since our method manipulates the position of particles included in the strip to achieve smooth connection between the shapes, it requires quasi-regular particle-set in the transition strip. Therefore, new particles are generated in the gap(s) of the transition strip in two steps. First, the closest neighbors of the particles are identified on the other polyline, and connected by so-called connectivity lines. Then, for each connecting line, the average distance of close neighbors is calculated with respect to the neighbors of the end particles of the connectivity line. Second, particles are generated along the connectivity lines dividing them into line-segments with quasi-regular length. The regular length means an approximation of the average distance between the neighboring particles. Finally, approximated surface normals are calculated in each generated particle based on the triangulated mesh. Figure 5 illustrates application of the algorithm.

Afterwards, the generated particles in the transition strip are manipulated based on the correction function. Finally, the second order discrete continuity condition of the constrained instancing is evaluated based on the Lai's method [3], and the transition strip can be partially modified at the insufficient places.

Knowing the transition strip between the surface patches, we can apply correction functions, which calculate the position of the particles based on the end conditions, such as the position $\mathbf{p}_i$ and the surface normals $\mathbf{n}_i$ of the particles at the border of the transition strip. Figure 6 shows the mathematical variables, and the principle.

To calculate the particles in the transitivity zone we

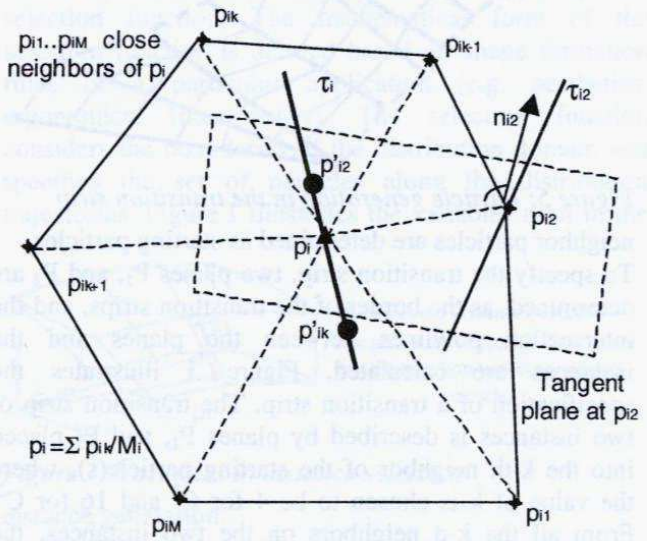


Figure 6: Principle of the correction function

applied an analytical method. The position $\mathbf{p}_i$ of particle π_i is described by the distribution trajectory τ_i at π_i, and the positions $\mathbf{p}_1..\mathbf{p}_{Mi}$ together with the normal vectors $\mathbf{n}_1..\mathbf{n}_{Mi}$ of all the neighbor particles ($\pi_1..\pi_{Mi}$) of π_i. In the first calculation, the distribution trajectory τ_i is approximated by the surface normal vector $\mathbf{n}_i$ at P_i, so as τ_i ($\mathbf{p}_i$, $\mathbf{n}_i$). From a neighbor particle $\pi_k \in (\pi_1..\pi_{Mi})$ of π_i, we compute the intersection-points $\mathbf{p'}_{ik}$ of the tangent plane at P_k and the approximated distribution trajectory τ_i of the π_i as follows:

$$\mathbf{p'}_{ik} = \frac{(\mathbf{p}_k - \mathbf{v}_i)\cdot\mathbf{n}_k}{\mathbf{n}_k^2 \cdot \left(\mathbf{n}_k \cdot \frac{\boldsymbol{\tau}_i}{|\boldsymbol{\tau}_i|}\right)} \cdot \mathbf{n}_k = \frac{(\mathbf{p}_k - \mathbf{v}_i)\cdot\mathbf{n}_k}{A_{ik}} \cdot \mathbf{n}_k \qquad (4)$$

The position $\mathbf{p}_i$ of the concerned particle is the average of all intersection-point $\mathbf{p'}_{ik}$ specified by the neighbors of a particle:

$$\mathbf{p}_i = \frac{1}{M_i}\sum_{k=1}^{k=M_i} \mathbf{p'}_{ik} \qquad (5)$$

The above equations can be applied for all the particles in the transition strip, and the result can be organized in a linear system of equations. The coefficient matrix of this linear system is a sparse matrix, which can be solved by a frontal method. The complexity of the matrix is linearly increasing with the number of elements.

$$\begin{bmatrix} 1 & 0 & \dots & & & & 0 \\ 0 & 1 & 0 & \dots & & & 0 \\ \dots & & & & & & \\ & -x_i \cdot M_i & \dots & \frac{x_{nk}^2}{A_{ik}} & \dots & & \\ 0 & & -y_i \cdot M_i & & \frac{y_{nk}^2}{A_{ik}} & \dots & \\ & & & -z_i \cdot M_i & & \frac{z_{nk}^2}{A_{ik}} & \\ & & & & & & 0 \\ & & & & & & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ y_1 \\ \dots \\ x_k \\ y_k \\ z_k \\ \dots \\ y_n \\ z_n \end{bmatrix} = \begin{bmatrix} x_1 \\ y_1 \\ \dots \\ \frac{x_{nk}^2 \cdot x_{vi}}{A_{ik}} \\ \frac{y_{nk}^2 \cdot y_{vi}}{A_{ik}} \\ \frac{z_{nk}^2 \cdot z_{vi}}{A_{ik}} \\ \dots \\ y_n \\ z_n \end{bmatrix}$$

First, we use the direction of the distribution trajectory as approximated surface normal, since it considers the vagueness in the surface normal direction [4]. Afterwards, the surface normal $\mathbf{n}_i$ at P_i is computed based on triangulation over the set of the generated instance particles. The calculation of correction function is repeated with the new surface normals, until the result fulfils the required discrete continuity conditions.

3. CONCLUSIONS

In this paper, we presented three methods of shape instance generation techniques. The techniques have been developed for selection of instances from domain distributed discrete models. Since the proposed techniques working by analytic manner, they provide real time calculation of shape instances. The monolithic instance generation by a single selection function naturally generates particle cloud instances with second order discrete continuity. The compound selection function applies fuzzy theory to specify weights of the selection functions for each particle in a transition strip. The constrained instancing is the most comprehensive method, by which the already selected instances can be modified. Using an approximated distribution trajectory in computation may cause inaccuracy in the continuity of the geometry, but it is acceptable for conceptual design.

REFERENCES

[1] Martin, R. R., "Modeling Inexact Shapes with Fuzzy Sets", *in Proceedings of CSG 94 Set-Theoretic Solid Modeling: Techniques and Applications*, 1994, pp. 73-98.

[2] Yamaguchi, Y., Nakamura, H., Kimura, F., "Probabilistic Solid Modeling: a New Approach for Handling Uncertain Shapes", *Geometric Modeling for Product Realization*, 1992, pp. 95-108.

[3] Lai, M. J., "Geometric interpretation of smoothness conditions of triangular polynomial patches", *Computer Aided Geometric Design*, 1997, pp. 191-199

[4] Rusák, Z., Horváth, I., Kuczogi, G., Vergeest, J.S.M., Jansson, J.," Discrete Domain Modeling for Shape Conceptualization ", *in Proceedings of the 4th International Conference of Engineering Design and Automation* - EDA 2000, 2000, pp. 228-233

Invariant topological characteristics of a geometrical object.

Ljudmila Sarycheva

National mining university of Ukraine, Dnepropetrovsk, K.Marx av.,19,
department of Geoinformatics, tel. (0562) 459-919, факс. (0562) 471-392,
e-mail: sarychev@prognoz.dp.ua, gis@nmuu.dp.ua

The work considers a method of calculation of invariant topological characteristics of a geometrical object. A geometrical object is considered as a simplicial complex. Curvilinear simplexes correspond to homeomorphic rectilinear simplexes. For each k-dimensional simplex incidence matrixes are calculated. The matrix are reduced to the canonical form. The set of incidence matrix's invariant multipliers of the simplicial complex is the characteristic of a geometrical object.

1. Introduction.

The visual perception first of all fixes the invariant, principal characteristics of an object, and then the unessential ones supplementing them with smaller details of a metric nature. This perception "stability" to the "white noise" is taken into account while describing an object when it is required to detect the major invariant characteristics of an object. For many problems "fine" metric properties of an object are unessential when topological characteristics invariant to continuous deformations of an object are important. This work is devoted to the development of invariant topological characteristics of a geometrical object.

2. Basic terms and definitions.

Let an object be represented as a set of points in the Euclidean space $\mathbf{R}^n$ (the number $\mathbf{n}$ of dimensions may be greater) not as a solid, but as a figure, allowing continuous deformation that preserves qualitative properties of an object. In this case homeomorphisms allow to eliminate unimportant details and disclose the major ones in a geometrical object [1].

Homeomorphism is a one-to one and mutually continuous to both sides mapping of one figure into another. For example rectilinear line segment and continuous arch (without self-intersections) are homeomorphic on the flat, quadrate and circle are homeomorphic, cube and tetrahedron are homeomorphic.

The set of points $\mathbf{A_0}$, $\mathbf{A_1}$, ..., $\mathbf{A_n}$ of the Euclidean space $\mathbf{R^{n+1}}$ forms $\mathbf{n}$-dimensional rectilinear simplex $\mathbf{G^n}$ (in other words $\mathbf{A_0}$, $\mathbf{A_1}$, ..., $\mathbf{A_n}$ are vertexes of simplex $\mathbf{G^n}$), if

$$\mathbf{OA} = \sum_{i=0}^{n} \mathbf{m_i OA_i} ,$$

where $\mathbf{O}$ is the origin of coordinates, $\mathbf{m_0}$, $\mathbf{m_1}$, ..., $\mathbf{m_n}$ are barycentric coordinates of a point $\mathbf{A}$ which is the center of gravity of points $\mathbf{A_0}$, $\mathbf{A_1}$, ..., $\mathbf{A_n}$;

$$\sum_{i=0}^{n} \mathbf{m_i} = 1, \quad \mathbf{m_i} \geq 0 .$$

A zero-dimensional simplex is represented as a point, one-dimensional simplex – as a rectilinear line segment on the flat, two-dimensional simplex – as a triangle in a three-dimensional space, three-dimensional simplex – as a tetrahedron (fig.1).

© UkrObraz'2000

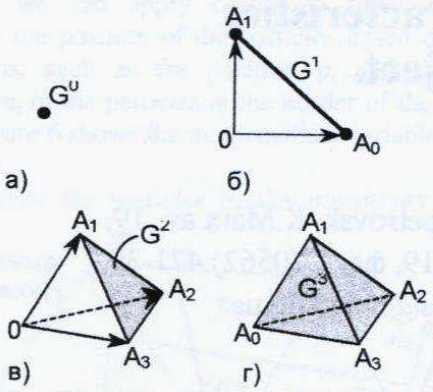


Fig.1. Examples of rectilinear simplexes: a) – zero-dimensional, b) – one-dimensional, c) – two-dimensional, d) – three-dimensional

The curvilinear simplexes are the reflexes of the rectilinear ones under homeomorphisms (fig.2)

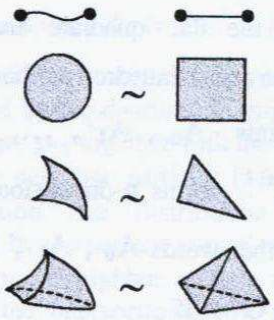


Fig.2. Examples of homeomorphic curvilinear and rectilinear simplexes.

If a set of points of **n**-dimensional simplex has one coordinate $\mathbf{m_i} = \mathbf{0}$, then this set forms **(n-1)**-dimensional simplex, that is called **(n-1)**-dimensional facet of simplex $\mathbf{G}^{\mathbf{n}}$ and is denoted as $\mathbf{G_i^{n-1}}$. Simplex $\mathbf{G}^{\mathbf{n}}$ has **(n+1)** facets of dimension **(n-1)**. If some **(n-k)** coordinates are equal to zero, then a **k**-dimensional facet of $\mathbf{G}^{\mathbf{n}}$ simplex is considered. The number of **k**- dimensional facets in $\mathbf{G}^{\mathbf{n}}$ simplex is equal to $\mathbf{C_{n+1}^{k+1}}$ - the number of combinations of **(k+1)** from **(n+1)**.

3. Constructing the matrix of incident elements characterizing a geometrical object.

Let's represent the **X** set (the object of recognition) of points of the Euclidean space as a union of finite number **N** of curvilinear simplexes having dimensions from zero to some **n** number, where

1) each point of the **X** set gets into some simplex;

2) two simplexes either have no common points or one of them is a facet of the other or they have common facet that is their intersection (fig.3).

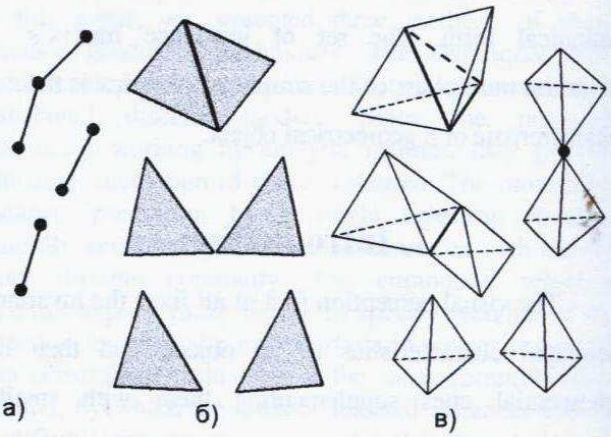


Fig.3. Variants of positional relationships of one-dimensional (a), two-dimensional (b) and three-dimensional (c) simplexes.

We will call this union a simplicial complex and denote it by **Q**.

With each curvilinear simplex some topological mapping of a rectilinear simplex to this curvilinear simplex is associated, the way the splice of two curvilinear simplexes by the common facet is corresponded by the splice of two corresponding rectilinear simplexes by some linear mapping of common facet.

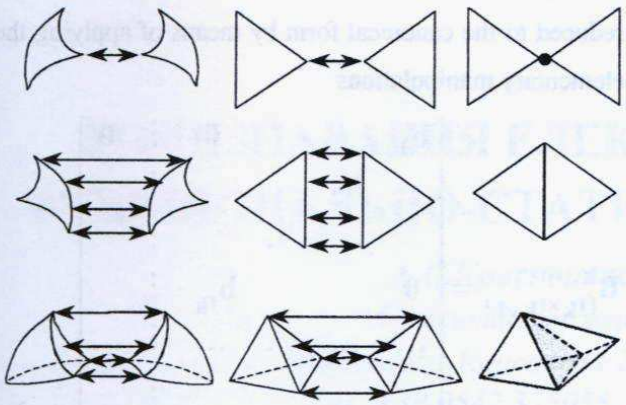


Fig. 4. Splice by zero-dimensional (a), one-dimensional (b), two-dimensional (c) facet.

When N is little, the X set is rather a "solid" body, if each simplex is considered as a solid and rotating simplexes around their common facets is allowed. With N increasing, the X set becomes more "pliable".

Let's consider the set of all k-dimensional simplexes of the X simplicial complex, number them,give each simplex an orientation and denote it as G_i^k, where i is a number, $1 \le i \le t_k$, $t_k < \infty$, t_k is the number of k-dimensional simplexes in X, $1 \le k \le n$, $n=\dim X$.

Two simplexes in the simplicial complex Q are called incidental if one of them is a facet of another.

The margin ∂G^k of oriented simplex G^k is the union of all it's $(k-1)$-dimensional facets taking into consideration their orientation:

$$\partial G^k = \sum_{i=0}^{k} (-1)^i G_i^{k-1} \quad .$$

For example, $\partial G^1 = A_1 - A_0$,

$$\partial G^2 = A_0A_1 + A_1A_2 - A_0A_2.$$

In this case

$$\partial G_j^{k+1} = \sum_{i=1}^{t_k} a_{ij}^k G_j^k \quad ,$$

$1 \le j \le t_{k+1}$, $1 \le i \le t_k$.

The number a_{ij}^k - is a coefficient of incidence of simplexes G_i^k and G_j^{k+1}. It is equal either to +1, or to -1, depending on the orientation with which the simplex G_i^k is included into the margin of the simplex G_j^{k+1} (+1, if the orientation assigned on G_i^k, coincides with the orientation induced on it by the G_j^{k+1} simplex; -1 otherwise), or to 0, if simplexes G_i^k and G_j^{k+1} are not incidental. Numbers a_{ij}^k, $1 \le j \le t_{k+1}$, $1 \le i \le t_k$, form k-dimensional incidence matrix of the size $t_k \times t_{k+1}$:

$$B^k = (a_{ij}^k)$$

Each simplicial complex Q associates with a suit of incidence matrixes: B^0, B^1, ..., B^{n-1}. For example for the simplex G^2 (fig.5)

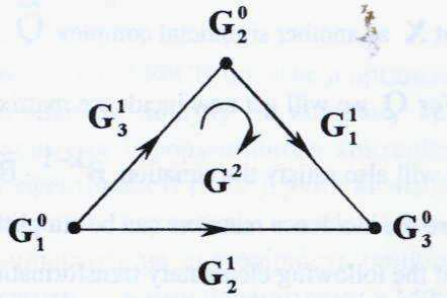


Fig. 5. Two-dimensional oriented simplex G^2

incidence matrixes look as follows:

$$B^0 = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \quad B^1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

The simplicial complex Q is completely and unambiguously determined by the set B^0, B^1, ..., B^{n-1} of incidence matrixes, i. e. all the properties of the Q complex are determined by its incidence matrixes.

4. Calculating invariant topological characteristics by means of incidence matrixes.

Since the square of the margin operator equals to zero: $\partial^2 = 0$ (the simplex margin has no margin itself), then

$$\partial\,(\partial G_j^{k+1}) = \sum_{i=1}^{t_k} a_{ij}^k\, \partial G_i^k =$$

$$= \sum_{m=1}^{t_{k-1}} \sum_{i=1}^{t_k} a_{ij}^k\, a_{mi}^{k-1}\, \partial G_m^{k-1} = 0.$$

Hence,

$$\sum_{i=1}^{t_k} a_{mi}^{k-1}\, a_{ij}^k = 0\,,\quad 1 \le j \le t_{k+1}\,,\quad \le m \le t_{k-1}\,.$$

At the settled k

$$B^{k-1} \cdot B^k = 0, \quad k=0, 1,\ldots, n-1.$$

This matrix equation will not change if we represent X as another simplicial complex $\widetilde{Q}$.

For $\widetilde{Q}$ we will get new incidence matrixes $\widetilde{B}^k$, but they will also satisfy the equation $\widetilde{B}^{k-1} \cdot \widetilde{B}^k = 0$. Therefore the incidence matrixes can be simplified by means of the following elementary transformations of simplexes.

If the G_i^k simplex is replaced by $G_i^k + G_m^k$, $i \neq m$, the matrix $\widetilde{B}^{k-1}$ is obtained from the matrix B^{k-1} by adding its m-th column to the i-th column of B^{k-1}. The matrix $\widetilde{B}^k$ is obtained from the matrix B^k by subtracting the i-th row from its m-th row. If the simplex G_i^k is replaced by $(-G_i^k)$, then the matrix $\widetilde{B}^{k-1}$ is obtained from B^{k-1} by multiplying the i-th column of B^{k-1} by (-1), and the matrix $\widetilde{B}^k$ is obtained by multiplying the i-th row of the matrix B^k by (-1). It is known [3], that any integer matrix can be reduced to the canonical form by means of applying the elementary manipulations

$$B^k_{(t_k \times t_{k+1})} = \begin{bmatrix} b_1 & & & 0 & \vdots & 0 \\ & b_2 & & & \vdots & \\ & & \ddots & & \vdots & \\ 0 & & & b_{r_k} & \vdots & \\ \cdots & \cdots & \cdots & \cdots & \vdots & \cdots \\ & & & & \vdots & \\ 0 & & & & \vdots & 0 \end{bmatrix}$$

where $r_k = \operatorname{rang} B^k$; $b_1, b_2, \ldots, b_{r_k}$ - nonzero invariant multipliers of a matrix.

Invariant multipliers are unambiguously determined by the matrix and can be calculated algorithmically. The suit of invariant multipliers of incidence matrixes of the simplicial complex Q is the characteristic of the geometrical object X. Characteristics are also the numbers

$\beta_k = t_k - r_k - r_{k-1}$,

$(r_{-1} = r_n = 0)$, $1 \le k \le n$, $n=\dim X$.

5. Conclusions.

A simple and evident method of calculating the invariant topological characteristics of a geometrical object was considered, based on representing the object as a simplicial complex and applying the incidence matrixes. Practical implementation of the described method involves cumbersome calculations, since it contains operations of transforming big matrixes.

References:

1. Понтрягин Л.С. Основы комбинаторной топологии. М., 1976.- 136 с.
2. Бусыгин Б.С., Мирошниченко Л.В. Распознавание образов при геолого-геофизическом прогнозировании.- Днепропетровск: Изд-во ДГУ, 1991.- 168 с.
3. Хорн Р., Джонсон Ч. Матричный анализ: Пер. с англ.- М.: Мир, 1989.– 655 с.

РОЗПІЗНАВАННЯ ЕЛЕКТРОНОГРАМ ЗА МЕТОДОМ ФУНКЦІОНАЛЬНО-СТАТИСТИЧНИХ ВИПРОБУВАНЬ

А.С.Краснопоясовський, С.О.Заговора
Сумський Державний Університет
Римського-Корсакова 2, м Суми Україна 40007
Тел. +38 0542 335055, факс +38 0542 334058
E-mail: pm@ssu.sumy.ua

Розглядається підхід до обробки та розпізнаванню електронограм, які одержують у режимі мікродіфракції на просвічуючому електронному мікроскопі з застосуванням методу функціонально-статистичних випробувань.

Вступ

Автоматизація розпізнавання електронограм все ще залишається однією з трудно формалізованих задач. Певну надію на її розв'язання пов'язують з інтенсивною розробкою методів автоматичної класифікації. Одним з них є метод функціонально-статистичних випробувань (МФСВ), що ґрунтується на оцінці інформаційної здатності системи розпізнавання образів [1,2]. Розглянемо підхід до розпізнавання електронограм за МФСВ.

1 Методологічні та теоретичні положення підходу

Нехай $\{X_m^0\}$, $m=\overline{1,M}$ - абетка класів розпізнавання. Реалізаціями відповідних образів є випадкові вектори $\{x_m^{(n)}\}$, $n=\overline{1,n^*}$, де n^* - достатній обсяг репрезентативної вибірки, які належать вибірковому простору X. Надалі простір X будемо вважати підпростором простору Хеммінга. Розв'язання задачі розпізнавання електронограм природно здійснюється в рамках дискримінантного класифікаційного аналізу, який передбачає на етапі навчання побудову нечіткого розбиття простору X на M областей $\{\tilde{X}_m^0\}$ і на етапі екзамену – прийняття рішення про належність образу, що розпізнається до певного класу із заданого алфавіту. При цьому для загального визначення нечіткого розбиття $\tilde{R}^{|M|}$, яке допускає перетин класів розпізнавання, необхідне виконання наступних умов:

$$\left(\forall \tilde{X}_m^0 \in \tilde{R}^M\right)\left[\tilde{X}_m^0 \neq 0\right];$$

$$\left(\exists \tilde{X}_k^0 \in \tilde{R}^M\right)\left(\exists \tilde{X}_l^0 \in \tilde{R}^M\right)\left[\tilde{X}_k^0 \neq \tilde{X}_l^0 \rightarrow \tilde{X}_k^0 \cap \tilde{X}_l^0 \neq 0\right];$$

$$\left(\forall \tilde{X}_k^0 \in \tilde{R}^M\right)\left(\forall \tilde{X}_l^0 \in \tilde{R}^M\right)\left[\tilde{X}_k^0 \neq \tilde{X}_l^0 \rightarrow Ke\tilde{X}_k \cap Ke\tilde{X}_l = 0\right];$$

$$\bigcup_{\tilde{X}_m^0 \in \tilde{R}^{|M|}} \tilde{X}_m^0 < X .$$

Ідея навчання за МФСВ полягає в організації ітераційної процедури пошуку на кожному краці навчання максимуму інформаційного критерію функціональної ефективності (КФЕ), який визначає екстремальні значення параметрів навчання, що безпосередньо впливають на достовірність прийняття рішень на екзамені. Такими параметрами в МФСВ є:

- параметри розділяючих гіперповерхней (РГП), в якості яких виступають еталонні вектори $\{x_m\}$, $m=\overline{1,M}$, вершини яких є геометричними і фізичними центрами класів розпізнавання $\tilde{X}_m^0$ і радіуси розділяючих гіперсфер d_m, які в кодовій відстані Хеммінга визначаються як

$d_m = \sum_{i=1}^{N} (x_{m,i} \oplus \lambda_{m,i})$, де $x_{m,i}$ - i-та координата вектора $x_m \in \tilde{X}_m^0$; $\lambda_{m,i}$ - і-та координата деякого вектора - реалізації образу, вершина якого знаходиться на РГП; N - кількість ознак розпізнавання; $\oplus$ - операція складання за модулем 2. Для спрощення будемо далі вживати означення $d_m = d(x_m \oplus \lambda_m)$;

© UkrObraz'2000

- система контрольних допусків на ознаки розпізнавання $\{\delta_{k,i}\}$, $i=\overline{1,N}$;

- рівні селекції первинних даних;

- крок квантування первинних даних як у часі, так і за рівнем;

- параметри плану навчання та інші.

В якості КФЕ розглянемо, наприклад, нормовану інформаційна міру Шеннона:

$$E=\frac{H_0-H(\gamma)}{H_0}, \qquad (1)$$

де $H_0=-\sum_{l=1}^{M}p(\gamma_l)\log_2 p(\gamma_l)$ - апріорна (безумовна) ентропія;

$H(\gamma)=-\sum_{l=1}^{M}\sum_{m=1}^{M}p(\gamma_l)p(\mu_m/\gamma_l)\log_2 p(\mu_m/\gamma_l)$ - апостеріорна (умовна) ентропія, що характеризує залишкову невизначеність після прийняття рішення; $p(\gamma_l)$ - безумовна ймовірність прийняття гіпотези γ_l; $p(\mu_m/\gamma_l)$ - апостеріорна ймовірність прийняття рішення μ_m за умови, що прийнята гіпотеза γ_l.

Для двохальтернативної системи оцінок (M=2) і рівноймовірних гіпотез, що характеризує найбільш важкий у статистичному сенсі випадок прийняття рішень, після заміни в (1) апостеріорних імовірностей на апріорні за формулою Байєса отримаємо [1]:

$$E=1+\frac{1}{2}\left(\frac{\alpha}{\alpha+D_2}\log_2\frac{\alpha}{\alpha+D_2}+\right.$$
$$+\frac{\beta}{D_1+\beta}\log_2\frac{\beta}{D_1+\beta}+\frac{D_1}{D_1+\beta}\log_2\frac{D_1}{D_1+\beta}+$$
$$\left.+\frac{D_2}{\alpha+D_2}\log_2\frac{D_2}{\alpha+D_2}\right), \qquad (2)$$

де α, β, D_1, D_2 - точнісні характеристики: помилки першого та другого роду, перша та друга достовірності відповідно. Отже критерій (2) є нелінійним функціоналом від точнісних характеристик процесу навчання. Крім того, він є неоднозначним, що потребує визначення в процесі навчання робочої області для його значень [2]. Оскільки навчальна вибірка є обмеженою за обсягом, то надалі замість точнісних характеристик будемо оперувати їх оцінками (емпіричними частотами):

$$\alpha=\frac{K_1}{n};\ D_1=\frac{K_2}{n};\ \beta=\frac{K_3}{n};\ D_2=\frac{K_4}{n}, \qquad (3)$$

де K_1, K_2 - кількість подій, які означають неналежність та належність реалізацій класу $\tilde{X}_m^0$ відповідно, якщо $\{x_m^{(n)}\}\in X_m^0$; K_3, K_4 - кількість подій, які означають відповідно належність і неналежність реалізацій класу $\tilde{X}_m^0$, якщо вони дійсно не належать класу X_m^0. Тоді після підстановки (3) в (2) одержимо робочу формулу для обчислення КФЕ

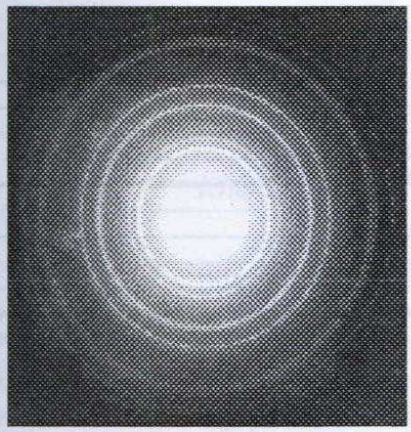
$$E_m=1-\frac{1}{2}\left(\frac{K_1}{K_1+K_4}\log_2\frac{K_1}{K_1+K_4}+\right.$$
$$+\frac{K_2}{K_2+K_3}\log_2\frac{K_2}{K_2+K_3}+\frac{K_3}{K_2+K_3}\log_2\frac{K_3}{K_2+K_3}+$$
$$\left.+\frac{K_4}{K_1+K_4}\log_2\frac{K_4}{K_1+K_4}\right). \qquad (4)$$

Отже на етапі навчання за МФСВ послідовно здійснюються процедури оптимізації наведених параметрів, починаючи з оптимізації РГП, оскільки ця процедура є обов'язковою частиною інших. Тобто процедури оптимізації інших параметрів спрямовані на наближення РГП до асимптотичної, яка забезпечує асимптотично максимальну достовірність рішень, що приймаються на екзамені. Після побудови оптимальних в інформаційному сенсі РГП МФСВ дозволяє застосовувати на екзамені відомі детерміновані алгоритми прийняття рішень, які обумовлюють виконання функції приналежності [1].

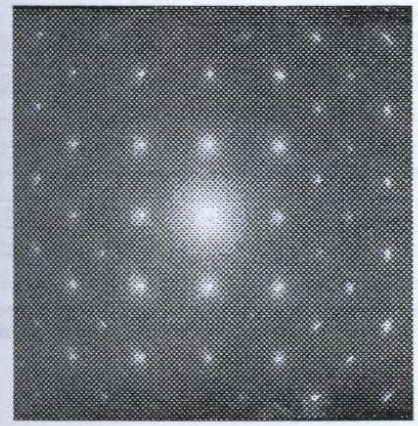
2 Обробка зображення

Обробку зображення розглянемо на прикладі електронограм зразків алюмінію (рис. 1,а) та золота (рис 1,б). Обробка зображення здійснюється з метою формування вхідних даних, які утворюють навчальну вибірку, в три етапи.

На першому етапі задається розмір рецепторного поля електронограми, здійснюється пошук центру симетрії зображення як точки відліку для подальшої обробки, і зчітування яскравості рецепторів.



а)



б)

Рисунок 1 - Електронограми: а) алюміній; б) золото.

На другому етапі формується діаграма розподілу яскравості L за радіусом R електронограми. При цьому для усереднення значень розподілу яскравості за кутом здійснюється перехід від прямокутних до полярних координат, що запобігає загубленню інформації. На рис.2 наведено розподіл яскравості за радіусом електронограми алюмінію, побудований на другому етапі. На третьому етапі з метою селекції піків яскравості застосовується спеціальний фільтр, який дозволяє усунути фоновий розподіл. На рис. 3,а наведено відфільтровані піки яскравості для електронограми алюмінію, а на рис. 3,б для електронограми золота.

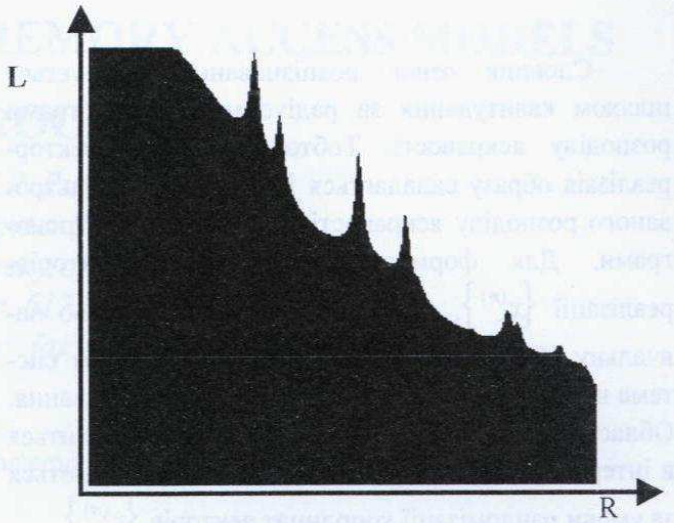
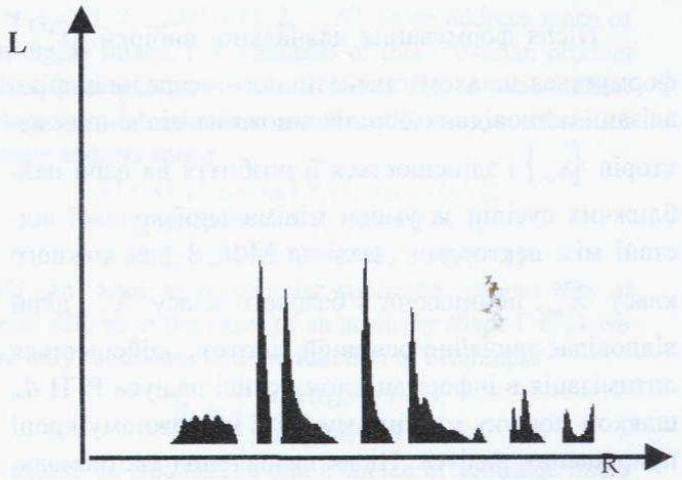
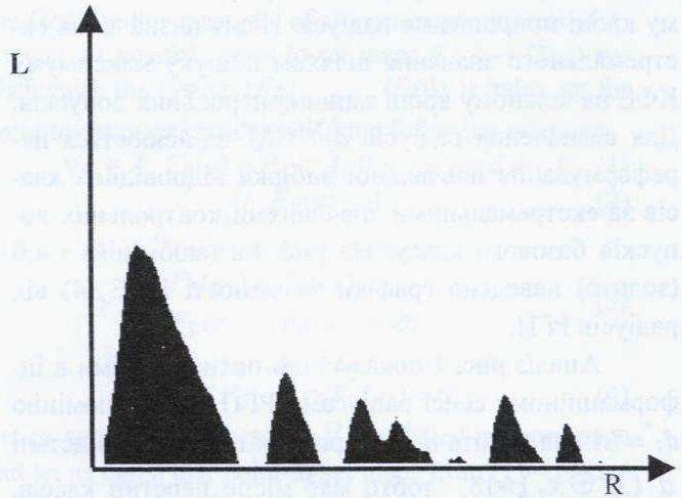


Рисунок 2. Вихідні дані другого етапу обробки зображення (алюміній).



а)



б)

Рисунок 3 - Вихідні дані третього етапу обробки зображення: а) алюміній; б) золото.

Словник ознак розпізнавання формується шляхом квантування за радіусом електронограми розподілу яскравості. Тобто первинний вектор-реалізаія образу складається із дискрет фідфільтрованого розподілу яскравості відповідної електронограми. Для формування вторинних векторів-реалізації $\{x_m^{(n)}\}$, які складають безпосередньо навчальну або екзаменаційну вибірку, задається система контрольних допусків на ознаки розпізнавання. Область значень контрольних допусків знаходиться в інтервалі 0-255 градацій яскравості і вибираються за умови рандомізації координат векторів $\{x_m^{(n)}\}$.

3 Реалізація алгоритму навчання.

Після формування навчальної вибірки $\{x_m^{(n)}\}$ формується шляхом статистичного усереднення реалізації відповідних образів множина еталонних векторів $\{x_m\}$ і здійснюється її розбиття на пари найближчих сусідів за умови мінімальної кодової відстані між векторами . далі за МФСВ для кожного класу $\tilde{X}_m^0$, починаючи з базового класу X_1^0, який відповідає звичайно основній гіпотезі, здійснюється оптимізація в інформаційному сенсі радіуса РГП d_m шляхом пошуку максимуму КФЕ на кожному кроці прирощення радіуса. Після визначення екстремального значення d_1^*, здійснюється оптимізація контрольних допусків пошуку максимуму КФЕ на кожному кроці прирощення радіуса. Після визначення екстремального значення шляхом пошуку максимуму КФЕ на кожному кроці зміни контрольних допусків. Для визначення радіусів $d_1^* \cdots d_M^*$ здійснюється переформування навчальної вибірки відповідних класів за екстремальними значеннями контрольних допусків базового класу. На рис. 4,а (алюміній) і 4,б (золото) наведено графіки залежності КФЕ (4) від радіусів РГП.

Аналіз рис. 4 показує, що оптимальними в інформаційному сенсі радіусами РГП є для алюмінію $d_1^*=7$ і для золота $d_2^*=9$ при між центровій відстані $d\ (x_1 \oplus x_2)=15$, тобто має місце перетин класів. Оптимальне значення контрольного поля допусків для номінального значення 30 дискрет складало10 градацій яскравості.

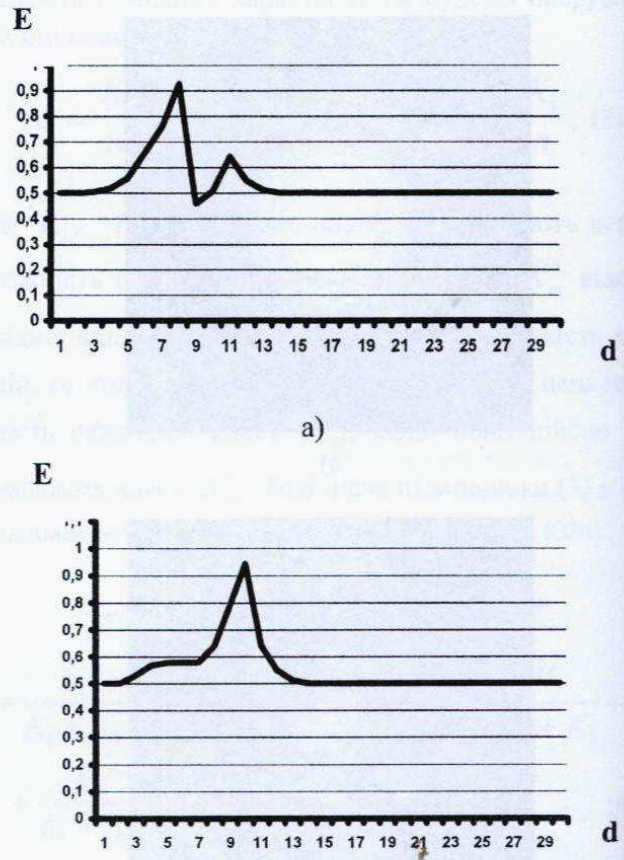


Рисунок 4 - Залежність КФЕ від радіусу РГП.

Висновок

Застосування МФСВ, який ґрунтується на оцінці інформаційної здатності системи розпізнавання, відкриває можливість розробки комп'ютерних систем класифікації електронограм різноманітних зразків матеріалів

Література

1. Краснопоясовський А.С. Технічна діагностика складних систем за методом функціонально-статистичних випробувань. // Обробка сигналів та розпізнавання образів: Праці третьої Всеукраїнської міжнародної конференції. - 1996.- С. 174-177.

2. Краснопоясовський А.С., Черниш А.В. Оцінка функціональної ефективності системи розпізнавання образів, що навчається. // Вісник Сум. держ. ун-ту. – 1997.- №2.- С. 112-118.

CONFLICTFREE PARALLEL MEMORY ACCESS MODELS FOR IMAGE PROCESSING

V.P. Mashtalir, E.P. Putyatin
Department of Computer Science,
Kharkov State Technical University of Radio Electronics,
14, Lenin Ave., Kharkov, 61726, Ukraine
e-mail: masht@skynet.kharkov.com; fax:(++038 0572) 409 113

Abstract. There are essential peculiarities when images of different nature are parallel processed. On numerous occasions subframe processing occupies a special role in filtering, restoration, correction etc. To decrease computing time, an approach to synthesis of memory structures is proposed. Single-cycle access is ensured for simultaneous processing of both all pixels belonging to arbitrary subframe and ensemble of pixels which are chosen per one from several overlapping subframes. Page scans are proposed as a toolkit to be exploited to obtain various memory models. All algorithms based on subframe processing can be formulated in page scans terms. Results presented can be easily extended to multidimensional data with parallel processing purpose.

INTRODUCTION

To date parallel and distributed computing (pipelined vector processors, transputers, PRAM (Parallel Random Access Machine), VSA (Virtual Systems Architecture), BSP (Bulk Synchronous Parallel) models etc.) continuously plays increasing role in image processing in general. If algorithms are based on the calculation of local image properties in all possible neighbourhoods of the image, there are essential peculiarities in processing of subframes (windows). From point of view of time overhead often memory access procedures occupy an essential long in the whole computing time. From earliest studies [1–3] the most steadfast attention was given to parallel memories for scanning fields or for data in matrix form. Classes of periodic, isotropic and linear functions were offered as toolkit for synthesis of conflictfree data exchange between the computing environment and memory modules [4–7].

We present point-to-set maps as toolkit for parallel memory structures creation for image processing. The subject in this study is 2-D signals but all got results can be easily extended to N-dimensional cases. Page scans are proposed to obtain memory modules with single-cycle access. Sequential subframe (all pixels of an arbitrary subframe are processed simultaneously) and parallel subframe (anywise pixels are chosen per one from a series of overlapping subframe and they are processed simultaneously) treatments can be formulated in terms of the proposed page scans.

I. PAGE SCANS AND MEMORY STRUCTURES

Let $A = \{1,2,...,M\} \times \{1,2,...,N\}$ be an address space of an digital image, i. e. elements of this Cartesian product identify displacement of all pixels B_{ij} in videobuffer. Moreover, let introduce subframe Ω as a subset of image address space

$$\Omega = \{i_1, i_2, ..., i_m\} \times \{i_1, i_2, ..., i_n\} \subseteq$$
$$\subseteq A,\ i_{s+1} = i_s + 1,\ j_{r+1} = j_r + 1,$$
$$s \in \{1,2,...,M-1\},\ r \in \{1,2,...,N-1\}.$$

We can analyse rectangular subframe without loss of generality as in the cases of an arbitrary shape $\Gamma \subset \Omega$ we are only to redetermine the function of brightness

$$B_{ij} = \begin{cases} B_{i,j}, (i,j) \in \Gamma, \\ 0, (i,j) \in \Omega/\Gamma. \end{cases}$$

It should be emphasised that a choice of subframe linear size substantially depends on task orientation and parallel execution environment.

Let $\Pi(A)$ be the aggregate of all non-empty subsets of A. Among all possible point-to-set maps $F: A \to \Pi(A)$ we distinguish the system $\{F_\beta\}_{\beta \in D(\beta)}$ ($D(\beta)$ is index set for point-to-set maps choice) satisfying following relations

$$\forall a \in A,\ F_\beta(a) = P_q \subset A,\ 0 \le q \le card\, A - 1; \quad (1)$$

$$\bigcup_{a \in A} F_\beta(a) = A; \quad (2)$$

$$\forall a_1, a_2 \in A : a_1 \ne a_2 \Rightarrow$$
$$\Rightarrow \begin{cases} F_\beta(a_1) = F_\beta(a_2), \\ F_\beta(a_1) \cap F_\beta(a_2) = \varnothing; \end{cases} \quad (3)$$

$$\forall \Omega \subseteq A,\ \forall a_1, a_2 \in \Omega : a_1 \ne a_2 \Rightarrow$$
$$\Rightarrow F_\beta(a_1) \cap F_\beta(a_2) = \varnothing \quad (4)$$

Let us agree that address set P_q is called memory page and let us agree that point-to-set maps from (1) – (4) are called page scans. The essence of a page scan lies in address relocation, when sampling, so that for given size relation all elements of an arbitrary subframe have to state in different real blocks of storage. Therefore, single-cycle access is attained without increasing of memory capacity.

For instance, let us indicate point-to-set maps as page scans pointed out

$$F_{\beta'}(a_{ij}) = P_q,\ a_{ij} \in P_q,\ q =$$

© UkrObraz'2000

$$= \mathrm{rem}\{(i+(j-1)m-1)/\mathrm{card}\ \Omega\},$$
$$F_{\beta''}(a_{ij}) = P_q,\ a_{ij} \in P_q,\ q =$$
$$= \mathrm{rem}\{(j+(i-1)n-1)/\mathrm{card}\,\Omega\},$$

where «rem(∘)» denotes a remainder under division.

To verify that $F_{\beta'}$ is the page scan, it suffices to check property (4) since relations (1) – (3) are obvious in this case. Hence, we have to examine implication

$$\forall a_{ij},\ a_{pq} \in \Omega \subseteq A \Rightarrow \begin{cases} F_{\beta'}(a_{ij}) = P_{t'} \\ F_{\beta''}(a_{ij}) = P_{t''} \end{cases} \Rightarrow t' \neq t''$$

where

$$t' = \mathrm{rem}\{(i+(j-1)m-1\},$$
$$t'' = \mathrm{rem}\{(p+(q-1)m-1\}.$$

Let us assume the contrary, namely, $t' = t''$. Then there is a divisibility of the value $|i+(j-1)m+(p-(q-1)m)|$ by $\mathrm{card}\,\Omega$ or, that is the same, the following equality holds

$$|(i-p)+m(j-q)| = r\,\mathrm{card}\,\Omega, \qquad (5)$$

where r is some natural number. In accordance with $\forall a_{ij},\ a_{pq} \in \Omega$ we have

$$|i-p| \le m-1, \quad |j-q| \le n-1. \qquad (6)$$

Taking into account (5) and (6) we get

$$|(i-p)+m(j-q)| \le |i-p|+m|j-q| \le .$$
$$\le m-1+m(n-1) = mn-1.$$

At the same time $mn-1 < \mathrm{card}\,\Omega$, consequently, $r=0$ and therefore $|i-p| = m|j-q|$. From inequalities (6) it follows $|m(j-q)| \le m-1$ that can be valid only under condition $|j-q|=0$, i.e. $q=j$ and similarly $p=i$. Thus we draw the conclusion about the contradiction to the initial assumption. As a consequence the point-to-set map $F_{\beta''}$ also generates the page scan.

Notice that according to (4) $\mathrm{card}\,\{F_\beta(a)\}_{a\in A} \ge \mathrm{card}\,\Omega$, hence we shall say: point-to-set map $F: A \to \Pi(A)$ induces an optimal page scan if $\mathrm{card}\,\{F_\beta(a)\}_{a\in A} = \mathrm{card}\,\Omega = mn$, in other words, if number of memory blocks equals number of subframe pixels.

Emphasise that point-to-set map satisfying (1) – (4) determines only the storage page therefore full address word is following diagonal product of maps

$$F_\beta \,\Delta\, \Im : A \to \Pi(A) \times \Re^\tau$$

where F_β is page scan, $\Im$ is some map which restriction

$$\Im\big|_{F_\beta^{-1}(a)\in A} : \{F_\beta(a)\}_{a\in A} \to \Re^\tau$$

is injection, i.e.

$$\forall a_1, a_2 \in \left\{F_\beta^{-1}(a)\right\}_{a\in A}, a_1 \neq a_2 \Rightarrow \Im(a_1) \neq \Im(a_2).$$

Introducing of τ-dimensional space into preceding map may be explained by hardware or system software environment specialities, for instance, by a hierarchical page segmentation, an insufficiency of digits per address word for direct access, masking of some pixels, etc.

Let us evaluate total number of all possible optimal page scans. Since property (4) is valid for arbitrary page scan, choose subframe as shown in Fig. 1. It is clear that there exist $mn!$ pixel reallocations for subframe chosen. Under displacement in direction ***a***) for one cell pages $P_{n+1}, P_{n+2}, \ldots, P_{2n}, \ldots, P_{(m-1)n+1}, P_{(m-1)n+2}, \ldots, P_{mn}$ are already fixed and it is necessary to fulfil pixel readdressing to pages $P_1, P_2, \ldots, P_n$. It is obvious that $n!$ variants are possible. But then to satisfy property (4) there is only one variant under displacement in direction ***b***). Judging

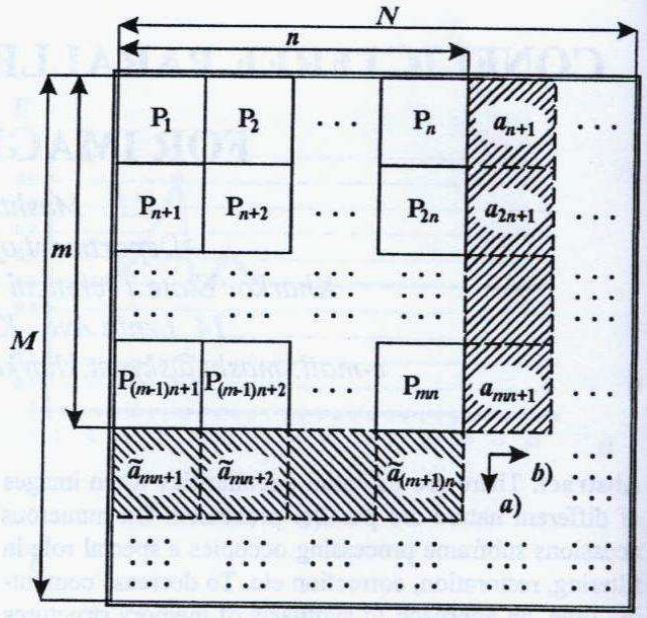


Fig. 1. Page scans number evaluation

similarly we have $m!$ variants of pixel readdressing to pages $P_{n+1}, P_{2n+1}, P_{3n+1}, \ldots, P_{(m-1)n+1}$ under initial displacement in direction ***b***) and we have one variant in direction ***a***). Hence, under condition $N \ge 2n$, $M \ge 2m$, implementing displacement $N-n$, $M-m$ times again, and repetition of initial subframe readdressing being accounted for, finally we get

$$K = (m!)^{N-n} + (r!)^{M-m} - 1.$$

Considering rest of possible relations between address space and subframes parameters, we summarise, up to $(mn)!$, total number of all possible optimal page scans

$$K = \begin{cases} (m!)^{N-n} + (n!)^{M-m} - 1,\ N \ge 2n,\ M \ge 2m; \\ (m!)^{N-n}[(2n-N)!]^{M-m} + \\ +(n!)^{M-m}[(2m-M)!]^{N-n} - 1, n \le N \le 2n, \\ \qquad\qquad m \le M \le 2m; \\ (n!)^{M-m}[(2m-M)!]^{N-n} + \\ +(m!)^{N-n} - 1, \qquad N \ge 2n,\ m \le M \le 2m; \\ (m!)^{N-n}[(2n-N)!]^{M-m} + \\ +(n!)^{M-m} - 1, \qquad n \le N \le 2n,\ M \ge 2m. \end{cases}$$

By virtue of what has already been said, we can draw conclusion that among all possible page scans it is desirable to choose only simplified (from point of view amount of address computation) multivalued maps.

II. PRODUCING PAGE SCANS BY ADDRESS INDICATORS

We shall focus attention on regular page scan in connection with that the calculation of pixel addresses should be realized as it is possible easier. Let us define parameter β, selecting page scans $\{F_\beta(a)\}_{a\in A}$, as vector-function $f(i,j) \in \Re^\tau, (i,j) \in A$. Let function $f(i,j)$ extend readdressing law to all address space under page scan given for arbitrary subframe. Let us agree to say that such function is address indicator and let us define it axiomatically:

i) restriction

$$f\big|_{\{1,2,\ldots,m\}\times\{1,2,\ldots,n\}} : \{1,2,\ldots,M\} \times \{1,2,\ldots,N\} \to \Re^\tau$$

is injection, i.e.

$$\begin{cases} \forall i_1, i_2 \in \{1,2,...,m\},\ \forall j_1, j_2 \in \{1,2,...,n\}: \\ : (i_1, j_1) \neq (i_2, j_2) \Rightarrow f(i_1, j_1) \neq f(i_2, j_2); \end{cases}$$

ii) address indicator has translation invariant relation of equality, i.e.

$$\forall (i_1, j_1), (i_2, j_2) \in \{1,2,...,m\} \times \{1,2,...,n\},$$
$$\forall (i_1 + \Delta i, j_1 + \Delta j), (i_2 + \Delta i, j_2 + \Delta j) \in A$$

if $f(i_1, j_1) = f(i_2, j_2)$, then

$$f(i_1 + \Delta i, j_1 + \Delta j) = f(i_2 + \Delta i, j_2 + \Delta j).$$

It is easily to prove that an address indicator $f(i,j)$ induced a page scan $F_\beta(a_{ij})$ determining by $A_{f(i,j)}$. Indeed, consider arbitrary subframe $\Omega \subset A$ with linear sizes $m \times n$. We are to check that from condition $\forall (i,j) \neq (p,q)$, $a_{ij}, a_{pq} \in \Omega$ it follows $f(i,j) \neq f(p,q)$. This inequality corresponds to assertion $F_\beta(a_{ij}) \neq F_\beta(a_{pq})$ (property (4) of page scans).

Let us assume the contrary, namely, $f(i,j) = f(p,q)$. Setting $\Delta i = \min(i,p) - 1$, $\Delta j = \min(j,q) - 1$ and always supposing $i' = i - \Delta i$, $p' = p - \Delta i$, $j' = j - \Delta j$, $q' = q - \Delta j$, we obtain $1 \le i' \le |p - i|$, $1 \le j' \le |q - j|$. Taking into account that a_{ij}, $a_{pq} \in \Omega$ we have

$$(i', j'), (p', q') \in \{1,2,...,m\} \times \{1,2,...,n\}$$

and from translation invariant relation of equality and the condition $f(i,j) = f(p,q)$ it follows $f(i', j') = f(p', q')$ which in accordance with property **i)** contradicts the assertion that $f(i,j)$ is an address indicator. Notice that since in this case a page scan is determined by relationship $F_\beta(a_{ij}) = A_{f(i,j)}$ then properties (1) – (3) are obvious. It now becomes clear that to find the regular page scan it suffices to check conditions **i), ii).**

Now let us consider an example of optimal address indicator (optimalness of address indicators is similar to that of page scan)

$$f(i,j) = n \operatorname{rem}\{i/m\} + \operatorname{rem}(j-1)/n\}. \qquad (7)$$

We first note that condition **i)** is obvious, indeed

$$f(\{1,2,...m\} \times \{1,2,...1n\}) =$$
$$= \begin{bmatrix} n & n+1 & ... & 2n-1 \\ 2n & 2n+1 & ... & 3n-1 \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ ... & ... & ... & ... \\ (m-1)n & (m-1)n+1 & ... & mn-1 \\ 0 & 1 & ... & n-1 \end{bmatrix}.$$

To complete the check it remains to show the validity of translation invariant relation of equality. Rewrite

$$f(i_1, j_1) = f(i_2, j_2)$$

in the form

$$n\left|\operatorname{rem}\left(\frac{i_1}{m}\right) - \operatorname{rem}\left(\frac{i_2}{m}\right)\right| = \left|\operatorname{rem}\left(\frac{j_1-1}{n}\right) - \operatorname{rem}\left(\frac{j_2-1}{n}\right)\right|.$$

The assumption: $\operatorname{rem}\left(\frac{i_1}{m}\right) - \operatorname{rem}\left(\frac{i_2}{m}\right) \neq 0$ implies that inequalities

$$n\left|\operatorname{rem}\left(\frac{i_1}{m}\right) - \operatorname{rem}\left(\frac{i_2}{m}\right)\right| \ge n,$$
$$\left|\operatorname{rem}\left(\frac{j_1-1}{n}\right) - \operatorname{rem}\left(\frac{j_2-1}{n}\right)\right| \le n-1$$

hold which contradicts $f(i_1, j_1) = f(i_2, j_2)$. Therefore, properties of remainders being accounted for, we have

$$\begin{cases} \operatorname{rem}\left(\frac{i_1+\Delta i}{m}\right) = \operatorname{rem}\left(\frac{i_2+\Delta i}{m}\right), \\ \operatorname{rem}\left(\frac{j_1-1+\Delta j}{n}\right) = \operatorname{rem}\left(\frac{j_2-1+\Delta j}{n}\right). \end{cases}$$

Multiplying the first equality by n and adding the second equality, we find

$$n \operatorname{rem}\left(\frac{i_1+\Delta i}{m}\right) + \operatorname{rem}\left(\frac{j_1-1+\Delta j}{n}\right) =$$
$$= n \operatorname{rem}\left(\frac{i_2+\Delta i}{m}\right) + \operatorname{rem}\left(\frac{j_2-1+\Delta j}{n}\right)$$

i.e. $f(i_1 + \Delta i, j_1 + \Delta j) = f(i_2 + \Delta i, j_2 + \Delta j)$, q.e.d.

It is easy to show that above address indicator is optimal. Indeed, from (7) it follows $f(A) \le mn - 1$. Further according to property **i)** the equality

$$\operatorname{card} f(\{1,2,...,m\} \times \{1,2,...,n\}) = mn$$

holds and since

$$\operatorname{card} f(A) \ge \operatorname{card} f(\{1,2,...,m\} \times \{1,2,...,n\})$$

taking into account zeroth page we find $\operatorname{card} f(A) = mn$.

Note that different problems require different subframe sizes. Therefore, rational compromise between relations of linear sizes of subframes and number of storage blocks can provide decreasing of time overhead for considerable number of image digital processing problems.

Let us agree to say that point-to-set map, satisfying (1) – (3), for which $\exists a_{ij}, a_{pq} \in \Omega$, and $F_\beta(a_{ij}) = F_\beta(a_{pq})$, is a quasipage scan. It is obvious that a quasipage scan F_β is optimal if

$$\forall \Omega \subset A,\ \forall F_\beta(a_{ij})\ \ \operatorname{card}\{\Omega \cap F_\beta(a_{ij})\} \le$$
$$\le \operatorname{Int}\{d/\delta + (d-1)/\delta\}, \qquad (8)$$

where $\delta = \operatorname{card}\{F_\beta(a_{ij})\}_{a_{ij} \in \Omega}$, $d = \operatorname{card}\Omega$.

Similarly, a map $\varphi : A \to \Re^\tau$, satisfying **i)**, is an address quasiindicator if

$$\forall \xi \in \varphi(\{1,2,...,m\} \times \{1,2,...,n\})$$
$$\operatorname{card}\varphi^{-1}\big|_{\{1,2,...,m\} \times \{1,2,...,n\}}(\xi) \le \operatorname{Int}(d/\delta + (d-1)/\delta).$$

It should be noted that a quasiindicator is more general readdressing model than an indicator (property **ii)** is weakened) and address quasiindicator generates quasipage scan. Moreover, their implementation is based on repeated use of page scans. Hence required point-to-set map is diagonal product

$$F_\beta \Delta \Phi_\beta : A \to \Re^\tau \times \{1,2,...,\operatorname{Int}(d/\delta + (d-1)/\delta)\}$$

for which equality $\operatorname{card}(F_\beta \Delta \Phi_\beta(a_{ij}) \cap \Omega) = 1$ holds.

Taking into account that in applications, as a rule, $N = 2^{\xi'}$, $M = 2^{\xi''}$, $n = 2^{\varsigma'}$, $m = 2^{\varsigma''}$ it is possible to find quasipage scans such that

$$d/\delta = \arg \min_{\beta \in D(\beta)} \operatorname{card}\{\Omega \cap F_\beta(a_{ij})\}.$$

In other words, each memory block contains the same quantity of pixels for arbitrary subframe.

III. GENERAL RELATIONSHIPS FOR SUBFRAME PROCESSING

To find general relationships for producing regular page scans, first let us analyse column subframe displacement in horizontal direction. Consider each column of an arbitrary subframe from address subspace $\{1,2,...,m\} \times \{1,2,...,n\}$ and connect with it the set of permutation of order m:

$$\pi_j = \begin{pmatrix} 1, & 2, & ..., & m \\ s_1, & s_2, & ..., & s_m \end{pmatrix}$$

where meaning of s_h, $h \in \{1,2,...,m\}$ lies in a number of memory pages. Moreover, let us consider a cyclic permutation $\pi_c = \begin{pmatrix} 1, & 2, & ..., & m-1, & m \\ 2, & 3, & ..., & m, & 1 \end{pmatrix}$. It is clear that permutation $\pi = \pi_c^{\alpha}\pi_j$ allows to get new columns from all other subframes. Here the degree α means quantity of shift cycles ($\alpha > 0$ up, $\alpha < 0$ down). Notice that there exists only the identity permutation of lines under displacement in vertical direction which satisfies (1) – (4) and saves the translation invariant relation of equality. Hence, the total quantity of regular page scans, up to $(mn)!$, is

$$K = n^{\text{sign}(M-m)} + m^{\text{sign}(N-n)} - \text{sign}(M-m)\text{sign}(N-n)$$

and, consequently, $K \le n + m - 1$.

From foregoing we are led to conclusion that general relationships for producing of regular page scans, up to bijection on $\{1,2,...,m\} \times \{1,2,...,n\}$, are written in the forms

$$F(a_{ij}) = n(\overline{\text{rem}}\{[i + \alpha_0(\lceil j/n \rceil - 1)]/m\} - 1) + \overline{\text{rem}}\{j/n\},$$
$$F(a_{ij}) = m(\overline{\text{rem}}\{[j + \alpha_0(\lceil i/m \rceil - 1)]/n\} - 1) + \overline{\text{rem}}\{j/m\},$$

where $\overline{\text{rem}}\{p/r\} = \begin{cases} \text{rem}\{p/r\}, & \text{if } \exists\omega \in \aleph : p \ne \omega r, \\ r & \text{if } \nexists\omega \in \aleph : p \ne \omega r; \end{cases}$

«$\lceil \circ \rceil$» denotes rounding up integer values; $\aleph$ is set of positive integer numbers. It should be noted that $\alpha_0(\lceil i/\underset{(\text{or } m)}{n} \rceil - 1) = \alpha$ is value from the basic permutation π given above.

Finally, we can indicate subframe processing in terms of page scans (point-to-set maps satisfying presented results).

Let $\{\Omega_\lambda\}_{\lambda\in\Lambda}$ be a set of subframes to be treated, where $\lambda \in \Re^\tau$ is any vector determining position of subframe in address space A, $\Lambda = \{\lambda_u\}_{u\in U}$, $\text{card}\,\Lambda = \text{card}\,U$, U is an index set. Then formula $\underset{u\in U}{\times} F_\beta\{\Omega_{\lambda_u}\}$, where «×» denotes Cartesian product of page scans

$$F_\beta : \Omega_{\lambda_u} \to \Pi(A),$$

describes optimal sequential subframe treatment of 2-D data. In the cases when inequality (8) is valid quasioptimal sequential-subframe processing is given in following form $\underset{u\in U}{\times}\ \underset{v\in V}{\times} F_\beta\left(\Psi\left(\Omega^v_{\lambda_u}\right)\right)$, where map

$$\Psi : \{\Omega_\lambda\}_{\lambda\in\Lambda} \to \Pi(A)$$

is optimal page scan for subframe chosen viz.

$$\Omega_{\lambda_u} = \bigcup_{v\in V} \Omega^v_{\lambda_u}, \quad V = \{1,2,...,\text{Int}(d/\delta + (d-1)/\delta)\},$$
$$\forall v_1, v_2 \in V\ \ \Omega^{v_1}_{\lambda_u} \cap \Omega^{v_2}_{\lambda_u} = \varnothing,$$
$$\forall v \in V\ \text{card}\,F_\beta\{\Omega^v_{\lambda_u}\} = \text{card}\,\Omega^v_{\lambda_u}.$$

Parallel subframe processing we can indicate as a set

$$\left\{\Phi_\gamma\left(\{\Omega_{\lambda_u}\}_{u\in U}\right)\right\}_7$$

of point-to-set maps (here $\gamma = \overline{1,\Gamma}$; $\Gamma = \text{card}\,\{\Omega_{\lambda_u}\}_{u\in U}$) for which

$$\forall\gamma \in \{1,2,...\Gamma\}\ \Phi_\gamma : \bigcup_{\lambda_u}(\Omega_{\lambda_u}) \to \Pi(A),$$
$$\Phi_\gamma\left(\bigcup_{\lambda_u}\Omega_{\lambda_u}\right) = \bigcup_{\lambda_u}\Phi_\gamma(\Omega_{\lambda_u}), \qquad (9)$$
$$\forall\gamma \in \{1,2,...\Gamma\}\ \text{card}\,\Phi_\gamma\{\Omega_{\lambda_u}\}_{u\in U} =$$
$$= \text{card}\,\{\Omega_{\lambda_u}\}_{u\in U},\ \forall\gamma, \forall\lambda \in \Lambda\ \Phi_\gamma(\Omega_\lambda) = P_q \in \Omega_q, \qquad (10)$$
$$\forall\gamma \in \{1,2,...\Gamma\}, \forall u_1, u_2 \in U : u_1 \ne u_2 \Rightarrow$$
$$\Rightarrow F_\beta\left(\Phi_\gamma\left(\Omega_{\lambda_{u_1}}\right)\right) \ne F_\beta\left(\Phi_\gamma\left(\Omega_{\lambda_{u_2}}\right)\right), \qquad (11)$$
$$\forall\lambda \in \Lambda, \forall\gamma_1, \gamma_2 \in \{1,2,...\Gamma\} : \gamma_1 \ne \gamma_2$$
$$\Rightarrow \Phi_{\gamma_1}(\Omega_\lambda) \ne \Phi_{\gamma_2}(\Omega_\lambda). \qquad (12)$$

Now explain properties (9) – (12). Essence of relationship (9) consists in additivness which, as a consequence, provides quasioptimal parallel-subframe processing. Formulae (10) describe structure of maps Φ_γ, namely, one element from given subframe is connected with this subframe. Implication (11) shows the choice of different storage pages for all subframes. Property (12) indicates action alteration of map Φ_γ when being passed to next set of elements to be processed.

CONCLUSION

Based on point-to-set maps models of conflictfree access to memory modules for parallel image processing have been proposed. Page scans are introduced as a toolkit to be exploited to obtain various memory models. Presented results provide reasonable choice of hardware, software and brainware to be used in task oriented computers for real time mode. All kinds of subframe processing can be formulated in terms of page scans. Single-cycle access is provided for simultaneous processing of both all pixels belonging to a subframe and set of pixels which are chosen per one from several overlapping subframes.

All results can be easily extended to multidimensional strutures of data with purpose of subset parallel processing.

REFERENCES

1. Kuck D. J. Illiac IV software and application programming // IEEE Trans-actions on Ccmputers. – vol. C-17, No 5, 1968. – P. 758-770.
2. Budnik P., Kuck D. J. The organization and use of parallel memories // IEEE Transactions on Computers. – vol. C-20, No 12, 1971. – P. 1566-1569.
3. Voorhis D.C., Morrin T.H. Memory system for image processing // IEEE Transactions on Computers. – vol. C-27, No 1, 1971. – P. 113-125.
4. Rebel B., Gösser M. On conflictfree parallel memory access // Computer and Artificial Intelligence. – vol. 2 , 1983. – P. 359-401.
5. Gösser M., Rebel B. Data structures and parallel memories // Proc. of the III International Workshop on Parallel Automata and Arrays. – Akademie-Verlag Berlin 1986. – P. 49-60.
6. Alexander W.E., Reeves D.S., Gloster C.S. (Yr) Parallel image processing with with the block data parallel architecture // Proc. of the IEEE (Special Issue on Parallel Architectures for Image Processing). – vol. 84, No 7, 1996. – P. 947-968.
7. Marion-Poty V., Miguet S. Data allocation strategies for parallel image processing algorithms // International Journal of Pattern Recognition and Artificial Intelligence. – vol 9, No 4 (Issue: Parallel Image Analysis: Theory and Applications), 1995. – P. 615-634.

Pattern recognition by method of probable algorithm quantums of knowledges (PAQK-method)

I.B.Sirodzha, B.Bar, I.I. Sirodzha
State Aero-Space University
named after Zhukovsky N. Yeg. «KhAI»
61070, Kharkov, Chkalova str.,17, chair 603,
phone (0572) 44-27-34

In the difference of widely known modeling of human's brain functions, using taught **neuron nets** for pattern recognition [1,2] quantum approach [3-5] is developing for modeling human generation of cause-consequence reasoning in processes of making decisions with help of taught on examples of **logical nets of probable reasoning (LNPR).**

The method of probable algorithm quantums of knowledges (PAQK-method) is proposed for pattern recognition using **v-quantum net of decisions' output** (v-QNDO), generated during the process of teaching LNPR on the totality of **scenario examples of taught knowledges (SETK).**

So instead of finding indexes of synoptical neurons' connections with given topology of net PAQK-method provides restoring beforehand unknown topology v-QNDO and definition of probable reliability's estimations of it's v-quantums as events. Taught and optimized on abundance v-QNDO is a base of probable quantums of knowledges (BPQK) and, at the same time, mechanism of the output of recognizing decisions with counting indexes of their reliability(IR).

1.INTRODUCTION

Great requirements in working out methods of decision poorly structured and unstructured problems of pattern recognition under uncertain conditions exist[6]. Let's distinguish a class $K(\alpha)$ of problems connected with α-undefinition, that slipulates for probable character, variety of types, uncompleteness, careless of initial information, with unknown informational signs and deciding rules, alternative and formal purposal criterions , and also with demand in knowledges of experts in concrete subject field.

The peculiarity of suggested PAQK-method of recognition concludes in attracting for outputing finding rules of making recognition decisions not only digital data of measurings and observations but portions (quantums) of knowledges of specialists in given subject field in the form of probable saying and cause-consequence conclusions of productional character, that have relations with finding decisions. With it using PAQK-method the task of indivisible formalization, presenting and algorithm manipulating of digital and lingvistic information in the form of variety of levels on complication **probable algorithm quantums of knowledges (vk-knowledges)** using vector-matrix operators is decided.

In the difference from traditional presenting teaching LNPR makes with given SETK and consists in definition of parameters net's vk-knowledges-knots, their quantity and cause-consequence connections between them. Teaching is finished when **structured completeness v-QNDO** is achieved, that is when equivalent count G, that has numeral function [7], which last level of summits keep only purpose v-quantums-conclusions, corresponds to the net.

Substantially it means, that with structured completeness v-QNDO admission of net's knots and logical connections between them, and from any v-quantum-knot exists way to any purpose v-quantum-consequence, is found.

So in structure completed v-QNDO **input** knots correspond to sending v-quantums of initial signs OMD or facts, **output** – v–quantums of purpose conclusions, and inside knots – v–quantums of intermediate confirmations or partial decisions.

Peculiarity vk–knowledges is in presence inbuild procedures for realization of exact countings of intermediate and concluding decisions' probables on the base of equivalent transformations of logical functions statisticallly depended v–quantums chains to independent by ortogonizating of corresponding logical functions .

2. ORGANIZATION OF THE TASK

It is necessary to decide two basic tasks of formaly presenting different-levels vk-knowledges (A_V**–task)** and algorithm synthesis of the rule of

© UkrObraz'2000

making recognition decisions on the base of using vk–knowledges (B_v **–task).**

Formally A_v –task for vk–knowledges let's present as multitude five

$$A_v = \langle S_v, K_v, \tilde{V}, \Pi_v, Q_v \rangle, \quad (1)$$

where S_v – symbol language v–quantums of knowledges, that consists of final multitude of letters, digits and symbols of the algorithm theory's operations;

K_v – final multitude of terminal v–quantums of knowledges ; set before experiment;

$\tilde{V}$ – probable multitude, that birth meanings a function $p_v(\bullet)$ of events reliable, described by different-levels v–quantums, that is multitude of indexes of reliability of vk–knowledges from the interval [0, 1];

Π_v – rules of designing of different-levels v–quantums;

Q_v – multitude of semantically codes and special structures of data, that determine level and maintenance v–quantums structures of knowledges with algoritmization.

In other words, in A_v –task it is necessary to create the formal mechanism of building class M_v of different-levels vk–knowledges in the language S with meanings of reliabiliti's degree from $\tilde{V}$ on the base of using rules Π_v to terminal v–quantums from K_v, and also with needed level and contents of v–quantums with corresponded semantic codes from Q_v.

Probable multitude $\tilde{V}$ on universal multitude U={u} is called totality of pairs of such type

$$\tilde{V} = \{\langle u \mid p_v(u) \rangle\}, \quad (2)$$

where $p_v(u): U \to [0,1]$.

Showed reflection U to segment [0, 1] we'll call **function of reliability** $p_v(u)$ probable multitude $\tilde{V}$, and meaning p(u) for concrete element $u \in U$ – **index of reliability** (IR) p(u) value u.

Different-**types characteristic** of the object of recognition (OR) in examples cause-consequence sayings of experts or judgment about outputing finding decisions with some appraisal of their probable from other sources, and also **quantity** and **ccontents requirements conclusions at the form SETK** in decided B_v –task from concrete subject field are given.

It is necessary to work out total method of algorithm decision of B_v –task, and also to synthesis and ground in terms of PAQK-method necessary basic algorithms: teaching LNPR (ALOBUCH), automatic quantuming (ALAKVA), optimization of the net (ALAKVA) , outputing recognition rule (AL(B_v)) and management of functioning v–QNDO (ALUPR).

3. METHOD OF DECISION A_v –, B_v –TASKS

Method of decision A_v –task is based on theoretic positions, given in works [3-6, 8], with using operators superposition (**P–operator**), line operator of concatenation (**CON <•> –operator**) and column (**CON[•] –operator**), and also new terminal v–quantums from multitude

$$K_v = \{vk_0a, vk_1y, vk_1b\}. \quad (3)$$

Here **terminal choosing v-quantum of knowledges of the nil level** vk_0a **with semantic**: «from ℓ-measured totality observed probable meanings of signs $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ to choose meaning α_k with reliable's index $p(\alpha_k) \in [0,1]$»
has such form:

$$vk_0a = [V_k^{(\ell)}(\alpha_1 \mid p(\alpha_1), \ldots, \alpha_k \mid p(\alpha_k), \ldots, \alpha_\ell \mid p(\alpha_\ell)) = \alpha_k \mid p(\alpha_k)], \quad (4)$$

where $vk_0 \in Q_v$, a , $V_k^{(\ell)} \in S_v$.

Terminal vector v–quantum of knowledges of the first level vk_1y with semantic : «observed OR $y \in \Omega$ j, has n different-types sings $x_j, (j = \overline{1,n})$, which meanings $\alpha_{k_j}^j \in \tilde{X}^j, (k_j = \overline{1, r_j})$ with correspondence IR $p(\alpha_{k_j}^j) \in [0,1]$» has a form

$$vk_1y = [d_1^{x_1} : d_2^{x_2} : \ldots : d_n^{x_n}] = [\alpha_1^1 \mid p(\alpha_1^1), \ldots, \alpha_{r_1}^1 \mid p(\alpha_{r_1}^1) : \alpha_1^2 \mid p(\alpha_1^2), \ldots, \alpha_{r_2}^2 \mid p(\alpha_{r_2}^2) : \ldots : \alpha_1^n \mid p(\alpha_1^n), \ldots, \alpha_{r_n}^n \mid p(\alpha_{r_n}^n)], \quad (5)$$

where $vk_1 \in Q_v$ name $y \in S_v$; $\alpha_i^j \mid p(\alpha_{r_i}^j) \in \tilde{V}$.

Terminal characteristic v–quantum of the first level vk_1b with the name b has a form

$$vk_1b = [\chi_{\tilde{Y}_j}(\alpha_k^j \mid p(\alpha_k^j))] = \begin{cases} 1, \text{if } \alpha_k^j \mid p(\alpha_k^j) \in \tilde{Y}_j, \\ 0, \text{if } \alpha_k^j \mid p(\alpha_k^j) \notin \tilde{Y}_j, \end{cases} \quad (6)$$

where $\tilde{Y}_j = \{\alpha_k^j\}$ - fixed probable multitude k's meanings j's sing, $vk_1 \in Q_v, b, \chi_{\tilde{Y}_j} \in S_v$ with **semantic**:

«meaning α_k^j j's sign OR in given moment is fixed by observing with IR $p(\alpha_k^j)$, if meaning of characteristic function $\chi_{\tilde{Y}_j}=1$; otherwise if $\chi_{\tilde{Y}_j}=0$ the result of observing are ignored as doubtful».

Definition 1. Algorithmic constructions, got from components of the multitude K_v (3) by final quantity of using to them P–operator, CON⟨•⟩–operator and CON[•]–operator in frameworks A_v–task (1), are called different-levels vk–knowledges. Class vk–knowledges M_v, made on definition1 1,let's call **standard class PAQK–models of** presenting of knowledges about object under recognition (OR).

For example, by expression

$$vk_1cl=$$

$$=[\overbrace{0,1|p(\alpha_2^1),1|p(\alpha_3^1),\varnothing:}^{x_1}\overbrace{1|p(\alpha_1^2),0,\varnothing:\varnothing|\varnothing:}^{x_2}\overbrace{\alpha_1^{3u}|p(\alpha_1^{3u}),\varnothing:}^{x_3=x_u}$$

$$\underbrace{:p(\to cl);A(cl)p(cl)}_{\text{output domain}}$$

(7)

is presented widen vector v-quantum of knowledges of the first level with name of purposed consequence c1 on purposed sign x_u, correspondence to three components sending sign x_1 and two components sign x_2 OR.

Semantic vk_1c1 (7) is: «**IF** observing the second knowledge α_2^1 sign x_1 with IR $p(\alpha_2^1)$ **OR** third knowledges α_3^1 with OR $p(\alpha_3^1)$ **AND** first meaning α_1^2 of sign x_2 **THEN** OR has one component purposed sign x_u with IR $p(\alpha_u^n)$; with it output IR p(cl) v–quantum is counted by algorithm A(cl), using set probability $p(\to cl)$ of implication, that is probability of joint making consequence c1 with sendings $\alpha_2^1|p(\alpha_2^1)$ **OR** $\alpha_3^1|p(\alpha_3^1)$ **AND** $\alpha_1^2|p(\alpha_1^2)$» for given v-quantum in multitude Q_v contended correspondent program and structure of data, that realize set sence and execution of the algorithm A(cl).

Inputed class M_v vk–knowledges has multitude present at the form of N–measured space $\tilde{X}^{(N)}$ PAQK–models. It guaranties comfortable for computer formalization of knowledges about decided rules, events and conformities to natural lows of the subject field by means of vector-matrix (look. view (7)) and analytical final-predicate presents[3-5].

The method of decision of B_v–task is based on definitions 2 and 3, and also on worked out algorithms ALOBUCH, ALAKVA, ALAKVA, $AL(B_v)$ and ALUPR

Definition 2. Teaching LNPR is called orientated count $G=(E,\Gamma)$, synthesis on given SETK using algorithm ALOBUCH and **had numeral function** $\Pi(X_i)$ $\forall X_i\in E$, determined on submultitudes-levels $N_1,N_2,\ldots,N_k\subset E$ of summits, and also next **properties**:

1) all summits (knots of the net) $X_i\in E,(1\le i\le n)$ correspond to sayings from SETK of concrete subject field, and **arcs** from $\Gamma: E\to E$ show on cause-consequence connections between knots with logical connections «AND», «OR», «NOT»;

2) all knots $X_i\in N_1\subset E$ with $\Gamma^{-1}N_1=\varnothing$ correspondent **output sending information** $e_i,(1\le i\le n_n)$ relatively some consequences $c_j,(1\le j\le m_c)$ with given IR $p(e_i)$ and $p(\to c_j)$;

3) all knots $X_i\in N_k\subset E$ with $\Gamma N_k=\varnothing$ are **purposing** (output) knots-conclusions C_s, $(1\le s\le S)$ with counted IR $p(C_s)$, and all summits intermediate levels between N_1 and N_k correspond to **partial consequence** c_j.

Syntheses LNPR at the mode of teaching has to be transformed in v–QNDO with help of special algorithm ALAKVA, that automatically determine **not given beforehand** quantity v–quantum **unknown beforehand levels** on **given rules** and **known knots of the** concrete levels of numeral function $\Pi(X_i)$ of the count G.

Definition 3. Purpose-oriented v-QNDO is called the result of count's transformation $G=(E,\Gamma)$ using algorithm ALAKVA to the count $G_{vk}=(E_{vk},\Gamma_{vk})$, that has following properties:

1) all summits $X_i\in E_{vk}$ correspond generated different-levels v-quantums, contended SETK of concrete subject field, and arcs $u_{ij}\in\Gamma_{vk}$ show at the logical connections v-quantums events;

2) all $X_i\in N_1\subset E_{vk},\Gamma^{-1}N_1=\varnothing$ correspond to **input vk-knowledges-sendings** with names $e_i,(1\le i\le n_n')$ relatively **vk-knowledges-consequences** with names $C_j,(1\le j\le m_c')$ with given indexis of reliability;

3) all $X_i\in N_k\subset E_{vk},\Gamma N_k=\varnothing$ are **purpose vk-knowledges-conclusionsи** with names и $C_s,(1\le s\le S)$ and counted IR $p(C_s)$, and all intermediate summits of

count G_{vk} corresponds to **partial vk-knowledges-consequences.**

Generated v–QNDO is optimized on criterion of abundance with the help of algorithm ALOPT and with management of algorithm ALUPR directly is used for outputing recognition decisions with counting correspondence probable appraisals of IR.

Peculiarity of manipulating vk-knowledges at the process of building and making decision by v-QNDO is **new way** of not approximate, but exactly **definition of probable appraisals** intermediate and purpose v–quantums conclusions.

It realized on the base of **ortogonalization** of logical functions, described connections between v–quantums events (sayings). With it the logic of sayings in v-QNDO is presented inпредставляется в **orthogonal dizunctional normal form**(ODNF).

Counting of meanings of these probabilities by ODNF is make by setting instead of its variables corresponded known probabilities with changing operations of logical multiplication and adding by analog arithmetic operators.

It is grounded like this. Let any event **e**:=«object $\omega \in \Omega$ has sign x» is made with probability $p(e)$, and event **c**:=«if ω ohas sign x, that ω corresponds to category **A**» is set with conditional probability $p(e \to c)$, that is called **probability of implication.** Then probability $P(e \wedge c)$ setting events at the same time **e** and **c** (as other event **z**) is determined by correlation

$$p(z) = P(e \wedge c) = p(e) \cdot p(e \to c). \qquad (8)$$

If now for showed events **e** and **c** to correspond logical variables α and β correlatively, then the probability **p(z)** (8) could be counted as probability of logical conjunction $\alpha \bullet \beta$ to be true meaning «1» on the any admission of meanings of its variables α and β with giving probabilities of appearing their true meanings.

For more complicated event **z1:=** «**if** object ω has sign x_1 **or** x_2, **or** ... **or** sign x_k, **then** ω belongs to class **A**» probability **p(z1)** can be counted on logical function $\alpha_1 \cdot \beta_1 \vee \alpha_2 \cdot \beta_2 \vee \ldots \vee \alpha_k \cdot \beta_k$, transformed to equivalent to it ODNF

$$\alpha_1\beta_1 \vee \overline{\alpha_1\beta_1} \cdot \alpha_2\beta_2 \vee \overline{\alpha_1\beta_1} \cdot \overline{\alpha_2\beta_2} \cdot \alpha_3\beta_3 \vee \ldots$$
$$\ldots \vee \overline{\alpha_1\beta_1} \cdots \overline{\alpha_{k-1}\beta_{k-1}} \cdot \alpha_k\beta_k.$$

At the result of deducting outputing **probable identified decisions** relatively OR ω with given sending vk–knowledges $vk_1Y_{B\omega}$ in B_v–task with help of algorithm АЛ(B_v) at the output v–QNDO purpose vk–knowledges are formed $vk_sR_{B\omega}, (s = 0,1)$ as one or some v–quantums with counted resulted IR $p(R_{B\omega})$. Semantic $vk_sR_{B\omega}$ shows at the **belonging identified OR** ω to definition category or **class** K from Ω with given reliability.

Let while decision B_v–task are used classified **SETK**(B_v) and at the result of studying was got correspondence v–QNDO as functional BPQK(B_v) for **recognition** (identification) of **situation** in conditions α–undefinition using IND–operator [3-4] as

$$BPQK(B_v) = \text{SETK}(B_v) \xrightarrow[\text{ALOBUCH,ALAKVA,ALOPT}]{\text{IND}} \text{BPQK} \qquad (9)$$

Then formally algorithm decision B_v–task concludes in getting purpose vk–knowledges $vk_sR_{B\omega}$ by DED–operator [3] as

$$vk_s \; R_{B\omega} = \text{DED}(\text{BPQK}(B_v), vk_1Y_{C\omega}; \text{AL}(B_v), \text{ALUPR}, vk_sR_{B\omega}):$$
$$= BPQK(B_v) \xrightarrow[vk_1Y_{B\omega}; \text{AL}(B_v), \text{ALUPR}]{\text{DED}} vk_sR_{B\omega}. \qquad (10)$$

Consequently, IND- and DED-operators are realized by algorithms, showed under the pointer in (9) and (10) formulas.

4. SOFTWARE AND EXPERIMENTAL INVESTIGATION OF PAQK-METHOD.

On the base of using adduced method in p. 3 interactive program complex (IPK) «V–KVANT» was created as investigational prototype intellectual (SSMD) on the base of using PC. Total architecture of IPK is presented on p. 1.

IPK successfully was tested during experimental decision of series tests and real tasks of patterns recognition.

Software «V–KVANT» is a kernel of new informational technology of knowledges' engineering, based on PAQK–method.

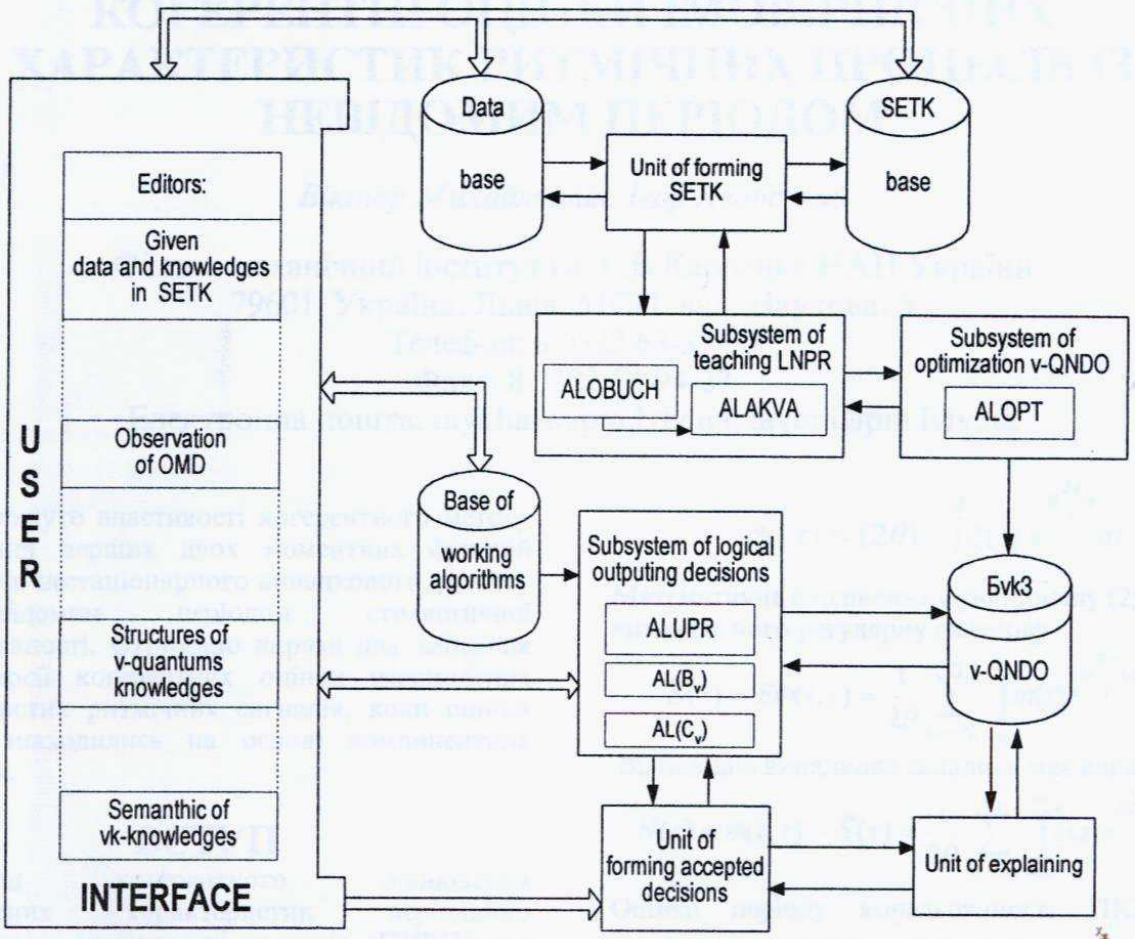


Fig. 1. Total architecture of IPK "V-KVANT"

CONCLUSION.

New PAQK–method of patterns recognition is worked out, on the base using which acting IPK «V–KVANT» was worked out for decision productional tasks of supporting making recognition and prognosing decisions. The results of experimental appraisal and comparement with other methods showed high level, affectivities and advanteges of PAQK–method. The last one conclude in decreasing average risk of recognition on the order, increasing speed of action of the teaching process in 2.5 times and decreasing volume of using memory on BPQK in 1,5 times.

LITERATURE

1. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика. – М.: Мир. –1992. –240с.

2. Горбань А. Н. Обучение нейронных сетей. – М.: СП Пара Граф.–1990. –160с.

3. Сіроджа І.Б. Знанняорієнтований підхід до проблеми розпізнавання образів // Праці Третьої Всеукраїнської Міжнародної конференції "Оброблення сигналів і зображень та розпізнавання образів" (УкрОбраз'96). – Київ.: Кіб. центр НАН України.–1996.– С. 30-33.

4. Сіроджа І.Б. Розпізнавання образів методом різнорівневих алгорітмічних квантів знань // Праці Четвертої Всеукраїнської Міжнародної конференції "Оброблення сигналів і зображень та розпізннавання образів" (УкрОбраз'98). – Київ.: Кіб. центр НАН України.–1998.– С. 17-19.

5. Сироджа И.Б. Квантовый подход к решению задач инженерии знаний // Проблемы бионики. – 1999.– Вып. 50. – С. 35-42.

6. Ларичев О.Н., Мошкович Е.М. Качественные методы принятия решений. – М.: Наук, Физмат.– 1996.–208с.

7. Кофман А. Введение в прикладную комбинаторику.–М.: Наука, 1975. – 480с.

8. Сироджа И.Б., Петренко Т.Ю. Квантовые модели нечетких знаний для принятия приближенных решений при недостатке данных // Проблемы бионики.– Харьков, 1998.– №48 – С. 129–140.

КОГЕРЕНТНІ ОЦІНКИ ІМОВІРНІСНИХ ХАРАКТЕРИСТИК РИТМІЧНИХ ПРОЦЕСІВ ІЗ НЕВІДОМИМ ПЕРІОДОМ

Віктор Михайлишин, Ігор Яворський

Фізико-механічний інститут ім. Г.В.Карпенка НАН України
79601, Україна, Львів, МСП, вул. Наукова, 5
Телефон: 8 0322 63-33-55
Факс: 8 0322 64-94-27
Електронна пошта: mykhai@ipm.lviv.ua, iavor@ipm.lviv.ua

Розглянуто властивості когерентного методу оцінювання перших двох моментних функцій періодично нестаціонарного випадкового процесу з невідомим періодом стохастичної повторюваності. Отримано вирази для зміщення та дисперсії когерентних оцінок імовірнісних характеристик ритмічних сигналів, коли оцінки періоду знаходились на основі компонентних статистик.

ВСТУП

Метод когерентного оцінювання ймовірнісних характеристик періодично корельованих випадкових процесів (ПКВП) дає змогу здійснювати оброблення сигналів безпосередньо у часовій області. Він грунтується на усередненні відліків реалізації через період корельованості. Попереднє оцінювання періоду впливає на властивості когерентних оцінок імовірнісних характеристик ПКВП.

КОГЕРЕНТНА ОЦІНКА МАТЕМАТИЧНОГО СПОДІВАННЯ

Когерентна оцінка математичного сподівання періодично корельованого випадкового процесу (ПКВП) при апріорі невідомому періоді корельованості визначається статистикою

$$\hat{m}(t) = \frac{1}{N}\sum_{n=0}^{N-1}\xi(t+n\tau_m), \qquad (1)$$

де N - кількість відліків , що усереднюються, τ_m - оцінка періоду корельованості.

У роботі [1] показано, що для визначення періоду корельованості можуть використовуватися функціонали, вид яких, аналогічний компонентним оцінкам відповідних імовірнісних характеристик ПКВП:

$$\hat{m}(t,\tau) = \sum_{l=-N_1}^{N_1}\hat{m}_l(\tau)e^{il\frac{2\pi}{\tau}t}, \qquad (2)$$

де

$$\hat{m}_l(\tau) = (2\theta)^{-1}\int_{-\theta}^{\theta}\xi(t)\,e^{il\frac{2\pi}{\tau}t}\,dt$$

Математичне сподівання функціоналу (2) визначає його регулярну складову

$$\tilde{S}(\tau) = E\hat{m}(t,\tau) = \frac{1}{2\theta}\sum_{l=-N_1}^{N_1}\int_{-\infty}^{\infty}m(s)e^{il\frac{2\pi}{\tau}(t-s)}\,ds.$$

Відповідна випадкова складова має вигляд

$$\tilde{N}(\tau) = \hat{m}(t,\tau) - \tilde{S}(\tau) = \frac{1}{2\theta}\sum_{l=-N_1}^{N_1}\int_{-\infty}^{\infty}o\xi(s)e^{il\frac{2\pi}{\tau}(t-s)}\,ds.$$

Оцінка періоду корельованості ПКВП τ_m є розв'язком нелінійного рівняння

$$\left[\frac{dS(\tau)}{d\tau} + \varepsilon\frac{dN(\tau)}{d\tau}\right]_{\tau=\tau_m} = 0, \text{ де } S(\tau) = \frac{\tilde{S}(\tau)}{\tilde{S}(T)},$$

$$N(\tau) = \frac{\tilde{N}(\tau)}{\sqrt{E\tilde{N}^2(\tau)}}, \quad \varepsilon = \frac{\sqrt{E\tilde{N}^2(\tau)}}{\tilde{S}(T)}.$$

При виконанні умови $\lim_{|u|\to\infty} b(t,u) = 0$, $\varepsilon \to 0$ при $\theta \to \infty$. Для достатньо великих θ параметр $\varepsilon << 1$ і розв'язок нелінійного рівняння знаходимо у вигляді ряду

$$\tau_m = \tau_0 + \varepsilon\tau_1 + \varepsilon^2\tau_2 + \ldots.$$

Невідомі величини першого наближення $\tau_0 = T + \Delta T$ і τ_1 знаходимо із системи лінійних рівнянь: $\Delta T = -\frac{s_1}{s_2}$, $\tau_1 = -\frac{n_1}{s_2}$, де

$$s_i = \left[\frac{d^i S(\tau)}{d\tau^i}\right]_{\tau=T}, \quad n_i = \left[\frac{d^i N(\tau)}{d\tau^i}\right]_{\tau=T}$$

Представимо функціонал (1) у вигляді

$$\hat{m}(t) = \frac{1}{N}\sum_{n=0}^{N-1}\int_{-\infty}^{\infty}\xi(t+nx)\,\delta(x-\tau_m)dx.$$

Скористаємося формальним розкладом δ - функції у ряд Тейлора

© UkrObraz'2000

$$\delta(x-\tau_m)=\delta(x-T)-\delta'(x-T)(\tau_m-T)+$$
$$+\frac{1}{2}\delta''(x-T)(\tau_m-T)^2-\ldots$$

Оскільки оцінка періоду корельованості подана у вигляді ряду по степенях малого параметра ε, обмежуючись першим наближенням, запишемо

$$\delta(x-\tau_m)=\delta(x-T)-\varepsilon\tau_1\delta'(x-T).$$

Тоді

$$\hat{m}(t)=\frac{1}{N}\sum_{n=0}^{N-1}\int_{-\infty}^{\infty}\xi(t+nx)\left[\delta(x-T)-\varepsilon\tau_1\delta'(x-T)\right]dx=$$

$$=\frac{1}{N}\sum_{n=0}^{N-1}\xi(t+nT)-\varepsilon\tau_1\left[\frac{1}{N}\sum_{n=0}^{N-1}\int_{-\infty}^{\infty}\xi(t+nx)\,\delta'(x-T)dx\right].$$

З властивостей похідної δ - функції маємо

$$\int_{-\infty}^{\infty}\xi(t+nx)\,\delta'(x-T)dx=-\xi_x'(t+nT).$$

Перетворимо статистику (1) до вигляду

$$\hat{m}(t)=\frac{1}{N}\sum_{n=0}^{N-1}\xi(t+nT)+\varepsilon\tau_1\frac{1}{N}\sum_{n=0}^{N-1}\xi_x'(t+nT). \qquad (3)$$

Отримане в [1] співвідношення, що визначає перше наближення

$$\tau_1=-\frac{n_1}{s_2}=-\frac{1}{\sqrt{EN^2(T)}\,s_2}\times$$
$$\times\left[\frac{i}{2\theta T}\sum_{l=-N_1}^{N_1}\Lambda_l\int_{-\theta}^{\theta}(s-t)\overset{o}{\xi}(s)e^{i\Lambda_l(t-s)}ds\right]$$

підставимо в (3), і в результаті маємо статистику

$$\hat{m}(t)=\frac{1}{N}\sum_{n=0}^{N-1}\xi(t+nT)-\frac{i}{2N\theta T\tilde{S}_2}\sum_{n=0}^{N-1}\sum_{l=-N_1}^{N_1}\Lambda_l\times$$
$$\times\int_{-\theta}^{\theta}(s-t)\xi_x'(t+nT)\overset{o}{\xi}(s)\,e^{i\Lambda_l(t-s)}ds$$

Знайдемо її математичне сподівання

$$E\hat{m}(t)=m(t)-\frac{i}{2N\theta T\tilde{S}_2}\sum_{n=0}^{N-1}\sum_{l=-N_1}^{N_1}\Lambda_l\times$$
$$\times\int_{-\theta}^{\theta}(s-t)E\overset{o}{\xi}(s)\left[\frac{\partial}{\partial x}\xi(t+nx)\right]_{x=T}e^{i\Lambda_l(t-s)}ds$$

і конкретизуємо вид співмножника

$$E\overset{o}{\xi}(s)\left[\frac{\partial}{\partial x}\xi(t+nx)\right]_{x=T}=\frac{\partial}{\partial x}b(t+nx,s-t-nx)_{x=T}=$$
$$=\left[\frac{\partial b(s,t-s+nx)}{\partial x}\right]_{x=T}$$

Таким чином, у першому наближенні оцінка математичного сподівання, отримана на основі статистики (1) із врахуванням того, що оцінка періоду корельованості знаходиться на основі функціоналу (2) виявляється зміщеною. Величина зміщення визначається таким співвідношенням:

$$\varepsilon_1\left[\hat{m}(t)\right]=-\frac{i}{2N\theta T\tilde{S}_2}\sum_{l=-N_1}^{N_1}\Lambda_l\times$$
$$\times\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}(s-t)\left[\frac{\partial b(s,t-s+nx)}{\partial x}\right]_{x=T}e^{i\Lambda_l(t-s)}ds$$

Легко переконатися у тому, що при $N\to\infty$ зміщення $\varepsilon_1\left[\hat{m}(t)\right]\to 0$. Тому така оцінка асимптотично незміщена.

Тепер дослідимо вплив випадкових флуктуацій на характеристики оцінки (1). Для цього визначимо її другий центральний момент

$$D\left[\hat{m}(t)\right]=E\hat{m}^2(t)-\left[E\hat{m}(t)\right]^2 \qquad (4)$$

Подамо статистику (1) у вигляді

$$\hat{m}(t)=\frac{1}{N}\sum_{n=0}^{N-1}\overset{o}{\xi}(t+nT)-\frac{i}{2N\theta T\tilde{S}_2}\sum_{l=-N_1}^{N_1}\Lambda_l\times$$
$$\times\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}(s-t)\left[\frac{\partial}{\partial x}\overset{o}{\xi}(t+nx)\right]_{x=T}\overset{o}{\xi}(s)e^{i\Lambda_l(t-s)}ds+$$
$$+m(t)-\frac{i}{2N\theta T\tilde{S}_2}\sum_{l=-N_1}^{N_1}\Lambda_l\times$$
$$\times\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}(s-t)\left[\frac{\partial}{\partial x}m(t+nx)\right]_{x=T}\overset{o}{\xi}(s)e^{i\Lambda_l(t-s)}ds$$

У припущенні гаусового ПКВП обчислюємо згідно (4) дисперсію оцінки математичного сподівання

$$D\left[\hat{m}(t)\right]=\frac{1}{N}\sum_{n=-N+1}^{N-1}(1-\frac{|n|}{N})b(t,nT)-$$
$$-\frac{1}{4N^2\theta^2T^2\tilde{S}_2^2}\sum_{l,k=-N_1}^{N_1}\Lambda_l\Lambda_k e^{i(l-k)\frac{2\pi}{T}}\times$$
$$\times\sum_{n,m=0}^{N-1}\int_{-\theta-\theta}^{\theta\,\theta}(s_1-t)(s_2-t)\times \qquad (5)$$
$$\times\left[\frac{\partial^2}{\partial x\partial y}\left[\begin{matrix}b(t+nx,my-nx)b(s_1,s_2-s_1)+\\ b(s_2,t-s_2+nx)b(s_1,t-s_1+my)\end{matrix}\right]_{x,y=T}\right]\times$$
$$\times e^{i\frac{2\pi}{T}(ks_2-ls_1)}ds_1ds_2-$$
$$-\frac{i}{N^2\theta T\tilde{S}_2}\sum_{l=-N_1}^{N_1}\Lambda_l e^{i\Lambda_l t}\times$$
$$\times\sum_{m,n=0}^{N-1}\int_{-\theta}^{\theta}(s-t)\left[\frac{\partial m(t+mx)}{\partial x}\right]_{x=T}b(t,s-t-nT)e^{i\Lambda_l(t-s)}ds$$

Аналіз співвідношення (5) показує, що при $N\to\infty$ та $\theta\to\infty$ другий центральний момент оцінки прямує до нуля. Таким чином, оцінка (1) слушна.

КОГЕРЕНТНА ОЦІНКА КОРЕЛЯЦІЙНОЇ ФУНКЦІЇ

Попереднє оцінювання періоду корельованості ПКВП впливає також на якість когерентної оцінки кореляційної функції. Припустимо, що період визначається за максимумом функціоналу математичного сподівання (2). Розглянемо когерентну оцінку кореляційної функції

$$\hat{b}(t,u) = \frac{1}{N}\sum_{n=0}^{N-1} \overset{o}{\xi}(t+n\tau_m)\,\overset{o}{\xi}(t+u+n\tau_m)\,. \quad (6)$$

Подамо статистику (6) в інтегральному вигляді, використовуючи фільтруючу властивість δ-функції і, розкладаючи її у ряд Тейлора. У першому наближенні по малому параметру ε отримаємо

$$\hat{b}(t,u) = \frac{1}{N}\sum_{n=0}^{N-1}\int_{-\infty}^{\infty} \overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\delta(x-\tau_m)dx =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \overset{o}{\xi}(t+nT)\,\overset{o}{\xi}(t+u+nT) -$$

$$- \varepsilon\tau_1 \frac{1}{N}\sum_{n=0}^{N-1}\int_{-\infty}^{\infty} \overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\delta'(x-T)dx =$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} \overset{o}{\xi}(t+nT)\,\overset{o}{\xi}(t+u+nT) +$$

$$+ \varepsilon\tau_1 \frac{1}{N}\sum_{n=0}^{N-1}\left[\frac{\partial}{\partial x}\left[\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right]_{x=T}\right].$$

Оскільки $\tau_1 = -\dfrac{n_1}{s_2}$, то для гаусового ПКВП має місце співвідношення

$$E\hat{b}(t,u) = b(t,u)\,.$$

Отже, в першому наближенні оцінка (6) незміщена.

Визначимо тепер дисперсію такої оцінки.

$$E\left[\hat{b}(t,u)\right]^2 =$$

$$= \frac{1}{N^2}\sum_{n,m=0}^{N-1}\left[\begin{array}{l} b(t,(m-n)T)b(t+u,(m-n)T) + \\ +\,b(t,u+(m-n)T)b(t,u+(n-m)T)\end{array}\right] +$$

$$+ \varepsilon^2 E\left[\begin{array}{l} \tau_1^2\,\dfrac{1}{N^2}\displaystyle\sum_{n,m=0}^{N-1}\left[\frac{\partial}{\partial x}\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right]_{x=T} \times \\ \times\left[\dfrac{\partial}{\partial y}\overset{o}{\xi}(t+my)\,\overset{o}{\xi}(t+u+my)\right]_{y=T}\end{array}\right].$$

Математичне сподівання другого доданка рівне

$$E\left[\begin{array}{l} \tau_1^2\,\dfrac{1}{N^2}\displaystyle\sum_{n,m=0}^{N-1}\left[\frac{\partial}{\partial x}\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right]_{x=T} \times \\ \times\left[\dfrac{\partial}{\partial y}\overset{o}{\xi}(t+my)\,\overset{o}{\xi}(t+u+my)\right]_{y=T}\end{array}\right] =$$

$$= -\frac{1}{4\theta^2 T^2\sqrt{EN^2(T)s_2^2}} \times$$

$$\times \sum_{k,l=-N_1}^{N_1}\Lambda_l\Lambda_k e^{i(l-k)\frac{2\pi}{T}t}\int_{-\theta-\theta}^{\theta}\int^{\theta}(s_1-t)(s_2-t)\times$$

$$\times E\overset{o}{\xi}(s_1)\,\overset{o}{\xi}(s_2)e^{i\frac{2\pi}{T}(ks_2-ls_1)}ds_1ds_2 \times$$

$$\times \frac{1}{N^2}\sum_{n,m=0}^{N-1}\left[\frac{\partial}{\partial x}\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right]_{x=T} \times$$

$$\times\left[\frac{\partial}{\partial y}\overset{o}{\xi}(t+my)\,\overset{o}{\xi}(t+u+my)\right]_{y=T}.$$

Враховуючи співвідношення

$$E\left[\begin{array}{l}\overset{o}{\xi}(s_1)\,\overset{o}{\xi}(s_2)\,\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\times \\ \times\,\overset{o}{\xi}(t+mx)\,\overset{o}{\xi}(t+u+mx)\end{array}\right] =$$

$$= b(s_1,s_2-s_1)G(t+nx,t+u+nx,t+mx,t+u+mx) +$$
$$+b(s_1,t-s_1+nx)G(s_2,t+u+nx,t+mx,t+u+mx) +$$
$$+b(s_2,t-s_1+u+nx)G(s_2,t+nx,t+mx,t+u+mx) +$$
$$+b(s_1,t-s_2+mx)G(s_2,t+nx,t+u+nx,t+u+mx) +$$
$$+b(s_1,t-s_2+u+mx)G(s_2,t+nx,t+u+nx,t+mx)\,,$$

де

$$G(s_1,s_2,s_3,s_4) = E\left[\overset{o}{\xi}(s_1)\,\overset{o}{\xi}(s_2)\,\overset{o}{\xi}(s_3)\,\overset{o}{\xi}(s_4)\right] -$$
$$- b(s_1,s_2-s_1)b(s_3,s_4-s_3)$$

запишемо остаточний вираз для дисперсії оцінки кореляційної функції:

$$D\left[\hat{b}(t,u)\right] =$$

$$= \frac{1}{N^2}\sum_{m,n=0}^{N-1}G(t+nT,t+u+nT,t+mT,t+u+mT) -$$

$$- \frac{1}{4N^2\theta^2T^2S_2^2}\sum_{k,l=-N_1}^{N_1}\Lambda_l\Lambda_k e^{i(l-k)\frac{2\pi}{T}t}\times$$

$$\times\int_{-\theta-\theta}^{\theta}\int^{\theta}(s_1-t)(s_2-t)e^{i\frac{2\pi}{T}(ks_2-ls_1)}\times$$

$$\times\left[\frac{\partial^2}{\partial x\partial y}\left[b(s_1,s_2-s_1)\times\right.\right.$$

$$\times G(t+nx,t+u+nx,t+mx,t+u+mx) +$$
$$+b(s_1,t-s_1+mx)G(s_2,t+u+nx,t+mx,t+u+mx) +$$
$$+b(s_1,t-s_1+u+nx)G(s_2,t+nx,t+mx,t+u+mx) +$$
$$+b(s_1,t-s_1+mx)G(s_2,t+nx,t+u+nx,t+u+mx) +$$
$$+b(s_1,t-s_2+u+mx)\times$$
$$\left.\left.\times\,G(s_2,t+nx,t+u+nx,t+mx)\right]\right]_{x,y=T}ds_1ds_2\,.$$

Метод оцінки періоду за статистикою (2) при $m(t)\approx 0$ непридатний. У цьому випадку слід використати кореляційний функціонал

$$\hat{b}(t,u,\tau)=\sum_{l=-N_2}^{N_2} B_l(u,\tau)e^{il\frac{2\pi}{\tau}t} \quad ,(7)$$

де

$$B_l(u,\tau)=(2\theta)^{-1}\int_{-\theta}^{\theta}\overset{o}{\xi}(t)\,\overset{o}{\xi}(t+u)e^{-il\frac{2\pi}{\tau}t}dt\,,$$

$$\overset{o}{\xi}(t)=\xi(t)-m(t)\,.$$

Якщо період визначати за абсолютним максимумом статистики (7), то перше наближення оцінки періоду має вигляд

$$\tau_1=-\frac{1}{s_2}\left[\frac{i}{2\theta T}\sum_{l=-N_2}^{N_2}\Lambda_l\int_{-\theta}^{\theta}(s-t)\,\overset{o}{\eta}(s)e^{il\frac{2\pi}{T}(t-s)}ds\right],\text{ де}$$

$\overset{o}{\eta}(s)=\overset{o}{\xi}(s)\,\overset{o}{\xi}(s+u)-b(s,u)$. Тоді із врахуванням співвідношення

$$E\left[\overset{o}{\eta}(s)\,\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right]=$$

$$=E\left[\overset{o}{\xi}(s)\,\overset{o}{\xi}(s+u)-b(s,u)\right]\times$$

$$\times E\,\overset{o}{\eta}(s)\left[\frac{\partial}{\partial x}\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right]_{x=T}ds$$

знайдемо математичне сподівання

$$E\left[\tau_1\frac{1}{N}\sum_{n=0}^{N-1}\left[\frac{\partial}{\partial x}\left(\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right)_{x=T}\right]\right]=$$

$$=-\frac{i}{2Ns_2\theta T}\sum_{l=-N_2}^{N_2}\Lambda_l e^{il\frac{2\pi}{T}t}\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}(s-t)e^{-il\frac{2\pi}{T}s}\times$$

$$\times E\,\overset{o}{\eta}(s)\left[\frac{\partial}{\partial x}\overset{o}{\xi}(t+nx)\,\overset{o}{\xi}(t+u+nx)\right]_{x=T}ds\,.$$

Зміщення когерентної оцінки кореляційної функції (6) у першому наближенні має вигляд

$$\varepsilon_1\left[\hat{b}(t,u)\right]=-\frac{i}{2Ns_2\theta T}\sum_{l=-N_2}^{N_2}\Lambda_l e^{il\frac{2\pi}{T}t}\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}(s-t)\times$$

$$\times\left[\frac{\partial}{\partial x}G(s,s+u,t+nx,t+u+nx)\right]_{x=T}e^{-il\frac{2\pi}{T}s}ds\,.$$

Дисперсія такої оцінки визначається восьмим моментом, котрий у випадку гаусового ПКВП виражається через кореляційну функцію. Тому й тут при виконанні умови $\lim_{|u|\to\infty}b(t,u)=0$ оцінка (6) є слушною.

КОГЕРЕНТНІ ОЦІНКИ ПРИ СПРОЩЕНІЙ ПРОЦЕДУРІ ВИЗНАЧЕННЯ ПЕРІОДУ

Нехай оцінка періоду ритміки ПКВП знаходиться шляхом максимізації статистики виду

$$\hat{m}_l(\tau)=\frac{1}{\theta}\int_{-\theta}^{\theta}\xi(t)\cos l\frac{2\pi}{\tau}tdt \qquad (8).$$

Тоді маємо перше наближення оцінки періоду корельованості у вигляді

$$\tau_1=-\frac{1}{s_2}\left[\frac{\Lambda_l}{\theta T}\int_{-\theta}^{\theta}s\,\overset{o}{\xi}(s)\sin\Lambda_l s ds\right].$$

Оскільки

$$E\left[\tau_1\sum_{n=0}^{N-1}\left[\frac{\partial}{\partial x}\xi(t+nx)\right]_{x=T}\right]=$$

$$=-\frac{\Lambda_l}{s_2^2\theta T}\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}s\left[\frac{\partial}{\partial x}E\xi(t+nx)\,\overset{o}{\xi}(s)\right]_{x=T}\sin\Lambda_l s ds=$$

$$=-\frac{\Lambda_l}{s_2^2\theta T}\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}s\left[\frac{\partial b(s,t-s+nx)}{\partial x}\right]_{x=T}\sin\Lambda_l s ds\,,$$

то зміщення оцінки (1) у даному випадку з точністю до величини першого порядку малого параметра ε рівне

$$\varepsilon_1\left[\hat{m}(t)\right]=-\frac{\Lambda_l}{N\theta T\tilde{S}_2}\sum_{n=0}^{N-1}\int_{-\theta}^{\theta}s\left[\frac{\partial b(s,t-s+nx)}{\partial x}\right]_{x=T}\times$$
$$\times\sin\Lambda_l s ds\,.$$

Дисперсія оцінки математичного сподівання (1) у першому наближенні для гаусового ПКВП має вигляд:

$$D\left[\hat{m}(t)\right]=\frac{1}{N}\sum_{n==N+1}^{N-1}(1-\frac{|n|}{N})b(t,nT)+\frac{\Lambda_l^2}{N^2\theta^2T^2\tilde{S}_2^2}\times$$

$$\times\sum_{n,m=0}^{N-1}\int_{-\theta-\theta}^{\theta}\int_{-\theta}^{\theta}s_1s_2\left[\frac{\partial^2}{\partial x\partial y}G(s_1,s_2,t+nx,t+my)\right]_{x,y=T}\times$$

$$\times\sin\Lambda_l s_1\sin\Lambda_l s_2 ds_1 ds_2-$$

$$-\frac{1}{N^2\theta T\tilde{S}_2}\sum_{m,n=0}^{N-1}\int_{-\theta}^{\theta}sb(s,t-s+nT)\times$$

$$\times\left[\frac{\partial}{\partial x}m(t+mx)\right]_{x=T}\sin\Lambda_l s ds\,.$$

Аналіз цього співвідношення показує, що оцінка (1) за умови оцінювання періоду корельованості за статистикою (8) також є слушною.

ВИСНОВКИ

Показано слушність когерентних оцінок імовірнісних характеристик ПКВП у випадку компонентного оцінювання періоду корельованості.

Отримані співвідношення для зміщення оцінок дають змогу контролювати величину систематичної похибки. Метод ефективний при вібродіягностиці обертових машин.

ЛІТЕРАТУРА

1.Михайлишин В.Ю., Яворский И.Н. "Компонентное оценивание периода стохастической повторяемости ритмических сигналов"//Автоматика.-1993.-№4.-С.43-47

SPEEDING-UP THE FRACTAL COMPRESSION WITH CLUSTERING

Nikolay N. Ponomarenko[1], Karen Egiazarian[2], Vladimir V. Lukin[1], Jaakko T. Astola[2]

[1]Dept 504, State Aerospace University (Kharkov Aviation Institute)
17 Chkalova Street, 61070, Kharkov, UKRAINE,
Tel/fax + 38 0572 441186, E-mail: lukin@mmds.kharkov.ua

[2]Tampere International Center for Signal Processing, Tampere University of Technology
P.O.Box-553, FIN-33101, Tampere, FINLAND, Tel. +358 3 365 3860,
Fax +358 3 365 3857, E-mail karen@cs.tut.fi, jta@cs.tut.fi

ABSTRACT

A problem of speeding-up the fractal compression of images is discussed. The techniques based on clustering are considered. An approach to clustering based on median split method is proposed. A procedure of reassigning the domain blocks between clusters is shown to improve the PSNR or to increase the compression ratio. The optimal number of clusters that minimizes the computation efforts for this procedure is derived. Some practical limitations of clustering are discussed and the particular ways out are given.

1. INTRODUCTION

Fractal image compression is an intensively developing topic in modern image processing [1]. Along with wavelet compression [2] it has drawn attention of many researchers and exhibited some advantages as a challenge to already used standard compression techniques.

The problems in fractal compression of still images are a very large amount of computations (and time) required and, thus, a necessity of speeding up the compression process without essential loss in quality [3], [4] as well as the recovered image enhancement. This paper mainly addresses the first problem dealing with speeding up the search of optimal domain blocks for range blocks by means of clustering.

There exist a lot of different strategies of domain search [3], but the strategy based on clustering appears to be the most reasonable. First, due to its application it becomes possible to obtain a rather great benefit in compression speeding up with reasonably small degradation of recovered image quality (or, if desirable, with small reduction of compression ratio). Second, this technique is easily adjusted to other methods of compression speeding up or image quality enhancement.

So below we first consider an approach described in [4] more in detail since our proposed methods are intended to improve it. Then an application of median split algorithm to data clustering is considered and its basic steps are given. Its advantages and drawbacks are analyzed. The procedure of reassigning the domain blocks between clusters is put forward and shown to be rather effective for PSNR increasing. All the benefits and drawbacks are illustrated using numerical simulation results for a set of traditional test images.

2. BASIC STEPS IN CLISTERING

In standard fractal compression a dominant part of time (almost 100%) is spent on finding the correspondence between the compressed range blocks and the domain blocks [1]. Denoting a number of range blocks as K and a number of domain blocks as N the complexity of compression becomes of $O(K*N)$ since it is required to compare each range block with each domain block. The use of clustering permits to avoid a linear search between all variants due to comparison of each range block with only the cluster centers and, then, with domain blocks of the cluster being the closest to the considered range blocks.

To perform these operations Wein and Blake [4] propose the following steps to be done in fractal compression: 1) the calculation of the number of clusters optimal for given number of domain and range blocks, 2) the division of the virtual codebooks (a domain block set) on clusters using the selected clustering method; 3) the search of the nearest cluster (according to its center) for each range block; 4) the search of the optimal domain block for each range block within the determined cluster.

The application of this scheme permits to decrease the complexity of domain search if the dimensions of virtual codebook clusters are equal. If initially it was $O(K*N)$, then, it becomes of order $O(K*M+K*N/M)+Ocl$ where M is the number of clusters and Ocl describes the complexity of clustering algo-

© UkrObraz'2000

rithm. As shown in [4], the value of the sum $K*M+K*N/M$ is minimized in case of $M=N^{0.5}$. The complexity is expressed then as $O(2*K*N^{0.5})+Ocl$.

Taking this into account one gets the following requirements to clustering algorithm:

1) it should provide the predetermined number of clusters desirably of equal size;

2) the complexity of clustering algorithm has to be considerably less than $O(K*N-K*M-N*M)$ in order to ensure considerable benefit in speeding up the fractal compression in comparison to methods based on linear (full) search;

3) the algorithms are to be able to operate with vectors having large dimensions (like 256 or 1024).

The first and the second requirements can be satisfied for several algorithms. Unfortunately, among these algorithms practically no one is able to satisfy the last condition. This results in noticeable degradation of recovered image quality in comparison to compression based on linear search. In particular, in [4] it is proposed to apply the *pnn*-algorithm with complexity $O(N*logN)$. However, to our opinion, the application of other methods could be more expedient.

3. MEDIAN SPLIT CLUSTERING

The median split clustering algorithm contains the following steps:

1) the calculation of the value variance for each vector coordinate (component);

2) the calculation of the median for the vector coordinate having the largest variance;

3) the division of domain field into two clusters having almost equal numbers of domains (the domains with the value for the considered coordinate less than median are referred to one cluster and otherwise);

4) until the number of clusters reaches the required value the steps 1-3 are repeated.

If the number of clusters is not the power of 2 the obtained clusters do not have equal sizes. However, this drawback is compensated by the higher speed of the considered algorithm. Actually, the median split algorithm complexity is $O(N*logM)$. That is why, if, for instance, the number of donain blocks is 65536 and the number of clusters is 256 the median split algorithm of clustering is approximately two times faster than *pnn*-algorithm.

The abovementioned technique of speeding up the fractal compression permits to get a very good (about $0.5N^{0.5}$) benefit in time needed. One more advantage is that the benefit almost does not depend on the range block number, but only upon the domain block number. Unfortunately, a considerable loss in recovered image quality (characterized by PSNR) is observed.

Let us confirm this by numerical simulation results. Table 1 presents the results of fractal compression for the test image (512x512) Lenna using the linear search method and two schemes of clustering – without and with reassigning the domain blocks (See next Section). The compression method performance is characterized by the values of PSNR for the same compression ratio (CR). It is seen (columns 3 and 4 of Table 1) that with the cluster number increasing the loss in PSNR becomes larger and for optimal value of M (256) the loss is of about 0.7 dB and it can be even larger. We have checked this for other traditional test images like "Mandrill" and "Peppers" and observed both the same tendency and the same order losses in PSNR.

One of the reasons is the necessity to perform the vector quantization for vectors with large dimensions. For median split algorithm the division of the vector field is done using only few vector coordinates (components) while the other ones remain unanalyzed. This leads to considerable inhomogeneity of the cluster content and the high probability of false decisions in clustering and comparison making. When the number of clusters increases and computations reduce, the PSNR decreases since the range blocks and the domain blocks the most close to them could occasionally drop into different clusters.

4. DOMAIN BLOCK REASSIGNING

Let us consider what can be done to get around this problem or, at least, to minimize its consequences. First, let us take into account the fact that for range blocks we find the closest cluster center and for domain blocks this procedure is not executed. They are considered to belong to that cluster to which they have been referred after initial clustering. But in practice a large percentage of domain blocks occur to be placed more close to the centers of another, neighbor clusters. So if after the initial clustering one makes for each domain the search of the closest cluster center the quality of finding the correspondence between the range blocks and the domain blocks can increase (although the complexity of search then increases by $O(N*M)$.

In this case the scheme of clustering for fractal compression is the following:

1) the calculation of cluster number optimal for given numbers of domain and range blocks;

2) the division of virtual codebooks (the domain block set) into clusters using the selected clustering algorithm;

3) the search of the nearest cluster (according to the distance to its center) for each range block;

4) the search of the nearest cluster (according to the distance to its center) for each domain block and translation of the domain block to this cluster if needed;

5) the search of the optimal domain block for each range block within the determined cluster.

The application of this scheme results in considerable decreasing of loss in the recovered image quality caused by clustering. Table 1 presents the corresponding

results (see columns 5 and 6). The range blocks for all the considered methods have the size 8x8, the domain block number is 62001, the clustering has been performed by median split algorithm.

It is seen from this Table that the increasing of the cluster number for the proposed clustering method with reassigning also results in reduction of PSNR in comparison to linear search. However, if the domain block reassigning is not used the loss in PSNR is about 0.3 dB larger than if the reassigning is applied (compare the results in columns 3 and 5 or analyze data in column 7). Moreover, the number (percentage) of domain blocks that have changed the cluster is large – from 60 to 72%.

Note that because of changing the clustering complexity – for clustering with domain block reassigning it became $O(K*M+K*N/M+K*N/M)+ Ocl$ – the optimal number of clusters is not equal to $N^{0.5}$. It depends on both the domain and range block numbers. In situations when $N>>K$ it is reasonable to select $M=K^{0.5}$ and if N and K are of the same order one should use $M=N^{0.5}$. The latter situation takes place in fractal compression using adaptive partition schemes [3] where the domain block number is restricted to get an opportunity of image partitioning on larger number of range blocks and its optimization. However, for majority of practical tasks one has $N>>K$ and the search complexity can be approximately expressed as $O(K*K^{0.5}+2*N*K^{0.5})+Ocl$.

Therefore, the procedure of domain block reassigning permits to partly solve the problem of recovered image enhancement. However, the dimensions of clusters become not equal and this can lead to some growth of computations to be performed. So one of problems to be considered in future is the application of methods providing better clustering, for instance, due to more perfect vector quatization approaches.

5. OTHER POSSIBLE SOLUTIONS

One heuristical way out is proposed below. Suppose we do not subtract the mean for every domain block before finding the median value and cluster splitting into two parts. Then the domains with similar (approaximately equal) coordinate values will fall down into the same cluster. If one deals with mosaic image compression then it is reasonable to assume the following. If the means for different domain blocks are rather close then these domain blocks are with larger probability placed in image not far away from each other than the domains with considerably different means. In turn, it means that the other vector coordinates of these domains can be highly correlated. If these assumptions are valid then the clustering by one coordinate at every step is "supported" by clustering for some other dependent coordinates of the vector.

Let us analyze the obtained numerical simulation results. Some of them are given in Table 2.The time required for finding the range and domain block correspondednce for equal size clusters is considered as T. The ratios of computation time in case of clustering with subtraction of mean and without subtraction are denoted as T_s/T and T_w/T, respectively. The number of domain blocks is given in the first column, the PSNR values are computed for range block size 8x8 and 64 clusters.

The test image is the same – Lenna 512x512 pixels. It is worth noting here that all the numerical simulation results in this paper are obtained for the following technique of color-to-gray scale image transformation $Gr=0.3R+0.59G+0.11B$ where R, G and B are the intensities of red, green and blue components of color image, respectively. We would like to mention also that the results (the PSNR values) can be noticeably (even by 1-2 dB) different for other variants of color-to-gray scale transformation.

As seen from Table 2 there is almost no difference in PSNR values for the considered clustering methods. The increase in computations with respect to ideal case is not too large, however, in the latter case (when the mean values are not subtracted) the increase is smaller due to better clustering with almost equal cluster dimensions.

Finally, let us give quantitative data concerning computational load for the linear search and the proposed clustering technique. Table 3 presents the results for three images - Lenna, Mandrill and Peppers. The PSNR values are given for the range block size 8x8, domain block size 16x16. The number of domain blocks is 62001. The optimal number of clusters for the considered case is equal to 62, but for suitability of the clustering algorithm operation it was selected to be 64. Three values of CR have been considered – 5.3, 21.3, and 85.3. We do not present here the time required for clustering expressed in seconds since the CPU depends upon many factors like computer type and rate, etc. That is why the ratio T_{ls}/T_c is given where T_{ls} and T_c are the time interval durations for making up the linear (full) search and search with clustering.

As seen from Table 3, far all the considered cases the difference in PSNR for compression based on linear search or clustering is not large. It does not extend 0.4 dB for small CR and reduces for larger CR values. Of course, the PSNR also depends upon image type and it is the worst (in case of equal CR) for the test image "Mandrill" which contains many textured regions. The most impressive result is the considerable speeding up the computations due to clustering. The benefit is especially large for CR values within the range of 5...20, where the computation time can be decreased by 30...60 times. Just this range of CR values is typical for many practical applications. For larger CR the ratio T_{ls}/T_c decreases but still remains large enough.

CONCLUSIONS

It is shown that due to clustering the computational complexity of the fractal compression can be significantly decreased. The median split method proposed for performing the clustering has been demonstrated to provide an appropriately small loss in PSNR with respect to the linear search fractal compression techniques for wide range of CR and different test images.

REFERENCES

[1] *Fractal Image Encoding and Analysis*, Edited by Yuval Fisher, Berlin Heidelberg, Germany: Springer Verlag, 1998.

[2] *Wavelet Image and Video Compression*, Edited by Pankaj N. Topiwala, Boston, USA: Kluwer Academic Publishers, 1998.

[3] B. Wohlberg, G. Jager, "A Review of Fractal Coding Literature", *IEEE Trans. On Image Processing*, Vol. 8, pp. 1716-1729, Dec. 1999.

[4] C. J. Wein, I.F. Blake, "On the Performance of Fractal Compression with Clustering", *IEEE Trans. On Image Processing*, Vol. 5, pp. 522-526, March 1996.

Table 1. Comparison of performance of different compression methods for test image Lenna, CR=21.3.

Cluster number	PSNR for linear search dB	PSNR for Clustering without Reassigning, dB	PSNR loss because of clustering, dB	PSNR for clustering with reassigning,dB	PSNR loss because of clustering, dB	Loss reduction due to reassigning, dB	Precentage of reassigned domains blocks
32	30.28	29.81	0.47	30.12	0.16	0.31	61.1%
64	30.28	29.75	0.53	30.04	0.24	0.29	65.7%
128	30.28	29.67	0.61	30.01	0.27	0.34	69.7%
256	30.28	29.58	0.7	29.9	0.38	0.32	72.7%

Table 2. Comparison of PSNR and computation time for two considered techniques of clustering.

Domain block number	PSNR, dB (Mean is not subtracted)	T_w/T	PSNR, dB (mean is subtracted)	T_s/T
62001	30.04	1.154	30.06	1.091
50000	30.00	1.126	29.98	1.053
38000	29.93	1.132	29.92	1.039
26000	29.80	1.097	29.80	1.045
20000	29.71	1.084	29.70	1.025
14000	29.57	1.101	29.54	1.006

Table 3. Comparison of PSNR and computation time for compression techniques based on linear search and clustering for different test images.

Test image	CR	PSNR, dB, linear search	Number of clusters	PSNR, dB, with clustering	Difference in PSNR, dB	T_{ls}/T_c
Lenna	5.3	36.34	128	35.95	-0.4	58.5
	21.3	30.28	64	30.06	-0.22	28.4
	85.3	26.00	32	25.80	-0.2	13.8
Mandrill	5.3	26.72	128	26.33	-0.4	57.6
	21.3	21.78	64	21.65	-0.13	28.0
	85.3	19.64	32	19.79	+0.15	14.9
Peppers	5.3	35.54	128	35.30	-0.24	59.0
	21.3	30.44	64	30.22	-0.22	29.4
	85.3	25.53	32	25.37	-0.16	14.8

MYRIAD FILTER PROPERTIES AND PARAMETER SELECTION

Sergey K. Abramov[1], Vladimir V. Lukin[1], Jaakko T Astola[2]

[1]Dept 504, State Aerospace University (Kharkov Aviation Institute)
17 Chkalova Street, 61070, Kharkov, UKRAINE,
Tel/fax + 38 0572 441186, E-mail: lukin@mmds.kharkov.ua

[2]Tampere International Center for Signal Processing, Tampere University of Technology
P.O.Box-553, FIN-33101, Tampere, FINLAND, Tel. +358 3 365 2923,
Fax +358 3 365 3857, E-mail jta@cs.tut.fi

ABSTRACT

Peculiar properties of myriad filters in case of their application to 1-D signal processing in mixed noise environment are considered. It is shown that the filter output bias and the filter noise suppression effectiveness considerably depend upon filter parameter *k*, signal behavior and noise characteristics. Some recommendations concerning parameter *k* selection are given.

1. INTRODUCTION

In processing the signals corrupted by mixed, i.e. additive and impulsive, noise the nonlinear filters have shown themselves to be a useful and efficient tool since they are able simultaneously to remove spikes, suppress additive noise component and preserve signal abrupt changes well enough [1], [2].

A new and recently introduced class of nonlinear data processing algorithms is a myriad filter and its modifications like weighted myriad filters [3], [4]. They have been proposed and designed as nonlinear filters for robust non-gaussian signal processing in impulsive noise environment, in particular, in case of α-stable distributions of noise having heavier tails than gaussian and, for some cases, Laplacian ones.

The performance characteristics of the standard myriad filter strongly depend on the so-called linearity parameter k that controls the outlier-rejection capability of an estimator. At the same time, this parameter determines the dynamic (output bias) and statistical characteristics of the myriad filter. Certainly, all these characteristics also depend upon a scanning window size.

Some of these properties have not been exploited in [3], [4] and for their analysis we apply approaches described in our papers [5], [6]. It is demonstrated, in particular, that myriad filters with rather small k (just under this condition an appropriate robustness is provided) distort piece-wise fragments of signal in peculiar manner and even "amplify" noise for rapidly increasing/decreasing signals. These drawbacks are undesirable and should be avoided. To get around this problem is possible due to proper, compromise, selection of parameter k. In order to perform it we present below the computer simulation results and give the corresponding recommendations based on test signal analysis.

2. MYRIAD FILTER DEFINITION

A sample myriad is an M-estimator of location. Given a set of samples $\{y(i)\}, i = 1,\ldots,N$, an M-estimator of location is described by expression $\hat{\theta} \stackrel{\Delta}{=} \arg\min_{\theta} \sum_{i=1}^{N} \rho(y(i)-\theta)$, where $\rho(.)$ is the cost function. The sample myriad is defined using the cost function $\rho(y) = \log(k^2 + y^2)$. Therefore, for standard myriad filter output the used cost function is $\sum_{i=1}^{N} \log(k^2 + (y(i)-\theta)^2)$ where $\{y(i)\}$ is formed by sample values within the scanning window of size N. In other words, the output $x(j)$ of the myriad filter applied to process $\{y(j)\}, j = 1,\ldots,J, J >> N$ is written as follows

$$x(j) = \arg\min_{\theta} \sum_{i=j-(N-1)/2}^{j+(N-1)2} \log(k^2 + (y(i)-\theta)^2). \quad (1)$$

Another, sometimes more convenient, way of the myriad filter output definition is expressed as

$$x(j) = \arg\min_{\theta} \prod_{i=j-(N-1)/2}^{j+(N-1)/2} (1 + (y(i)-\theta)^2 / k^2). \quad (2)$$

It can be easily shown that the sample myriad reduces to the sample mean as $k \to \infty$. At the same time, as $k \to 0$ the sample myriad tends to the distribution mode and this is a rather important property. That is why

© UkrObraz'2000

we used earlier the terms "rather small" and "rather large" values of parameter k. However, one should keep in mind that these relative definitions depend upon signal scale.

Other details concerning the myriad and weighted myriad filter peculiarities can be found in [3], [4]. Our purpose is to attract the reader's attention to specific aspects that could be important in practical applications.

3. CONSIDERED SIGNAL/NOISE MODEL

Let us consider the following signal/noise model of a 1-D sampled data sequence

$$y(j)=\begin{cases} S(j)+n_a(j), \text{ with probability } 1-P_{imp} \\ S(j)+n_{imp}(j), \text{with probability } P_{imp}, \end{cases} \quad (3)$$

where $S(j)$ denotes the true signal value of the j-th sample; $n_a(j)$ is the zero mean gaussian additive noise with the variance σ_a^2; $n_{imp}(j)$ defines the amplitude of impulsive noise that occurs with the probability P_{imp} and it is supposed that $|n_{imp}| > \sigma_a$.

The considered test signal is presented in Fig.1, **test**. The behavior of the myriad filter output was studied for different fragments of this signal: 1) a constant signal (for example, the fragment with indices from 10 to 40); 2) a step edge (indices 40-60); 3) piecewise linear curves (indices 90-110 and 190-210); 4) linearly increasing and decreasing signals (110-140 and 160-190); 5) peak maximum (140-160); 6) polynomial maximum (265-285); 7) a specific fragment (240-260). In other words, this set of test fragments covers a rather wide variety of situations taking place in practice. Thus, the analysis performed in this paper characterizes the myriad filter output peculiarities from quite different points of view.

4. OUTPUT BIAS

Myriad filter mean output considerably depends upon noise characteristics even if the noise is gaussian. This property is, in general, typical for many nonlinear filters [1], [2], [5]. However, since for rather large k the myriad filter characteristics approximate the standard mean filter ones, it is not worth considering the myriad filter bias for this case. The parameter k selection to be rather small when essentially nonlinear and robust properties of the myriad filter are provided is of more theoretical and practical interest.

The studies done have shown that in case of small k, considerably less than the step edge height Δh, the myriad filter preserves it better than a standard median filter with the same N. This conclusion is valid for wide range of additive noise variances $\sigma_a << \Delta h$.

In general, the properties of all nonlinear filters also depend upon scanning window size N. Below we analyze only the case N=9, however, the character of main dependences is similar for other N within the limits from 5 to 11.

However, very peculiar effects take place in the neighborhoods of indices 100, 200, 250, 300, i.e. for piecewise curve junction points. These effects are illustrated in Fig. 2. When k is rather small (0.01) the myriad filter output bias derived as $\Delta_{myr}(j) = E[x(j)] - S(j)$ can have sign that is opposite to the output bias sign in case of large k. For instance, for the neighborhood of the sample with index 100 the myriad filter output bias is positive but very small for j<100, and for j=101, 102, and 103 it becomes negative (if k=0.01). Finally, with further j increasing the myriad filter mean $E[x(j)]$ quickly "jumps" and becomes approximately equal to $S(j)$ (See Fig. 2,a). Similar effects are observed for k=0.01 in the neighborhoods of sample indices 200, 250 (See Fig.2,b), and 300. At the same time, for k=1.0 the myriad filter output behavior is much more typical and it is quite similar to many other nonlinear (for example, α-trimmed) and linear filters [5] (Consider curves for k=1.0 in Fig.2,a and 2,b). All these effects are explained by the mode selection property of myriad filters with small k.

The Table 1 presents the numerical simulation results for the test signal and different noise characteristics. The values of mean square errors (MSE) have been evaluated for the entire signal (χ_t) and locally for its fragments. For example, χ_{10-40} is the local MSE evaluated for the signal fragment from the index 10 to the index 40 and the MSE χ_{40-60} just corresponds to the step edge neighborhood. The MSEs do not characterize the output bias directly since they also take into account the variance of noise remained after filtering. However, the contribution of output bias into local MSEs is considerable in step edge and junction point neighborhoods. That is why the MSE values give reliable information for analysis of filter output bias for many fragments except constant, linearly increasing/decreasing signal (where bias is zero for case of spike absence) and polynomial maximum fragments. In general, polynomial maximum is preserved by myriad filter well enough if k is not rather large.

The dynamic errors (output bias) in the neighborhood of peak extremum (index 150) depend upon k and usually they slightly increase with parameter k increasing. It should be also mentioned that except of output bias the considerable growth of residual noise variance can be observed in the neighborhoods of junction points (indices 100, 150, 200, 250, 300) if k is rather small.

And this is one more drawback of the myriad filters. As it can be seen from Table 1 for *Case A* (small variance of gaussian noise and spikes are absent), the smallest MSE values in the step edge neighborhood are provided by the myriad filter with k=0.01. At the same time for other junction point neighborhoods the better results are observed for myriad filters with k=1.0 and k=10.0. This means that a trade-off between noise suppression efficiency and detail preservation ability of the myriad filter should be found and it is reachable by means of parameter k selection (variation).

5. NOISE SUPPRESSION EFFICIENCY AND ROBUST PROPERTIES

In case of rather small k the peculiar effects can be observed for constant and linearly increasing/decreasing signal (LIDS) fragments. The noise suppression efficiency of nonlinear filter can be assumed sufficient if it does not differ a lot from the corresponding efficiency of the standard mean filter with the same N [6]. For example, for the best nonlinear noise suppressing filters like Wilcoxon and Hodges-Lehman ones the residual noise variance is not larger than by 30% in comparison to mean filter for LIDS [6]. The myriad filter with rather small k satisfies this requirement for constant signal (analyze χ_{10-40} in Table 1 for *Case A*). But note that in this situation the value of k is comparable with σ_a. If the ratio σ_a / k increases the difference between noise suppression efficiency of the myriad and mean filters for constant signal radically becomes larger. This means that the parameter k selection can be scale dependent and for constant signal the sample scale is determined by noise standard deviation σ_a (if the spikes are absent).

Another tendency is observed for LIDSs. Results in Table 1 (consider $\chi_{110-140}$ and $\chi_{160-190}$ for *Case A*) show that the variance of "residual" noise for k=0.01 is larger than even σ_a^2. That is the myriad filter with rather small k can "amplify" noise. And this is a very serious drawback to be avoided. Now let us take into account that the sample scale in this situation is mainly determined not by noise standard deviation but by parameter $\Delta S_{sc} = N\Delta S$ where $\Delta S = |S(j+1) - S(j)|$. Thorough studies have demonstrated that by selecting $k > 0.3\Delta S_{sc}$ it is possible to appropriate noise suppression efficiency of the myriad filter for LIDS. And the selection k=0.01 does not satisfy this inequality for LIDS of the considered test signals. This is also the reason why large residual noise variances are observed for k=0.01 in the neighborhoods of samples with indices 150 (peak maximum), 250 and 300 (junction points of polynomial maximum with constant signals). Therefore, the conclusion is that for almost all the signal fragments except step edge neighborhood the proper selection of parameter k can be done on basis of local estimation of sample scale. But one should keep in mind that k increasing leads to worse robustness with respect to spikes and also results in larger output bias.

To confirm the fact that the myriad filter robustness sufficiently depends upon k let us briefly analyze the numerical simulation data for *Case B* (spike presence) in Table 1. As seen, the myriad filter with k=0.01 out performs the myriad filters with k=1.0 and 10.0 for some signal fragments in the sense of smaller local MSE. In particular, this takes place for constant signal fragment, step edge neighborhood and even the fragment with indices from 90 to 110 (first piecewise curve junction neighborhood). The main reason are the good robust properties of the myriad filter with k=0.01. However, the best, i.e. the smallest MSEs χ_t for entire signal are provided by the myriad filter with k=1.0 for both *Case A* and *Case B*. These interesting results serve as a motivation for locally adaptive approach application to signal processing on basis of myriad filters with varied parameters. Several approaches to providing such filter robustness are worth exploring. One of them is the use of robust estimators of scale.

CONCLUSIONS

The definition of the myriad filter is given and its advantages and drawbacks are considered. The basic characteristics of the myriad filter can be tuned by selecting the only parameter k but it should be a compromise decision depending upon the required trade-off between filter robustness, noise suppression efficiency and output bias. Numerical simulation results and performed analysis results give a starting point for further studies.

REFERENCES

[1] J. T. Astola, P. Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*, New York, NY: CRC Press, Boca Raton, 1997.

[2] I. Pitas, A. N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, New York: Kluwer Academic Publishers, 1990.

[3] S. Kalluri, G. Arce, "Adaptive weighted myriad filter algorithms for robust signal processing in α-stable noise environments", *IEEE Trans. On Signal Processing*, Vol. 46, pp. 322-334, Feb. 1998.

[4] S. Kalluri, G. Arce, "Fast algorithms for weighted myriad computation by fixed point search", *IEEE Trans. On Signal Processing*, Vol. 48, pp. 159-171, Jan. 2000.

[5] S. Abramov, N. Tulyakova, V. Lukin, V. Melnik, "Dynamic characteristics of some nonlinear 1-D signals", *Proceedings of 1999 Finnish Signal Process-*

ing Symposium FINSIG'99, Oulu, Finland, May 1999, pp. 184-188.

[6] K. Willner, P. Kuosmanen, V. Lukin, A. Po-grebniak, "Nonlinear filters and rapidly increasing/decreasing signals corrupted with noise, *CD-ROM Proceedings of Nonlinear Signal/Image Processing (NSIP) Conference*, Michigan, USA, Sept. 1997, 5 pages.

Table 1.The MSEs for the considered filters

	χ_t	χ_{10-40}	χ_{40-60}	χ_{90-110}	$\chi_{110-140}$	$\chi_{140-160}$	$\chi_{160-190}$	$\chi_{190-210}$	$\chi_{240-260}$	$\chi_{265-285}$
Case A: σ_a^2=0.001; P_{imp}=0.00, n_{imp}=0.00										
Noisy	0.00099	0.00102	0.00100	0.00097	0.00098	0.00094	0.00101	0.00095	0.00101	0.00098
K=0.01	0.00742	0.00013	0.00015	0.00058	0.00116	0.00244	0.00276	0.00152	0.05298	0.00076
K=1.0	0.00357	0.00011	0.02462	0.00030	0.00011	0.00122	0.00011	0.00051	0.01457	0.00042
K=10.0	0.00631	0.00011	0.03525	0.00029	0.00011	0.00125	0.00011	0.00051	0.03135	0.00043
Case B: σ_a^2=0.01; P_{imp}=0.03 n_{imp}=1.00										
Noisy	0.03961	0.04662	0.03825	0.04058	0.03804	0.03711	0.04083	0.03457	0.03641	0.03998
K=0.01	0.01097	0.00218	0.01291	0.00316	0.00435	0.00637	0.00773	0.00506	0.05543	0.00365
K=1.0	0.00670	0.00274	0.02879	0.00332	0.00288	0.00345	0.00308	0.00373	0.02041	0.00257
K=10.0	0.01134	0.00515	0.04034	0.00541	0.00501	0.00551	0.00508	0.00596	0.03738	0.00455

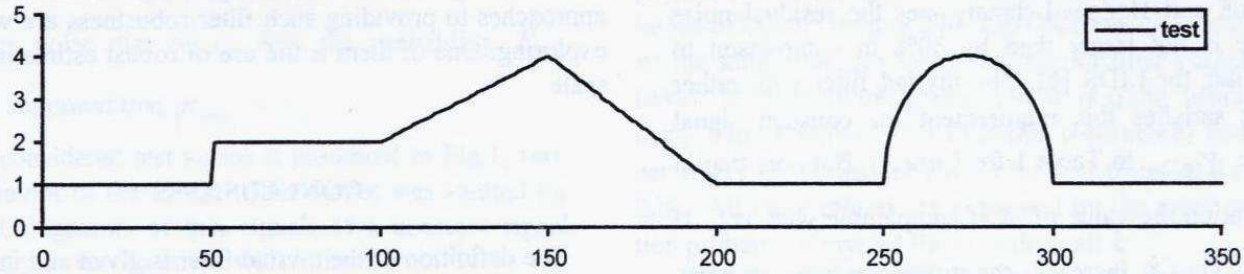
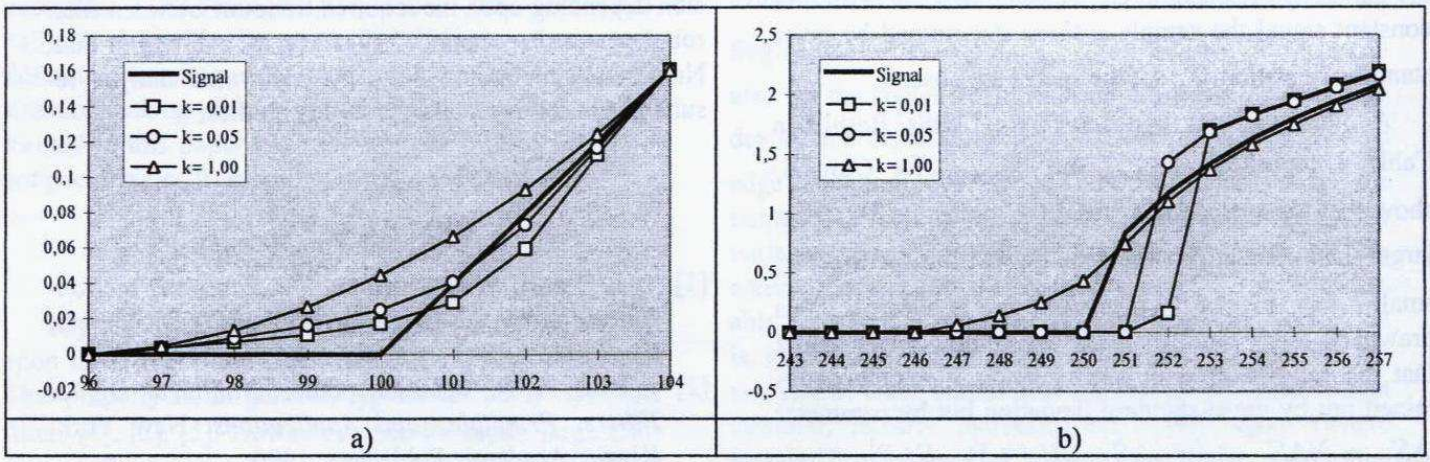


Figure 1. The considered noise-free test signal.



a) b)

Figures 2. Myriad filter mean outputs for piecewise curve fragments (only gaussian noise, k=0.01, 0.05 and 1.0): a) neighborhood of sample 100; b) neighborhood of sample 250.

Методи стискання даних за допомогою перетворення сплесків

Р.В.Лотоцький

Інститут кібернетики ім. Глушкова НАН України, т.(044)211-56-54, roman_l@i.com.ua,

Розглядається застосування перетворення сплесків для однієї з найважливіших проблем обробки інформаціїї – стискання графічних зображень. Наводяться основні факти теорії сплесків в об'ємі, достатньому для конструктивного опису побудованого алгоритму. Результати роботи розробленного програмного комплексу проаналізовані на прикладах стискання зображень реальних відбитків пальцівю. Обговорюються основні підходи до удосконалення розглянутого методу стискання.

1.Вступ

Методи стискання даних можна розділити на два класи: стискання із втратами і без втрат. Алгоритми з другого класу зазвичай застосовуються для перетворення текстової інформації, тому що в цьому випадку потрібне точне відновлення даних. На відміну від текстової інформації, для графічних напівтонових зображень допускаються втрати при стисканні. У цьому випадку відновлене зображення не буде точною копією оригіналу, але часто така умова і не потрібна, оскільки людське око не сприймає невеликі зміни окремих пікселів. Застосування алгоритмів із втратами дозволяє істотно підвищити коефіцієнт стискання в порівнянні з алгоритмами без втрат. У даній роботі розглядається метод стискання з втратами графічних зображень, основним етапом якого є перетворення сплесків (Wavelet). Результати роботи реалізованого алгоритму показано на прикладах обробки реальних фотографій відбитків пальців.

2.Основні відомості з теорії перетворення сплесків

2.1. Неперервне перетворення сплесків. Неперервне перетворення сплесків визначається таким чином:

$$WF(a,b)=\frac{1}{\sqrt{|a|}}\int_{-\infty}^{\infty} f(t)\overline{\psi\left(\frac{t-b}{a}\right)}dt=$$

$$=\int_{-\infty}^{\infty} f(t)\overline{\psi_{a,b}}dt;$$

Тут, функція ψ задовільняє наступним умовам:

$$\psi\in L_2(R),\ \int_{-\infty}^{\infty}\frac{|\hat{\psi}(\xi)|^2}{|\xi|}d\xi<\infty. \qquad (1)$$

Тут і далі $\hat{\psi}$ - перетворення Фур'є функції ψ. Базисна система сплесків $\{\psi_{a,b}\}$ складається з усіх комбінацій зсувів і стиснень функції ψ, що у термінах теорії сплесків називається "утворюючим сплеском"(mother wavelet).

$$\psi_{a,b}=\frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right), \qquad (2)$$

Зворотнє перетворення:

$$f(t)=C_{\psi}^{-1}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}WF(a,b)\psi_{a,b}(t)\frac{dadb}{a^2}, \qquad (3)$$

де $C_{\psi}=2\pi\int_{-\infty}^{\infty}|\hat{\psi}(\xi)|^2\frac{1}{|\xi|}d\xi$.

2.2. Кратно-масштабний аналіз. Кратно-масштабний аналіз є математичним апаратом для конструктивного опису і побудови різноманітних базисів функцій сплесків. Вперше це поняття було введено Маллатом [1]. Основні положення теорії кратно-масштабного аналізу можна знайти також у [2,3].

Кратно-масштабний аналіз на $L_2(\mathbf{R})$ визначається як послідовність вкладених підпросторів V_j на $L_2(\mathbf{R})$ з такими властивостями:

1) $V_j\subset V_{j+1}, j\in Z$;
2) $v(x)\in V_j \Leftrightarrow v(2x)\in V_{j+1}$;
3) $v(x)\in V_0 \Leftrightarrow v(x+1)\in V_0$;
4) $\overline{\bigcup_{j=-\infty}^{+\infty}V_j}=L_2(\mathbf{R})$, $\bigcap_{j=-\infty}^{+\infty}V_j=\varnothing$, де $\overline{\bigcup_{j=-\infty}^{+\infty}V_j}$ -- замикання підпростору $\bigcup_{j=-\infty}^{+\infty}V_j$.
5) Існує функція φ така, що система функцій $\{\varphi(x-k)\,|\,k\in Z\}$ утворює ортонормований базис простору V_0. Функція φ називається масштабуючою функцією.

Із властивостей 2) і 5) випливає, що множина функцій $\varphi_k^j(x) = \sqrt{2^j}\varphi(2^j x - k)$, $j \in \mathbf{Z}$, $k \in \mathbf{Z}$, утворює ортонормований базис підпростору V_j.

Оскільки $V_0 \subset V_1$, то будь-яку функцію з V_0 можна записати як лінійну комбінацію базисних функцій $\sqrt{2}\varphi(2x-k)$. Отже, маємо

$$\varphi(x) = \sum_k h(k)\sqrt{2}\varphi(2x-k),$$

де коефіцієнти $h(k)$ внаслідок ортогональності $\varphi(x)$ визначаються таким чином:

$$h(k) = \left(\varphi(x),\ \sqrt{2}\varphi(2x-k)\right). \qquad (4)$$

Нехай W_j -- ортогональне доповнення підпростору V_j до V_{j+1}, $V_{j+1} = V_j \oplus W_j$. З останньої рівності і властивості 1) можна записати

$$\forall j, j_1, j < j_1: \quad V_{j_1} = V_j \oplus \bigoplus_{k=0}^{j_1 - j - 1} W_{j_1 - k}.$$

Таким чином, враховуючи, що $\overline{\bigcup_{j=-\infty}^{+\infty} V_j} = L_2(\mathbf{R})$, маємо $\bigoplus_{j \in Z} W_j = L_2(\mathbf{R})$.

Головною властивістю кратно-масштабного аналізу є той факт, що якщо послідовність вкладених підпросторів V_j задовольняє умовам 1)-5), то існує можливість побудови такого ортонормованного базису $\{\psi_k^j; j,k \in Z\}$ простору $L_2(\mathbf{R})$, $\psi_k^j \in W_j$, що $\forall f \in L_2(\mathbf{R})$:

$$P_{j+1}(f) = P_j(f) + \sum_k (f, \psi_k^j)\psi_k^j,$$

де $P_j(f)$ -- ортогональна проекція функції f на V_j. Базис $\{\psi_k^j; j,k \in Z\}$ називається базисом сплесків. Так як $Wj \subset Vj_{+1}$, то має місце таке подання функції $\psi^j(x)$:

$$\psi^j(x) = \sum_k g(k)\varphi_k^{j+1}(x).$$

Доведено [2], що умова ортонормованності базису $\{\psi_k^j; j,k \in Z\}$ виконується при наступному виборі коефіцієнтів $g(k)$: $g(k) = (-1)^k h(-k+1)$. Таким чином, формули залежності $\psi^j(x)$ від $\{\varphi_k^{j+1}(x)\}$ і $\varphi^j(x)$ від $\{\varphi_k^{j+1}(x)\}$ мають такий вид:

$$\varphi^j(x) = \sum_k h(k)\varphi_k^{j+1}(x), \qquad (5)$$

$$\psi^j(x) = \sum_k (-1)^k h(-k+1)\varphi_k^{j+1}(x). \qquad (6)$$

2.3. Швидке перетворення сплесків. Кратно-масштабний аналіз дозволяє побудувати швидкий алгоритм перетворення сплесків. Так як $V_{j+1} = V_j \oplus W_j$, то функцію $v_{j+1} \in V_{j+1}$ можна записати

$$v_{j+1} = \sum_k \lambda_k^{j+1}\varphi_k^{j+1}(x) = v_j(x) + w_j(x) = \sum_l \lambda_l^j \varphi_l^j(x) +$$

$$+ \sum_l \gamma_l^j \psi_l^j(x), \qquad (7)$$

де $v_j \in V_j$ і $w_j \in W_j$

Встановимо зв'язок між коефіцієнтами $\lambda_k^{j+1}, \lambda_l^j$ і $\lambda_k^{j+1}, \gamma_l^j$. З формули (5) випливає:

$$\varphi_k^j(x) = 2^{\frac{j}{2}}\varphi(2^j x - k) = 2^{\frac{j+1}{2}} \sum_l h(l)\varphi(2^{j+1}x -$$

$$-2k - l) = \sum_l h(l)\varphi_{2k+l}^{j+1}(x) = \sum_l h(l-2k)\varphi_l^{j+1}(x) \qquad (8)$$

Аналогічно:

$$\psi_k^j(x) = \sum_l g(l-2k)\varphi_l^{j+1}(x) \qquad (9)$$

Звідси маємо

$$\lambda_k^j = \sum_l h(l-2k)\lambda_l^{j+1}, \qquad (10)$$

$$\gamma_k^j = \sum_l g(l-2k)\lambda_l^{j+1}. \qquad (11)$$

Обернене перетворення отримуємо аналогічно. З (7)-(9) випливає, що:

$$\lambda_k^{j+1} = \sum_l \lambda_l^j(\varphi_l^j, \varphi_k^{j+1}) + \sum_l \gamma_l^j(\psi_l^j, \varphi_k^{j+1}) =$$

$$= \sum_l h(l-2k)\lambda_l^j + \sum_l g(l-2k)\gamma_l^j. \qquad (12)$$

Таким чином, якщо відомі коефіцієнти розкладу даної функції по базисних функціях у деякому підпросторі V_j, то, скориставшись (10), (11), можна знайти розклад цієї функції також в усіх підпросторах $W_J, J < j$, а скориставшись (12), можна відновити вихідний розклад.

Формули (10), (11) можна записати в більш компактній і зручній для практичного застосування формі, використовуючи так зване операторне подання фільтрів.

Отже, нехай вектор $\boldsymbol{a}=(a_l)$, $l \in [0,N]$ складається з коефіцієнтів розкладу деякої функції

по базису підпростору V_j. Визначимо оператори H і G:

$$(Ha)_k = \sum_{l=0}^{N} h(l-2k)a_l, \qquad (13)$$

$$(Ga)_k = \sum_{l=0}^{N} g(l-2k)a_l, \forall k \in \left[0, \left[\frac{N}{2}\right]\right], \qquad (14)$$

[•] - ціла частина числа. Ці оператори відповідають перетворенням (10), (11) і виконують один крок у перетворенні сплесків. Відповідні оператори для оберненого перетворення (12) мають такий вигляд:

$$\left(H^* a\right)_l = \sum_{k=0}^{\left[\frac{N}{2}\right]} h(l-2k)a_k, \qquad (15)$$

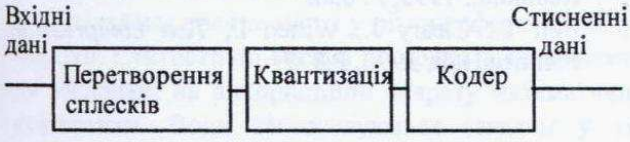
$$\left(G^* a\right)_l = \sum_{k=0}^{\left[\frac{N}{2}\right]} g(l-2k)a_k, \ \forall l \in [0,N]. \qquad (16)$$

3. Алгоритм стискання зображень методом сплесків.

На сьогодні найбільш відомим алгоритмом стискання графічних файлів з втратами є алгоритм JPEG [4], в якому використовується косинусоїдальне перетворення Фур'є (КПФ). Цей алгоритм дозволяє стискати графічні зображення в 25-30 разів без суттєвого погіршення якості. Проте якщо зміна інтенсивності пікселів зображення не є плавною (як, наприклад, у відбитках пальців), то коефіцієнт стиску різко падає. Пропонований алгоритм має структуру, подібну алгоритму JPEG, але використовує замість КПФ перетворення сплесків.

У наступних викладах будемо вважати, що зображення є двомірною дискретною функцією, $f(x,y) \in [0,255]$ значеннями якої є 256 відтінків сірого (0 -- чорний, 255 -- білий)

Процес стискання складається з таких трьох кроків: перетворення сплесків, квантизація, кодування (Рис. 3).



Вхідне зображення розбивається на блоки 8×8 пікселів. Наведені вище операції застосовуються окремо до кожного блоку.

На першому кроці виконується перетворення сплесків. Вище було розглянуто одномірне перетворення сплесків для векторів. У двомірному випадку для перетворення матриці даних застосовуємо одномірне перетворення для кожного рядка, і для отриманої в результаті матриці знову застосовуємо одномірне перетворення для кожного стовпчика.

Оскільки різні коефіцієнти перетворення по-різному впливають на відновлення вихідного зображення, то, для підвищення ступеня стискання, отримані коефіцієнти w_i квантизуються, тобто діляться на коефіцієнти квантизації: $k_i . r_i = w_i / k_i$ Потім, округлені до найближчого цілого, вони надходять на вхід ентропічного кодера.

При підборі правильної таблиці квантизації $Q = (k_{i,j})$ необхідно враховувати вплив кожного коефіцієнта на відновлення зображення. У даній роботі пропонується такий алгоритм для знаходження матриці квантизації. Нехай E - деякий рівень похибки, тобто $E=\left\| f(x,y) - \tilde{f}(x,y) \right\|$, де $\tilde{f}(x,y)$ -- відновлене зображення. Для кожного коефіцієнта перетворення знаходимо таке максимальне значення k_j, що $\left\| f(x,y) - \tilde{f}(x,y) \right\| \le E$. Нехай $k_{\min} = \min_j k_j$ і $q_j = \frac{k_j}{k_{\min}}$. Тоді шукана матриця має такий вигляд: $Q = (q_j k_{\min})_j$. За допомогою коефіцієнта $k_{\min}$ будемо регулювати коефіцієнт стискання - при більшому $k_{\min}$ стискання буде більшим. Так як невигідно знаходити матрицю квантизації для кожного зображення окремо в зв'язку із значним часом роботи алгоритму, то достатньо побудувати одну матрицю для деякого класу зображень. Такими класами можуть бути, наприклад, відбитки пальців, фотографії обличчя людини.

На заключному етапі алгоритму стискання квантизовані коефіцієнти кодуються за допомогою одного з алгоритмів ентропійного кодування. У *даній роботі використовується динамічний* алгоритм кодування Хаффмана. Цей алгоритм добре відомий у теорії кодування ([4, 5]).

3.1. Особливості розробленої програми.

Розроблена програма дозволяє стискати зображення з градацією сірого. У якості базису сплесків використовується система функцій Добеші [2].

Таблиця квантизації, побудована для класу зображень відбитків пальців, має такий вид:

5k, 6k, 7 *k*, 7k, 8k, 10k, 10k, 8 *k*,
6k, 7k, 7 *k*, 7k, 10k, 10k, 10k, 9 *k*,
7k, 7k, 8 *k*, 7k, 11k, 11k, 10k, 10k,
7k, 7k, 8 *k*, 7k, 10k, 10k, 10k, 10k,
8k, 9k, 10 *k*,10k,11k, 11k, 11k, 10k,
9k, 10k, 10 *k*,10k,11k, 11k,11k, 10k,
9k, 9k, 10 *k*,10k,11k, 11k,11k, 11k,

8k, 8k, 9 *k*, 9k, 10k, 10k,10k, 10k

3.2. Отримані результати. Результати роботи програми будемо розглядати на прикладі фотографії відбитка пальця. У наступній таблиці наведені деякі значення "відстані" між оригінальним зображенням і зображенням, відновленим після стискання. "Відстань" обчислюється по формулі середньоквадратичного значення різниць між інтенсивностями відповідних пікселів двох зображень:

$$D = \sqrt{\frac{\sum_{i=0}^{width-1} \sum_{j=0}^{height-1} \left(f(i,j) - \tilde{f}(i,j)\right)^2}{width * height}},$$

де $f(i,j)$ і $\tilde{f}(i,j)$ -- функції інтенсивності пікселей відповідно оригінального і відновленого зображень, $width \times height$ -- розміри зображення (Таблиця).

Таблиця

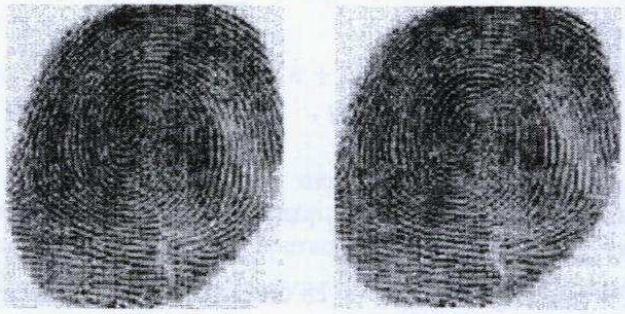
Коефіцієнт стискання	Відстань
3	2.93
5	4.15
10	6.84
15	8.10
20	9.51
25	11.53
30	12.59
35	13.52
40	14.19
45	15.15

Нижче приведено результати роботи програми на прикладі фотографії відбитку пальця

Оригінальне зображення: Коефіцієнт стискання 10:



Коефіцієнт стискання 20: Коефіцієнт стискання 40:



ВИСНОВКИ

Основним вдосконаленням аналізованого методу є, звичайно, використання кращих систем сплесків, тобто систем, що краще відбивають специфіку вхідної функції, в даному випадку зображення. На сьогодні найкращими базисами сплесків, що застосовуються для стискання зображень, є сплески Добеші [2]. Кращі результати можна одержати, використовуючи для перетворення суперпозицію різних базисних функцій сплесків. Також важливим чинником, що дуже сильно впливає на коефіцієнт стиску, є правильний вибір таблиці квантизації. У даній роботі при знаходженні коефіцієнтів квантизації враховується вплив кожного коефіцієнта окремо, хоча в загальному випадку втрати при квантизації двох коефіцієнтів не будуть рівні сумі втрат для кожного коефіцієнта окремо. Найкращим методом знаходження цих коефіцієнтів буде використання класичних методів оптимізації, але час роботи таких алгоритмів може бути дуже великим.

Література

1. Mallat S. *Multiresolution approximation and wavelets.* //Trans. AMS – 1989 --315, -- P. 69-88.
2. Daubechies I., *Ten lectures on wavelets.* // SIAM Philadelphia, 1992. -- P.377
3. Новиков И.Я., Стечкин С.Б. *Основы теории всплесков* // Успехи мат. наук. – 1998 -- 53. № 6 (324) -- С. 53-128
4. Pennebaker W.B., Mitchell J.L., *JPEG Still Image Data Compression Standard.* Van Nostrand Reinhold., 1993, P. 656.
5. Bell T., Cleary J., Witten I., *Text compression.* Prentice Hall, 1990, P.337

COMMON APPROACH TO SOLUTION OF SOME PROBLEMS OF DISCRIMINANT AND CLASTER ANALYSIS

O.V.Babak[1], *A. S. Gasanov*[1], *T.A. Babayev*[2], *J. V. Backlan*[3]

1. International Research and Training Centre of Information Technologies and Systems,
40, prospect Academica Glushkova, 03680, Kiev, Ukraine.
Ph.: (38044) 266-41-87 fax: (38044) 266-15-70,
E-mail: vig@tel.dlab.kiev.ua

2. Baku Scientific and Training Centre
11, I. Idayat-Zade, 370154, Baku, Azerbaijan,
Ph.: (99412) 66-38-81, fax: (99412) 66-39-95
E-mail: tbabayev@azeri.com

3. Education-Scientific complex "Institute for Applied System Analysis",
37, Peremoghy ave, 01056, Kiev, Ukraine.
Ph: (38044) 243-59-20, fax: (38044) 243-59-20
E-mail: kopernik2000@mail.ru

Abstract In this work a common approach is proposed, that simplifies solution of some problems of the discriminant and claster analysis. The approach is based on restoration of an objectively existing relation, hidden in input data. The relation can be synthesized as a multiplicative function in accordance with obtained information on direction of gradient components of linear model of object classification. The revealed relation projects input information onto an axis and makes it possible to get an enough simple decision rule of dividing objects onto classes, or to establish, that in a given class of uniform objects there exist subclasses of objects, similar in a sense.

ВСТУП

Статистичний підхід щодо розпізнавання образів, методами якого є дискримінантний та кластерний аналіз, займає значне місце у різних сферах науки та техніки. Статистичні моделі алгоритмів розпізнавання засновані на використанні апарату математичної статистики. Вони застосовуються загалом у тих випадках, коли відомі або можуть бути просто визначені ймовірні характеристики класів, що, наприклад, відповідають функції розподілу.

З математичної точки зору засоби розв'язання задач у вказаних двох галузях розпізнавання образів мають істотні відмінності. Тому створення єдиного підходу до їх розв'язання навіть щодо окремих випадків, викликає певний практичний та теоретичний інтерес. Така можливість виникає, якщо у вихідній інформації об'єктивно існують приховані залежності одного характеру, виявлення яких дозволяє здобути бажаний результат. Наслідком цього є істотне спрощення розв'язання задачі, що одержується за допомогою одного і того ж алгоритмічного та програмового забезпечення.

Методологія розпізнавання при цьому виглядає приблизно так. Допустимо, наприклад, що є дані, які отримано в результаті фізичного або імітаційного експерименту. Ці дані в деякому дуже обмеженому змісті, що характеризують досліджуваний об'єкт або явище, необхідно спробувати звести разом для того, щоб встановити, які закономірності відбуваються в наявному матеріалі. Для цього висувається деяка робоча гіпотеза, що породжується евристикою, якій придається математичний образ – потім покладений в основу моделі розпізнавання тієї чи іншої задачі дискримінантного або кластерного аналізу.

1. ЕЛЕМЕНТИ ТЕОРІЇ ТА ОСОБЛИВОСТІ АЛГОРИТМІВ

Перш ніж перейти до особливостей побудови єдиного підходу до розв'язування деяких завдань дискримінантного та кластерного аналізу відзначимо:

При дискримінантному аналізі, у разі розпізнавання, наприклад, двох класів об'єктів, за наявними даними або відновлюються, або виявляються дискримінантні залежності y^* від n – числа ознак (змінних) x. Доречно відзначити, що тут відновлення та виявлення – це принципово різні задачі. Якщо відновлення – це вибирання за заданим

© UkrObraz'2000

набором ознак її структури з певного класу залежностей [1], то при виявленні використовується не весь набір ознак, а лише кращі з них або їх комбінації. Особливо важливим є виявлення комбінацій ознак, тому що вже давно спостережено, що часто "розрізнювання" приховане не у самих ознаках, а в їх сполученнях [2]. Таким чином, здійснюється перехід до нового, лінійного за параметрами й нелінійному за змінними простору, вимірність якого завдяки редуціювання вхідної інформації можна істотно знизити (аж до одновимірного).

Під кластеризацією об'єктів зазвичай вважається виявлення у певній матриці даних звичайного, об'єктивно існуючого порядку, завдяки чому стає можливим виділення кластерів – деяких підмножин множин, що досліджуються.

У праці [2] наведено основні різновиди кластерного аналізу, які зустрічаються найчастіше. Найбільш розповсюджений метод виявлення кластерів пов'язаний із завданням на багатьох спостереженнях належної метрики, наприклад, евклідової та вирахуванням за допомогою її відстаней між усіма точками вказаної множини.

Доречно відмітити, що особливістю процесу виділення кластерів є те, що він може розглядатися й як геометричне завдання про виявлення у певному просторі "щільних" накопичень точок. Вказана особливість обумовлює можливість розвинути зовсім специфічні методи її вирішення. Один з таких методів, що дозволяє не лише спростити рішення, але й візуально "побачити" підсумки кластеризації розглядається у поданій роботі. Ідея запропонованого методу базується на редукції інформації, в результаті чого вихідну, як правило, складну багатомірну задачу перетворюють у більш просту, наприклад, трьохмірну, двомірну або одномірну. При цьому існують два принципово різних напрямки скорочення розмірності простору наявних даних.

В основі першого напрямку лежить стиснення набору вихідних факторів (незалежних змінних) максимум до трьох за рахунок відсіювання решти, якщо вдається встановити, що останні практично не впливають на виділення кластерів. Вказана процедура носить назву цілеспрямованого проектування й, зазвичай, опирається на метод головних компонент [5], технічна реалізація якого, відрізняючись складністю, і є можливою лише тоді, коли фактори корельовані між собою. Другий напрямок, що розвивається нами, базується на пошукові деякої компактної функціональної залежності, що об'єднує усі значні і некорельовані (корельовані) між собою фактори, число яких може бути більше трьох, у вигляді так званої узагальненої змінної (поняття введено у працях О. Г. Івахненко [4]), завдяки чому стає можливим перетворення вихідної багатомірної задачі в зручну одномірну.

Звичайно, відображення багатомірного простору на одномірне теоретично не може зменшити мінімально досяжний рівень помилки виявлення кластерів. Однак, з успіхом можна вважати, що якщо дані, які породжуються певним джерелом, підпорядковуються нормальному закону розподілу, то ці жертви будуть мінімальними [7]. Відрізняючись, як буде вказано нижче, простотою технічної реалізації, другий напрямок є плідним для виділення кластерів, практично некорельованих між собою значущих факторів.

Оскільки при кластерному аналізі об'єкти поділяються на групи за їх подібністю у певному понятті, що можна інтерпретувати й так: дані, що належать до певного класу однорідних об'єктів поділяються на підкласи. Саме таке розуміння розбиття класу подібних об'єктів на підкласи разом з проектуванням вхідної інформації на одновимірний простір створює базу єдиного підходу до розв'язання деяких задач дискримінантного та кластерного аналізу.

Виникає питання: як знайти оптимальні у певному сенсі сполучення ознак, не вдаючись до комбінаторного перебору, що потребує великого обсягу обчислень, не гарантуючи вірності вибору того чи іншого критерію. Відповідь на це питання базується на таких положеннях:

1. Більшість явищ довкілля і відповідно їх моделей нелінійні. При цьому найважливіша інформація про характер структури такої моделі є в лінійній за параметрами й змінними моделі явища, що може бути представлена в виді (1) в якій є інформація про складові градієнта функції відгуку:

$$\hat{y} = \sum_{i=1}^{n} a_i x_i , \qquad (1)$$

де $\hat{y}$ - прогнозоване значення функції відгуку y, a_i - оцінки коефіцієнтів, x- число ознак (змінних).

2. Більшість моделей класичної фізики є мультиплікативні і можуть бути представлені у виді (2):

$$y = k\prod_{i=1}^{n} x_i^{p_i} , \qquad (2)$$

де $k>0$ – коефіцієнт (раціональне число), p_i – показник степені позитивне (негативне) раціональне число.

3. Розв'язання деяких задач дискримінантного та кластерного аналізу зводиться до виявлення прихованих у вихідних даних закономірностей, що об'єктивно існують, у вигляді мультиплікативних моделей типу (2).

В роботі [3] показано: якщо у лінійній моделі (1) коефіцієнти a є оцінки напрямку складових градієнту функції y, то можна синтезувати узагальнену змінну.

$$v = \prod_{i=1}^{n} x_i^{p_i}, \qquad (3)$$

де p_i дорівнюють ±1, залежно від знаку відповідної компоненти градієнта.

Функції (1) і (3) у певному змісті подібні. Подібність їх є у тому, що із зменшенням (зростання) будь-якої змінної відповідно змінюватимуться (зростатимуть або зменшуватимуться) значення y та v. Саме це, зберігаючи фізичні особливості явища при переході від його лінійної моделі до нелінійної, забезпечує вибір сполучень ознак, не удаючись до комбінаторного перебору. Цей результат закладено в основу єдиного підходу до розв'язання деяких задач дискримінантного та кластерного аналізу. Схема алгоритмів їх рішення включає такі етапи:

1. За даними навчальної вибірки, де для позначення належності об'єктів до двох класів або одного, використовується, відповідно, значення ± 1 або +1, відновляється лінійна функція (1) у результаті чого знаходяться оцінки коефіцієнтів a, що є оцінками напрямку складових її градієнту.

2. Знаючи напрямок складових градієнту функції відгуку, синтезується узагальнена змінна v (3).

3. Значення v_j, $j = \overline{1,l}$, де l – довжина навчальної вибірки, якщо це необхідно – масштабуються і потім розміщуються на одній координатній осі.

4. Характер розміщення v_j аналізується в залежності від роду задачі.

Примітка. За обчислення v можуть бути використані, як нормовані, так і ненормовані, значення незалежних змінних. При цій ситуації, значення, що включають їхнє нульове не розглядаються, оскільки в даному випадку класифікація стає неможливою.

Далі розглядатимуться особливості алгоритмів розв'язання таких задач.

2. ЗАДАЧА ДИСКРИМІНАНТНОГО АНАЛІЗУ

Нехай задано навчаючу вибірку довжиною l

$$\{x_{ij}, y_j^*\}, \quad i = \overline{1,n}, \quad j = \overline{1,l}, \qquad (4)$$

причому $\forall x_{ij} \neq 0$, де y_j^* приймає значення ±1 в залежності від належності класу A або B елементів вибірки за якою відбудовується лінійна дискримінантна функція y^*. При цьому стають відомими оцінки напрямку складових її градієнту і можливий синтез v [3]. Оскільки значення v_j, $j = \overline{1,l}$, розташовані на одній координатній осі, нескладно встановити межі належності їх до класів A і B та одержати у даному простому випадку розв'язальне правило.

При побудові його необхідно врахувати важливе поняття дикримінантного аналізу – "відмова від розпізнавання" [2]. Сутність його полягає в тому, що в сумнівних випадках (існує область, що включає об'єкти як класу A, так і класу B), або при поганій згоді (існує область в якій об'єкти розташовані далеко від середніх значень класів) класифікація неможлива. Отже, при необхідності в таких випадках вона повинна здійснюватися іншими засобами, зокрема, наприклад, виходячи з евристичних розумінь. Природно, використовуючи відмову від класифікації, можна різко підвищити її якість при зменшенні числа правильних класифікаційних рішень. Таким чином, з урахуванням сумнівних випадків (область з межами v_a і v_b, $v_a < v_b$, де a, b деякі числа з вибірки $\overline{1,l}$) і при поганій згоді (область з межами $-\infty$, v_c і v_d, $+\infty$ де $v_c < v_a$ і $v_b < v_d$ і c, d теж деякі числа з вибірки $\overline{1,l}$) вирішальне правило матиме вигляд:

$$v_c < v < v_a, \quad v \in \text{класу } A,$$
$$v_b < v < v_d, \quad v \in \text{класу } B.$$

При цьому оцінка вирішальних правил повинна бути виконана на основі оцінювання класифікаційної помилки. Найпростіша можливість полягає у застосуванні його до об'єктів навчання і підрахунку числа невірних рішень. Інша можливість більш суворої оцінки пов'язана з поділом множини об'єктів на навчальне і контрольне. При цьому навчальна множина використовується при побудові вирішального правила, а контрольна - для оцінки його якості. Найбільш сильний результат може бути отриманий із застосуванням методу контролю, що стежить і який передбачає вилучення з навчальної множини лише одного об'єкта, який потім використовується для оцінки якості вирішального правила [1].

3. ЗАДАЧА КЛАСТЕРНОГО АНАЛІЗУ

Нехай задано навчаючу вибірку

$$\{x_{ij}\}, \quad i = \overline{1,n}, \quad j = \overline{1,l},$$

при чому $\forall x_{ij} \neq 0$, що репрезентує клас однорідних об'єктів. Аналогічно до вибірки дискримінантного аналізу її можна записати як (4) з

тією різницею, що усі v прийматимуть значення +1 (тільки один клас). Далі, як і раніше, відбудовується лінійна функція (1) і синтезується v. У зв'язку з тим, що v_j розміщені на одній осі, легко побачити чи є на ній "густі" скупчення точок (підкласи подібних об'єктів або кластери). Необхідно відмітити, що проблема остаточного виявлення кластерів зовсім не проста, оскільки оцінка міри подібності вельми суб'єктивна.

Пояснюється це тим, що в кластерному аналізі не існує однозначного кількісного критерію подібного помилці класифікації в дискримінантному аналізі. Такий критерій не можна сформувати, оскільки в різноманітних прикладних задачах різноманітними можуть бути і цілі аналізу, спрямовані на одержання нової інформації. У аналізованому випадку усе ж доцільно ввести ряд понять, що дозволяють якоюсь мірою порівняти і сприяти виділенню кластерів. До таких понять ставляться:

по-перше, довжина кластера L визначається наступним способом:

$$v_{j\min} = \min\{v_j; \overline{1,k}\},$$
$$v_{j\max} = \max\{v_j; \overline{1,k}\},$$
$$L = (v_{j\max} - v_{j\min}),$$

де k - число кластерів;
по-друге, щільність кластера:

$$d_j = \frac{m_j}{L_j}, \quad j \in \overline{1,k},$$

де m - число елементів j-го кластера.

З огляду на те, що кластер є утворенням, де внутрішні структурні зв'язки в декілька разів більші порівняно з зовнішніми зв'язками одного кластера з іншим, для виділення кластерів також корисно ввести поняття міри близькості елементів кожного кластера у вигляді

$$\forall(v_{i+1} - v_i) \le \varepsilon,$$

де i - елемент кластера, ε - деяке чисельне значення, що може бути задане відповідно до характеру тієї чи іншої задачі кластеризації.

ВИСНОВКИ

У відповідності з викладеним можна зробити наступні висновки.

1. Запропоновано єдиний підхід до вирішення деяких задач дискримінантного та кластерного аналізу, особливостями яких є наявність у вхідних даних прихованих об'єктивно існуючих залежностей, які можна відобразити у вигляді мультиплікативних функцій [3].

2. В основі єдиного підходу до вирішення деяких задач дискрімінантного і кластерного аналізу лежить у першому випадку розбиття багатьох різного роду об'єктів на класи, а у другому випадку розбиття багатьох подібних об'єктів (класу об'єктів) на підкласи. Ця обставина дозволяє розглядати вищевказані задачі з однієї і тієї ж математичної точки зору, що спирається на побудову лінійних моделей, які утримують інформацію про спрямування складових градієнта лінійних функцій.

3. Наявність інформації про спрямування складових градієнта лінійної моделі при рішенні деяких задач дискримінантного і кластерного аналізу дозволяє здійснити синтез узагальненої змінної у вигляді мультиплікаційної функції, зберігаючи фізичні особливості явища.

4. Значення мультиплікаційної функції у задачах дискримінантного і кластерного аналізу можуть розташовуватися на одній координатній осі, що суттєво спрощує процедуру побудови алгоритмів класифікації та кластеризації і робить її доступною для широкого кола користувачів.

ЛІТЕРАТУРА

1. Алгоритмы и программы восстановления зависимостей (под ред. В. Н. Вапника). - М.: Наука. - 1984. - 816 с.
2. Распознавание образов: состояние и перспективы. Пер. с англ. (К. Верхаген, Р. Дейн и др.). - М.: Радио и связь. - 1985. - 104 с.
3. Бабак О. В. Об одном подходе к решению задачи восстановления зависимости в классе кусочно-линейных функций // Проблемы управления и информатики. - 1995. - N6. - С. 134-141.
4. Ивахненко А. Г., Мюллер Й. А. Самоорганизация прогнозирующих моделей. - К.: Техника, 1985: Берлин: ФЕБ Ферлаг Техник, 1984. - 233 с.
5. Прикладная статистика: Классификация и снижение размерности: Справ. Изд./С.А.Айвазян и др., под ред. С. А. Айвазяна. - М.: Финансы и статистика, - 1989. – 426 с.
6. Распознование, классификация, прогноз/ Отв. ред. Ю. И. Журавлев. – М.: Наука, 1989. – 280 с.
7. Дуда Р., Харт П. Распознавание образов и сцен. – М.: Мир, 1976. – 511 с.8

Принципи зовнішнього доповнення у методі групового урахування аргументів (МГУА) та в методі граничних спрощень (МГС)

Васильєв В.І., Ланге Т.І., Кобець Н.М.

В. І. Васильєв: Міжнародний науково-навчальний центр ЮНЕСКО/МПІ інформаційних технологій і систем НАН і МО України (МННЦ ІТІС), Україна, 252022, Київ, пр. Академіка Глушкова 40, т. (044)2664187, факс (044)2661570
Н. М. Кобець: Національний технічний університет України "Київський політехнічний інститут", пр. Перемоги 37,т. (044)4411892
Т. І. Ланге: Fachhochschule Merseburg, Fachbereich Elektrotechnik, FH Merseburg,Geusaer Straβe, D-06217 Merseburg, Deutschland, т.(03461)462528, факс (03461)462919

Анотація - Розглядаються та аналізуються два індуктивних методи відновлення функції - метод групового урахування аргументів (МГУА) і метод граничних спрощень (МГС). Вказано їхні переваги перед іншими методами та можливість їхнього поєднання з метою поліпшення їхніх екстраполяційних властивостей.

ВСТУП

Всі індуктивні методи, які засновані на неповній індукції, а до них безумовно відносять МГУА [1] та МГС[2], відрізняються тим, що в них загальні висновки робляться на підставі часткових фактів, а це може призвести як до вірних, так і до помилкових рішень. Причина такої невизначеності полягає в тому, що окремі факти, на яких засновуються загальні висновки, не завжди достатньо добре характеризують досліджуване явище. Разом з цим отримані загальні висновки повинні пояснювати не тільки окремі факти, але і все досліджуване явище в цілому, тобто ці загальні висновки не повинні змінюватися при практично нескінченному збільшенні кількості експериментів. Тому якість індуктивного висновку визначається не тільки і не стільки поясненням окремих, отриманих у процесі експериментів фактів, скільки екстраполяційними спроможностями цих висновків, їхньою здатністю до експансії в область явища, що не охоплена експериментами.

Кожний раз, коли модель вибирається із надто складного класу, здебільшого не вистачає емпіричних даних для її однозначного пояснення, тобто модель виявляється складніше того, що несуть у собі накопичені факти, і ці факти просто не в змозі відтворити таку модель.

У задачах відновлення багатовимірних залежностей будь-яке спрощення моделі приводить до згладжування тих або інших деталей. Навпаки, надмірне ускладнення моделі без урахування обсягу експериментальних даних призводить до непомірної свободи поводження апроксимуючої функції в області, що не була охоплена експериментом, у той час як більш прості моделі в цій області поводяться більш "обережно".

Далі будуть розглядатися індуктивні методи стосовно до вирішення задачі відновлення функцій, суть якої у виявленні та моделюванні деякої закономірності, що зв'язує основні характеристики досліджуваного явища у функціональну залежність наступного виду

$$y = F(X) \quad . \qquad (1)$$

Зробити це потрібно на підставі обмеженої кількості експериментів, тому тут постає центральна проблема усіх індуктивних методів, яка полягає в правильному співвідношенні складності моделі, що синтезується, з кількістю емпіричних даних.

Серед деяких методів, у яких особлива увага приділяється пошуку такого співвідношення, виділяються метод групового урахування аргументів (МГУА) [1] та метод граничних спрощень (МГС) [2], [3]. Тому тут будуть розглянуті ці два методи з погляду на можливість їх взаємодоповнення.

Особливості методу групового урахування аргументів (МГУА) [1]

У основу методу покладено декілька основних принципів: неостаточність проміжних рішень, зовнішнього доповнення, самовідбору проміжних рішень, єдиності остаточного рішення. Одночасно з оптимізацією коефіцієнтів апроксимуючого полінома відбувається й оптимізація складності, що дозволяє не піклуватися про вибір структури цього полінома. Спеціальний критерій (зовнішнє доповнення) є перешкодою переускладненню моделі. Цей критерій відкидає усе, що не в змозі відтворити коротка навчальна вибірка (відкидання

© UkrObraz'2000

«сміття»). В результаті залишається тільки те, що може бути надійно «підтверджено» конкретною вибіркою, а відкидаються «фантазії» моделі.

Задача розв'язується у декілька етапів. Спочатку для всіх незалежних змінних (аргументів) створюються усі можливі групи комбінацій. У практичних алгоритмах до кожної комбінації належить лише два аргументи (попарне урахування аргументів). Для кожної пари (групи) створюється частковий опис, тобто деяке просте рівняння не вище другого порядку, аргументами якого є обрана пара. Вид часткового опису однаковий для всіх груп протягом усього процесу навчання. А всю вибірку поділяють на дві частини: навчаючу та перевірочну. Таким чином утворюється зовнішнє доповнення (перевірочна вибірка), що відіграє роль сита, яке відсіває усе надмірно складне, що не має права на існування при таких обмеженнях інформації. Коефіцієнти часткових описів визначаються за даними навчаючої вибірки. В результаті утворюються множини вирішень, оскільки часткове рівняння кожної пари розглядається як деяка спрощена модель функції, яку відновлюють.

У якості апроксимуючого полінома дуже часто використовується поліном Колмогорова-Габора [1]:

$$y = \alpha_0 + \sum_{i=1}^{m} \alpha_i x_i + \sum_{i=1}^{m}\sum_{j=1}^{m} \alpha_i x_i x_j + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m} \alpha_i x_i x_j x_k + \ldots \quad (2)$$

тому що за допомогою такого полінома можна досягнути достатньо точної апроксимації будь-якої диференційованої функції F.

Цю складну залежність заміняємо множиною простих функцій:

$$y_1=f(x_1,x_2);\ y_2=f(x_1,x_3);\ y_s=f(x_{m-1},x_m), \quad (3)$$

де $s = C_m^2$, причому функція f усюди однакова.

Дуже часто в якості функції f вибираються прості залежності.

$$y(x_i,x_j) = a_0+a_1x_i+a_2x_j+a_3x_ix_j \quad (4)$$

У другому ряду розглядаються моделі виду

$$z_1= f(y_1,y_2);\ z_2= f(y_1,y_3);\ldots;\ z_s= f(y_{s-1},y_s) \quad (5)$$

Складність полінома зростає від ряду до ряду. Так, наприклад, у другому ряду будуть отримані поліноми, що містять нелінійні члени виду x_1^2, $(x_1\ x_3)$, $(x_2\ x_3)$, $(x_1^2\ x_3)$, $(x_1^2 x_2\ x_3)$ і т.д., коефіцієнти цих поліномів можуть бути визначені по тих же точках навчальної послідовності і не потребують додаткової інформації, хоча їхня складність увесь час зростає. При цьому кількість коефіцієнтів, що визначаються, значно перевищує кількість точок навчальної послідовності. Якби не було зовнішнього доповнення, тобто перевірочної вибірки, то алгоритми МГУА, завдяки складності синтезованих поліномів, могли б абсолютно апроксимувати функцію (1) у всіх точках навчальної вибірки, але при цьому не залишилося б ніяких гарантій задовільного поводження функції, яку відновлюють, на нових точках.

Зупинка алгоритму відбувається відразу ж по досягненні єдиного мінімуму відхилень, отриманих на перевірочній вибірці. Тим самим вибирається модель оптимальної складності, що встановлює компроміс між складністю й обсягом інформації, що використовується при синтезі моделі. Як тільки складність перевищить можливості навчальної інформації, процес синтезу моделі припиняється.

Особливості методу граничних спрощень (МГС)

Будь-яка індукція містить у собі елементи дедукції і з нею пов'язана, та, навпаки, дедукція пов'язана з індукцією. Такий діалектичний взаємозв'язок індукції і дедукції наводить на думку про те, що будь-які індуктивні методи повинні містити в собі елемент дедукції, тобто спиратися на деякі твердження, перевіряти на них правильність індуктивного висновку на всіх етапах його формування. Саме ці міркування були прийняті до уваги при створенні МГС. В цьому методі зовнішнім доповненням є дедуктивне твердження (теорема), що не дозволяє надмірно ускладнювати результуючу модель, побудовану методами індукції. У цьому методі, як і у МГУА, зовнішнє доповнення є фільтром, що відкидає "сміття" і залишає тільки те, що формує аналізовану закономірність (1).

Задача відновлення багатовимірної функції (1) у МГС спочатку зводиться до стандартної задачі навчання розпізнаванню образів (НРО), а потім рішення задачі НРО забезпечує рішення задачі відновлення із заданою якістю і надійністю. Нехай задана навчальна вибірка пар $x_1y_1, \ldots, x_ly_l$, де x_v - вектор, а y_v - скаляр, який визначає значення функції F (1) у точці, що відповідає вектору x_v. Потрібно відновити багатовимірну неперервну функцію F так, щоб для будь-якого x_v, виконувалася нерівність

$$|y_v-F(x_v\alpha)|\le\xi, \quad (6)$$

де α - параметри функції, що відновлюється.

Хай, як і в [3], кожному x_v відповідає два значення y:

$$y_{v1}= y_v+\xi; \qquad y_{v2}= y_v - \xi; \quad (7)$$

При цьому навчальна вибірка збільшиться вдвічі і розділиться на дві підмножини V_1 і V_2, одна з яких містить елементи (x_{v1}, y_{v1}), а інша (x_{v2}, y_{v2}). Підмножини V_1 і V_2 можна розглядати як образи, які є явно відокремленими, і якщо

вдасться розділити ці образи, то тим самим буде відновлено функцію $F(x, \alpha)$, а це гарантує виконання (6) для всієї навчальної вибірки. Порушення співвідношення (6) буде відбуватися з частотою помилкового розпізнавання образів V_1 і V_2. Для поділу на образи можна використовувати будь-який алгоритм навчання розпізнаванню образів, у тому числі й альфа-процедуру [3], що, у свою чергу, є однією з реалізацій методу граничних спрощень. Як вирішальна, тобто, і апроксимуюча функція $F(x, \alpha)$ використовується поліном (2), у якому m-розмірність вектора $X=x(x_1,\ldots,x_m)$, а α - параметри, що настроюються.

З геометричної точки зору поліном (2) є гіперплощиною у спрямляючому просторі узагальнених координат x_i; x_ix_j; $x_ix_jx_k$, і т.д. Процедура відбере тільки ті доданки, які не будуть суперечити нерівності (6), тобто процедура як би пропускає поліном (6) через сито, залишивши тільки те, що визначає залежність (1), а все «сміття» буде відкинуте.

У процесі навчання частина членів полінома відкидається, а частина використовується як аргументи функції $F(x,\alpha)$. Таке сортування членів полінома здійснюються за допомогою зовнішнього доповнення, що є деяким дедуктивним твердження, яке засноване на передумовах теорії емпіричного ризику. Основний зміст цієї теорії - це вказування тих умов, при виконанні яких емпіричний ризик (тобто ризик, що обчислюється по емпіричним даним) збігається або майже збігається із середнім ризиком, обчисленим за нескінченними вибірками.

Одне з центральних тверджень теорії емпіричного ризику, що використовується в рамках проблеми НРО, міститься в теоремі [4], в якій доводиться те, що якщо з N вирішальних правил вибирається одне, що безпомилково розділяє випадкову і незалежну вибірку довжини l, то з ймовірністю $(1-\eta)$ можна утверджувати, що ймовірність помилкової класифікації за допомогою цього правила не перевищує наперед задане число

$$\varepsilon = \frac{\ln N - \ln \eta}{l}. \qquad (8)$$

Ця теорема лягла в основу теорії редукції [2], на якій базується метод граничних спрощень. Відповідно до цієї теорії задача синтезу складних поверхонь, що розділяють, у вихідному "засміченому" просторі величезної розмірності заміняється задачею синтезу такого простору малої розмірності ($n_0 << m$), у якому образи, які подані навчальною вибіркою, легко розділяються простим (частіше усього лінійним) вирішальним правилом. Іншими словами, насамперед приводиться у відповідність складність конструкцій, які індукують, із наявним обсягом інформації. Вихідна задача зводиться до більш простої таким чином, щоб її складність не перевищувала можливостей емпіричних даних, тобто складність повинна бути не вище тієї, що може упевнено проявитися в рамках наявної інформації.

Основою редукції при індуктивному синтезі простору є дедуктивне твердження приведеної вище теореми. Розширення простору обмежено дедуктивним твердженням, що не припускає ніяких ускладнень без істотного поліпшення результату. Дедуктивне твердження теореми як би зважує кожне ускладнення і визначає його вартість. Здійснюється послідовний синтез простору, у якому можливий лінійний поділ. При цьому дедуктивне твердження дуже просто перевіряється на навчальній послідовності.

Різновидом МГС є альфа-процедура [3], основна особливість якої полягає в тому, що кожна з m властивостей піддається особливій перевірці за допомогою критерію зовнішнього доповнення, який визначається дедуктивним твердженням теореми. При цьому, перед тим, як будувати індуктивну конструкцію, кожна властивість перевіряється на відповідність теоремі, і якщо практика не суперечить вимогам теорії, то продовжується ускладнення індуктивного висновку.

Якщо проводити аналогію між МГУА і МГС, то в МГУА зовнішнє доповнення формується на основі індуктивного висновку і на кожному кроку порівнюються дві індуктивних конструкції, побудовані на двох різноманітних вибірках: на навчальній і перевірочній. І якщо ці конструкції близькі, то ускладнення продовжується. У МГС порівнюються дві моделі: індуктивна і дедуктивна. Ускладнення є припустимим, якщо відокремлююча силу цього ускладнення, яка легко обчислюється за навчальною вибіркою, перевищує мінімально припустиму відокремлюючу силу, що обчислюється відповідно до теореми і виконує роль зовнішнього доповнення. Як у тому, так і в іншому методі ускладнення продовжується доти, поки воно дає позитивний результат, тобто поки моделі, які порівнюються, не суперечать одна одній.

Можливості поєднання МГУА та МГС

Обидва методи базуються на застосуванні принципу зовнішнього доповнення, щоб усунути некоректності у постановці задачі. У МГУА - це додаткова перевірка проміжних моделей, які отримані на перевірочній вибірці, а в МГС індуктивна модель будується з урахуванням її відповідності з теоретичними рекомендаціями, отриманими в результаті дедуктивних міркувань.

В обох випадках застосовуються засоби проти зайвого переускладнення моделі, але в МГУА це здійснюється шляхом додаткового контролю екстраполяційних властивостей, а в МГС - шляхом забезпечення умов, при яких ці екстраполяційні властивості гарантовані за теорією.

Можна сподіватися, що поєднання цих методів дозволить помітно поліпшити розв'язання задач відновлення залежностей в умовах коротких вибірок шляхом послідовного застосування відразу двох критеріїв зовнішнього доповнення.

Можна використовувати проміжні поліноми МГУА як вихідні дані альфа-процедури, які уже пройшли додаткову перевірку на перевірочній вибірці. Це дає явну перевагу при ускладненні результуючого полінома. У звичайній альфа-процедурі (МГС) після перевірки всіх лінійних членів полінома (2) використовуються всі нелінійні члени цього полінома як додаткові змінні, що складаються з двох аргументів (x_i, x_j), а в комбінованому методі для цієї мети можна використовувати моделі, отримані в першому ряду МГУА, які пройшли самовідбір за будь-яким критерієм зовнішнього доповнення (наприклад, за критерієм середньоквадратичного відхилення на перевірочній вибірці). У цьому випадку альфа-процедура здійснить перевірку тільки тих моделей $y(x_i,x_j)$, які були обчислені за (5), що пройшли поріг самовідбору за критерієм зовнішнього доповнення МГУА. У результаті кожна проміжна модель, тобто кожне значення y пройде подвійну перевірку: на спроможність працювати на нових даних (перевірочна вибірка) і на відповідність теорії, тобто на спроможність взяти на себе навантаження, обумовлене мінімально припустимою відокремлюючою силою. Таким чином, у цьому випадку МГУА відіграє роль відбору нелінійних змінних для альфа-процедури. Лінійні змінні відбираються звичайним алгоритмом альфа-процедури, а якщо серед них не знайдеться достатньої кількості змінних, що мають відокремлюючу силу, більшу за мінімально припустиму, то починає працювати МГУА, який на першому ряду генерує, а потім перевіряє по зовнішньому критерію проміжні поліноми другого ступеня, частина з яких використовуються як вхідні в альфа-процедуру. Якщо ж і це не призводить до повного лінійного поділу навчальної вибірки, то використовуються поліноми більш високих ступенів, що пройшли перевірку на перевірочній вибірці.

Можна поміняти місцями МГС і МГУА, використовуючи МГС для попередньої перевірки змінних, що використовуються у подальшому в МГУА. У цьому випадку відповідно до МГС будується лінійна модель по загальній схемі. Якщо ця модель виявиться не повною, тобто не дасть безпомилкового поділу вибірки, то серед усіх $\mathbf{x_i}$ **(i=1,2,...,m)** вибираються тільки m*, відокремлююча сила яких максимальна, і ці змінні використовуються в першому ряду МГУА для побудови квадратичних часткових поліномів. Коефіцієнти проміжних поліномів визначаються по МНК, а самі поліноми перевіряються на перевірочній вибірці. З побудованих поліномів вибирається m* найкращих і з них формується $C^2_{m^*}$ пар, кожна з яких перевіряється і ранжирується за алгоритмом альфа-процедури по дедуктивному критерію мінімально припустимої відокремлюючої сили. Усі пари, побудовані на змінних, які обчислюються алгоритмом МГУА, ранжируються за дедуктивним критерієм відокремлюючої сили. Кожній парі буде відповідати узагальнена змінна $\tilde{x}_{ij}$ [3]. Кращі пари, тобто відповідні їм змінні $\tilde{x}_{ij}$, використовуються для формування пар нового ряду МГУА. Далі працює звичайна схема МГУА, тобто всім парам приписується свій поліном, і знаходяться по навчальній послідовності його коефіцієнти. Всі поліноми перевіряються на перевірочній вибірці і m* кращих із них формуються в пари, кожна з яких перевіряється за критерієм відокремлюючої сили. Серед усіх пар вибирається m* найкращих і відповідні їм узагальнені змінні $\tilde{x}$ використовуються для формування нових пар для алгоритму МГУА. Така схема спільної роботи двох методів припускає паралельну роботу двох алгоритмів - алгоритму МГУА й алгоритму МГС. Алгоритм МГС припиняє свою роботу тоді, коли на якомусь етапі або відбудеться безпомилковий поділ вибірки, або алгоритм МГУА призведе до погіршення зовнішнього критерію. Як і в попередній комбінованій схемі правило, згідно з яким зупиняється робота алгоритму, можна лишити за конструктором, що вибере момент зупинки роботи або в МГС, або - МГУА.

ЛІТЕРАТУРА

1. Ивахненко А. Г., Степашко В.С. *Помехоустойчивость моделирования*.- Киев: Наукова думка, 1985 - 215 с.

2. Васильев В.И. *Теория редукции в проблемах экстраполяции.* //Проблемы управления и информатики. 1996 - №1-2, с.

3. Васильев В.И., Суровцев И.В. *Индуктивные методы обнаружения закономерностей, основанные на теории редукции* //Управляющие системы и машины - 1998 - №5 с.3-14

4. Вапник В.Н. *Алгоритмы и програмы восстановления зависимостей.* - М.: Наука, 1984 - 815 с.

5. Вапник В.Н., Червоненкис А.Я. *Теория распознавания образов*.- М.: Наука, 1974. -416 с.

Ітеративна процедура, яка використовує принцип дуальності в задачах навчання розпізнаванню образів

Васильєв В.І., Струк С.П.
МННЦ ІТІС, Україна, Київ

Розглядається ітеративна процедура навчання розпізнаванню образів, побудована на основі теорії редукції з використанням принципу дуальності. В основі запропонованої процедури лежить альфа-процедура, яка дозволяє послідовно корегувати простір в процесі його розширення і вирішальне правило, яке працює в ньому. Наводиться приклад розв'язку конкретної задачі.

ВСТУП

Принцип дуальності вперше був сформульований в теорії оптимального управління для умов, коли керуючий вплив носить двоїстий характер: з одного боку він використовується для приведення об'єкта управління в необхідний стан, а з іншого боку — для вивчення й уточнення математичної моделі цього об'єкта. Не можна керувати об'єктом оптимально, не знаючи його характеристик. Але можна, тимчасово відклавши убік оптимальність, керувати, уточнюючи при цьому характеристики моделі, покращуючи тим самим управління, постійно наближаючи його до оптимального. Недостатність вихідної інформації приводить до необхідності об'єднувати вивчення об'єкта з його управлінням. У системах дуального управління завжди є протиріччя між цими двома функціями управління. Успішне управління можливе лише при достатній близькості об'єкта і його моделі, але близькість їхніх характеристик можлива лише при достатньому об'ємі інформації про об'єкт, яка накопичується під час керування. Можна почати керувати, не чекаючи уточнення всіх особливостей моделі, але за це доводиться платити тимчасовим погіршенням якості управління, одержуючи натомість нову інформацію і поліпшуючи модель.

Аналогічно вищесказаному, при навчанні розпізнаванню образів часто в міру надходження навчальної інформації на кожному кроці уточнюється вирішальне правило. Спочатку імовірність помилкового розпізнавання нових об'єктів може бути великою, але зате інформація про неї дозволяє корегувати вирішальне правило. При цьому корекція звичайно проводиться тільки тоді, коли правило помиляється. Так працює багато алгоритмів послідовного навчання, наприклад – *перцептронні*, але в них розв'язки приймаються тільки після завершення навчання. Крім того, ці алгоритми діють весь час в одному і тому ж просторі властивостей об'єктів.

Алгоритми навчання часто втрачають привабливість тільки тому, що для одержання необхідної якості і надійності розпізнавання їм потрібна навчальна вибірка занадто великого об'єму, що особливо неприємно, коли її об'єкти з'являються рідко, і для кожного з них треба проводити дорогі вимірювання. Доводиться довго чекати і витрачати великі матеріальні ресурси поки буде отримана вибірка, об'єм якої дозволить одержати потрібні характеристики розпізнавання. У цьому випадку розумно використовувати принцип дуальності: нехай погано навчена система починає розпізнавати відразу ж при надходженні об'єктів різних класів, причому кожна помилка розпізнавання буде сигналом до корегування вирішального правила. Таким чином можна використовувати будь-які ітераційні алгоритми навчання розпізнаванню образів, але тут будуть розглянуті тільки алгоритми, побудовані на принципах теорії редукції [1], які мають в даному випадку значні переваги.

Для алгоритмів, побудованих на теорії редукції, загальним є те, що при навчанні вони синтезують простір ознак, розмірності $n_0 << m$ (m – кількість вихідних властивостей, які вимірюються на об'єктах), в якому лінійне вирішальне правило безпомилково розділяє навчальну вибірку. Однак на практиці часто зустрічаються завдання, в яких навчальна вибірка формується поступово, один об'єкт за іншим. У цьому випадку бажано організувати навчання так, щоби відбір ознак починався відразу ж після появи перших об'єктів. При цьому подвійність (дуальність) будуть виявляти відразу дві найважливіших характеристики навчання: вирішальне правило і простір, у якому це правило діє.

© UkrObraz'2000

Опис базової альфа-процедури

Процедура навчання для випадку бінарних ознак вже розглянута в [2]. Тут же буде розглянута альфа-процедура (аналогічна описаній в [2]), і яка успішно працює у випадку неперервних і змішаних просторів, де принцип дуальності використовується з таким же успіхом. Як і у випадку бінарних ознак, тут простір формується послідовно — одна ознака за іншою, а *кроками ітерації* будемо називати моменти відбору нової, додаткової ознаки, що дозволяє в розширеному просторі розділити вже накопичену вибірку лінійно і безпомилково. Навчання починається, коли вперше одночасно виконуються відразу два співвідношення:

$$V_1[l] \neq \varnothing, \quad V_2[l] \neq \varnothing, \qquad (1)$$

де $V_1[l]$ і $V_2[l]$ — множини першого і другого образу на навчальній вибірці довжиною l.

Співвідношення (1) показують, що множини обох образів не порожні і містять хоча б по одному об'єкту. Побудовані на першому кроці простір і вирішальне правило в ньому відразу ж використовуються для розпізнавання. Діє це правило до першої помилки, після чого починається наступний етап навчання, що полягає в розширенні простору. Сигналом для продовження процесу навчання може служити одна, чи кілька помилок, або задане відношення кількості помилок до довжини накопиченої навчальної вибірки. На кожному кроці навчання простір розширюється шляхом додавання такої нової ознаки, яка дозволяє побудувати в новому просторі лінійне вирішальне правило, що безпомилково розділяє накопичену навчальну вибірку.

На першому кроці одна з множин (V_1 або V_2) містить тільки один елемент, що приводить до виконання умови (1). Як тільки виконується умова (1), створюється C_m^2 пар із всіх m вимірюваних на об'єктах властивостей, і на площині кожної пари для кожного об'єкта відповідно до альфа-процедури [1] обчислюється направляючий параметр

$$\tilde{x}_{ik} = \rho_i \cos(\beta_i + \alpha_{si}), \qquad (2)$$

де i та k – індекси однієї з пар координат x_i та x_k, вибраних з C_m^2 властивостей, $\tilde{x}_{ik}$ – деяка пряма, що проходить через початок координат під кутом α_{si} до координати x_i; $\rho_i = \sqrt{x_i^2 + x_k^2}$; $\beta_i = \arctan\dfrac{x_i}{x_k}$ — кут, що характеризує кожен об'єкт в координатах x_i та x_i, $\alpha_{si} = 0°\ldots360°$ – кут повороту $\tilde{x}_{ik}$; s – змінний параметр, що визначає величину кута α_{si}.

На Рис.1 показано, як визначаються параметри α-процедури для одного об'єкта X_j і для одного значення кута α_{si}. Величина $\tilde{x}_{ik}$ є проекцією точки X_j на напрямок $\tilde{x}_{ik}$ і визначається координатами x_i та x_k, які відповідають об'єкту X_j. Кут α_{si} змінюється на певну величину (наприклад, $\Delta\alpha = 1°$) від 0° до 360° з відліком від напрямку x_i. В процесі обертання вибирається такий кут α_{si}, для якого $\alpha_{si}^* = \arg\max_s(\omega_{is})$. На першому кроці ω_{is} визначається відстанню між єдиним об'єктом одного з образів і найближчим до нього об'єктом другого образу. Кут α_{si} відповідає максимальній відстані, що ділиться навпіл, а на середині відрізка ставиться перпендикуляр, що визначає вирішальне правило для пари x_i и x_k. Із всіх C_m^2 пар вибирається така, для якої відстань ω_{is} максимальна, і тоді відповідний перпендикуляр буде визначати вирішальне правило першого кроку навчання. При появі нових об'єктів щоразу корегується α_{si}^* для обраної найкращої пари, так щоби відповідна відстань ω_{is} прийняла максимальне значення. Якщо при появі деякого об'єкта не вдалося знайти α_{si}, при якому образи вибірки розділяються безпомилково, то переходять до другого кроку навчання. Пара, вибрана на першому кроці, визначається напрямком $\tilde{x}_{ik}[1]$, побудованим на цьому кроці. На другому кроці навчання проглядаються всі $(m–2)$ властивостей, що залишилися, і кожна з яких утворює з напрямком $\tilde{x}_{ik}[1]$ площину, в якій потрібно організувати обертання відрізка $\tilde{x}_2$ навколо початку координат. Для кожного кута і кожного об'єкта вибірки знаходиться направляючий параметр

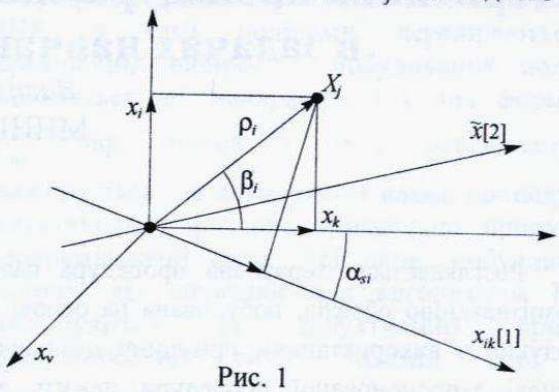


Рис. 1

$$\tilde{x}_2 = \rho_2 \cos(\beta_2 + \alpha_{s2}), \qquad (3)$$

де $\rho_2 = \sqrt{x^2[1] + x_v^2}$; x_v– одна з $(m–2)$ властивостей, ортогональних до площини двох вже вибраних ознак x_i, x_k; $\beta = \arctan\dfrac{x_v}{\tilde{x}_2}$.

При фіксованому положенні кута α_{s2} всі точки накопиченої навчальної вибірки проектуються на напрямок $\tilde{x}_2$ і вибирається значення кута α^*_{s2}, яке максимізує відстань між найближчими точками різних образів на напрямку $\tilde{x}_2$. З $(m-2)$ властивостей вибирається таке, для якого ця відстань максимальна. Вибраній властивості буде відповідати напрямок $\tilde{x}$ [2], що визначається кутом α^*_{s2}, і яка буде використовуватися в наступних кроках ітеративної процедури. На другій ітерації відрізок ω_{is} на напрямку $\tilde{x}$ [2] ділиться навпіл, а вирішальним правилом на цьому кроці буде тривимірна площина, яка перпендикулярна цьому відрізку і проходить через його середину.

При появі нових об'єктів щоразу в площині $\tilde{x}$ [1] $-\tilde{x}$ [2] корегується кут α^*_{s2}, так щоби максимізувати на визначеному ним напрямку відстань між найближчими точками різних образів. Якщо при цьому не вдасться розділити накопичену навчальну вибірку, то організується наступний (третій) крок навчання. На цьому кроці вищеописана процедура проводиться в площині $\tilde{x}$ [2] $- x_t$, де x_t – одна із не задіяних властивостей. Така процедура повторюється доти, поки ріст навчальної вибірки не перестане приводити до помилок розпізнавання нових об'єктів. Отримане в результаті ітеративної процедури правило, яке дає розв'язок (ми називаємо його вирішальним) — це площина, перпендикулярна до останнього напрямку $\tilde{x}$ [φ], яка проходить через середину відрізка, що з'єднує найближчі точки різних образів.

Якщо спробувати використовувати для реалізації принципу дуальності інші процедури навчання, що не використовують теорію редукції, то на кожному кроці прийшлося би розв'язувати до кінця задачу в m-мірному просторі властивостей об'єктів. При цьому такі процедури програвали б α-процедурі, що на будь-якому кроку лише корегує лінійне вирішальне правило не в просторі всіх m властивостей, а на площині. Це дозволяє спостерігати весь ітеративний процес навчання на екрані комп'ютера. Крім того, описана процедура на кожному кроці здійснює відразу два види корекції — корегування простору (вибір додаткової ознаки) і корегування вирішального правила в ньому.

Підсилення вирішального правила

Раніше була описана ітеративна α-процедура, яка будує єдине, оптимальне в розумінні [3] вирішальне правило, яке проходить на першому кроці через середину відрізка ω_{is}, що визначає максимальну відстань між єдиним об'єктом одного образу і найближчим до нього об'єктом іншого образу. На наступних кроках відрізок ω_{is} визначає максимальну відстань між найближчими об'єктами різних образів. На Рис. 2 об'єкти різних образів позначені кружечками і хрестиками. Вирішальне правило $R(X)$, що проходить перпендикулярно до

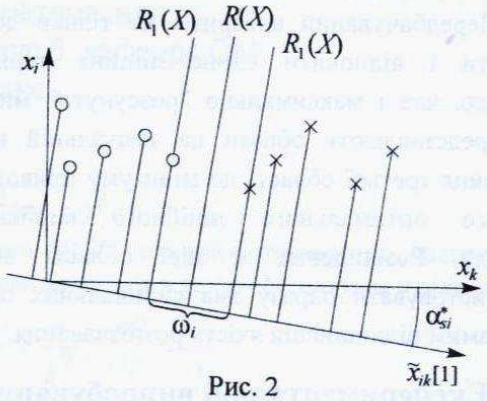


Рис. 2

відрізка ω_{si} і розділяє його навпіл, буде оптимальним в розумінні [5] щодо накопиченої вибірки.

Для будь-якого вирішального правила $R(X)$ (в тому числі й оптимального), що розділяє навчальну вибірку безпомилково, справедлива теорема, доведена в [3]. Основне твердження цієї теореми полягає в тому, що якщо з N вирішальних правил вибирається одне, яке безпомилково розділяє випадкову і незалежну вибірку довжиною l, то із ймовірністю $(1-\eta)$ можна стверджувати, що ймовірність помилкової класифікації при використанні цього правила не перевищить величини

$$\varepsilon = \frac{\ln N - \ln \eta}{l}. \qquad (4)$$

Проведемо паралельно до $R(X)$ дві площини: $R_1(X)$ та $R_2(X)$ (Рис.2). Кожна із цих площин буде перпендикулярна відрізку ω_{si}, але площина $R_1(X)$ проходить через точку x_1, а площина $R_2(X)$ — через точку x_2. Точки x_1 та x_2 – є найближчими точками різних образів.

Рішення про належність кожного об'єкта навчальної вибірки може прийматися по двох правилах:

$$\begin{cases} X \in V_1, & \text{якщо } R_1(X) > 0; \\ X \in V_2, & \text{якщо } R_1(X) \le 0. \end{cases} \qquad (5)$$

або

$$\begin{cases} X \in V_1, & \text{якщо } R_2(X) \ge 0; \\ X \in V_2, & \text{якщо } R_2(X) < 0. \end{cases} \qquad (6)$$

Обидва правила безпомилково розділяють випадкову вибірку, а тому теорема в однаковій

степені визначає їхню якість (ε) і надійність (η). Весь простір розбивається цими правилами на три області:

1. Область, яка містить тільки об'єкти $X \in V_1$, обидва правила в цій області діють безпомилково.
2. Область, що містить тільки об'єкти $X \in V_2$, обидва правила в цій області діють безпомилково.
3. Область, в якій немає об'єктів, але обидва правила – $R_1(X)$ та $R_2(X)$ – суперечать одне одному: ($R_1(X) > 0$ та $R_2(X) < 0$).

Передбачуваний алгоритм не тільки дозволяє виявити і відновити єдине лінійне вирішальне правило, але і максимально "розсунути" множини, що представляють образи на навчальній вибірці. Звуження третьої області до мінімуму приводить до єдиного оптимального лінійного вирішального правила. Розширення ж цієї області дозволяє використовувати одразу два вирішальних правила, тим самим підвищивши якість розпізнавання.

Експериментальні випробування

Описана процедура була апробована для розв'язку конкретної задачі. В якості повної навчальної вибірки використовувалась тестова випадкова множина, що складалася із 100 об'єктів, серед яких було 50 з одного образу, і 50 об'єктів іншого. На кожному об'єкті вимірювалось m=18 властивостей. Серед них було 4 ознаки $\{X_1, \ldots X_4\}$. В ході експерименту об'єкти появлялись послідовно – один за одним. Процедура починала працювати відразу ж після виконання співвідношень (1). На першому кроці аналізувались всі властивості до знаходження першої ознаки. На другому кроці аналізувалось $(m-1) = 17$ пар, і вибиралась найкраща.

Було проведено ряд експериментів, кожен із яких відрізнявся послідовністю об'єктів, що появлялися. Причому, окрім випадкових послідовностей вибиралися ще й такі, які для навчання системи були критичними. На цих експериментальних даних була в повному об'ємі перевірена ітеративна α-процедура. В результаті її роботи були синтезовані простори, в яких в гіршому випадку містилося до двох не оптимальних ознак (оптимальними при синтезі вибірки ми взяли ознаки $\{X_1, \ldots X_4\}$). Цікавим фактом є те, що в одній з вибірок при максимально коректному навчанні система знайшла розв'язок, оптимальніший за передбачений під час генерації вибірки (замість чотирьох ознак було відібрано три, і побудовано вирішальне правило, яке безпомилково розділяло навчальну вибірку). Для досягнення повного розділення навчальної вибірки системі достатньо було не більше десяти ітерацій. Наведемо два приклади, які ілюструють адаптацію системи в критичних для неї ситуаціях. Тут наводиться послідовність правил, які будуються системою в кожному з випадків (до повного розділення навчальної вибірки).

I приклад:

1. $x_3 x_2 = 0{,}29$.
2. $0{,}73*x_2 + 0{,}68*x_0 = 0{,}70$.
3. $0{,}68*x_2 + 0{,}59*x_0 - 0{,}44*x_4 = 0{,}34$.
4. $0{,}73*x_2 + 0{,}64*x_0 - 0{,}20*x_4 + 0{,}12*x_5 = 0{,}60$.
5. $0{,}71*x_2 + 0{,}62*x_0 - 0{,}20*x_4 + 0*x_5 - 0{,}28*x_6 = 0{,}48$
6. $0{,}71*x_2 + 0{,}62*x_0 - 0{,}20*x_4 + 0*x_5 - 0{,}10*x_6 + 0{,}26*x_1 = 0{,}62$.

II приклад:

1. $x_4 x_1 = 0{,}21$.
2. $0{,}74*x_1 + 0{,}67*x_0 = 0{,}57$.
3. $0{,}56*x_1 + 0{,}69*x_0 + 0{,}47*x_3 = 0{,}78$.
4. $0{,}51*x_1 + 0{,}63*x_0 + 0{,}47*x_3 + 0{,}36*x_2 = 0{,}93$.
5. $0{,}42*x_1 + 0{,}52*x_0 + 0{,}39*x_3 + 0{,}35*x_2 - 0{,}53*x_0 = 0{,}54$.

…

10. $0{,}45*x_1 + 0{,}56*x_0 + 0{,}41*x_3 + 0{,}37*x_2 - 0{,}31*x_0 - 0{,}02*x_4 - 0{,}08*x_2 + 0{,}03*x_5 - 0{,}07*x_3 + 0{,}28*x_2 = 0{,}77$.

Можна побачити, що на 5 кроці в I прикладі система, знайшовши кращий варіант, ніж X_5, анулює його. На 5 кроці в II прикладі вибирає повторно властивість, яка вже присутня в правилі. Це означає, що система корегує згенероване нею правило в доволі широких межах. Тобто, відмічається досить високий рівень гнучкості в процесі навчання.

ВИСНОВКИ

Підсумувавши результати роботи, можна сказати, що альфа-процедура може в повній мірі використовуватися для побудови ефективних ітеративних алгоритмів.

ЛІТЕРАТУРА

1. Васильев В.И. Теория редукции в проблемах экстраполяции //Проблемы управления и информатики. – 1996. № 1,2.— С. 239–251.
2. Васильев В.И. Итеративная процедура обучения имитационному поведению //Кибернетика и выч. техника. – 1989. Вып. 84.— С. 27–33.
3. Вапник В.Н., Червоненкис А.Я. Теория распознавания образов (статистические проблемы обучения). – М.: Наука, 1974.— 416 с.

ПОБУДОВА ЕЛІПСА МІНІМАЛЬНОЇ ПЛОЩІ НА ПЛОЩИНІ ТА ЕЛІПСОЇДА МІНІМАЛЬНОГО ОБ'ЄМУ В m-ВИМІРНОМУ ЕВКЛІДОВОМУ ПРОСТОРІ ДЛЯ СКІНЧЕННОГО НАБОРУ ТОЧОК.

Ю. Ю. Милейко, Б. В. Рубльов.
Київський національний університет імені Тараса Шевченка,
кафедра обчислювальної математики.
Київ, проспект Глушкова 2, корпус 6, кафедра ОМ.
Тел.: 266-41-64, факс.:
Email:

ВСТУП.

Актуальність теми. В наш час теорія розпізнавання образів як прикладна наука знайшла своє використання в багатьох математичних та прикладних дослідженнях, таких як обчислювальна геометрія, ядерна фізика, геохімія, біологічна та медицинська кібернетика, соціологія, криміналістика та в багатьох інших, і з кожним роком можливості застосування розпізнавання образів все збільшуються.

Добре відомо, що задача побудови еліпсоїда мінімального об'єму (ЕМО) і еліпса мінімальної площі (ЕМП) має надзвичайно важливе значення і викликає величезний інтерес, не лише у зв'язку із задачами розпізнавання. Все це обумовлено надзвичайно широкою областю її застосувань. Побудова ЕМП є не лише випадком побудови ЕМО для евклідової площини. Вона й особисто є дуже актуальною та має велике практичне значення, при застосуваннях в різних галузях народного господарства. Зокрема ці алгоритми небхідні для комп'ютерної диагностики злоякісних пухлин, які дозволяють розпізнавати злоякісни пухлини на ранніх стадіях, що є дуже важливим для здоров'я людини. Крім того методи побудови ЕМП можна буде в подальшому узагальнити на випадок ЕМО. Зрозуміло, що можливість застосувань алгоритмів побудови ЕМО є набагато ширшою.

Крім того межа ЕМП та ЕМО може виступати в якості еліптичної дискримінантної функції (чи її наближення) ([1]) для розділення двох точкових множин в m-вимірному евклідовому просторі. Загальних ефективних методів побудови подібних функцій поки що не існує.

ГЕОМЕТРИЧНІ ВЛАСТИВОСТІ ЕМП.

Задача 1. Знайти рівняння ЕМП для довільного трикутника.

Нехай ABC — довільний трикутник, вершини якого на декартовій плщині мають такі кординати: A(-a,0), B(a,0), C(b,c). Рівняння ЕМП, що є розв'язком задачі 1 має вигляд:

$$c^2x_0^2 - 2bcx_0z_0 + \left(b^2+3a^2\right)z_0^2 - -2a^2cz_0 = a^2c^2 \qquad (1)$$

Нехай $ABCD$ опуклий чотирикутник, діагоналі якого перетинаються в точці O і діляться у відношеннях:

$$\frac{AO}{OC} = \frac{1}{k^2},\ \frac{BO}{OD} = \frac{n^2}{1}.$$

Клас усіх опуклих чотирикутників, афінно інваріантних заданому чотирикутнику $ABCD$ позначимо $W(k,n)$.

Задача 2. Побудувати ЕМП для чотирикутника з класу $W(k,n)$.

Задача 3. Побудувати ЕМП для чотирикутника з класу $W(k,n)$, якщо відомо, що всі чотири вершини чотирикутника знаходяться на цьому еліпсі.

Остання задача буде еквівалентна такій проблемі:

Задача 4. В задане коло вписати чотирикутник з класу $W(k,n)$ найбільшої площі.

Чотирикутник, що є розв'язкм задачі 4 має такий кут між діагоналями, що задовольняє рівнянню

$$g(\varphi) = L\cos^3\varphi + 2J\cos^2\varphi + (2L-3I)\cos\varphi + J = 0, \qquad (2)$$

де

$$\begin{cases} I = \left(n^2k^2+1\right)\left(n^2+k^2\right) \\ J = 2nk\left(1-n^2\right)\left(1-k^2\right) \\ L = 4n^2k^2. \end{cases}$$

Лема 1. Нехай $M = A_1A_2\ldots A_n$ — опуклий многокутник; BCD — трикутник найбільшої площі, з вершинами вибранними з вершин многокутника M; $e(BCD)$ — ЕМП для трикутника BCD. Якщо при цьому многокутник $M \subset e(BCD)$, тоді є ЕМП для усього многокутника. Якщо ж трикутник $A_iA_jA_k$ $1 \le i < j < k \le n$ не є трикутником найбільшої площі серед усіх трикутників, вершини яких вибираються з вершин многокутника $A_1, A_2, \ldots, A_n$, то ЕМП для цього трикутника не може бути ЕМП для усього многокутника M.

© UkrObraz'2000

Теорема 1. ЕМП для опуклого многокутника $M = A_1A_2 \ldots A_n$, на якому знаходиться рівно три вершини M, є ЕМП для трикутника, з вершинами в цих точках.

Зауваження 1. Якщо ЕМП e опуклого многокутника $M = A_1A_2A_3 \ldots A_n$ проходить рівно через чотири вершини M, то цей чотирикутник не обов'язково має найбільшу площу серед усіх чотирикутників вигляду $A_iA_jA_kA_l$, де $1 \le i < j < k < l \le n$.

Задача 5. Для довільного опуклого чотирикутника $ABCD$, діагоналі якого в точці перетину поділяються у відношеннях $1:k$ та $1:m$, сформулювати критерій, який дозволить по значенням k та m визначити кількість вершин (тобто три чи чотири), що знаходяться на його ЕМП.

Нехай $ABCD \in W\left(\frac{1}{\sqrt{k}}, \frac{1}{\sqrt{m}}\right)$. Тоді ЕМП проходить рівно через три вершини $ABCD$, якщо виконується умова:

$$m + 1 + \frac{1}{m} < k < +\infty \qquad (3)$$

ПОБУДОВА ЕМП ДЛЯ ЕЛЕМЕНТАРНИХ БАГАТОКУТНИКІВ

Цей алгоритм дозволяє будувати ЕМП для трикутника, чотирикутника, п'ятикутника та шестикутника, які визначаються координатами своїх вершин.

Крок 1. Для трикутника ЕМП знаходиться безпосередньо за формулою (1).

Крок 2. Для чотирикутника $BCDE$ перевіряємо умову (3). Якщо вона виконується, то розв'язуємо кубічне рівняння (2), знаходимо кут φ_0, а далі робимо афінне перетворення заданого чотирикутника $BCDE$ в чотирикутник $B_1C_1D_1E_1$, що є вписаним, та має кут між діагоналями φ_0. ЕМП для нього є описане коло, і зворотнім афінним перетворенням A^{-1} знаходимо ЕМП для заданого чотирикутника $BCDE$. Якщо умова (3) не виконується, то серед чотирьох трикутників BCD, BCE, BDE, CDE знаходимо той, що має найбільшу площу, ЕМП для нього (крок 1) і буде ЕМП для $BCDE$.

Крок 3. Для заданого п'ятикутника $BCDEF$ знаходимо трикутник найбільшої площі (нехай це буде BCD), будуємо для нього ЕМП e_1 (крок 1), якщо він містить решту точок (E і F), то він і є шуканим ЕМП для усього п'ятикутника. Інакше, знаходимо точку, найбільш "еліптично" віддалену від еліпса e_1, (нехай це буде точка E), далі (крок 2) будуємо ЕМП e_2 для чотирикутника $BCDE$, який обов'язково проходить через всі його вершини, якщо e_2 містить п'яту вершину F, то він є шуканим, інакше будуємо ЕМП для решти чотирикутників, які можна утворити з вершин B, C, D, E, F, з обов'язковим проходженням його через усі чотири вершини. Якщо один з них містить п'яту вершину, то він є шуканим, інакше ЕМП для п'ятикутника $BCDEF$ проходить через усі його п'ять вершин (крок 1).

Крок 4. Для шестикутника $BCDEFG$ робимо аналогічно випадку п'ятикутника.

ПОБУДОВА ЕМП ДЛЯ ДОВІЛЬНОЇ СУКУПНОСТІ ТОЧОК.

Нехай нам задано N точок $C_1, C_2, \ldots, C_N$.

Крок 1. На першому кроці будуємо опуклу оболонку заданих точок і одержуємо опуклий многокутник $M = A_1 \ldots A_n = conv(C_1, \ldots, C_N)$. Це робиться за час $O(N \log N)$ або $O(nN)$ (n — кількість точок на опуклій оболонці) в залежності від того, який саме алгоритм ми вибираємо.

Крок 2. Серед точок $A_1, A_2, \ldots, A_n$ знаходимо трикутник найбільшої площі, нехай це трикутник BCD — це робиться за час $O(n \log^2 n)$. Всі три вершини трикутника утворюють множину Z.

Крок 3. Будуємо ЕМП e для множини Z — вона є елементарним многокутником, а тому при побудові використовуємо алгоритм 1. Час його роботи є лінійним відносно числа n — кількості вершин многокутника M. Якщо e містить всі інші точки многокутника M, то він є шуканим і алгоритм завершує свою роботу. Якщо ж є вершини, що не попадають всередину e, то переходимо до наступного кроку алгоритму.

Крок 4. ЕМП e для множини Z, яка містить не більше як шість точок побудовано, як ЕМП для чотирикутника або ж п'ятикутника, а тому решту точок викидаємо з множини Z, а далі знаходимо серед точок, що не попали всередину ЕМП для множини Z найбільш "еліптично" віддалену від еліпса e і додаємо її до множини Z, при цьому множина Z містить не більше як шість точок, і переходимо на крок 3 алгоритма.

Тепер перелічемо інші привабливі риси нашого алгоритму та його переваги над іншими існуючими алгоритмами.

1. Ще раз наголосимо той факт, що за скінченну кількість кроків будується справжній ЕМП для заданої сукупності точок, при цьому кількість операцій після побудови опуклої оболонки є пропорційною кількості точок, що є вершинами опуклої оболонки, бо кожен крок є побудовою ЕМП для елементарного многокутника і викнується за сталий час. Щодо інших достоїнств запропонованого алгоритма слід відзначити, досить тісний зв'язок між площею елементарних многокутників та ЕМП, чим більше площа елементарного многокутника тим більше ймовірність того, що ЕМП побудований для нього й буде шуканим ЕМП усієї множини. Особливо цей зв'язок ми бачимо у випадку трикутників, хоча для решти елементарних многокутників нема такого безпосереднього зв'язку, слід сказати, що площа того елементарного многокутника для якого будується

справжній ЕМП не може за площею суттєв відрізнятися від максимальної площі відповідного многокутника з вершинами в заданих точках. Легко побачити, як цей факт використовує наш алгоритм.

2. Крім того, що наш алгоритм за скінченну кількість кроків будує справжній ЕМП, треба сказати, що їх кількість може бути дуже малою; так в досить багатьох випадках вже на першому кроці ми можемо знайти шуканий ЕМП, (це буде у випадку, коли ЕМП для опуклого многокутника проходить рівно через три точки).

З впевненістю можна сказати, що і в інших випадках кількість кроків буде мінімально можливою.

3. На будь-якому з кроків ми можемо зупинити процес і прийняти одержаний еліпс в якості наближення (знизу) бажаного ЕМП, причому навіть для одного кроку та найгіршого випадку щодо розташування заданих точок похибка наближення не буде перевищувати двадцяти відсотків (цей випадок детально розглянуто в попередньому розділі).

4. При переході від кроку до кроку не відбувається ніякого накопичення похибки, тому що з попереднього кроку ми беремо лише інформацію про точки множини M, а обчислений (можливо з деякою похибкою) еліпс на попередньому кроці ми забуваємо, бо на новому кроці ми одержимо новий точний ЕМП для деякого елементарного многокутника. Будь-який існуючий ітераційний процес, що залишає попередню ітерацію накопичує похибку, яка може врешті решт дуже суттєво погіршити знайдене наближення, і навіть взагалі бути далеким від істини, а сам процес дослідження поведінки похибки є досить складною справою.

ГЕОМЕТРИЧНІ ВЛАСТИВОСТІ ЕМО.

Серед існуючих методів їх знаходження майже всі є наближеними та ітераційними. В цій главі був запропонований підхід, що спирається на геометричні властивості ЕМО, подібна ідея дала змогу в попередній главі одержати досить гарні за характеристиками алгоритми побудови ЕМП. Дещо подібне, хоча й з превними обмеженнями вдалося досягти і для еліпсоїдів.

Теорема 2. Для довільного сімплекса E в m-вимірному евклідовому просторі відношення $\xi_0 = \dfrac{V(E)}{V(e(E))}$ є сталою величиною і не залежить від конкретного сімплекса E.

Теорема 3. Якщо для опуклого многогранника $F = A_1A_2 \ldots A_N$ з $\boldsymbol{R}^m$ ЕМО $e(F)$ містить рівно $m+1$ вершину F, а саме точки $\{B_1, B_2, \ldots, B_{m+1}\} \subset \{A_1, A_2, \ldots, A_N\}$, то $e(F) = e(G)$, де G — многогранник, що утворюється вершинами $B_1, B_2, \ldots, B_{m+1}$.

Теорема 4. Якщо для опуклого многогранника $F = A_1A_2 \ldots A_N$ з $\boldsymbol{R}^m$ ЕМО $e(F)$ містить рівно $(m+1)$ вершину F, а саме точки $\{B_1, B_2, \ldots, B_{m+1}\} \subset \{A_1, A_2, \ldots, A_N\}$, то з усіх многогранників, що утворюються вершинами $A_{i_1}, A_{i_2}, \ldots, A_{i_{m+1}}$, де $1 \le i_1 < i_2 < \ldots < i_{m+1} \le N$ найбільший m-вимірний об'єм має той многогранник, що має вершини в точках $B_1, B_2, \ldots, B_{m+1}$.

Теорема 5. ЕМО для сукупності точок $M = A_1A_2 \ldots A_N$ в m-вимірному евклідовому просторі, який має з цією множиною точок рівно K спільних точок (які ми позначимо $B_1, B_2, \ldots, B_k$), де $m+1 \le k \le \dfrac{1}{2}m(m+3)$ то він є й ЕМО для сукупності точок M.

Теорема 6. Нехай для множини точок $M = A_1A_2 \ldots A_N$ в m-вимірному евклідовому просторі існує підмножина $B = B_1B_2 \ldots B_k$, що складається з K точок, для якої ЕМО $e(B)$ містить всі точки множини M, то він є ЕМО для всієї множини M, тобто має місце рівність $e(M) = e(B)$.

ПОБУДОВА ЕМО ДЛЯ ДОВІЛЬНОЇ МНОЖИНИ ТОЧОК.

Для випадку довільного m-вимірного простору до елементарних ми віднесемо опуклі многогранники, що містять від $m+1$ до $\dfrac{1}{2}m(m+3)+1$ вершини. Для подібних множин ЕМО шукається з допомогою окремого алгоритму, при цьому, використовуючи дуже малу кількість точок в елементарному многограннику, цей алгоритм може використовувати навіть будь-який відомий алгоритм побудови ЕМО (при цьому навіть простий перебір вершин не буде вимагати дуже багато часу). Найголовніший висновок дуже простий — за сталий час можна знайти ЕМО для елементарного многогранника, або ж навіть наближення цього еліпсоїда з будь-якою точністю. Найголорвніше, що нам треба знати — це сукупність точок, на які спирається ЕМО для елементарного многогранника. При достатньо визначеній точності їх можна знайти точно (не сам ЕМО для елементарного многогранника, а лише точки, на які той спирається, сам ЕМО може бути знайдений наближенно).

Алгоритм 2.4.1.

(Побудова ЕМО для довільної множини точок).

1 крок. Вибираємо довільну $m+1$ точку з заданої множини M, які утворюють множину Z (елементарний опуклий многогранник F).

2 крок. Будуємо ЕМО e для множини Z (яка в свою чергу є елементарним многогранником F); всі вершини, що не попадають на межу еліпсоїда вилучаються з множини Z.

3 крок. Якщо всі вершини множини M розташовани всередині еліпсоїда e, то цей еліпсоїд є бажаним

ЕМО, а тому алгоритм закінчує свою роботу; інакше переходимо на наступний крок алгоритма.

4 крок. Вибираємо серед точок мнгожини M, що не попали всередину e "еліптично найбільш віддалену" від еліпсоїда e і додаємо її до множини Z. Переходимо на крок 2.

Якщо провести деякий аналіз запропонованого алгоритма, то слід сказати, що він зберігає всі привабливі риси аналогічного алгоритма для випадку площини. Нагадаємо їх тому, що вони притаманні й нашому алгоритму.

1. Цей алгоритм вимагає скінченну кількість операцій на відміну від багатьох інших ітераційних алгоритмів, де процес завершується при одержанні певної наперед визначенної точності (як правило без урахування похибки обчислень), кожний раз при цьому одерпжується лише наближення ЕМО.

2. В багатьох випадках досить швидко знаходиться справжній ЕМО, що слідує з теореми 6.

3. На будь-якому кроці ми можемо зупинити процес і прийняти одержанний еліпсоїд в якості наближення ЕМО, об'єм якого буде не перевищувати об'єм справжнього ЕМО.

4. При переході від кроку на крок не відбувається ніякого накопичення похибки, тому що з попереднього кроку ми залишаємо лише інформацію про точки множини M, а знайденний черговий еліпсоїд можемо забути.

Все перелічене дає нам змогу констатувати той факт, що цей алгоритм є одним з найкращих для побудови ЕМО за витратами та точністю наближення, а тому безпосередньо знайде своє використання при багатьох застосуваннях.

І останнє чи не найважливіша властивість наведеного алгоритма така. Як показує тестування роботи цього алгоритма більше як семи випадках з десяти (приблизно сімдесят відсотків від загальної кількості) справжній ЕМО спирається на "максимальну можливу кількість точок" (так для випадку площини це п'ять точок, для тривимірного простору — 9 точок, для довільного m — $\frac{1}{2}m(m+3)$ точок). Але для цього випадку ми зможемо точно вказати всі ці точки (навіть використовуючи деякий наближенний метод), ну а далі проблема точного знаходження ЕМО зводиться до розв'язку системи лінійних рівнянь, де невідомими виступають коефіцієнти квадратичної форми, що визначає ЕМО, а в якості координат слід підставити координати знайдених $\frac{1}{2}m(m+3)$ точок. Таким чином більше як в семидесяти відсотках всіх практичних прикладів ми можемо знайти ЕМО точно, що не має місця в інших подібних алгоритмах. Саме завдяки останній властивості цей алгоритм вже знайшов своє використання, і ще знайде подальше застосування в багатьох прикладних проблемах.

ВИСНОВКИ.

Зрозуміло, що загальні методи, які дозволяють розв'язувати великий клас задач дуже корисні і необхідні, як теоретичне дослідження і для вимог практичного застосування. Так методи недиференційовної оптимізації, що були розроблені у відділу Шора, зокрема пристосовані і для побудови наближень ЕМО.

Ми вважаємо, що детальне вивчення природи об'єкта, дозволяє менш загальні проблеми розв'язувати ефективніше (інколи набагато) ніж методи. Так наш підхід до проблеми побудови ЕМП: ЕМО, як до задач обчислювальної геометрії, а не до оптимізаційних, дозволило розробити алгоритм їх точної побудови, а не наближень, що пропонують інші підходи. Порівняння нашого алгоритму побудови ЕМП з алгоритмом, розробленим у відділі Шора показав перевагу нашого на всіх параметрах — точність побудови ЕМП, швидкість, стійкість до різних розташувань точок. Ми сподіваємось, що подібне досконале вивчення суті кожної проблеми дозволить знаходити більш ефективні алгоритми їх розв'язання, що в наш час особливо цінується в обчислювальній геометрії.

ЛІТЕРАТУРА.

1. Петунин Ю.И., Рубльов Б.В. Распознавание образов с помощью квадратичных дискриминантных функций // Обчислювальна та прикладна математика. –1996. – вып. 80. – с. 89-104.
2. Рубльов Б.В., Петунин Ю.И. Элипс минимальной площади, содержащий конечное множиство точек // УМЖ. –1998. – 50. – часть 1: - №7. – с. 937 – 981; часть 2: №8. – с.1098 – 1105.
3. Рубльов Б.В. Різні модифікації оптимального алгоритму побудови еліпса мінімальної площі. // Вісник Київського університету 1999. – вип. 1. – с. 223 – 228.
4. Ляшко И.И., Рубльов Б.В. Геометричні властивості еліпсоїда найменшого об'єму в m-вимірному евклідовому просторі. // ДАН НАН України. – 1999. - №9. – с. 17 – 19.
5. Милейко Ю.Ю., Рубльов Б.В. Побудова еліпсоїда мінімального об'єму для елементарних багатогранників. // Вісник Київського університету. – 1999. – вип. 3. – с. 236 – 242.
6. Рубльов Б.В. Алгоритму побудови еліпсоїда мінімального об'єму // Журнал обчислювальної та прикладної математики. – 1998. - №1(83). – с. 105 – 107.
7. Shor N.Z., Berezovski O.A. New algorithm for constructiny optimal circumscribed and incribed ellipsoid // Optimization methods and Software. - 1992. - Vol. 1. – p. 283 – 299.

МНОЖИНИ ПРИНАЛЕЖНОСТІ В ЗАДАЧАХ КЛАСИФІКАЦІЇ СИГНАЛІВ.

М.Ф.Кириченко,
Інститут кібернетики НАНУ,031876,Київ,проспект академіка Глушкова,40
Тел.(044)266-3725,E-mail kir@dept115.icyb.kiev.ua
Роман Куц (США),Yale University,USA
М.П.Лепеха,Київський національний університет, факультет кібернетики

АНОТАЦІЯ.

У даній роботі подані нові математичні засоби для розробки ефективних алгоритмів класифікації і розпізнавання просторових об'єктів. Вони базуються на побудові в просторі ознак множин належності об'єктів до того або іншого класу, а також умов неперетину цих множин. Запропоновані засоби використовується для розпізнавання просторових об'єктів на підставі аналізу ультразвукових ехо-сигналів.

Умови класифікації сигналів є достатньо конструктивними, допускають побудову ефективних алгоритмів, а також програмних засобів для перевірки роздільмості множин у просторі ознак і розпізнавання об'єктів. Результати роботи є розвитком ідей, поданих у раніше опублікованих статтях авторів [3]-[5].

ВСТУП.

Запропоновані в роботі математичні засоби розвиваються з метою використання їх при створенні мобільних сонарних систем, що були б здатні розпізнавати просторові об'єкти на підставі прийнятих ультразвукових ехо-сигналів. Ці засоби базуються

- на виявленні в ультразвукових ехо-сигналах характерних для даного класу розпізнаваних об'єктів групи ознак,
- формуванні в просторі виділених ознак для кожного з об'єктів спеціальних множин належності у формі канонічних геометричних тіл (еліпсоїди, паралелепіпеди і т.п.),
- формулюванню умов роздільмості множин приналежності між собою.

Утворені множини приналежності повинні бути достатньо компактно натягнуті на множину характерних точок кожного об'єкта. Геометричне тіло з достатньо компактним включенням у себе точок множини ознак одного об'єкта будемо називати контейнером. Канонічність форми контейнерів дозволяє в аналітичній формі записати умови неперетину тих контейнерів, що відносяться до різних об'єктів розпізнавання. Постановки задач класифікації й ідеї побудови контейнерів будемо ілюструвати на практичній задачі розпізнавання за даними ультразвукової локації просторових об'єктів [4]-[5].

Математичні результати статті є фактично розвитком ідеї розкладу по Карунnену-Лоеву з використанням в задачах класифікації просторових об'єктів [2].

1. ПОСТАНОВКИ ЗАДАЧ.

Як вхідна інформація в задачі розпізнавання N об'єктів розглядаються сигнали $x_j(t),\ j=\overline{1,M}$, $t\in[0,T]$. Кожний сигнал $x_j(t),\ t\in[0,T]$ перетворюється в вектор $a(j)\in R^m$ простору ознак, який будемо іменувати *зображенням об'єкта* в просторі ознак. Множину всіх векторів $a(j),\ j=\overline{1,n}$ для деякого конкретного об'єкта будемо позначати

$$\Omega_a=\left\{a: a=a(j),\ j=\overline{1,n}\right\} \qquad (1)$$

із матрицею

$$A=\left(a(1)\vdots\ldots\vdots a(n)\right). \qquad (2)$$

Задача 1. Визначення центру множини Ω_a.

Визначити деяку точку $\bar{a}\in R^m$, що будемо називати центром для Ω_a, в ефективному для наступного розпізнавання об'єктів виді. Звична, але не завжди ефективна форма представлення центру має вигляд

$$\bar{a}=\frac{1}{n}\sum_{j=1}^{n}a(j)\ . \qquad (3)$$

Позначимо $\tilde{a}=a(j)-\bar{a}$, $\tilde{A}=\left(\tilde{a}(1)\vdots\ldots\vdots\tilde{a}(n)\right)$.

Задача 2. Конструювання множин приналежності (контейнерів).

Сформувати множину $\Omega_a(p)$ з достатньо простим математичним описом (структурою), що задовольняє умові приналежності

$$\Omega_a(p)\supset\Omega_a\ , \qquad (4)$$

де p - деякий параметр, що визначає тип цієї множини.

Звичайні форми для множин приналежності: многомірний еліпсоїд, циліндричний еліпсоїд,

© UkrObraz'2000

паралелепіпед і їхні об'єднання при відповідним чином вибраних центрах.

Задача 3. Умова неперетину для множин приналежності.

Сформулювати необхідні і достатні умови неперетину для різноманітних множин приналежності Ω_a. Сформулювати тільки достатні умови неперетину контейнерів для різноманітних множин, але в зручній для використання формі.

Задача 4. Оптимальний вибір простору ознак для робастності умов неперетину контейнерів.

Розробити алгоритм вибору простору ознак для оптимізації відстаней між контейнерами для різноманітних множин Ω_a.

Задача 5. Правила класифікації об'єктів.

Сформулювати правило класифікації об'єктів із використанням контейнерів і умов їхнього неперетину.

2. РОЗПОДІЛ ЗОБРАЖЕНЬ В ПРОСТОРІ ЇХ ВЛАСНИХ ВЕКТОРІВ.

Для конструювання множин-контейнерів і формулювання правил класифікації об'єктів у наступних розділах цієї роботи будуть використані система власних векторів у просторі ознак і підпростори, обумовлені цією системою. Тому приведемо деякі властивості власних векторів і власних значень для матриці $\widetilde{A}$. Відповідно до [1] матриця $\widetilde{A}$ має таке сингулярне представлення:

$$\widetilde{A} = \sum_{i=1}^{r} y_i x_i^T \lambda_i \ , \quad r = rank\,\widetilde{A}, \qquad (5)$$

де

$$\widetilde{A}\widetilde{A}^T y_j = \lambda_j^2 y_j \ , \quad y_i^T y_j = \delta_{ij} \ , \qquad (6)$$

$$\widetilde{A}^T \widetilde{A} x_j = \lambda_j^2 x_j \ , \quad x_i^T x_j = \delta_{ij} \ , \qquad (7)$$

$$\widetilde{A}^T y_j = \lambda_j x_j \ , \quad \widetilde{A} x_j = \lambda_j y_j \ . \qquad (8)$$

$i, \ j = \overline{1,r}$, $\lambda_1^2 \geq \ldots \geq \lambda_r^2$.

Формулу (5) можна розписати в співвідношеннях по стовпчиках $\widetilde{a}(j)$ матриці $\widetilde{A}(j)$ таким чином

$$\widetilde{a}(j) = \sum_{i=1}^{r} \lambda_i y_i x_{ji} \ , \ j = \overline{1,n} \ , \qquad (9)$$

де $x_i^T = (x_{1i}, \ldots, x_{ni})$. Тоді

$$\widetilde{a}(j) = \sum_{i=1}^{s} \lambda_i y_i x_{ji} + \Delta_{js}, \quad s < r \qquad (10)$$

$$\Delta_{js} = \sum_{i=s+1}^{r} \lambda_i y_i x_{ji} \ . \qquad (11)$$

При цьому мають місце оцінки

$$\left\| \Delta_{js} \right\|^2 = \sum_{i=s+1}^{r} \lambda_i^{\,2} \left| x_{ji} \right|^2 \leq \lambda_{s+1}^2 \sum_{i=s+1}^{r} \left| x_{ji} \right|^2 \leq \lambda_{s+1}^2 \qquad (12)$$

згідно таким співвідношенням

$$\left(x_1 \vdots \ldots \vdots x_n\right)^T \left(x_1 \vdots \ldots \vdots x_n\right) = I_n \ ,$$

$$\left(x_1 \vdots \ldots \vdots x_n\right)\left(x_1 \vdots \ldots \vdots x_n\right)^T = I_n \ ,$$

$$\sum_{i=s+1}^{r} \left| x_{ji} \right|^2 \leq \sum_{i=1}^{n} \left| x_{ji} \right|^2 = 1 \ .$$

3. КОНСТРУЮВАННЯ МНОЖИН ПРИНАЛЕЖНОСТІ – КОНТЕЙНЕРІВ.

Розглянемо декілька варіантів опису множин-контейнерів.

Еліптичні циліндри.

З (10), (11) і (12) випливають такі вирази для контейнерів у формі еліптичних циліндрів:

$$\Omega_a^1(s,\lambda) = \left\{ a : \left(a - \bar{a}\right)^T \sum_{i=1}^{s} y_i y_i^T \lambda_i^{-2} \left(a - \bar{a}\right) \leq \alpha^2(s,\lambda \right.$$

$$\left. \lambda_{s+1}^{-2} \left\| \left(I_m - \sum_{i=1}^{s} y_i y_i^T \right)(a - \bar{a}) \right\|^2 \leq \beta^2(s,\lambda) \right\}, \ s < r,$$

$$\alpha^2(s,\lambda) = \max_{j=\overline{1,n}} \tilde{a}_j^T \sum_{i=1}^{s} y_i y_i^T \lambda_i^{-2} \tilde{a}_j \leq 1, \qquad (13)$$

$$\beta^2(s,\lambda) = \max_{j=\overline{1,n}} \lambda_{s+1}^{-2} \sum_{k=s+1}^{r} x_{jk}^2 \lambda_k^2 \leq \max_{j=\overline{1,n}} \sum_{k=s+1}^{r} x_{jk}^2 \leq 1$$

і

$$a_j \in \Omega_a^1(s,\lambda), \ \ \forall\, j = \overline{1,n}. \qquad (14)$$

Якщо $s = r$, то формула (13) матиме вигляд

$$\Omega_a^1(r,\lambda) = \left\{ a : (a - \bar{a})^T \left(\widetilde{A}^+\right)^T \widetilde{A}^+ (a - \bar{a}) \leq \right.$$

$$\leq \max_{j=\overline{1,n}} e_j^T \widetilde{A}^+ \widetilde{A} e_j \ , \qquad (15)$$

$$\left. (a - \bar{a})^T Z(\widetilde{A}^T)(a - \bar{a}) \leq \varepsilon^2 \right\}, \ \varepsilon > 0,$$

де $\widetilde{A}^+ = \sum_{j=1}^{r} x_j y_j^T \lambda_j^{-1}$, $Z(\widetilde{A}^T) = I_n - \widetilde{A}\widetilde{A}^+$,

$$a_j \in \Omega_a^1(r,\lambda), \ \ \forall\, j = \overline{1,n}. \qquad (16)$$

Варто зауважити, що

$$\lambda_j^2 = \sum_{i=1}^{n} (\tilde{a}_i^T y_j)^2, \ \ j = \overline{1,r} \qquad (17)$$

і λ_j^2 має тенденцію зростати з порядком n.

Для тих випадків, коли при $s << n$ виконуються умови

$$s \max_{i=\overline{1,n}} \left|\tilde{a}_i^T y_j\right| < \sum_{i=1}^{n} (\tilde{a}_i^T y_j)^2, \quad j=\overline{1,s},$$

множина-контейнер буде мати вигляд

$$\Omega_a^1(s,d) = \left\{ a:\ (a-\bar{a})^T \sum_{i=1}^{s} y_i y_i^T d_i^{-2} (a-\bar{a}) \le s, \right.$$

$$\left. \left\| (I_m - \sum_{i=1}^{s} y_i y_i^T)(a-\bar{a}) \right\| \le d_{s+1} \right\}, \quad s<r, \quad (18)$$

де

$$d_i = \max_{i=\overline{1,n}} \left|\tilde{a}_i^T y_j\right|,$$

$$d_{s+1} = \max_{i=\overline{1,n}} \left\| (I_n - \sum_{j=1}^{s} y_j y_j^T)\tilde{a}_i \right\|, \quad j=\overline{1,s},$$

і може досягати менших розмірів ніж $\Omega_a^1(s,\lambda)$, тобто

$$\Omega_a^1(s,d) \subset \Omega_a^1(s,\lambda).$$

Відповідно при $s=r$ множина-контейнер $\Omega_a^1(r,d)$ має вигляд

$$\Omega_a^1(r,d) = \left\{ a:\ (a-\bar{a})^T \sum_{i=1}^{r} y_i y_i^T d_i^{-2} (a-\bar{a}) \le r, \right.$$

$$\left. \left\| (I_m - \sum_{i=1}^{s} y_i y_i^T)(a-\bar{a}) \right\|^2 \le \varepsilon^2 \right\}, \ \varepsilon^2 > 0, \quad (19)$$

Тоді з побудови цих множин випливає $a_j \in \Omega_a^1(s,d)$, $a_j \in \Omega_a^1(r,d)$, $\forall j=\overline{1,n}$.

Еліпсоїди.

Із співвідношень(10) - (12) випливає така форма представлення:

$$\Omega_a^2(s,\lambda) = \left\{ a:\ (a-\bar{a})^T \left[\sum_{i=1}^{s} y_i y_i^T \lambda_i^{-2} + \right.\right.$$

$$\left.\left. + \lambda_{s+1}^{-2} \left(I_m - \sum_{i=1}^{s} y_i y_i^T \right) \right] (a-\bar{a}) \le \alpha^2(r,\lambda) \right\}$$

$$s=\overline{1,r-1},\ a_j \in \Omega_a^2(s,\lambda),\ j=\overline{1,n} \quad (20).$$

При $s=r$ множина-контейнер має вигляд еліпсоїда в R^m:

$$\Omega_a^2(r,\lambda) = \left\{ a:\ (a-\bar{a})^T \left[\sum_{i=1}^{r} y_i y_i^T (\lambda_i^{-2}\alpha^{-2}(r,\lambda) - \varepsilon^{-2}) + \right.\right.$$

$$\left.\left. + \varepsilon^{-2} I_m \right] (a-\bar{a}) \le 1 \right\},\ \varepsilon^2 > 0 \quad (21)$$

Еліпсоїдальні множини-контейнери відповідні еліптичним циліндрам (18), (19) мають таке представлення

$$\Omega_a^2(s,d) = \left\{ a:(a-\bar{a})^T \left[\sum_{i=1}^{s} y_i y_i^T (d_i^{-2} s^{-1} - d_{s+1}^{-2}) + \right.\right.$$

$$\left.\left. + d_{s+1}^{-2} I_m \right] (a-\bar{a}) \le 1 \right\},\ s=\overline{1,r-1} \quad (22)$$

$$\Omega_a^2(r,d) = \left\{ a:\ (a-\bar{a})^T \left[\sum_{i=1}^{r} y_i y_i^T (d_i^{-2} r^{-1} - \varepsilon^{-2}) + \right.\right.$$

$$\left.\left. + \varepsilon^{-2} I_m \right] (a-\bar{a}) \le 1 \right\},\ \varepsilon^2 > 0 \quad (23)$$

Паралелепіпеди.

Як і для еліптичних циліндрів і еліпсоїдів введемо до розгляду такий вид цих множин:

$$\Omega_a^3(r,d) = \left\{ a: \left|(a-\bar{a})^T y_i\right| \le d_i,\ i=\overline{1,r}, \right.$$

$$\left. \left|(a-\bar{a})^T y_j\right| \le \frac{\varepsilon}{\sqrt{m}},\ j=\overline{r+1,m} \right\} \quad (24)$$

Очевидно, що

$$\Omega_a^3(r,d) \subset \Omega_a^2(r,d).$$

5. УМОВИ НЕПЕРЕТИНУ МНОЖИН КОНТЕЙНЕРІВ.

Нерівності Коші й умови неперетину множин.

Будемо використовувати нерівність Коші

$$\left(l^T a\right) \le l^T K^{-1} l a^T K a,$$

що має місце при $\forall\, l \in R^m$, $a \in R^m$ і довільній додатньо-визначеній матриці $K \in R^{m\times n}$. При цьому

$$\max_{a\in\{a:\ a^T K a \le 1\}} l^T a = \sqrt{l^T K^{-1} l} \quad (26)$$

Для двох еліпсоїдальних множин

$$\Omega_1 = \left\{ a:\ (a-\bar{a})^T K_1 (a-\bar{a}) \le 1 \right\}, \quad (27)$$

$$\Omega_2 = \left\{ a:\ (a-\bar{b})^T K_2 (a-\bar{b}) \le 1 \right\}, \quad (28)$$

де $K_1 > 0$, $K_2 > 0$ - додатньо-визначені матриці, необхідне і достатня умова їх неперетину відповідно до (26) має вигляд:

$$\exists\, l \in R^m,\ \sqrt{l^T K_1^{-1} l} + \sqrt{l^T K_2^{-1} l} < l^T (b-a), \quad (29)$$

а при $\bar{l} = \bar{b} - \bar{a}$ ця умова перетворюється лише в достатнє, але з більш зручною для використання формою

$$\sqrt{(\bar{b}-\bar{a})^T K_1^{-1} (\bar{b}-\bar{a})} + \sqrt{(\bar{b}-\bar{a})^T K_2^{-1} (\bar{b}-\bar{a})} < \left\|(\bar{b}-\bar{a})\right\|^2 \quad (30)$$

При такому представленні матриць

$$K_1 = \sum_{j=1}^{m} y_j(1) y_j^T(1) \lambda_j^{-2}(1),$$

$$K_2 = \sum_{j=1}^{m} y_j(2) y_j^T(2) \lambda_j^{-2}(2),\ \lambda_j(k) \ge \lambda_{j+1}(k),$$

$$y_i^T(k)y_j(k)=\delta_{ij},\ i,j=\overline{1,m},\ k=\overline{1,2}$$

умови (29) і (30) одержують відповідно таку форму

$$\exists\, l\in R^m,\quad \sqrt{\sum_{j=1}^{m}(l^T y_j(1))^2\lambda_j^2(1)}+$$

$$+\sqrt{\sum_{j=1}^{m}(l^T y_j(2))^2\lambda_j^2(2)}<l^T(\bar b-\bar a),$$

$$\sqrt{\sum_{j=1}^{m}\left[(\bar b-\bar a)^T y_j(1)\right]^2\lambda_j^2(1)}+$$

$$+\sqrt{\sum_{j=1}^{m}\left[(\bar b-\bar a)^T y_j(2)\right]^2\lambda_j^2(2)}<\left\|\bar b-\bar a\right\|^2.$$

Для формулювання умов неперетину паралелепіпедів нерівність Коші

$$\left|l^T a\right|\le\max_{j=\overline{1,m}}\left|a_j\right|\sum_{i=1}^{m}\left|l_i\right|$$

будемо використовувати у вигляді

$$\left|l^T a\right|\le\max_{j=\overline{1,m}}\left|a^T y_j\right|\sum_{i=1}^{m}\left|l^T y_i\right|,\ \ y_i^T y_j=\delta_{ij},\ \ i,j=\overline{1,m},$$

при цьому

$$\max_{a\in\left\{a:\ \left|a^T y_j\right|\le 1,\ j=\overline{1,m}\right\}} l^T a=\sum_{i=1}^{m}\left|l^T y_i\right|. \qquad (31)$$

Тоді для двох множин

$$\Omega_1=\left\{a:\ \left|y_j^T(1)(a-\bar a)\right|\le d_j(1),\ \ j=\overline{1,m}\right\},$$

$$\Omega_2=\left\{a:\ \left|y_j^T(2)(a-\bar b)\right|\le d_j(2),\ \ j=\overline{1,m}\right\} \qquad (28)$$

необхідна і достатня умова їхнього неперетину буде такою

$$\exists\, l\in R^m,\quad \sum_{j=1}^{m}\left(\left|l^T y_j(1)\right|d_j(1)+\left|l^T y_j(2)\right|d_j(2)\right)<$$

$$<l^T(\bar b-\bar a), \qquad (33)$$

а більш проста достатня умова неперетину множин (32) має форму

$$\sum_{j=1}^{m}\left[\left|(\bar b-\bar a)^T y_j(1)\right|d_j(1)+\left|(\bar b-\bar a)^T y_j(2)\right|d_j(2)\right]<$$

$$<\left\|\bar b-\bar a\right\|^2. \qquad (34)$$

При дослідженні умов неперетину множин типу еліптичних циліндрів будемо використовувати таку формулу

$$\max_{a\in\left\{a:\ \sqrt{\sum_{j=1}^{s}\left(a^T y_j\right)^2}\le 1,\ \sqrt{\sum_{j=s+1}^{m}\left(a^T y_j\right)^2}\le 1\right\}} l^T a=$$

$$=\sqrt{\sum_{j=1}^{s}\left(l^T y_j\right)^2}+\sqrt{\sum_{j=s+1}^{m}\left(l^T y_j\right)^2}. \qquad (35)$$

Тут для двох множин

$$\Omega_1=\left\{a:\ (a-\bar a)^T K_1(1)(a-\bar a)\le 1,\right.$$

$$\left.(a-\bar a)^T K_1(2)(a-\bar a)\le 1\right\}, \qquad (36)$$

$$\Omega_2=\left\{a:\ (a-\bar b)^T K_2(1)(a-\bar b)\le 1,\right.$$

$$\left.(a-\bar b)^T K_2(2)(a-\bar b)\le 1\right\},$$

де

$$K_i(1)=\sum_{j=1}^{s_i}y_j(i)y_j^T(i)\lambda_j^{-2}(i),$$

$$K_i(2)=\sum_{j=s_{i+1}}^{m}y_j(i)y_j^T(i)\lambda_j^{-2}(i),$$

необхідною і достатньою умовою їх неперетину буде

$$\exists\, l\in R^m,\quad \sqrt{l^T K_1^+(1)l}+\sqrt{l^T K_1^+(2)l}+$$

$$+\sqrt{l^T K_2^+(1)l}+\sqrt{l^T K_2^+(2)l}<l^T(\bar b-\bar a), \qquad (37)$$

а достатня умова буде мати форму

$$\sum_{i,j=1}^{2}\sqrt{(\bar b-\bar a)^T K_i^+(j)(\bar b-\bar a)}<\left\|\bar b-\bar a\right\|^2. \qquad (38)$$

Тут

$$K_i^+(1)=\sum_{j=1}^{s_i}y_j(i)y_j^T(i)\lambda_j^2(i),$$

$$K_i^+(2)=\sum_{j=s_{i+1}}^{m}y_j(i)y_j^T(i)\lambda_j^2(i).$$

Умови неперетину в просторі ознак при класифікації сигналів визначаються застосуванням вище отриманих умов неперетину множин до множин-контейнерів (18)-(24).

Література

1. Гантмахер Ф.Р. Теория матриц.-М.:Наука,1967.-287 с.
2. Hiroshi Murase, Shree K. Nayar, Visual Learning and Recognition of 3-d Objects from Appearance, - International Journal of Computer Vision, 14, 5-24 (1995).
3. R. Kuc. Biomimetic sonar locates and recognizes objects. IEEE Journal of Oceanic Engineering, vol. 22 (4), 1997.
4. Кириченко Н.Ф., Куц Р.(США), Лепеха Н.П. Распознавание трехмерных объектов по ультразвуковым эхо-сигналам. //Проблеми управления и информатики.-1999.-№5.-С.110-122.
5. Кириченко М.Ф., Куц Роман (США), Лепеха М.П. Алгоритми розпізнавання об'єктів ультразвуковими сонарами. Теорія обчислень: Збірник наукових праць. Інститут кібернетики НАНУ. Київ, 1999, с.196-200

КОНСТРУКТИВНІ АЛГОРИТМИ РОЗБИТТЯ СКЛАДНИХ СИСТЕМ НА ОСНОВІ ОПТИМАЛЬНОГО ЗГОРТАННЯ

Роман Базилевич
ДУ "Львівська політехніка"; вул.С.Бандери, 12; Львів, 79646; Україна
Тел: +380322 398877; Факс: +380322 744300; E-mail:rbaz@polynet.lviv.ua

Абстракт. Розглядається задача розбиття складних систем на частини, яка з математичної точки зору відноситься до неполіноміальних (важкорозв'язних) комбінаторних проблем. Запропоновано нові конструктивні алгоритми на основі методу оптимального згортання схеми. Виділено основні макропроцедури для кількох різних стратегій. Дано оцінку їх ефективності.

1. Вступ

Постійне зростання складності систем опрацювання даних не тільки не знімає, а підвищує актуальність однієї з найскладніших комбінаторних задач аналізу даних – оптимального розбиття системи на частини за заданими критеріями та з забезпеченням необхідних обмежень. Звичайно знайти оптимум для задач даного класу у випадку їх великої розмірності, не говорячи вже про реальні задачі надвеликих розмірностей, надзвичайно важко. Тому з практичної точки зору задовольняються "добрими" розв'язками, наближеними до оптимальних. З математичної точки зору проблема відноситься до неполіноміальних задач дискретної оптимізації, які отримали назву важкорозв'язних (клас *NP*). Існує багато різноманітних методів та алгоритмів розбиття [1]. Ефективним виявився метод оптимального згортання схеми (ОЗС) [2]. Важлива ідея підходу полягає в тому, щоб в операторах алгоритмів замість базових елементів системи як основних "цеглинок" для формування розв'язку використовуються кластери довільної розмірності з заданою структурою їх взаємного ієрархічного входження. Задача має різноманітні прикладні застосування, зокрема для аналізу даних, ієрархічної класифікації об'єктів, автоматизованого проектування та інших прикладних проблем.

2. Формулювання задачі

Заданою є система $S=(P, E)$, де $P=\{p_1,..., p_n\}$ – множина базових елементів та $E=\{e_1,..., e_m\}$ - множина зв'язків. Необхідно отримати розбиття $\widetilde{P}^*=\{P_1^*,...,P_k^*\}$ множини P на k підмножин, яке за вибраним критерієм якості наближається до оптимального: $Q(\widetilde{P}^*)\rightarrow opt_i Q(\widetilde{P}^i)$, де $\widetilde{P}^i$ – довільне розбиття. Як критерій якості найчастіше розглядається число зовнішніх зв'язків $Q=|E^{ex}|$, що зв'язують між собою частини розбитої системи.

3. Класифікація алгоритмів

Суть конструктивних алгоритмів розбиття полягає у формуванні розв'язку шляхом покрокового визначення його окремих елементів. Кожна частина розбиття є одним елементом розв'язку. Перш за все перед розбиттям необхідно виділити окремо незв'язані частини схеми. Таку задачу може виконати метод ОЗС. Якщо схема має такі частини, то в процесі згортання утвориться не одне дерево, а ліс $F^R=\{T_1^R,..., T_f^R\}$ з двох чи більшого числа дерев. Кожне дерево $T_i^R \in F^R$, $i=1,..., f$, відповідає одній зв'язаній частині схеми. Першою при декомпозиції повинна бути спроба отримання окремих частин розбиття на основі аналізу лісу F^R. Виділяємо дві групи підсхем: що задовольняють бажані обмеженням і що їх не задовольняють. Підсхеми, що задовольняють бажані обмеження, заносимо в розв'язок та усуваємо з подальшого розгляду. Серед решти підсхем, що не задовольняють обмеження, виділимо окремо більші за розмірами від потрібних та менші. Розгляд решти частини схеми розпочинаємо з підсхем, найбільших за розмірами. З кожної такої підсхеми формуємо ціле число частин розбиття, що задовольняють обмеження, а решту елементів усуваємо для подальшого включення в інші підсхеми. Виникає часткова задача розбиття великої зв'язаної підсхеми на задане число частин з залишком, який буде доданий до підсхем, менших за розміром ніж потрібно для частин розбиття. Саме ця задача є базовою і подальший розгляд буде їй присвячений.

Менші ніж потрібно за розмірами підсхеми разом з "кусками", відрізаними від великих підсхем, які розглядалися спочатку, стараємося згрупувати разом в потрібні за розміром частини

розбиття, включити в розв'язок та усунути їх з подальшого розгляду. Тут виникає задача, подібна до типової комбінаторної задачі пакування "рюкзака". Відмінність типової задачі полягає в тому, що "рюкзак" заповнюється неподільними предметами заданого об'єму, причому в "рюкзаку" може залишатися певна частина незаповненого простору. У нашому випадку такий вільний простір не залишається, оскільки заданими є число частин розбиття і відповідно число складових базових елементів, які необхідно "упакувати" в рюкзак. Якщо число елементів виділених частин жорстко не задається, а можливим є певне відхилення від середнього значення, то це сприятиме отриманню кращого розв'язку, оскільки може не завжди вимагати розрізання певного кластера на дві частини, одна з яких входить у виділену частину, а друга переноситься для включення в решту ще несформованих частин. Проблема оптимального розв'язування цієї задачі потребує окремого додаткового дослідження.

Розглянемо базову задачу розбиття без залишку однієї зв'язаної частини схеми на підсхеми, що задовольняють бажані обмеження. Доцільно виділити три основні групи алгоритмів: послідовного, паралельно-послідовного та дихотомічного формування початкових розв'язків. Спільним для всіх них є побудова дерева згортання T^R методом ОЗС для формування кластерів. Послідовний алгоритм покроково виділяє окремі частини. Основними його кроками є такі: 1) будуємо дерево T^R до моменту виділення першого кластера Cl^*, розмір якого є рівним бажаному значенню однієї частини розбиття або більшим за нього; 2) якщо розмір кластеру Cl^* відповідає бажаному значення, то переходимо до п.4, у протилежному випадку визначаємо число елементів Δn, які необхідно усунути з кластера для отримання частини розбиття бажаного розміру; 3) переносимо Δn елементів з кластера Cl^* в решту частину схеми, утворюючи бажану частину $P^* \in \widetilde{P}$; 4) від всієї схеми відтинаємо частину P^* і, якщо ще залишаються непризначені у частини елементи, то оновлюємо дані та повертаємося на п.1 для формування наступної частини розбиття. У протилежному випадку робота алгоритма закінчена.

Паралельно-послідовний алгоритм на першому кроці на основі дерева T^R виділяє кластери, число яких дорівнює бажаному значенню та які за своїми розмірами найкраще наближаються до бажаних. Далі відбувається коригування розмірів частин.

Дихотомічний алгоритм є низхідним. Ідея його полягає в формуванні частин, кратних за числом елементів до бажаних, на основі дерева згортання T^R шляхом дихотомічного ділення зверху-вниз. Алгоритм має наступні кроки:1) будуємо повне дерево згортання T^R; 2) виділяємо на дереві T^R ребро v^*, що розділяє всю множину елементів P на дві підмножини P' та P'', найбільш доцільні для їх подальшого рекурсивного дихотомічного дроблення на частини; 3 визначаємо число частин k_1 та k_2 ($k_1 + k_2 = k$), які можуть бути утворені з обох підсхем.; 4) визначаємо число елементів Δn, які необхідно перенести з однієї підсхеми в іншу для формування підсхем з числом елементів, кратним до чисел k_1 та k_2 та переносимо їх процедурами, описаними нижче; 5) формуємо дві підмножини P_1 та P_2, утворені відповідно з підмножин P' та P''. Якщо розмір однієї з підсхем (або обох) відповідає бажаному, то вона заноситься в розв'язок як виділена частина розбиття та з подальшого розгляду усувається. В протилежному випадку більшу за розміром підсхему заносимо в чергу для розгляду та ділення на наступних кроках, а меншу розглядаємо як початкову схему та дробимо поверненням на п.1; 6) беремо наступну з черги підсхему та повертаємося на п.1, розглядаючи її як початкову схему. Якщо черга порожня, то робота алгоритма закінчена.

Відзначимо основні переваги та недоліки описаних алгоритмів. Послідовний алгоритм є найпростішим, проте йому найбільш властивими є "жадібні" риси, а тим самим і незадовільні результати при великій розмірності задачі. В процесі відтинання окремих частин від кластерів залишаються "хвости", з яких формуються останні частини. Ці частини є неякісними, нещільними, з великим числом зовнішніх зв'язків. Недоліком паралельно-послідовного алгоритму є більша складність. Не завжди на дереві згортання вдається знайти розріз, який добре наближається до бажаного розв'язку за числом кластерів та їх розмірністю, тобто виділити добрі початкові ядра для подальшого їх зменшення чи збільшення. Початкові ядра можуть суттєво відрізнятися за розмірами між собою та від бажаного значення, що ускладнює роботу алгоритма. Алгоритм на основі дихотомічного ділення вважаємо найбільш ефективним, придатним для задач великих та надвеликих розмірностей. Він є складнішим від першого, проте на відміну від нього не утворює "хвостів", що погіршують якість останніх виділених частин розбиття. Він обирає найбільш раціональні розрізи схеми для подальшого дроблення на частини. Йому є найменш властивим принцип "жадібності", оскільки виділяє не по одній частині за найкращими умовами, а здійснює паралельний розділ схеми з урахуванням потреб всіх частин одночасно.

4. Основні макропроцедури

Побудова дерева згортання є найбільш важливою макропроцедурою, яку використовують всі описані алгоритми. Основні особливості цієї процедури розглянуто в [2,3]. Якщо в результаті згортання схеми утворюється не одно дерево, а два чи більше їх число, то це вказує на наявність в

схемі більше однієї зв'язаної підсхеми. Кожне дерево відповідає одній такій підсхемі.

Для оптимального виділення та перенесення визначеного числа Δn елементів з однієї підсхеми в іншу пропонуються два підходи. Перший передбачає використання дерева згортання, другий – оціночних критеріїв. Вхідною інформацією для даної макропроцедури є кластер Cl^*, з якого необхідно усунути елементи, та решта частина схеми, яку умовно позначаємо через R. Решта частина може складатися з однієї або кількох підсхем. Це суттєвого значення не має, оскільки R розглядається як єдине ціле. Вихідною інформацією повинна бути множина з Δn елементів, які усуваються з кластера Cl^* та переносяться в R.

Розглянемо можливі підходи детальніше. При використанні першого підходу кластер Cl^* розсипають та розглядають його на рівні базових елементів. Решта частини схеми R розглядається як один елемент. На такій множині будується дерево T^R. Визначаємо число елементів, що приєднуються до дерева на шляху від вершини R до його кореня (найвищої вершини). Можливими є чотири випадки: 1) це число дорівнює бажаному значенню Δn елементів; 2) воно є більшим за нього; 3) є меншим за нього; 4) дорівнює нулю.

У першому випадку задача розв'язується безпосередньо шляхом включення всіх виділених елементів в розв'язок. У другому випадку йдучи від вершини R вверх по дереву до кореня знаходимо вершину, при відтинанні якої разом з попередніми вершинами утворюється бажана множина з Δn елементів. Якщо ж такої вершини не знаходимо, але є дві суміжні, при відтинанні яких виникають підсхеми з меншим та більшим від бажаного значення числа елементів, то можливими є дві процедури продовження. Для першої процедури беруть вершину, яка разом з попередніми утворює множину Δn^- з меншим від бажаного числа елементів та відповідні елементи на шляху від неї до вершини R приєднують до R . Далі процедуру повторюємо виділяючи вже множину $\Delta n^{'} = \Delta n - \Delta n^-$ на редукованій множині початкових елементів без приєднаних Δn^- елементів до R . Процедура рекурсивно продовжуємо до отримання бажаного розв'язку або до моменту, коли першою утворюється вершина з більшим за бажане число елементів.

Для другої процедури, яка виникає також тоді, коли першою на шляху до кореня дерева згортання є вершина, що відповідає кластеру з більшим ніж потрібно числом елементів, доцільною є наступна реалізація. Елементи цього кластера розсипаються, решта елементів з однієї сторони вершини утворюють нову підсхему $R^{'}$, куди входить також частина початкових елементів, а з другої – підсхему $R^{''}$, куди входять також елементи множини R. На утвореній множині будується нове дерево згортання. В процесі згортання забороняється об'єднуватися підсхемам, що включають одночасно $R^{'}$ та $R^{''}$. Далі на шляху між вершинами $R^{'}$ та $R^{''}$ визначають найкращу для розрізання схеми вершину і т.д. аж поки не буде знайдено бажаний розв'язок. На кожному такому кроці редукується множина невизначених елементів, частина елементів призначається або в одну або в другу підсхему. Проте в процесі реалізації першої та другої процедур може виникнути ситуація, коли на шляху є одна або кілька вершин, що відповідають кластерам, елементи яких не приєднуються ні до однієї з двох вершин $R^{'}$ чи $R^{''}$, а згортаються при побудові дерева T^R повністю окремо. У цьому випадку використовуємо описаний нижче підхід, що реалізує призначення елементів та кластерів на основі оціночних критеріїв для всіх непризначених базових елементів та утворених з них кластерів.

У третьому випадку, коли число елементів, що приєднуються до дерева на шляху від вершини R до його кореня є меншим за бажане значення Δn елементів, проте частина необхідних елементів виділилася, приєднуємо ці елементи до решти схеми R та рекурсивно продовжуємо пошук подібним чином. Якщо ж число всіх елементів, що приєднуються визначеними процедурами є недостатнім і на шляху від R до кореня не появляються при черговій рекурсивній реалізації названої процедури нові вершини, або вони не виникали з самого початку, що відповідає четвертому з наведених вище випадків, необхідно іти від кореня вниз по дереву суміжною гілкою. Проте дати однозначно рекомендації яким чином виділяти тут нові елементи чи кластери важко. В даному разі краще продовжити виділення решти потрібних для перенесення елементів на основі другого підходу, що використовує оціночні критерії.

Розглянемо процедуру призначення елементів на основі оціночних критеріїв, яка може мати як самостійне значення, так і допоміжне в описаному вище другому випадку. На елементах підсхеми, з якої необхідно виділити для перенесення бажане число елементів, будується дерево згортання T^R. Решта частина схеми розглядається як єдине ціле у вигляді одного базового елемента R у четвертому випадку або як два базових елементи $R^{'}$ та $R^{''}$ у другому випадку. При цьому беруться до уваги зв'язки з рештою частиною схеми. На дереві виділяються всі кластери, число елементів яких є рівним або меншим від числа елементів, що необхідно перенести. Для кожного такого кластера та базового елемента визначається значення критерію його перенесення в бажану частину (приріст числа зовнішніх зв'язків або інші) та вони відповідно впорядковуються. Наступними можливі дві стратегії – поодинокого призначення (кластера або елемента) з наступною переоцінкою значення критерію або групового призначення, при якому

один раз визначається множина елементів бажаної потужності.

Задача призначення виділених з однієї підсхеми елементів в решту (декілька) зв'язаних з нею підсхем полягає у призначенні заданого числа Δn елементів, виділених з кластеру Cl^*, в решту підсхем $R_1,\ldots, R_r$. Можливими як і попередньому випадку є два підходи – на основі дерева згортання схеми та оціночних критеріїв. При використанні дерева згортання елементи множини Δn розглядаються як базові нарівні з окремими підсхемами $R_1,\ldots, R_r$. Будується ліс дерев F^R , в якому кожне окреме дерево відповідає одній частині з множини $R_1,\ldots, R_r$. Згортання схеми відбувається з обмеженням, при якому підсхемам, що включають різні частини, об'єднуватися не дозволяється, що й викликає утворення лісу замість одного дерева. В результаті отримуємо r окремих дерев з приєднаними до них Δn елементами. Другий підхід полягає у використанні оціночних критеріїв та призначенні кластерів та елементів в ту чи іншу підсхему за найкращими їх значеннями.

Для визначення найкращого ребра для розрізання схеми на дві частини можливими є декілька підходів. У першому випадку вибираємо два найвищих кластери. Тут можливою є поява ситуації, коли число елементів одного з кластерів є набагато меншим за потрібне середнє значення $n_{сер}$ елементів в одній частині ($n_{сер}= n/k$). Тоді необхідно опуститися по дереву до ребра так, щоб число елементів однієї з підсхем, що виділяється, не було меншим за $\alpha n_{сер}$, де α - наперед заданий коефіцієнт, наприклад α =0,7 – 0,9.

Другий випадок передбачає вибір такого ребра розрізу, коли число Δn елементів, що необхідно переносити з однієї підсхеми в іншу для забезпечення кратності чисел їх елементів до числа $n_{сер}$, є найменшим. Третій випадок передбачає вибір такого розрізу, коли числа елементів обох підсхем, що утворюються внаслідок розділу, відрізняються найменше. У четвертому випадку вибирається такий розріз, коли число зв'язків розрізу, поділене на число елементів меншої підсхеми, є найменшим.

5. Експериментальні дослідження

Обчислювальні затрати для основної процедури - побудови дерева згортання є близькими до лінійних від числа базових елементів [3]. Для розробленої програми розглянуто вісімнадцять тестів фірми IBM, що стали типовими для задач розбиття [4]. Час згортання схеми з 12752 елементів (тест ibm01) склав 14 сек., а для схеми з 210613 елементів (тест ibm18) - 320 сек. при 100% об'єднанню на кожному кроці (жорстка кластеризація) та 50 сек. і 1253 сек. відповідно при об'єднанні тільки найкращих пар (вільна кластеризація). Розрахунки виконано на ПК з процесором Intel-Celeron 450.

Експерименти [5], підтвердили добру якість отриманих результатів на множині відомих тестів при сумісному використанні конструктивних та оптимізаційних алгоритмів [6]. Для дванадцяти розглянутих тестів результати не є гіршими ніш отримані іншими методами, у п'яти випадках одержані кращі від відомих (тести C5315, C7552, s13207, s38417, s38584), у восьми випадках результати є оптимальними (тести C499, C1355, C1908, C3540, C7552, C6288, s15850, s38584).

6. Висновки

Розглянуті алгоритми конструктивного розбиття дозволяють задачу великої розмірності рекурсивно зводити до розв'язування аналогічних задач меншої розмірності. Як простішим для реалізації вважаємо метод послідовного типу; кращі результати слід очікувати від дихотомічного методу. Алгоритми використовують однакові базові макропроцедури, що дозволяє побудувати єдиний пакет програм, який реалізує всі зазначені підходи. Це створить умови для розв'язування кожної задачі кількома алгоритмами з подальшим вибором найкращого результату. Бажаним є їх сумісне використання з оптимізаційними алгоритмами. Для цього можуть бути використані будь-які ітераційні алгоритми оптимізації розбиття, проте доцільнішим вважаємо використання підходів, основаних на методі оптимального згортання схеми [6], що дозволить застосувати основні виділені тут макропроцедури, які є спільними в обох випадках.

7. Література

1. Sadiq M.Sait, Habib Youssef. *VLSI physical design automatic, Theory and practice.* New York, IEEE Press, 1995.
2. Базилевич Р.П. *Декомпозиционные и топологические методы автоматизированного конструирования электронных устройств.* Львов: Вища школа. Изд-во при Львов. гос. ун-те, 1981. -168с.
3. Базилевич Р.П., Подольський І. *Ієрархічна кластеризація складних схем.* Вісник ДУ "Львівська політехніка", № 392, "Комп'ютерна інженерія та інформаційні технології", 2000.
4. Alpert C.J. *The ISPD-98 Circuit Beanchmark Suit. Proc.* ACM/IEEE Intern. Symposium on Physical Design, April 1998, pp.80-85.
5. Bazylevych R.P., Melnyk R.A., Rybak O.G. *Circuit partitioning for FPGAs by the optimal circuit reduction method.* VLSI Design. Vol 11, No 3, 2000.
6. Базилевич Р. *Ітераційні алгоритми оптимізації розбиття складних систем на основі оптимального згортання.* В даному збірнику.

ІТЕРАЦІЙНІ АЛГОРИТМИ ОПТИМІЗАЦІЇ РОЗБИТТЯ СКЛАДНИХ СИСТЕМ НА ОСНОВІ ОПТИМАЛЬНОГО ЗГОРТАННЯ

Роман Базилевич
ДУ "Львівська політехніка"; вул.С.Бандери, 12; Львів, 79646; Україна
Тел: +380322 398877; Факс: +380322 744300; E-mail:rbaz@polynet.lviv.ua

Абстракт. Розглядається задача розбиття складних систем на частини, яка з математичної точки зору відноситься до неполіноміальних (важкорозв'язних) комбінаторних проблем. Запропоновано нові ітераційні алгоритми оптимізації розбиття на основі методу оптимального згортання схеми. Виділено основні макропроцедури для кількох стратегій. Вперше реалізовано алгоритми обміну кластерами довільної розмірності. Дано оцінку ефективності запропонованих алгоритмів.

1. Вступ

Існує багато різноманітних ітераційних алгоритмів оптимізації розбиття складних схем [1]. Знайти оптимум для задач даного класу у випадку їх великої розмірності, не говорячи вже про реальні задачі надвеликих розмірностей, надзвичайно важко. Тому з практичної точки зору задовольняються "добрими" розв'язками, бажано максимально наближеними до оптимальних. Задача розбиття відноситься до неполіноміальних проблем дискретної оптимізації, які отримали назву важкорозв'язних (клас *NP*). Переважна більшість з них оперує окремими базовими елементами схеми, і тільки деякі – кластерами однаковими розмірностями [2]. Такі алгоритми мало придатні для задач великих та надвеликих розмірностей, вимагають значного комп'ютерного часу та заводять розв'язок в локальні екстремуми, далекі від оптимальних. Ідея пропонованих алгоритмів полягає у використанні кластерів довільної розмірності на рівні з базовими елементами схеми для ітераційного обміну між частинами початкового розв'язку. Для підвищення ефективності алгоритмів застосовуються принципово нові підходи, основані на перепризначенні кластерів та окремих базових елементів з використання методу оптимального згортання схеми (ОЗС) [3,6] та оціночних критеріїв. Метод ОЗС ефективно використовується для виявлення кластерів схеми, що надзвичайно важко вирішувалося іншими методами.

2. Формулювання задачі

Заданою є схема $S=(P, E)$, та початкове розбиття множини її елементів $\widetilde{P}^0=\{P^0_1,...,P^0_k\}$ на k частин. Тут $P=\{p_1,..., p_n\}$ – множина базових елементів та $E=\{e_1,..., e_m\}$ - множина зв'язків схеми. Необхідно отримати оптимізоване розбиття $P^*=\{P^*_1,...,P^*_k\}$, яке за вибраним критерієм якості Q наближається до оптимального: $Q(\widetilde{P}^*)\to opt_iQ(\widetilde{P}^i)$, де $\widetilde{P}^i=\{P^i_1,..., P^i_k\}$ – довільне i–те можливе розбиття. Тут $P^i_j=\{p^i_{j1},..., p^i_{jkj}\}$, $j=1,...,k$; $P^i_j\in P^i$ – частина розбиття (підмножина). Для кожного розбиття виконуються вимоги:

$(\forall P^i_j\in\widetilde{P}^i)[(P^i_j\subset P)\&(P^i_j\neq\varnothing)]$;

$(\forall P^i_r, P^i_s\in\widetilde{P}^i)[P^i_r\cap P^i_s=\varnothing]$; $\bigcup^k_{j=1}\bigcup^{kj}_{t=1}P^i_{jt}=P$.

Як критерії якості найчастіше розглядаються число зовнішніх зв'язків $Q_1=|E^{ex}|$, тобто елементів множини $E^{ex}\in E$, що зв'язують між собою окремі частини розбитої схеми. Другим часто вживаним критерієм є сумарне число зовнішніх виводів всіх частин: $Q_2=\Sigma^k_{i=1}|E^{ex}_i|$, де E^{ex}_i – множина зовнішніх зв'язків i–ої частини розбиття. Тобто, якщо зв'язок з'єднує дві частини розбиття, він береться з коефіцієнтом два, якщо три – тоді з коефіцієнтом три і т.д. Часто вживаним критерієм є також число частин k, яке необхідно мінімізувати при виконанні двох обмежень – на число елементів кожної частини та число її зовнішніх зв'язків. Така задача розбиття отримала назву пакування. Звичайно початкове розбиття формується конструктивними методами або випадково.

Оптимізація здійснюється шляхом багатокрокового покращання початкового розв'язку: $\widetilde{P}^0\to\widetilde{P}'\to\widetilde{P}''\to\widetilde{P}'''\to ...\to\widetilde{P}^*$. Число кроків оптимізації звичайно визначається в процесі самої оптимізації – вона виконується поки наступає покращання або до моменту, коли воно є вже незначним, або в межах виділеного часу. В деяких випадках може наперед задаватися число кроків оптимізації. Відсутність покращання, що наступає

© UkrObraz'2000

на певному кроці оптимізації, не значить про знаходження оптимального результату (хоча й не виключає таку можливість. У більшості випадків реального проектування процес заходить в певний локальний екстремум, для знаходження виходу з якого з продовженням оптимізації необхідно використовувати спеціальні алгоритми. Більшість відомих ітераційних алгоритмів оптимізації розбиття реалізує попарну оптимізацію. Практично невідомими є алгоритми, в яких оптимізуються одночасно декілька частин розбиття. Деякі такі можливості групової та попарної оптимізації, розглядаються нижче.

3. Попарна оптимізація

Важливими чинниками, що впливають на результат при попарній оптимізації є: порядок вибору пар для оптимізації, тобто її послідовність, безпосередньо алгоритми та критерії оптимізації. Доцільнішим вважаємо спершу оптимізацію частин, які є сильніше зв'язаними, тобто що мають більше спільних зв'язків. Для визначення таких пар необхідно побудувати матрицю зв'язності між частинами розбиття. Далі можливими є два шляхи: покрокова оптимізація окремих частин зі всіма рештою, тобто спершу вибирається одна частина, яка почергово оптимізується зі всіма рештою, потім друга зі всіма рештою, крім першої, і т.д. Такий повний цикл включає $n(n-1)/2$ попарних оптимізацій. Після першого циклу виконується другий і т.д. до моменту, коли покращання вже не наступає, або воно є незначними. Для іншої стратегії на основі матриці зв'язності для першого кроку вибирається $n/2$ незалежних пар, що мають найбільше спільних зв'язків. На другому кроці розглядається така ж кількість незалежних пар, що не оптимізувалася на першому кроці і т.д. Перший цикл закінчується, коли розглянуті всі можливі $n(n-1)/2$ пари. Після цього цикл повторюється до завершення оптимізації як і в попередньому випадку. Можливе спільне (почергове) використання обох стратегій послідовно по одному циклу. Перша стратегія має більш "жадібний" характер, оскільки вона створює найкращі умови для першої частини. Решта частин мають гірші стартові умови. Друга стратегія створює сприятливіші умови для рівномірної оптимізації всіх частин.

Пропонуються два типи алгоритмів попарної оптимізації – на основі методу оптимального згортання схеми та оціночних критеріїв. Перший випадок є придатним для двох та більшого числа частин. Тут реалізується "розсипання" на базові елементи двох вибраних частин та здійснюється спроба нового розбиття конструктивними методами з наступною оптимізацією. Якщо спроба є позитивною, то приймається новий розв'язок для пари частин. При побудові дерева згортання зв'язки з рештою частиною схеми можуть не братися до уваги. Підхід може бути ефективним для великого числа частин, оскільки початкове розбиття здійснювалося на великій множині базових елементів, а в даному випадку вона є суттєво меншою – розглядається множина елементів тільки двох частин, і тому ймовірність знаходження кращого розв'язку є високою. Тобто, реалізується алгоритм, розроблений для знаходження початкового розміщення, тільки на суттєво звуженій множині елементів з наступною оптимізацією.

Другий алгоритм оптимізації використовує оціночні функції. Для обох частин будуються окремі дерева згортання. Тут зв'язки з рештою частиною схеми беруться до уваги. Для всіх кластерів та базових елементів обох дерев визначаються оціночні функції їх перенесення в з однієї в іншу частину. Формується впорядкований за значенням критерію список кластерів та елементів, де вони не розділяються. В обох списках виділяються кластери та елементи, для яких значення критерію є сприятливим для переносу. Розглядаюся можливості їх попарного та групового обміну між частинами. Виділяються всі пари, для яких сумарне значення критерію є позитивним та обмін яких не порушує обмежень на числа елементів в частинах чи інші. Розглядаються також варіанти чистого перенесення кластера чи елемента з однієї частини в іншу без обміну, якщо це не викликає порушень обмежень. Можливими є стратегії обміну, в яких дозволяється деяке (наперед задане за величиною) порушення обмежень на число елементів з наступним поверненням до вимог без порушення обмежень. Тобто, реалізується коливальний ітераційний процес, коли за певних умов дозволяється виходити за обмеження з обов'язковим поверненням до необхідних вимог. Це може сприяти виходу з локальних оптимумів, в які зайшов процес оптимізації.

Для кожної спроби обміну визначається реальне значення критерію. Воно може відрізнятися від вже визначеного оціночного (як суми оціночних критеріїв для окремих елементів чи кластерів), оскільки елементи (кластери), що обмінюються, мають зв'язки між собою. Для точної оцінки розробляються спеціальні економні процедури, які визначають спільні зв'язки та враховують їх при визначенні значення критерія. Можлива й оцінка значення критерію обміну на основі повного перерахунку потужностей множин зовнішніх зв'язків для новоутворених після обміну частин. Вибирається пара, яка дає найкращі результати обміну. Процес визначення наступної пари для обміну повторюється аналогічним чином з побудовою нових дерев згортання для покращених на попередньому кроці частин. Оптимізація припиняється, коли коли обмін не дає покращання.

Якщо в списку значень оціночних критеріїв є декілька незалежних кластерів (елементів), що

мають його позитивне значення, то розглядається можливість "групового" обміну з групою незалежних кластерів (елементів) відповідної потужності другої підсхеми. Такий підхід суттєво прискорює оптимізацію, особливо на перших кроках, коли можлива поява великого числа кластерів (елементів) з позитивним значенням оціночного критерію. Проте не виключеним є отримання гірших результатів ніж при попарному обміну, при якому коригуються значення критеріїв після обміну кожною парою.

Можливим є ще один підхід до оптимізації на основі згортання схеми. Зменшення числа зовнішніх зв'язків може бути викликано тільки обміном кластерів (елементів), що мають спільні зв'язки в обох частинах. Тому виділивши такі кластери (елементи) необхідно зробити спробу їх обміну. Тут доцільно використати алгоритм перенесення певної множини елементів з однієї частини в іншу. Виділені спільні елементи "розсипають". Всі решта невиділені елементи обох частин (окремо по частинах) об'єднуються в групи R' та R'' На утвореній множині будується дерево оптимального згортання. На шляху дерева між вершинами R′ та R″ визначається вершина, що задовольняє бажані обмеження та має найкращі значення критерію розбиття. Ця вершина приймається за точку розділу схеми на дві нові покращені частини. Після першого кроку процес оптимізації можна продовжувати поки наступає покращання. Основними макропроцедурами в описаних підходах є наступні:

1. "Розсипання" двох частин на елементи.
2. Формування нового розв'язку на розсипаних елементах конструктивними методами з наступною оптимізацією.
3. Побудова на елементах двох виділених для оптимізації частин окремих дерев згортання.
4. Визначення значення оціночних критеріїв при перенесенні елемента чи кластера з однієї частини в іншу.
5. Формування двох списків елементів та кластерів, впорядкованих за значеннями вибраних оціночних критеріїв перенесення (окремо для кожної частини для базових елементів та кластерів, виділених на деревах згортання в попередньому пункті).
6. Визначення елементів чи кластерів зі сприятливим для перенесення з однієї частини в іншу значенням оціночного критерія.
7. Визначення груп елементів та кластерів зі сприятливим для перенесення з однієї частини в іншу значенням оціночного критерія.
8. Визначення пар елементів та кластерів зі сприятливим для обміну між частинами значенням оціночного критерія.
9. Визначення груп елементів та кластерів зі сприятливим для обміну між частинами значенням оціночного критерія.
10. Оцінка результату (значення критерію) після перенесення чи обміну на основі визначенням множини зовнішніх зв'язків між оптимізованими парами частин двома шляхами:
- на основі інформації про їх складові елементи;
- на основі інформації про початкові до оптимізації множини зв'язків пар елементів, кластерів чи їх груп, що обмінюються, та кластерів, з яких вони усуваються, та інформації про перетин зовнішніх множин зв'язків між парами (елементами, кластерами та групами), що обмінюються. Цей підхід дозволяє скоротити обчислення, проте вимагає зберігання інформації про множини зв'язків всіх кластерів, що формуються при побудові дерев згортання.

4. Групова оптимізація

Як вже відзначалося, метод ОЗС дозволяє виконувати оптимізацію не тільки в межах двох частин, але й більшого їх числа. Тут можливими є багато різних стратегій оптимізації. Для цього вибирається група з трьох чи більшого число частин, вони "розсипаються" на базові елементи, на їх основі знаходиться новий розв'язок з використанням алгоритмів як конструктивного, так і описаних тут ітераційного типів. Важливим є належний відбір частин в групу, визначення їх числа та послідовності оптимізації. Вплив цих чинників на якість кінцевих результатів вимагає спеціальних досліджень. Задача є непростою та неоднозначною. Вважаємо доцільним для отримання певних рекомендацій побудову на множині частин як базових елементів дерева згортання, для чого попередньо формується матриця зв'язності частин. Дерево згортання визначає згустки по дві, три, та більше число частин розбиття. Для таких згустків можна виконати повторне розбиття для отримання кращих результатів, а також оптимізація розбиття між такими групами. Проте групами для перерозподілу елементів можуть бути не тільки вказані згустки, а будь-які зв'язані групи частин розбиття, включно з повним переглядом всіх можливих зв'язаних груп. Оптимізацію перерозподілу елементів необхідно здійснювати зверху-вниз, послідовно знижуючи число частин, що включаються в групи. На найнижчому рівні здійснюється попарна оптимізація всіх частин за вибраною стратегією.

Основними макропроцедурами тут є наступні:

1. Формування матриці зв'язності або списку зв'язків між частинами. Для кожної пари частин в найпростішому випадку визначається число спільних зв'язків, яке заноситься у відповідну клітку матриці. Можна також повні небінарні зв'язки моделювати бінарними, наприклад, n-арний зв'язок між n частинами моделюється повним графом з вагою ребер $2/n$. Матриця є симетричною. Можна враховувати й реальні n-арні зв'язки між

частинами, та як вхідну інформацію використовувати список зв'язків між частинами, в якому з початкового списку зв'язків усуваються всі внутрішні зв'язки частин, тобто такі, елементи яких повністю входять в одну частину, а записи зв'язків між частинами редукуються шляхом заміни імен базових елементів іменами частин, в які вони входять.

2. Побудова дерева згортання на основі матриці зв'язності чи списку зв'язків між частинами. Вхідною інформацією може бути матриця зв'язності, що відображає бінарні зв'язки між частинами.

3. Формування на основі дерева згортання груп по 2, 3 та більше число частин для наступної взаємної оптимізації.

4. Визначення стратегії (послідовності) оптимізації. Оптимізація повинна бути низхідною, тобто спочатку взаємно оптимізуються групи з кількох частин, останніми оптимізуються пари з двох частин. Бажано оптимізувати всі пари, що мають між собою спільні зв'язки.

5. Експериментальні дослідження

Експерименти, приведені в [5,6], підтвердили добру якість отриманих результатів на множині відомих тестів при сумісному використанні конструктивних [4] та описаних тут ітераційних алгоритмів. Розглядалася одна з найскладніших реальних задач розбиття – пакування складної електронної схеми в програмовані логічні матриці. Особливістю задачі є те, що необхідно задовольнити два жорсткі обмеження – на число зовнішніх зв'язків та число складових елементів для кожної частини. Ці величини не повинні перевищити заданого значення. Критерієм оптимізації є число частин розбиття, яке необхідно мінімізувати. Для всіх дванадцяти розглянутих тестів результати не є гіршими ніш отримані іншими методами, у п'яти випадках одержані кращі від відомих результати (тести C5315, C7552, s13207, s38417, s38584), у восьми випадках результати є оптимальними за числом виділених частин (тести C499, C1355, C1908, C3540, C7552, C6288, s15850, s38584).

6. Висновки

Вперше для оптимізації розбиття складних схем реалізовані процедури обміну кластерами довільної розмірності між частинами початкового розбиття. Обмін здійснювався між кластерами як однакових так і різних розмірностей. Великі кластери могли обмінюватися з малими; в окремих випадках, коли це покращувало значення критерію, здійснювався перенос кластерів без обміну; були реалізовані також процедури як одиночних, так і групових обмінів кластерів. Все це сприяло підвищенню ефективності оптимізаційного процесу. Інші методи чи алгоритми, не маючи таких процедур, заводять оптимізацію в локальні екстремуми, пошук виходу з яких є дуже проблематичним. Запропоновані підходи суттєво розширюють область пошуку екстремуму, причому пошук цей здійснюється цілеспрямовано в найбільш перспективних напрямках. Суттєвим також є процес виявлення самих кластерів, що забезпечувався застосуванням методу оптимального згортання схеми. В інших методах задача виявлення кластерів в схемі вирішується надзвичайно складно, вимагає значних обчислювальних затрат та орієнтована, в основному, на кластери однієї наперед заданої невеликої розмірності.

Розглянуті алгоритми рекурсивно зводять задачу великої розмірності до аналогічних задач меншої розмірності. Вони використовують однакові базові макропроцедури, що дозволяє побудувати єдиний пакет програм для всіх описаних підходів і розв'язувати кожну задачу кількома алгоритмами з вибором найкращого результату. Описані алгоритми є ітераційними, бажаним є їх сумісне використання з конструктивними. Можуть бути використані будь-які конструктивні алгоритми, проте доцільнішим вважаємо застосування підходів, основаних на методі оптимального згортання схеми [4], що дозволить застосувати однакові базові макропроцедури в обох випадках.

7. Література

1. Sadiq M.Sait, Habib Youssef. *VLSI physical design automatic, Theory and practice.* New York, IEEE Press, 1995.
2. Yousseb Saab. *Past-analysis-based clustering dramatically improves the Fiduccio-Mattheyses algorithm,*.Proc. of European DAC with EURO-VHDL`93, 1993, pp. 22-27.
3. Базилевич Р.П. *Декомпозиционные и топологические методы автоматизированного конструирования электронных устройств.* Львов: Вища школа. Изд-во при Львов. гос. ун-те, 1981. -168с.
4. Базилевич Р., *Конструктивні алгоритми розбиття складних систем на основі оптимального згортання.* В даному збірнику.
5. Bazylevych R.P., Melnyk R.A., Rybak O.G. *Circuit partitioning for FPGAs by the optimal circuit reduction method.* VLSI Design. Vol 11, No 3, 2000.
6. Базилевич Р.П., Рибак О.Г. *Оптимізація пакування програмованих логічних матриць на основі оптимального згортання схеми.* Вісник ДУ "Львівська політехніка", 1999, № 385.

Generalized Automatic Phonetic Transcribing of Speech Signals

Taras K. Vintsiuk

UNESCO/IIP International Research-Training Centre for Information Technologies and Systems,
40 Academician Hlushkov Avenue, Kyjiv 252022 Ukraine
Тел.: +380 44 266-4356
vintsiuk@uasoiro.freenet.kiev.ua

ABSTRACT

A so-called generalised phoneme recognition problem for the two-level speech understanding system is being solved. It means that under free phoneme order it is being found the $N>>1$ best phoneme sequence recognition responses. The method is based on constructive description of diverse realisations of a speech signal. A stochastic generative automata grammar, which is assigned to synthesise the speech signal prototypes, serves for it. This grammar composes all possible speech signal prototypes with allowance for non-linear rate of pronouncing in general, and of the pronouncing of individual phonemes in particular, as well as co-articulation and reduction of sounds and non-linear variation of the speech signal intensity along the time axis. To make deeper the earlier fulfilled research, phoneme-threephones (PT) signal prototypes are introduced. Rules for joining of PT signal prototypes into sequences are evident: the output and input phonemes of joining PT have to coincide. The problem is being solved using new compu-tational scheme of dynamic programming, based on (for substantial reduction in both memory and calculation requirements) the concepts of potentially optimal index and phoneme response.

АБСТРАКТ

Тарас Вінцюк. Узагальнене автоматичне фонетичне транскрибування усномовного сигналу. Розв'язується проблема так званого узагальненого автоматичного транскрибування усномовного сигналу, яка виникає при створенні дворівневої системи розміння мови. Вона полягає у знаходженні $N>>1$ найкращих послідовностей фонем, які складають відповідь розпізнавання. Метод грунтується на конструктивному описі (заданні) всього розмаїття мовних сигналів. Для цього використовуються стохастичні автоматні породжувальні граматики, які синтезують модельні ("еталонні") сигнали зв'язної мови, що відрізняються нелінійно змінюваними в часі темпом та інтенсивністю вимовляння, враховують коартикуляцію та редукцію звуків, індивідуальні особливості голосу. Щоб більш адекватно врахувати змінюваність мовних сигналів, введені поняття фонем-трифонів та їх модельних сигналів, індивідуального усномовного файлу (паспорта). Правила об'єднання модельних сигналів фонем-трифонів в послідовності є очевидними: вихідне ім'я та вхідне ім'я двох сусідніх фонем-трифонів повинні збігатись. Проблема узагальненого автоматичного фонетичного транскрибування розв'язується за допомогою ефективної процедури динамічного програмування, в якій, з метою значного скорочення обсягів обчислень та пам'яті, використані поняття потенційно-оптимальних індексів та потенційно-оптимальних фонемних відповідей розпізнавання.

INTRODUCTION

Still it is retained popular such approach in automatic speech recognition and understanding. It assumes that firstly continuous speech must be recognised as phoneme sequence, and then this phoneme sequence must be recognised and understood as word sequence and meaning to be transmitted by a speech signal [1, 2].

Though this approach seems to be erroneous, since the best method of finding of phonemes to be transmitted is both to recognise and to understand a speech signal, however it shows a preference for simplifying the research job distribution between specialists in acoustics, phonetics, linguistics, informatics.

To get better this approach it was proposed to introduce significant decisions in phoneme recognition procedures [2, 3]. The next step consists in making improvements to used generative automata grammars, for example instead of phoneme-diphones speech model [1, 3] to put into operation a phoneme-threephones one.

In this paper it is proposed a so-called generalised phoneme-threephone recognition problem for the two-level speech understanding system. The structure of this system is shown in Fig. 1. A generalised phoneme recognition problem means that under free phoneme order it is being found the $N>>1$ best phoneme sequence recognition responses. Then a Speech Interpreter analyses these phoneme sequences through Natural Language Knowledge filter.

PHONEME RECOGNITION IN CONTINUOUS SPEECH. GENERAL IDEA

The general idea is, taking into account inertial properties of articulation apparatus and language phonetics only, to construct some PT generative automata grammar which can synthesise all possible continuous speech model signals (prototypes) for any phoneme sequence. This grammar has to reflect such phenomena of speech signal variety as non-linear change of pronouncing both rate and intensity, sound co-articulation and reduction, sound duration statistics, phonemeness, and so on. Then the phoneme-by-phoneme recognition of unknown continuous speech signal will be involved in a synthesis of the most likely speech model signal and a determination of the phoneme structure of the latter.

© UkrObraz'2000

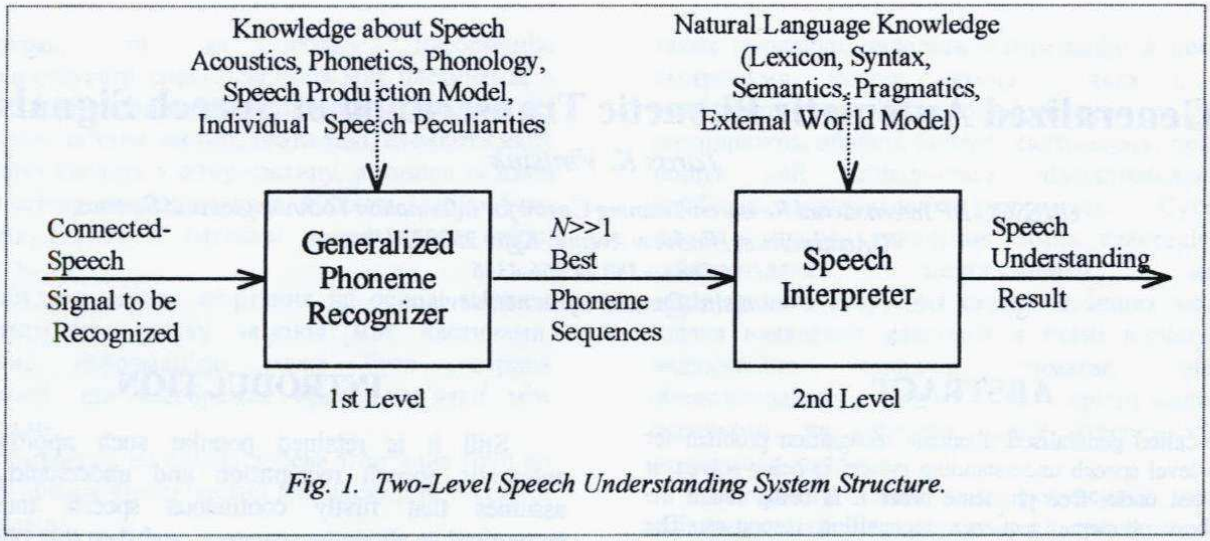


Fig. 1. Two-Level Speech Understanding System Structure.

The problem of directed synthesis, sorting out and formation of a phoneme sequence recognition response is solved by using a new computational scheme of dynamic programming, in which (for a substantional reduction in memory and calculation requirements) the concepts of potentially optimal both index and phoneme are used.

At first the phoneme-by-phoneme continuous speech recognition problem will be considered. Then this statement will be generalised for $N>>1$ best phoneme sequences.

GENERAL FREE PHONEME-THREEPHONE SEQUENCES GENERATIVE GRAMMAR

This mentioned generative grammar for free phoneme sequences will be given under PT interpretation.

Let be given the finite set K of the phonemes $k \in K$. The phoneme alphabet includes the phoneme-pause #. In K there will be distinguished stressed and non-stressed vowels, hard and soft consonants, stationary phonemes like $k \in$ {A, O, U, E, I, Y [all stressed and non-stressed], V, V' [the symbol ' denotes soft-ness], ZH, Z, Z', J, L, L', M, M', N, N', R, R', F, F', KH, KH', SH, #} $\equiv K^{st} \subset K$, which change their duration, and transitive phonemes $k \in$ {B, B', G, G', D, D', K, K', P, P', T, T'} $\equiv K^{tr} \subset K$.

Then there are considered all possible PT or about 2,000—3,000 basic PTs $t \in T$. Each PT t from the PT alphabet T besides the name t has also the triple name $t=uWv$ where $u,W,v \in K$ and u,v are input and output phoneme names for PT t, respectively. So the PT $t=uWv$ is the phoneme W that is considered under influence of neighbouring phonemes u and v in context, they are the first u which precedes W and the second v which follows W.

From now on we will assume that besides phoneme and PT alphabets there are given such knowledge:

A. A finite set E of elementary speech signal prototypes or typical one-quasiperiodical segments $e(j) \in E$ where $j \in J$ is a $e(j)$ name in the name alphabet J. E.g. there are $|J| = |E| = 2^{16}$ elements in E and J. So the set J makes the microphoneme level of speech patterns and the pair (J, E) is the code book for one-quasiperiods.

B. A finite set T of PT $t \in T$. The PT t is specified by its acoustical transcription in the alphabet J: $t = (j_{t1}, j_{t2}, \ldots, j_{ts}, \ldots, j_{tq(t)})$, where s indicates the ordinal place in the transcription t and $q(t)$ is the transcription duration for t.

C. Distributions $P(x/j)$ of observed elements (quasiperiods) x for all $j \in J$, particularly $P(x/j)=P(x/e(j))$.

The knowledge mentioned in A, B and C are found at training mode [1, 2]. For each speaker they form a so-called Speech Speaker file.

After the preprocessing a speech signal to be recognised is presented by the sequence X_{ol} of observed one-quasiperiodical segments or elements x_i: $X_{ol} = (x_1, x_2, \ldots, x_i, \ldots, x_l)$, where l is the quantity of observed quasiperiods. The segment $X_{mn} = (x_{m+1}, x_{m+2}, \ldots, x_n)$, $0 \le m < n \le l$ is considered as a signal realisation of the PT t with the probability which is calculated as the convolution on microphonemes bounds $\{r_s\}$:

$$P(X_{mn} / t) = \max_{\{r_s\}} \prod_{s=1}^{q(t)} \prod_{i=r_{s-1}+1}^{r_s} P(x_i / j_{ts}), \qquad (1)$$

where $r_0 = m$, $r_{s-1} < r_s$, $r_{q(t)} = n$. The respective stochastic generative automata grammar (graph) for both PT model signals generating and comparison of the signal segment X_{mn} with all generated ones accordingly to (1) is shown in Fig. 2a. That graph has $q(t)$ states. To each state s it is ascribed the microphoneme $j(s) = j_{ts}$ with the distribution $P(x/j_{ts})$. The transitions between states are doing in accordance to arrows and during 0 or 1 discrete time steps. It is forbidden to remove microphonemes here. The grammar shown in Fig. 2b forbids to remove more than two microphonemes running. Schematic notes for PT graph $t=uWv$, $u,W,v \in K$ are given in Fig. 2c, where only the input $s=u$ and the output $s=v$ states are distinguished.

Let us unite all PT graphs into common one. It is permissible to connect PT into phoneme sequences so

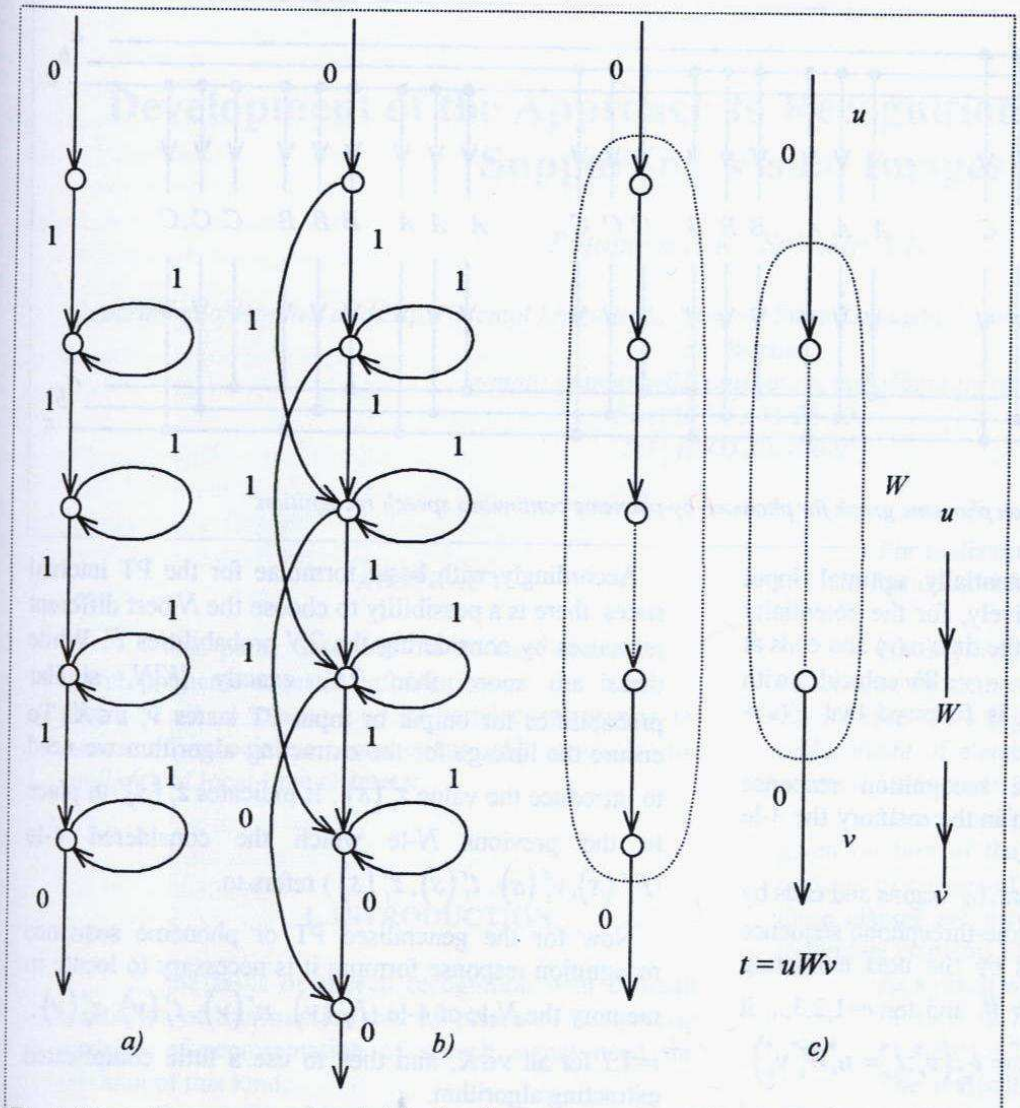


Fig. 2. Generative grammars (graphs) for the phoneme-threephone: a) no microelement omis-sion; b) no two microelements running omission; 3) schematic notes of the PT graph t = uWv.

that the output phoneme name of preceding PT coincides with the input phoneme name of following one. It means that the input and output states for different but permissible for linking PT have to coincide.

Going such a way it will be received the common phoneme graph (CPG) for continuous speech signal generation. This full CPG for three phoneme alphabet $K=\{a, b, c\}$ is shown in Fig. 3. It is distinguished the input state $s=u$, the output one $s=v$ and internal states s for each PT $t=uWv$, $u,W,v\in K$. One of the states $s=k\in K$ in Fig. 3 is associated with the phoneme-pause #. Let us introduce the overall enumeration of states in the CPG accordingly with a permissible movement along the arrows.

Looking into CPG the best phoneme sequence recognition response or, that is the same, the best permissible PT sequence recognition response is defined by maximisation of the expression (2):

$$P\left(X_{0l}/\left(t_1,\ldots,t_s,\ldots,t_Q\right)\right)=\max_{\{r_s\}}\prod_{s=1}^{Q}P\left(X_{r_{s-1}r_s}/t_s\right), \quad (2)$$

where $\{r_s\}$ are the bounds between phonemes-threephones in X_{0l}.

PHONEME SEQUENCE RECOGNITION ALGORITHM

Let be designated by $\Omega_i(s)$ a set of continuous speech prototypes of duration i which are generated by the CPG as a result of movement from state $s=\#$ to state s within i time steps. Let be denoted by $F_i(s)$ the best probability (2) which is reached on the set $\Omega_i(s)$ but for the initial segment $X_{0i}=(x_1, x_2,..., x_i)$, and by $n_i(s)$ the potentially optimal beginning of the last PT $t_i(s)$ in the best PT sequence for $\Omega_i(s)$.

Let $F_r(s)$, $n_r(s)$, $t_r(s)$ have been calculated for all states s and for all time steps $r<i$ which precede i. Then after the next observed element x_i appearance simultaneously (in parallel) for all states s new values $F_i(s)$, $n_i(s)$, $t_i(s)$ are calculated in order a), b), c):

a) for all internal PT states $s\in t=uWv$, besides PT first states, and for all t (see Fig. 2a):

$$F_i(s)=\max\{F_{i-1}(s\text{-}1),\ F_{i-1}(s)\}\cdot P(x_i/j(s)),$$

$$n_i(s)=\begin{cases}n_{i-1}(s-1), & if\ \ F_{i-1}(s-1)\geq F_{i-1}(s);\\ n_{i-1}(s), & if\ \ F_{i-1}(s-1)<F_{i-1}(s);\end{cases}$$

b) for all first internal states $s=s_1(t)\in t=uWv$ and for all $t\in T$ (see Fig. 2a):

$$F_i(s_1(t))=\max\{F_{i-1}(u(t), F_{i-1}(s_1(t))\}\cdot P(x_i/j(s_1(t))),$$

$$n_i\left(s_1(t)\right)=\begin{cases}i-1, & if\ \ F_{i-1}\left(u(t)\right)\geq F_{i-1}\left(s_1(t)\right);\\ n_{i-1}\left(s_1(t)\right), & if\ \ F_{i-1}\left(u(t)\right)<F_{i-1}\left(s_1(t)\right);\end{cases}$$

c) for all common output states v of all PT $t=uWv$ with the same $v(t)=v$:

$$F_i(v)=\max_{t=uWv:v(t)=v}F_i\left(v(t)\right),$$

$$t_i(v)=u_i(v)W_i(v)v=\arg\max_{t=uWv:v(t)=v}F_i\left(v(t)\right),$$

$$n_i(v)=n_i\left(v\in t_i(v)=u_i(t)W_i(t)v\right),$$

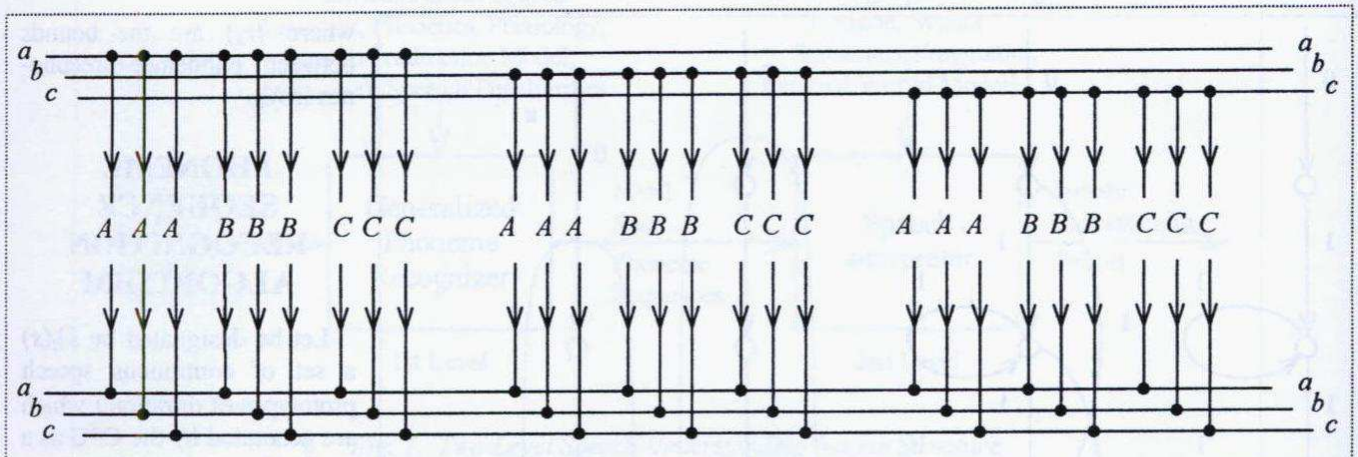


Fig. 3. Common phoneme graph for phoneme-by-phoneme continuous speech recognition.

where $u_i(v)$ and $W_i(v)$ are potentially optimal input phoneme and phoneme, respectively, for the potentially optimal PT $t_i(v)$ which begins at the time $n_i(v)$ and ends at the time i ; since each input state $u \in t=uWv$ coincides with the same output state u then it is followed that $F_i(u)=F_i(v=u)$.

For the phoneme sequence recognition response forming it is sufficient to remain in the memory the 3-le array $F_i(v)$, $n_i(v)$, $t_i(v)$, $i=1:l$, $v \in K$.

Since a continuous speech signal X_{0l} begins and ends by a full PT $t=\#\#\#$ then the phoneme-threephone sequence recognition response is formed by the next extracting algorithm. Let be $n_1^* = l$, $t_1^* = \#$ and for $r=1,2,3,...$ it will be extracted $t_{r+1}^* = t_{n_r^*}\left(v_r^* : t_r^* = u_r^* W_r^* v_r^*\right)$, $n_{r+1}^* = n_{n_r^*}\left(v_r^* : t_r^* = u_r^* W_r^* v_r^*\right)$ until $n_{r+1}^* = 0$ will be reached. Then the PT sequence t_r^*, $r=1,2,3,...$ will be the PT recognition response in the opposite direction and n_r^*, $r=1,2,3,...$ will be the respective PT bounds in the signal X_{0l}. The phoneme sequence recognition response will be W_r^*, $r=1,2,3,...$.

To begin the recognition process it is assigned $F_0(u=\#)=1$ and $F_0(s)=0$ for all other states $s \neq 0$.

THE GENERALISED ALGORITHM

To find $N>>1$ best phoneme or PT sequences in the signal X_{0l} let us modify the basic algorithm.

Now for all states s in the CPG and for any time step i it will be calculated N-le of not 3-le but 4-le $(F_i^r(s), n_i^r(s), t_i^r(s), z_i^r(s))$, $r=1:N$ which is composed of N best probabilities $F_i^r(s)$ that correspond to N best but different PT sequence recognition responses for X_{0i}..

Accordingly with basic formulae for the PT internal states, there is a possibility to choose the N best different responses by considering the $2N$ probabilities F. While there are more than $2N$, exactly $|K|N$, similar probabilities for output or input PT states v, $u \in K$. To ensure the linkage for the extracting algorithm we need to introduce the value $z_i^r(s)$. It indicates $z_i^r(s)$ -th place in the previous N-le which the considered 4-le $(F_i^r(s), v_i^r(s), t_i^r(s), z_i^r(s))$ refers to.

Now for the generalised PT or phoneme sequence recognition response forming it is necessary to locate in memory the N-le of 4-le $(F_i^r(v), n_i^r(v), t_i^r(v), z_i^r(v)$, $i=1:l$ for all $v \in K$, and then to use a little complicated extracting algorithm.

CONCLUSION

There exists such an opinion that it is possible to design a machine for automatic phoneme recognition in continuous speech without any appealing to speech understanding. Here it is proposed one effective robust algorithm for this problem solving which guarantees $N>>1$ best phoneme sequence responses finding.

ЛІТЕРАТУРА

1. Vintsiuk T.K. , *Avtomatyka* 6, 40 - 49 (1972); 1, 63 - 72 (1973).
2. Vintsiuk T.K., *Analysis, Recognition and Understanding of Speech Signals*, Kiev: Naukova dumka, 1987, 264 p.
3. Vintsiuk T.K. , "Generelized Problem for Automatic Phoneme Recognition", *Proceedings of the Workshop SPECOM'97*, Cluj-Napoca, Romania, pp 115 - 118, 1997.

Development of the Approach to Recognition of Speech With a Support on Visual Images

Potapova R.K., Sobakin A.N.

Department of Applied and Experimental Linguistics, Moscow State Linguistic University, 125445, Moscow, Ostozenca, 38, Russia
e-mail: potapova@linguanet.ru; mglu@online.ru
Fax: (095) 246-28-07;
Tel.: (095) 201-56-97

ABSTRACT

This paper describes the concept of receipt of useful phoneme-acoustic information on the bases of nature of visual images (and its parts) configuration by means of associative connections [1] specially for similarity of local-time character.

1. INTRODUCTION

The tasks of speech recognition with difficult conditions of transmission and by means of not precise methods of representation of speech signal need the research of this kind.

Except for a hierarchical way of the account of feature units the parallel analysis of the quantitative characteristics of offered attributes is possible. **The method** realizing parallel way of classification of words, is based on use of a matrix of affinity received for the relation of belonging to classes of words. The given descriptions of two words concern to one class, if taxonomic distance between the descriptions in space is least.

2. METHOD AND EXPERIMENT

With the purpose of demonstration of serviceability of an offered technique we shall consider procedure of classification of the descriptions of ten words (for figures) above mentioned. As the initial data we use values of related features for them [2].

In connection with that a number of features used for the description of the images of words, is given by several values minimal and maximal, for several figures of the image of a word etc.), with the purpose of simplification of the further calculations we shall take advantage only of their average values.

According to the offered attributes we shall receive a matrix of distances (tab. 1).

For realization of procedure of classification we use the feature of a belonging for threshold value equal 10.

Having applied the chosen threshold value concerning all elements of a matrix of distances (see tab. 2), we received a logic matrix of the given relation, by replacement of elements exceeding 10: "zeros", and all others – "ones" (tab. 2).

Having analysed a logic matrix, we see, that the given relation of the belonging are derivated by some variants of classes of equivalence (groups) of words. In these classes are included words, to which correspond individual elements of a matrix of the relation.

As a result we received:
{1.4,8}, {2,7,9,0}, {3,5}, {6} or
{1.4,8,0}, {2.7,9,}, {3}, {5.6}.

The multialternativeness of classification is a consequence of intransity of logic matrix of the given relation. It is easy to be convinced of it, having applied to it criterion of check of property of transitivity. It is necessary to apply to elimination of multialternativeness a method of transitive short circuit consisting in multiplication of a logic matrix on themselves so long as its elements will cease to change.

Division of initial set of the images of words into groups more large, than the separate word, is received by virtue of that circumstance, that we used in example only four features from seven, used at direct classification.

Received preliminary results of use offered in the paper attributes of the approached description of speech images of separate words confirm an opportunity of their division with the help of this system attributes.

At the same time it is necessary to note preliminary character of such conclusion owing to limited volume of sample of images used for experiment both by amount of realizations, and on number of the speakers.

For a substantiation of more universal character offered or not how many modified system of features will be carried out more extensive experiment: the experimental and number of the subjects increased.

© UkrObraz'2000

Besides the preliminary conclusion about an opportunity of development of the automated technique of classification of the images of words and its realization can be made on the basis of a parallel way of classification.

Processing of parametrical representation of words received with use of band-spectrograph can be considered in quality of one of tasks of statistical processing of the graphic information having two levels of "brightness" in each "point" of the image.

Each realization of a word of one speaker in such graphic representation has its own frame, the beginning and the ending of this one is defined accordingly by beginning and ending of pronouncing of the given speech formation.

The beginning of a word in all experiments was determined on occurrence of the first not "zero" in one of frequency channels from F1 up to F2 (in frequency interval it corresponds with frequencies from 100 Hz up to 4000 Hz). Not "zero" values of each of the specified ranges correspond to excess of accumulation for 20 ms of energy of a speech signal of some threshold value.

Thus, size of threshold value, its choice in relation to level of a signal are very important for steady and reliable finding of the left border of the frame. From the further description of a way of statistical processing of frame it becomes clear, as far as this parameter is important for all procedure of statistical recognition in whole.

The right border of frame of the image of a word was determined similar on last value distinct from zero, of a level of a signal, in same frequency ranges also depend on the same threshold value.

Each realization of a word frame was put in conformity the contrast (graphic) image, on the basis of which statistically were created model sequences.

Creation of model sequences: for creation of model sequences each frame with image of a word was exposed to preliminary transformation which consist in the following.

The image was transformed into a sequence of value of brightness (zero and one) by means of transformation of everyone vertical frequency section of a speech signal in a horizontal vector. Such transformation was carried out by "turn" of a vertical vector on 90 degrees clockwise, that transformed a vertical vector in horizontal. Its left values corresponded with the bottom frequency ranges, and right – high frequency ranges. Each subsequent vector incorporated with previous on the right and, thus, long horizontal vector – was formed line (sequence from "zero" and "ones").

After the described preliminary transformation of each frame of the image we have a set of sequences for set of pronouncings of ten words by ten speakers. This set of sequences serves a statistical material for formation of model sequences.

Formation of a model sequence for each word (class) was carried out by calculation for each position of a sequence of probability of occurrence in it of unit on ten speaker realizations. In practice not probability of occurrence of unit, and size appropriate to numerator of this probability was used. In other words, was counted up amount of units in each position of sequences for the various speakers.

Sequence, received on the basis of statistics, was accepted for reference for the given word (class).

Reference sequences for everyone thus were created words (class). In the given experiment number of classes of recognition coincided with amount of words of recognition and was equaled to ten. As much was created of reference sequences.

At a stage of recognition the showed(presented) sample of a word as a sequence of "zero" and "ones" is compared to the standards and the measure of similarity " of a represented sample for each standard is calculated ". The algorithm of calculation " measures of similarity " creates in set of binary sequences metric space being base for each algorithm of races cognition.

In considered space of binary sequences the metrics as the sum of conditional probabilities of occurrence of "one" or "zero" is offered in each position of a sequence.

This sum of conditional probabilities in view of the made remarks is counted up as follows. If in the presented standard in some position there is a "one", in the sum the number of "ones" from reference is added from sequences of a checked class in the same position. If in the presented standard in a considered position there is a "zero", in the sum the difference is added. Last action corresponds to addition in the sum of probability of occurrence of "zero" (additional probability) in the given position.

The complete summation is made on length of the presented word. In the volume a case, when the word comes to an end before the standard (model), procedure of summation comes to an end. If the standard is shorter than a word, it is supplemented in "zero", that corresponds to "zero" probability of occurrence of "ones" in these positions.

The received sum is normalized on the length of presented word, and this sum is accepted for a measure of similarity of the presented word.

The greater value of a measure of similarity corresponds to greater "similarity" of the presented word to the given class. The maximal value of measures of similarity on all classes will correspond hypothetically to belonging of showed word to this class.

3. CONCLUSION

The described algorithm of recognition was examined on a material to ten words realized by ten speakers.

Experimental results of comparison of ten words pronounced by the speaker C., with reference sequences are given in the table 3. It is necessary to note,

that the showed speech samples of the speaker C. not participated in reception of the samples for comparison.

As it is visible from the given table, the greatest value of a measure of similarity are located on the main diagonal of a matrix. It means in the whole serviceability of such approach in recognition of the images of words.

At the same time, there are separate difficult cases of comparison for offered measures of similarity. In particular, the measure of similarity of a word "eight" with various sample sequences has no the brightly expressed maximal value, that can be explained as large variability of separate speaker realizations of the word which has served with a basis for creation of the sample of the appropriate class ("eight").

Variability of speaker realizations in this class is explained, on the one hand, by various degree of a reduction of past-stressed vowel, on the another – by various degree of coarticulation, penetrating all word as a whole.

The possible ways of increase of reliability of algorithm of recognition consist in expansion of statistical base of creation of the samples and improvement of techniques of comparison of speaker realizations with the samples.

REFERENCES

1. R. K. Potapova. "Ob odnom podkhode k klassifikatsiji reche-zritelnych obrazov na baze assoziativnykh svjazej. Mater. VI Vseros. Konf. "Nejrokompjutery i ikh primenenije". M., 2000.
2. R.K. Potapova. "Rech: kommunikatsija, informatsija, kibernetika." Radio i Svjax. M., 1997.

APPENDIX

Table 1. A matrix of distances

Word	1.	2.	3.	4.	5.	6.	7.	8.	9.	0
1.	0	11,25	24,29	1,02	21,4	16,05	14,09	7,28	12,17	6,20
2.		0	21,26	12,03	16,28	10,68	3,61	13,91	9,30	8,21
3.			0	23,85	6,64	11,05	21,01	18,5	13,18	18,00
4.				0	21,77	16,08	14,86	7,24	12,50	6,65
5.					0	6,35	15,34	17,85	9,73	15,24
6.						0	10,45	12,88	4,95	9,46
7.							0	16,43	9,63	10,35
8.								0	10,60	6,44
9.									0	6,06
0										0

Table 2. A logic matrix of distances

Word	I	2	3	4	5	6	7	8	9	0
I	1	0	0	I	0	0	0	1	0	I
2		1	0	0	0	0	I	0	I	I
3			1	0	I	0	0	0	0	0
4				I	0	0	0	I	0	I
5					I	I	0	0	I	0
6						I	0	0	I	I
7							I	I	0	I
8									0	I
9									I	I
0										I

Table 3. Experimental results of comparison of ten words, pronounced by the speaker C., with reference sequences.

Word	"0"	"1"	"2"	"3"	"4"	"5"	"6"	"7"	"8"	"9"
"0"	6,11	4	4,69	3,7	5,37	3,96	4,18	3,35	5,02	4,99
"1"	4,43	6,54	3,25	2,63	5,65	3,12	2,94	2,05	4,31	4,1
"2"	5,28	3,07	6,77	4,55	3,48	5,73	4,84	3,79	4,11	5,08
"3"	2,84	3,3	3,81	7,84	4,08	5,35	5,47	5,93	4,53	5,65
"4"	4,41	5,28	3,69	4,5	7,51	3,77	4,44	3,45	4,74	5,74
"5"	3,98	3,47	4,37	5,34	3,32	7,15	4,32	4,26	3,10	5,01
"6"	4,51	4,57	5,46	6,7	6,4	5,77	7,54	5,52	5,8	6,21
"7"	2,73	2,89	3,41	6,63	3,86	4,65	3,91	7,06	3,48	5,02
"8"	4,98	4,86	4,04	4,76	5,62	4,25	4,47	3,52	6,06	6,05
"9"	4,31	3,31	4,48	5,59	5,96	4,89	5,93	4,73	5,42	7,43

IDENTIFYING TRANSIENT PATTERNS IN IDIOMATIC GREEK AND SLAVONIC MUSIC

*Dionysios Politis** *Alexander Dolia*† *Panagiotis Linardis***

{*Multimedia Lab, **VLSI Lab},
Computer Science Dept., Aristotle University of Thessaloniki, GR-540 06, GREECE.
e-mail: {dpolitis, linardis}@csd.auth.gr
† Dept.504, State Aerospace University,
17 Chkalova St., Kharkov, 61070, Ukraine.
e-mail: lukin@mmds.kharkov.ua

Abstract—**This paper describes a method for determining the content affinity of musical patterns in vocal reproduction in Greek and Slavonic by using wavelet decomposition techniques. The goal of the analysis is to determine the exact musical description of the patterns in order to reproduce them by musical synthesizers acting as synthetic singers. The exemplar pattern of this paper is that of "petasti", which was selected because it is a transient phenomenon particular in traditional Greek songs and chants, in both instrumental and vocal performances. A comparison of the pattern is attempted with Slavonic chants by determining time and frequency localizations of the pattern.**

Index Terms—**Wavelet decomposition, musical patterns, "petasti", voice synthesis.**

I. INTRODUCTION

This paper copes with the identification of musical patterns that literally exist in Greek and Slavonic musical traditions. The origin of these patterns can be found in manuscripts that are at least 10 centuries old. In Fig. 1.(a) a fragment of a manuscript is shown which implicitly denotes these patterns [1]. In Fig. 1.(b) explicit marks of "petasti" are recorded in Byzantine music notation [2].

The creation of a musical notation system proved to be a painful and long procedure in the West and especially in the East, where as the manuscripts that have survived indicate (Fig. 1), we have had different notations evolving one from the other in a shorthand like form. This musical system was not confined only to ecclesiastical music; it was a generalized musical system originating directly from ancient Greek Music and was used as the usual music surface by all the people living in the vast areas of the Byzantine empire, from Southern Italy and the Balkans up to Ukraine and Russia and down to Middle East and Egypt.

Finally the Byzantine notation prevailed in the East as the dominant musical notation and was reformed in 1814 and 1881 from committees of the Ecumenical Patriarchate of Constaninople to an analytical notation

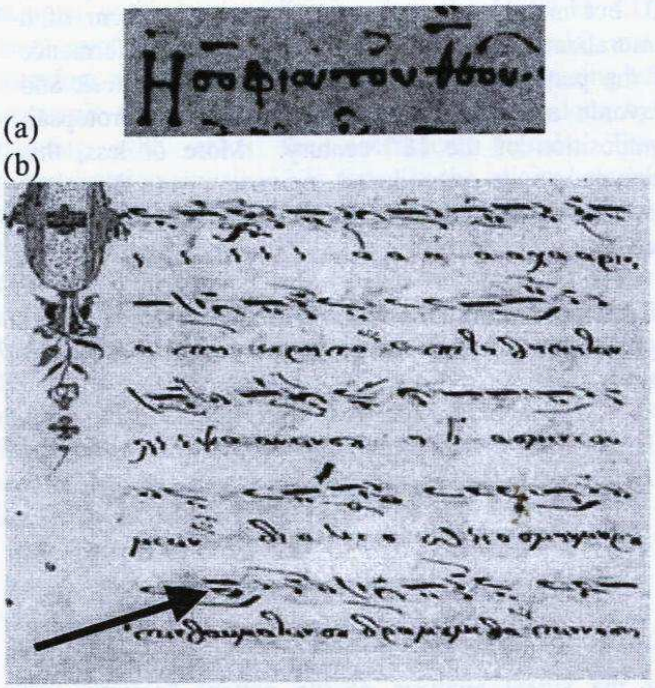


Fig. 1. Byzantine Music Manuscripts: (a) Detail from an early manuscript, the so-called 'Chartres fragment', with musical notation, beginning of the sticheron "Η σοφια του Θεου, Mode plagal D", Copenhagen. (b) "Petasti" pointed out in a Doxastarion composed by Jacov Protopsalt, Mount Athos, 1805 AD.

system whose symbols came out of the numerous symbols of the earlier shorthand-like notations [3].

Major differences between the polyphonic Western music tradition and the Byzantine system are:

- the intervalistic system (system of musical scales and their microtonal distribution) that contains predominantly musical intervals that are smaller than the well tempered ones. Successors of the ancient Greek Dorian, Lydian, Phrygian, Mixolydian Modes and their plagal ones are encountered [3][4]. This characteristic introduces a subjective criterion of orientalization in the psycho-acoustics of the hearings, and
- the use of larger, formalized transitory patterns as main elements of the musical structure that are

© UkrObraz'2000

inherent in vocal performance and render a qualitative nuance.

A musical pattern that has a quantitative and qualitative as well character is that of "petasti". This pattern has been detected in instrumental performance of traditional Greek songs [5] and in Byzantine music vocal performance [6]. According to Byzantine Music theory [2], "petasti" is a transient phenomenon that assigns a phonation quality to the raising or lowering of pitch. In this paper, analysis of this pattern is performed in order to decipher its qualitative nature. By the term qualitative we mean the characteristics that enable the specialized listener to recognize the "petasti" pattern not simply as a pitch fluctuation, a fluctuation of fundamental frequency F0, but mainly as a parametric transitory form of a gutturalization. Analysis is focused on the performance of the pattern in Mode A (Dorian) [4] in Greek and Slavonic according to the classical Jacov Protopsalt composition of the 18th century. More or less, this performance is reproduced nowadays in the slow rendition of the vesper hymn "Lord, I have cried unto thee".

The exact description of this pattern in terms of microtonal distributions will enable the correct musical reproduction by computer programs acting as synthetic singers.

II. EXTRACTION OF AUDIO FEATURES USING THE DISCRETE TIME WAVELET TRANSFORM

For the signal analysis of the pattern "petasti" the Discrete Time Wavelet transform (DTW) will be used in conjunction with the classical Short Time Fourier Transform (STFT). One major advantage afforded by wavelets is the ability to perform local analysis. This is useful near the discontinuity areas of the signal, namely the consonants for vocal signals. A second advantage is that with each stage DWT the signal is seperated into *approximations* and *details* while downsampling is performed. The DWT can be seen as an equivalent of a tree-structured multirate filter bank (Fig. 2) [7].

The mother wavelet used for decomposition was db4, the fourth of the Daubechies family wavelets, and the wavelet decomposition tree is presented in Fig. 2, albeit for convenience only two levels are drawn. D(z) is the outcome of a high pass filter and downsampling, thus a detail component, and A(z) is the outcome of a low pass filter and downsampling, thus an approximation component. This procedure produces at each stage DWT coefficients which are denoted in Fig. 2 as $\mathbf{cA_i}$ for the approximation components and as $\mathbf{cD_i}$ for the detail components for the **i**-th level analysis of signal S.

In order to analyze the transient nature of "petasti", in both Greek and Slavonic reproductions of classical

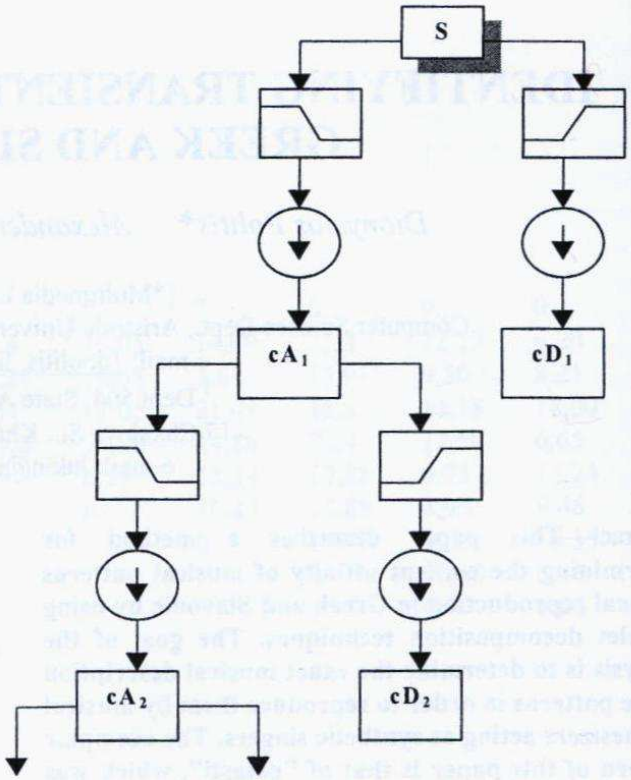


Fig. 2. The 2-level binary tree structured filter bank for signal S.

musical phrases, prototypal performances of these melodies have been recorded.

The waveform of the prototypal performance in Greek is shown in Fig. 3(a) and its spectrogram is shown in Fig. 3(b). In Fig. 3(b) the segment around the first phoneme /i/, the one that is performed with "petasti", fluctuates around note E3 in a manner close to wavelet db4 (Fig. 4). This is the reason for choosing the Daubechies family of wavelets.

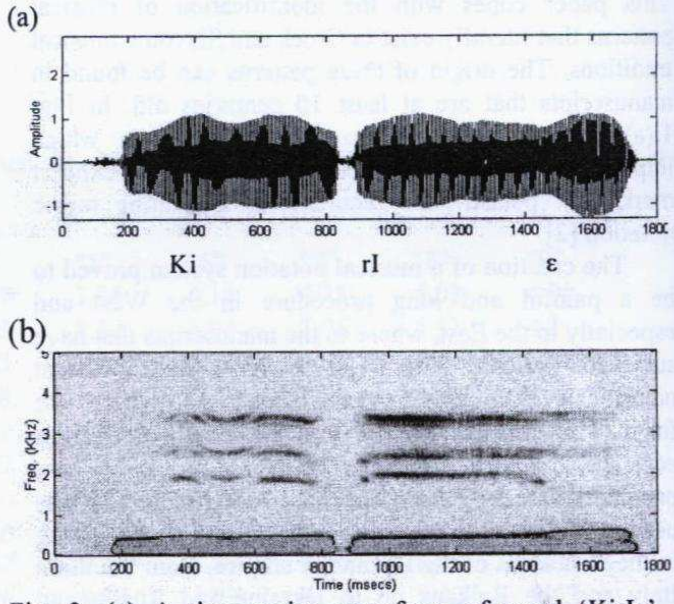


Fig. 3. (a) A time series waveform of word 'KirIε' consisting of three morphemes and a diphone segment. The first phoneme /i/ is performed as a "petasti" pattern. (b) The corresponding spectrogram with emphasized formants F1, F2 and F3.

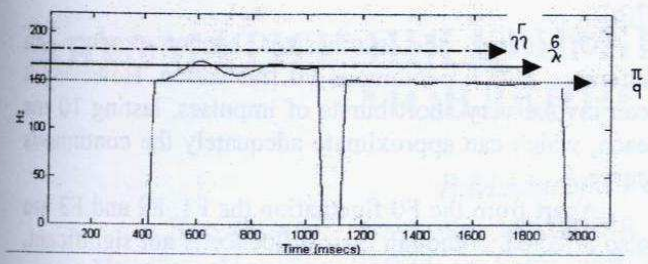


Fig. 4. F0 fluctuations around note Βου, the equivalent of note E3. Note Βου corresponds to 158 Hz, while the pitch levels of notes Πα and Γα, shown with dotted lines, correspond to 144 Hz and 170 Hz respectively. The interval Βου-Γα is slightly bigger than a semitone.

The pattern we examine consorts the first morpheme of the utterance in Fig. 3(a). By using a cepstrum based approach, fundamental frequency F0 is estimated, and the "petasti" fluctuation around note Βου is estimated. Βου is the equivalent of note E3 of the well tempered scale and it is assigned the frequency of 158 Hz [2]. The subject N. Kougias who has performed the musical pattern has achieved a remarkable tuning with the resonant frequencies.

As seen from the formant curves of Fig. 3(b), it is clear that the musical pattern affects not only F0 but also the first three formants, F1, F2 and F3.

The waveform of an equivalent prototypal performance in Slavonic is presented in Fig. 5(a). The word uttered is 'Gospodi' with the first /o/ performed with "petasti". For this utterance, the cepstrum based technique for the determination of F0 yields poor results.

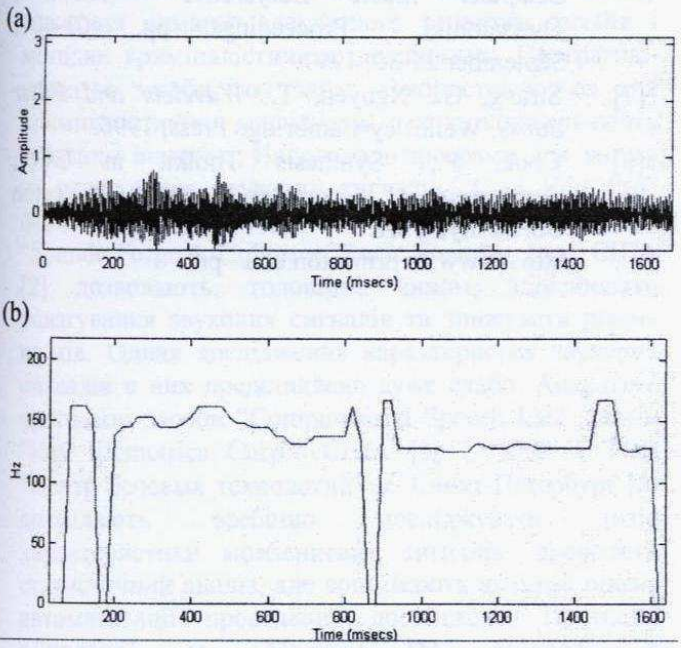


Fig. 5. (a) Time series of morpheme /Go/ from the slow rendition of the word 'Gospodi' having the first /o/ performed with "petasti". (b) F0 cepstrum based estimation of approximation cA3 of the original signal.

In this case, the signal was decomposed by using the fourth wavelet of the Daubechies family and was analyzed using the third-level approximation of it. Prior to estimating F0, upsampling takes place. Then F0 is readily estimated. The final F0 contour is drawn in Fig. 5(b).

Apart from F0 fluctuations, the first three formants of the utterance also have important information. Following the same decomposition scheme, we obtain the detail coefficients of the analyzed signal in order to estimate the first three formants of the utterance.

We focus our analysis on signals cA3 and cD3. cA3 consists of the approximation coefficients downsampled and decomposed from the original signal by the Discrete Wavelet Transform three times. Both signals are thus sampled at 5512.5 Hz (deriving from the 44100 Hz CD level quality recordings of the original signals) and therefore a resampling procedure takes place aiming to upsample them to at least 11025 Hz. This is done by padding with zeros the coefficient signals. The approximation signal cA3 is virtually the original signal that has become less noisy after the successive decompositions, and the high frequency information concerning formants F1, F2 and F3 is filtered out of the signal in the form of the detail coefficients signal cD3.
In Fig. 6 we see an estimation of the formant fluctuation for the utterance with "petasti" of /o/ (Fig. 6(a)). It is evident that in this case the performed analysis is inadequate. Consequently, we dilate the signal estimating its approximation and detail coefficients according to the binary-tree structure of Fig. 2.

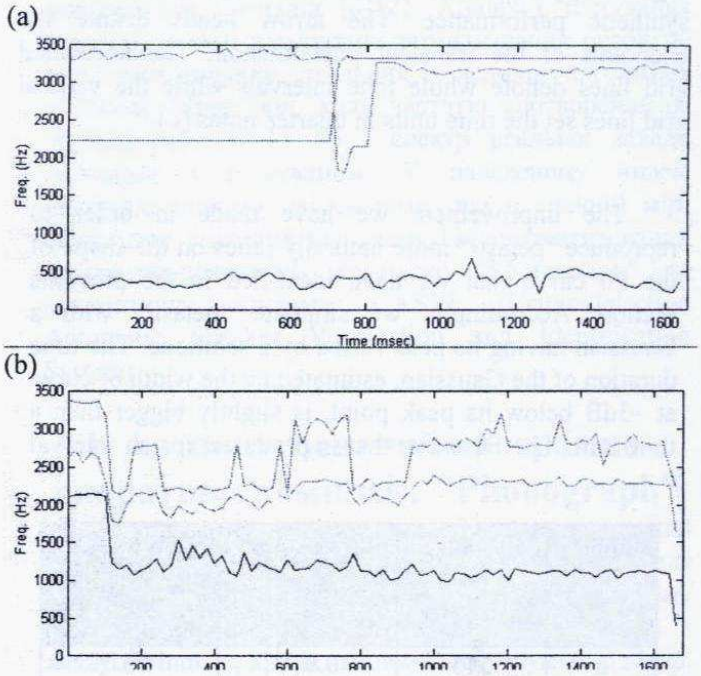


Fig. 6. (a) Formants F1, F2 and F3 estimated from the prototypal signal in Slavonic (b) Formants F1, F2 and F3 estimated from cD3, the detailed third level high frequency component of the signal.

III. CONCLUSIONS: ADVANCES IN SYNTHETIC VOICE REPRODUCTION

The synthetic musical performer we have used to reproduce melodic themes with "petasti" was a formant synthesizer that had physically modeled the glottis and shape files of a Greek singer. This program was compiled using P. Cook's Synthesis Toolkit in C++ [8].

This voice synthesizer, in order to produce a "petasti"-like phonation quality, produced discrete sounds like the ones shown in Fig. 7(a). If we consider that phoneme /o/ has the typical formant distribution described in Table I, then the conventional synthetic reproduction of /o/ with "petasti" would merely yield three intermediate notes, with the second note raised by a semitone. This means that a quarter-note influenced by "petasti" is analyzed as the sequence [sixteenth-note, eight-note sharp, sixteenth-note.]

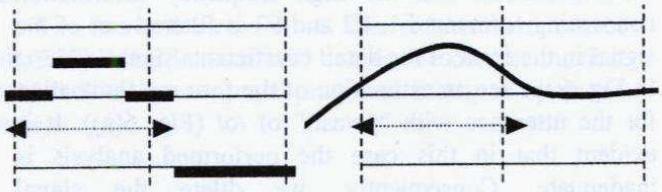


Fig. 7. (a) Quantum leap versus (b) continuous F0 synthetic performance. The arrow heads denote the endpoints of the transient phenomenon. The horizontal grid lines denote whole tone intervals while the vertical grid lines set the time units in quarter-notes (♪).

The improvement we have made in order to reproduce "petasti" more naturally relies on the shape of the F0 curve that we have identified in the previous section. Accordingly, we simulate "petasti" with a Gaussian having its peak raised by a semitone. The time duration of the Gaussian, estimated by the width of curve at –3dB below its peak point, is slightly bigger than a time unit. The F0 curve that is produced for an interval of a major second (i.e. a whole tone) is shown in Fig. 7(b).

Of course, the synthesizer cannot produce an utterance with a continuous F0 fluctuation. However, it can invoke very short bursts of impulses, lasting 10 ms each, which can approximate adequately the continuous curve.

Apart from the F0 fluctuation the F1, F2 and F3 are also affected. Although their influence is not significant, yet they add naturalness to the synthesized phonemes.

Phoneme	formants (Hz)	Freq. Width	Relative amplitudes (dB)
/o/	515	0.977	0
	1805	0.810	-10
	2526	0.875	-10

Table I. Estimated mean values for phoneme /o/. The frequency width of the formants is not denoted in Hz but as it is inserted into the synthesizer, i.e. exp(2W/Fp).

REFERENCES

[1] "Monumenta Musicae Byzantinae", Institute for Greek and Latin, University of Copenhagen. http://www.igl.ku.dk/MMB.

[2] Manuscript from Iveron Monastery of Mount Athos, cod. 437, 1805 AD, reproducing the Doxastarion of Jacov protopsalt (†1800).

[3] Panagiotopoulos, D., *Theory and Praxis of Byzantine Ecclesiastical Music*, 3rd ed, Athens, 1982 (In Greek).

[4] Jacobs, A., *The Penguin Dictionary of Music*, Penguin, USA, 1991.

[5] Pikrakis, A., Theodoridis, S., Kamarotos, D., "Recognition of Isolated Musical Patterns in the context of Greek Traditional Music", *Third IEEE International Conference on Electronics, Circuits and Systems ICECS '96*, Rhodes, pp. 1223-1226, October 13-16, 1996.

[6] Politis, D., Tsoukalas, A., Linardis, P., Bakalakos, A., "VIDI-A Voice Instrument Digital Interface for Byzantine Music", *International Computer Music Conference ICMC97*, Thessaloniki, Proceedings, pp. 403-407, September 25-30, 1997.

[7] Strang, G., Nguyen, T., *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996.

[8] Cook, P., "Synthesis Toolkit in C++", Department of Computer Science, Princeton University, USA. http://www.cs.princeton.edu/~prc

АВТОМАТИЗАЦІЯ КРИМІНАЛІСТИЧНИХ ДОСЛІДЖЕНЬ МОВЛЕННЄВИХ СИГНАЛІВ

В.М.Магера І.І.Горбань, С.В.Лєвий

Державний науково-дослідний експертно-криміналістичний центр МВС України, Україна, 01024, м. Київ, вул. Богомольця, 10, тел./факс.: 291-39-53.

Іститут проблем математичних машин і систем Національної Академії наук України, Україна, 01187, м. Київ, проспект Глушкова, 42, електронна пошта: gorban@immsp.kiev.ua, Тел.: 38(044) 266-61-74, факс:.38(044) 446-8129.

Галузева науково-дослідна лабораторія методів і засобів спеціального призначення радіотехнічного факультету Національного технічного університету України, кафедра радіотехнічних приладів і систем, Україна, 01056, м. Київ, пр-т Перемоги, 37 (2103), електронна пошта: Leviy@lmmfr.kiev.ua, тел./факс.: 274-89-84.

АНОТАЦІЯ

Описано програмно-апаратний комплекс та програму, що дозволяють досліджувати мовленнєві сигнали у автоматичних та автоматизованих режимах. Висока ефективність досліджень досягається на сигналах із рівнем спотворення до 35 дБ та відношенням сигнал-шум – до 10-12 ДБ. Наводиться методика і результати тестування програм.

ВСТУП

Сучасний стан розвитку апаратури запису мовлення вимагає адекватного розвитку засобів і методик криміналістичних досліджень. Програмно-апаратні засоби, які зараз використовуються для криміналістичних досліджень, вже не задовольняють сучасним вимогам. Наприклад, програми для запису та редагування звукових сигналів "Cool Edit Pro" фірми "Syntrillium Software Corporetion" США [1] і "Sounde Forge 4.5" фірми "Sonic Foundry, Inc." США [2] дозволяють, головним чином, здійснювати редагування звукових сигналів та знижувати рівень шумів. Однак дослідження характеристик звукових сигналів в них представлено дуже слабо. Апаратно-програмні засоби "Computerized Speech Lab" фірми "Kay Elemetrics Corp." США [3] і "SIS" – ТОВ "Центр речевых технологий" м. Санкт-Петербург [4] дозволяють всебічно досліджувати різні характеристики мовленнєвих сигналів, проводити статистичний аналіз, але вони мають низький рівень автоматизації проведення досліджень. Програма "Диалект" м. Москва [5] дозволяє в автоматизованому режимі проводити ідентифікаційний аналіз мовленнєвих сигналів. Але великий обсяг підготовчих операцій і оцінка деяких проміжних результатів обчислень покладається на оператора.

Зазначені програмні засоби та деякі інші демонструють гарні результати лише при роботі в умовах малих частотних спотворень і низького рівня перешкод. Коли деструктивні фактори значні, ці системи не ефективні.

Намагання створити нове покоління систем дослідження мовлення істотно більш ефективних, ніж існуючі, привело до появи нових ідей, методів і алгоритмів обробки, орієнтованих на підвищення стійкості роботи в умовах спотворень і зашумлення мовленнєвих сигналів [6-12]. Аналіз і тестування декотрих нових алгоритмів розпізнавання особи по мовленню виявили загальний їх недолік: алгоритми достатньо ефективні, коли частотні спотворення не перевищують 10-15 дБ і спектр реальної завади співпадає з очікуваним. У наведеному нижче матеріалі описані дві системи, що в значній мірі позбавлені вказаних недоліків. Це автоматизований програмно-апаратний комплекс "Phonograph" та автоматична система CASVI (Crime-detection Automatic Speaker Verification and Identification System).

1. Автоматизований програмно-апаратний комплекс "Phonograph"

Автоматизований програмно-апаратний комплекс "Phonograph" призначений для криміналістичних досліджень в мовленнєвому діапазоні частот: голосу, сигналів акустичної і неакустичної природи виникнення, шумів і завад. Він входить до складу автоматизованої криміналістичної системи "Логос" для дослідження матеріалів та засобів звуко- і відеозапису [13].

Комплекс складається із персонального комп`ютера типу Pentium II, звукової карти ExpertColor MED 3201 InterWave AMD

/AM78C210KC/, узгоджувального пристрою "Sound card satellite" (SCS), комутатора зовнішніх пристроїв КВУ-54, принтера, головних телефонів "AKG K141", магнітофонів (котушкового "Электроніка-004К-Стерео", мікрокасетного "Panasonic RR-930" і компакт-касетного "TEAC V-2030S"), відеомагнітофона "Panasonic NV-F55", підсилювача "Kenwood", акустичної системи "Амфітон 100АС-022" і програмного забезпечення.

Структурна схема автоматизованого програмно-апаратного комплексу "Phonograph" наведена на рис.1.

Відтворення і запис звукових файлів в(із) пам`яті комп`ютера проводиться з використанням спеціально розробленого пристрою SCS, який узгоджує по рівню вихідний канал магнітофону і лінійний вхід звукової плати. Пристрій дозволяє прослуховувати лівий, правий стереоканали і здійснює їх змішування. Конструктивно він розміщується в великому (5,25") зовнішньому відсіку, який виходить на лицьову сторону персонального комп`ютеру, що забезпечує вільний доступ до всіх органів регулювання.

Комутатор зовнішніх пристроїв здійснює: комутацію 16 звукових лінійних стерео входів з 16 звуковими стерео виходами; комутацію джерела звукового сигналу для прослуховування на головні телефони через вбудований стерео підсилювач низьких частот в дихотонічному режимі (комутація лівого телефону навушника незалежна від комутації правого); перетворення стерео виходів в моно виходи; комутацію одного із 4-х відео входів з одним відео виходом; включення/виключення мереженої напруги 220 В, 50 Гц 16-ти зовнішнім пристроям. Управління всіма режимами роботи комутатору проводиться від персонального комп`ютеру. Створення комутатору дозволила значно підвищити рівень автоматизації і зв`язати всі периферійні пристрої в функціонально єдиний комплекс.

Програмне забезпечення "Phonograph" дозволяє здійснювати: ввід-вивід аналогового сигналу в(із) ПЕОМ; функції цифрового магнітофону; фільтрацію і компандування фонограм в реальному часі; виміри часових, амплітудних, спектральних і кепстральних характеристик мовленнєвих і інших сигналів; виконання статистичного аналізу характеристик сигналів; автоматизований ідентифікаційний аналіз і отримання оцінки співставлення голосів різних дикторів; встановлення номеру копії фонограми по частоті мережі 50 Гц; автоматичне встановлення телефонного номеру по тональним посиланням АТС, які зафіксовані на фонограмі та інше.

Комплекс забезпечує виконання досліджень у відповідності до створених та затверджених методик.

Ідентифікаційні дослідження мовлення представлені методикою аналізу акустичних характеристик голосу і лінгвістичною методикою аналізу мовлення. Дослідження акустичних характеристик голосу виконується за традиційним (дослідження окремих часових, амплітудних, спектральних, кепстральних та інших характеристик мовленнєвого сигналу) і автоматизованим аналізами. Кількість ознак, що досліджуються, при традиційному аналізі становить 27, при автоматизованому – більше 100. Автоматизований аналіз виконується на довільному тексті та на текстозалежному матеріалі. Лінгвістичний аналіз виконується за просодичними та фонетичними характеристиками мовлення. Кількість ознак, що досліджуються, при лінгвістичному аналізі - 32.

У комплексі "Phonograph" автоматизовано технологічні операції визначення характеристик мовлення на придатність для проведення ідентифікаційних досліджень, відбір для аналізу подібних голосних звуків, розмітку мовленнєвого сигналу, розрахунок характеристик по окремим ознакам мовленнєвого сигналу, формування рішення і інше.

Комплекс розрахований на кваліфікованих операторів, які працюють у сфері дослідження матеріалів та засобів звуко- і відеозапису.

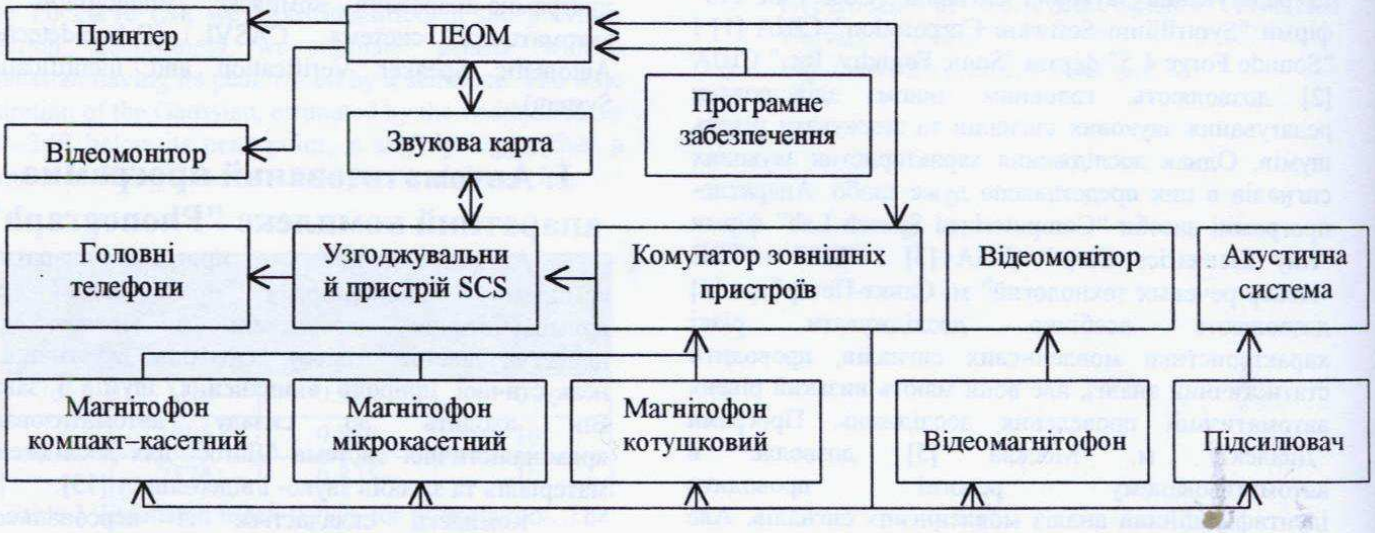


Рис.1. Структурна схема автоматизованого програмно-апаратного комплексу "Phonograph"

2. Система CASVI

Система CASVI забезпечує ідентифікацію і верифікацію особи за голосом у автоматичному режимі. Загальний опис цієї програми і її можливостей наведено в роботах [14-17]. Система ефективного вирішує задачі, що на неї покладаються, при наявності великих частотних спотворень і зашумленні мовленнєвих сигналів. Головним її елементом є підсистема ідентифікації (рис. 2). Вона складається із системи функціональних програм (СФП), системи управління (СУ) і бази даних (БД).

СФП забезпечує обробку і порівняння сигналів і включає програму розрахунку спектрів (РС), програму розрахунку інформаційних ознак (РІО) і програму порівняння інформаційних ознак (ПІО).

СУ виконує задачі взаємодії з користувачем, управління процесами обробки і порівняння фонограм, формування і поповнення бази даних. Вона містить програму, яка забезпечує інтерфейс для користувача (ІК), програму управління функціональними програмами (УФП) і програму управління базою даних (УБД).

Час, необхідний для прийняття рішення підпрограмою, визначається режимами роботи і можливостями комп`ютеру і може становити від декілька хвилин в простих випадках до десятків хвилин.

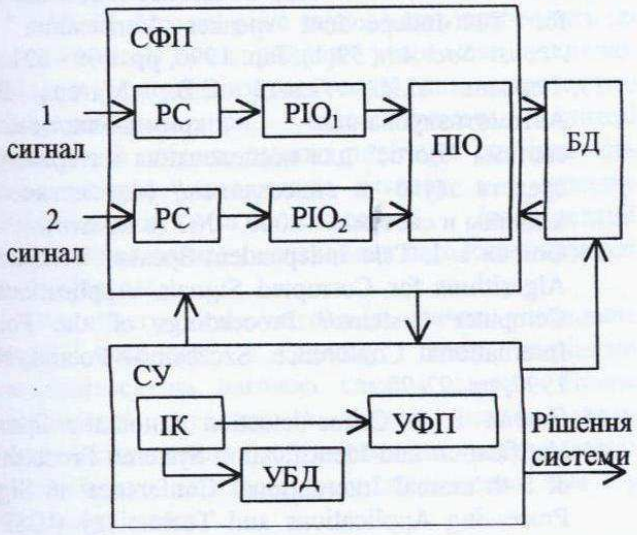


Рис. 2. Структура підпрограми ідентифікації CASVI

Функціональні програми реалізують кепстральні алгоритми.

РС проводить розділення запису на окремі короткі фрагменти (фрейми), для кожного фрейму розраховує спектр, статистичну характеристику корисного сигналу і завади та сортує фрейми на сигнальні і завадові. РІО реалізує формування сигнальних блоків фреймів і розрахунок інформаційних ознак: центрування кепстрів, а також перші і другі похідні кепстрів.

ПІО проводить розрахунок відстані між записами по кожній інформаційній ознаці окремо, розрахунок оцінок щільностей імовірності для однакових і різних голосів із БД, формування підсумкової відстані між голосами, формування порогу з використанням отриманих оцінок щільностей імовірності, порівняння підсумкової відстані з порогом, формування рішення про ідентичність голосів і розрахунок оцінки імовірності помилкового рішення.

В СУ головними є програми, які забезпечують ІК і УФП. Інтерфейс користувача СУ складається із вікон головного меню, поточного результату та бази даних.

Головне меню дозволяє вибрати файли, оцінити рівень шумів фонограм, прослухати їх і виконати ідентифікацію по одному із описаних вище варіантів.

Вікно поточного результату надає результат ідентифікації у вигляді рішення системи Yes/No (да/нет) і розрахункову імовірність помилки прийнятого рішення. Воно інформує користувача про рівень шумів фонограм.

Вікно бази даних дозволяє відібрати файли для БД, прослухати їх і запустити розрахунок *спектрів і кепстральних ознак, які використовуються* при ідентифікації.

Підсистема ідентифікації CASVI розрахована на роботу з тексто-незалежним матеріалом. Але, враховуючи низьку якість запису реальних фонограм, які надходять на експертизу, в окремих випадках доцільно використовувати для порівняння заздалегідь підготовлений матеріал. У цьому випадку з допомогою комплексу "Phonograph" у автоматизованому режимі може проводитись відбір подібних наголошених голосних, що стоять між приголосними звуками. Із них формуються масиви для порівняння з допомогою підсистеми ідентифікації CASVI.

Програма орієнтована на операторів з нижчим ступенем кваліфікації, ніж для роботи з комплексом "Phonograph", так як більшість технологічних операцій повністю автоматизована. Це зменшує витрати на спеціальну підготовку експертів.

3. Тестування працездатності комплексу "Phonograph" та системи CASVI

Тестування програмно-апаратного комплексу "Phonograph" здійснювалось згідно з програмою, затвердженою Секцією експертизи матеріалів та засобів звуко- і відеозапису Науково-методичної ради Міністерства юстиції України. Для кожного окремого режиму чи розрахунку були сформовані спеціальні сигнали та вхідні дані, на яких перевірялась працездатність комплексу. Тестування показало відповідність усіх параметрі розрахунковим і дієздатність всієї системи в цілому. Комплекс "Phonograph" був розглянутий на Науково-координаційній раді МВС України та рекомендований до впровадження у експертну практику.

Алгоритми ідентифікації CASVI тестувались на 32 п`ятихвилинних записах чоловічих голосів. Відношення сигнал-завада для всіх записів складало

не менше 30 дБ. Записи були розділені на дві однакові групи, до яких входили по два комплекти записів восьми дикторів. Записи першої групи використовувались як зразки, а другої - як робочі. Ефективність роботи досліджувалась при сильних частотних викривленнях сигналу і зашумленні, типових при передачі повідомлення по телефонному каналу. Для цього перший комплект робочих записів викривлявся по частоті і зашумлявся на кожній частоті до фіксованого рівня.

Експерименти проводились в два етапи: навчання і тестування. На етапі навчання формувались дані для БД з використанням зразкових записів. Ефективність алгоритмів перевірялась на етапі тестування з використанням робочих записів. Число спроб в разі однакових дикторів було 72, у випадку різних - 504. Тривалість записів дорівнювалась 12 секундам.

Результати експериментів показали, що в умовах частотних викривлень до 35 дБ і зашумленні до відношення сигнал-завада 12 дБ імовірність правильної ідентифікації становить 90% при імовірності помилкової тривоги 10%.

Висновки

Результати тестування програмно-апаратного комплексу "Phonograph" та системи CASVI свідчать про їх ефективну роботу при дослідженні мовленнєвих сигналів.

Програмно-апаратний комплекс "Phonograph" є більш дослідницькою програмою і дозволяє виконувати всі види акустичних та просодичні і фонетичні лінгвістичні дослідження. Значна частина технологічних операцій та режимів автоматизована, що значно скорочує час виконання експертиз.

Практично всі функції підпрограми ідентифікації CASVI працюють у автоматичному режимі при значних частотних викривленнях (до 35 дБ) і відношенню сигнал-шум 12 дБ. Наведені результати суттєво перевищують ті, що до цього часу було досягнуто. Значне покращання технічних параметрів стало можливим завдяки комплексному підходу до проблеми і використанню нових робастних алгоритмів обробки, оптимізованих для роботи в несприятливих умовах.

Комплекс "Phonograph" та система CASVI за результатами порівняння записів голосів наводять не тільки рішення про їх ідентичність, але і оцінку імовірності помилки прийнятого рішення.

ЛІТЕРАТУРА

1. http://www.syntrillium.com
2. http://www.sfoundry.com/index.html
3. Новосельский А.Ф. Компьютерная система для всесторонней работы с речью CLS MODEL 4300B// Компьютеры+Программы. – 1995. – №7(22). – С.66-71.
4. Рекомендации по эффективному использованию возможностей АРМЭФ SIS при выполнении криминалистических экспертиз. – Центр речевых технологий. – Санкт-Петербург, 1995.
5. Идентификация лиц по фонограммам русской речи на автоматизированной системе "Диалект" / Попов Н.Ф., Линьков А.Н., Кураченкова Н.В., Байчаров Н.В. / Под ред. Фесенко А.В. – М.: В/ч 34435, 1996. – 102 с.
6. Hansen J. H. L, Mammone R. J., Young S. editors// IEEE Transactions on Speech and audio Processing. - 1994. – October.
7. Mansour D. and Juang B. H. The short-time modified coherence representation and noisy speech recognition// IEEE Trans. Acoust., Speech, Signal Processing. –1989. –37. – June. – P. 796-804.
8. Newneyer L. and Weintraub M. Probabilistic Optimum Filtering for Robust Speech Recognition// Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing. – 1994. –1. – P. 417-420.
9. Reynolds D. A. and Rose R. C. Robust Text-independent Speaker Identification Using Gaussian Mixture Speaker Models// IEEE Trans. Speech, Audio Processing. –1995. – 3. – January. – P. 72-83.
10. Nolazco Flores J. A., Young S. J. Continuous Speech Recognition in Noise Using Spectral Subtraction and HMM Adaptation// Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing. – 1994. –1. – P. 409-412.
11. J. Mammone, X. Zhang, R. P. Ramachandran, "Robust Speaker Recognition," *IEEE Signal Processing Magazine,* Sept. 1996, pp. 58 - 71.
12. M. A. Lund, C. C. Lee, "A Robust Sequential Test for Text-Independent speaker Verification," *J. Acoust. Soc. Am.* 99(1), Jan. 1996, pp. 609 - 621.
13. Горбань И.И., Левый С.В., Магера В.Н. Автоматизированная криминалистическая система "Логос" для исследования материалов и средств звуко- и видеозаписи// Математические машины и системы. –2000. –№3 (в печати).
14. Gorban I. I. Text Independent Speaker Verification Algorithms for Corrupted Signals. Applications of Computer Systems// Proceedings of the Fourth International Conference. Szczecin – Poland, Nov, 1997, pp. 92-98.
15. Gorban 1. 1. Crime-detection Automatic Speaker Verification and Identification System// Proceedings of 9-th annual International Conference on Signal Processing Applications and Technology (ICSPAT 98). - 1998. - August.
16. Горбань И.И., Горбань Н.И., Клименко А.В., Хазанович М.С. Подсистема верификации новой криминалистической системы автоматической верификации и идентификации личности по голосу (CASVI)// Математические машины и системы. – 1997. –№2. –С.61-64.
17. Горбань И.И., Клименко А.В. Робастні алгоритми верифікації особи за голосом, що призначені для роботи в умовах сильних завад та спотворень мовних повідомлень// 4 Всеукраїнська міжнародна конференція по обробці та розпізнаванню образів "Укробраз 98". – 1998.

ГЕНЕРУВАННЯ ПРАВИЛ РОЗСТАВЛЯННЯ НАГОЛОСІВ У БАГАТОМОВНОМУ АСПЕКТІ

Микола Сажок

Міжнародний науково-навчальний центр інформаційних технологій та систем
40 просп. Академіка Глушкова, Київ 252022
Електронна пошта: mykola@uasoiro.freenet.kiev.ua

Mykola Sazhok. **Generating Stress Rules in Multilingual Aspect.** The risen problem of stresses, particularly in Ukrainian, often confuses text-to-speech and recognition engines. Given analysis of stress in Ukrainian generalizes knowledge about this important aspect. Cases when stress moving in a word changes its meaning and/or its role in sentence are considered. Revealed properties allow for building heuristics and statistics based rules predicting a stress position in most cases and resolving several types of ambigual stresses in words. Within entire paper cross-lingual stress aspects are discussed on example of Ukrainian and English. It is shown how parts of word influent on stress position and how narrow context helps to resolve stress ambiguities.

1 Вступ

Роль наголосів в усній мові важко переоцінити. Властивість наголосу коригувати та й зовсім змінювати зміст промовленого широко використовується у спілкуванні людей.

Хибний наголос ріже слух і часто призводить до непорозуміння при сприйнятті усної мови людиною. Тому в усномовних системах для практичного використання при автоматичному синтезі правильний наголос значно поліпшує розбірливість синтезованого усномовного сигналу, а у випадку розпізнавання та смислової інтерпретації усномовного сигналу нехтування наголосами взагалі неприпустиме.

Розмаїтість типів наголосів значно ускладнює аналіз наголосу в цілому, тому вирішено зосередитися на наголосі саме у словах. Попри особливості наголосу в різних мовах, можна виокремити чимало спільних рис, притаманних розставленню наголосів як у споріднених, так і у віддалених мовах.

Використання знань про наголос в системах автоматичного синтезу, розпізнавання та інтерпретації усномовного сигналу завжди було на часі. Проте, для мов, де наголос є надто нерегулярним і не позначається на письмі, зокрема для української та англійської, дослідження з цього питання потребують більше передумов та ретельності. Особливо це стосується для мов, що допускають велику кількість словоформ, до яких відноситься українська.

2 Загальний аналіз природи наголосу

Що ж таке по суті є наголос? Чіткого визначення цієї просодичної характеристики напевно не існує. Очевидно, що це явище є чимось принципово важливим в усній мові. Також помічено, що під час наголошення вимовляння найбільш відповідає фонетичним нормам та є найближчим до написання. Тобто в наголошеній позиції фонема-алофон звучить (реалізується) найбільш чітко, найменшою мірою піддаючись спотворенням та редукції.

В загальному випадку, наголос може бути сильнішим та слабшим, виокремлюючись на тлі як цілої фрази, речення, так і лише окремого слова або сусідніх складів. Отже, наголоси формуються на трьох рівнях: фрази, слова або синтагми та на субсловарному мікрорівні або на рівні складів.

Зосереджуючись на рівні не вище слова, зазначимо, що тут наголоси бувають основні та другорядні, які балансують на межі мікрорівня.

Позиція основного наголосу у слові рахується мовознавцями в залежності від мови частіше з кінця слова, але буває і з початку. Тут у деяких мов спостерігається регулярність. Хоча при вимовлянні фрази в цілому слово може втратити наголос, набуваючи другорядного наголосу зовсім не там, де розташовується основний. Типовою ілюстрацією такої властивості є приклад з французької "m*ou*len r**ou**ge", де жирні літери позначають основний наголос, а похилі літери – другорядний.

Простежується також залежність наголосу від типу мовлення. В першу чергу ритм, потім темп, розмовність впливають на формування наголосів, їх плавання, перехід у слабкіший стан. Вплив ритму на наголоси чудово простежується на зразках поезії, пісенної творчості, коли диктат ритму інколи зовсім не зважає на граматично правильний наголос.

Прикладом цього є уривок з української народної пісні: "Ой у лісі калина, калина, калина", у той час, як словник подає "калина"

В українській мові основний наголос може перебувати на останньому або другому з кінця складі. Досить поширений також дактилічний наголос, тобто третій з кінця. Але цим не обмежується і в деяких словоформах позиція наголосу може перебувати аж на сьомому з кінця складі як-от у слові "спізнюватимемося".

Англійська мова також має нерегулярний наголос. Відлік позиції наголосу ведеться з початку слів, як для більшості германських мов, і припадає основний наголос переважно на перший склад. Помічено також тенденцію англійською мови до перероблення наголосів запозичених слів, наприклад французьких, на свій лад.

Наголос тісно пов'язаний з іншими просодичними характеристиками, особливо інтона-

ційними. Простежується чітка залежність різких змін інтонаційного контуру саме одночасно з початком наголошеного складу. Безумовно, кожна усна мова своєрідно реалізує наголос, проте крос-лінгвістичні порівняння в багатьох випадках дозволяють адаптувати знання про наголос з однієї мову на іншу.

3 Автоматичне розставлення наголосів

Побудувати алгоритм розставлення наголосів для мов з нерегулярним наголосом у словах і просто і складно. Здавалося б, що може бути простішого, ніж знайти слово з тексту в словнику, де завчасно проставлені наголоси та поставити наголос у шуканому слові. Проте такий шлях вирішення не може задовольнити з багатьох причин.

Для мов з розмаїтістю словоформ, якою є українська, де наголоси від форми до форми часто змінюються, саме створення словників, які враховують позицію наголосу є проблематичним, не кажучи вже про обсяги пам'яті, потрібні для такого словника, та час доступу до певного слова.

Потреба у використанні величезних обсягів пам'яті обмежує використання усномовних технологій як на сучасних комп'ютерах, так і в портативних виробах особливо, коли обмеження на обсяги пам'яті та потужності процесора суттєві. Вихід з цього закуту вбачається у винайденні правил, які б дозволяли розставляти наголоси в довільному слові для певної мови.

Спроби формування таких правил вперше були зроблені для англійської мови. Як було зазначено, велика кількість англійських слів містить основний наголос на першому складі. Водночас були виявлені суфікси, що "перетягають" наголос. Це "-ion", "-ize", "-ate" та деякі інші. Послабленими в багатьох випадках залишаються також префікси. Так на прикладі слів "expect" і "expectation" видно, що префікс "-ex" залишається ненаголошеним, а додавання суфіксів "-ate", "-ion" веде до зсуву наголосу в напрямку до суфікса.

В українській мові на роль найбільш поширеного наголосу претендують як перший, так і другий і третій від останнього. Остаточно визначити це можливо лише проаналізувавши ужиток слів, керуючись текстовим корпусом, а також з міркувань більш економного задання правил. Також виокремлюються нечисленні суфікси, за якими можна визначити наголос ("-еньк", "-есенък"). Крім того, спостерігається чітка тенденція до послаблення закінчень, за винятком прикметників.

Складені та довгі слова видаються найбільш проблематичними для розставлення наголосу, особливо в англійській мові.

4 Зсув наголосу при творенні слів і форм

Зміна наголосу широко використовується при словотворенні. Так англійські "record" і "record" ілюструють перехід іменника у дієслово, під час якого відбувається зсув наголосу.

Українська мова більш широко використовує зміни позиції наголосу і не лише при творенні нових слів, але і при творенні форм слова. Так в іменниках наголос може зсунутись при переході у множину (рука-руки, поле-поля), рідше при звичайному відмінюванні під впливом певних суфіксів.

Дієслова досить вільно поводяться з наголосами, хоча простежуються деякі закономірності з урахуванням сукупного впливу на наголос усіх частин слова та кількості складів у слові.

Прикметники найбільш консервативно зберігають наголошену голосну, при тому, що утворені від них прислівники уникають наголосів на останньому складі (значний-значно), інколи – на передостанньому (веселий-весело).

5 Проблема неоднозначності наголосів

Зміна наголосу в слові призводить, як було сказано, найчастіше до помилки у вимові. Випадки, коли зміна наголосу веде до переведення слова в іншу словоформу, або змінює його значення на інше, часом безпосередньо не пов'язане з початковим семантично, призводять до помилки у смисловій інтерпретації як окремого слова, так і цілого вислову.

У випадках, коли наголос змінює форму слова слід витримати непорушність узгодження слів у реченні. Для цього достатньо визначити керування в якому перебуває це слово за допомогою слів-орієнтирів з найближчого контексту. Такими орієнтирами є прийменники або їх відсутність, прикметники, як найбільш консервативні у наголошенні, та слова з інших частин мови, які вимагають певного узгодження, і чиє значення не викликає сумніву.

Так, у словосполученні "знімати з руки" прийменник "з" вимагає після себе родовий або твірний відмінок, тому й наголос має відповідати одному з цих відмінків. У вислові "потисли один одному руки" слово "руки" керується дієсловом "потисли", яке вимагає знахідного відмінку.

У випадках, коли можливі варіянти наголосів, що відповідають різним частинам мови, постає проблема визначення частини мови цього слова в контексті. Варіянти наголосів, які не змінюють частину мови, змінюючи при цьому значення слова (плачу, замок), слід розглядати окремо.

6 Підсумок

Проведений аналіз дозволяє виробити евристичні та статистичні підходи до формування правил розставлення наголосів не лише в контексті слова, а і словосполучення та цілої фрази, розв'язуючи у багатьох випадках проблему неоднозначності інтерпретації слів при усномовному синтезі та розпізнаванні. Для багатомовних систем важливим напрямком подальших досліджень вбачається використання знань про наголоси зі споріднених мов.

СТАТИСТИЧНИЙ АНАЛІЗ РИТМІЧНИХ СИГНАЛІВ ДЛЯ ДІАГНОСТИКИ ТЕХНІЧНОГО СТАНУ РЕЙОК

Ігор Ісаєв, Георгій Трохим, Ігор Яворський

Фізико-механічний інститут ім. Г. В. Карпенка Нан України
79601, Україна, Львів, МСП вул. Наукова, 5
телефон: 8 (0322) 63-33-55
факс: 8 (0322) 64-94-27
електронна пошта: isayev@ipm.lviv.ua, george@ipm.lviv.ua, iavor@ipm.lviv.ua

Представлено новий підхід до проблеми діагностики технічного стану рейок, що базується на дослідженні сигналу від електромагнітного збурення рейок під час проходження вагона-дефектоскопа методами нестаціонарної статистичної обробки. Основуючись на моделі сигналу у вигляді періодично корельованого випадкового процесу проведено обробку корисного сигналу від ідеальної бездефектної рейки та рейки з дефектом. Результати обробки порівняно з результатами стаціонарної обробки.

ВСТУП

При аналізі технічного стану залізничної колії основну увагу приділяють стану рейок, від якого залежить допустима швидкість та безпека руху потягів. Технічний стан рейок оцінюють за наявністю в них дефектів, тобто структурних неоднорідностей, що не передбачені конструкторсько-технологічною документацією і незалежні по впливу на рейку від інших неоднорідностей [1].

Основними постачальниками засобів дефектоскопії рейок залізницям Північної Америки є компанії Herzog Services, Pandrol Jackson і Sperry Rail Service, Європи - Speno International. Вони постійно вдосконалюють вимірювальні системи, обладнують дефектоскопи комп'ютерами, що підвищує їх продуктивність та дозволяє обробляти більший об'єм інформації. Програмні засоби можуть визначати рисунок дефекта і за ним проводити класифікацію.

Європейська компанія Speno International в своїх останніх розробках використовує комп'ютерні технології для прискорения процесу дефектоскопії, приділяючи особливу увагу покращенню давачів та підвищенню ефективності інтерфейсу з оператором. Одним з напрямів комп'ютеризації є створення бази даних, що дозволятиме визначати тенденції виникнення, розвитку дефектів та прогнозувати зміну стану рейок для оптимального планування їхньої заміни. На Львівській залізниці, для порівняння, комп'ютеризація звелась до заміни фотозапису діагностичних сигналів в магнітному вагоні-дефектоскопі розробки початку 50-тих років на комп'ютерне протоколювання та архівацію даних з низкою сервісних можливостей відображення на екрані монітора та видруку на прінтер. Весь тягар виявлення та класифікації дефекту лежить і надалі на операторі [2].

1. Принцип роботи магнітного дефектоскопа

1.1. На рис.1. висвітлено принцип роботи магнітного дефектоскопа, а на рис.2. приведено ідеалізоване представлення розподілу магнітного

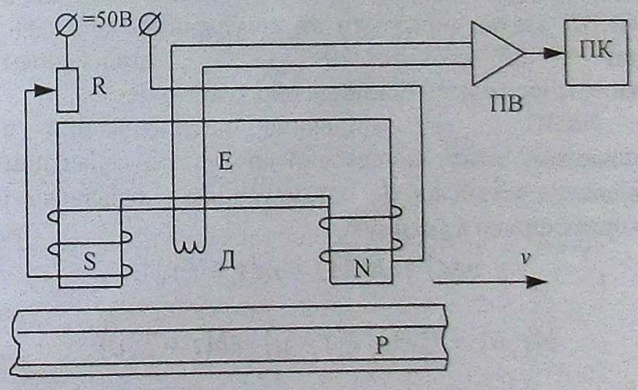


Рис.1. Блок -схема вагона дефектоскопа: *ЕМ - електромагніт з живленням від джерела постійного струму напругою 50 В; R - реостат регулювання струму підмагнічування; Д - давач, утворений індукційною котушкою, розміщеною на немагнітній лижі, що притискається до поверхні котіння та переміщується вздовж рейки Р з швидкістю v; ПВ-пристрій вводу інформації в персональний комп'ютер ПК*

потоку та напруги в котушці при поздовжньому скануванні рейки. З рисунка видно, що дослідження вихідної напруги котушки давача при відомій

© UkrObraz'2000

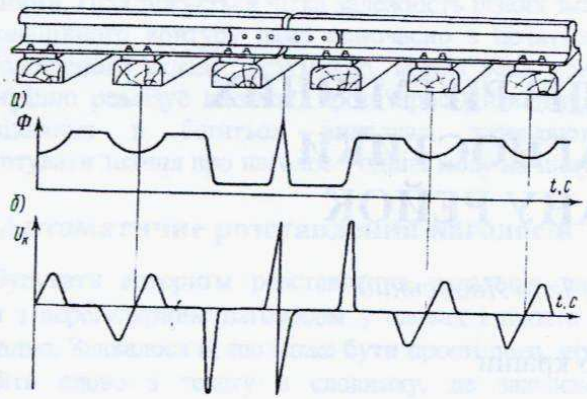


Рис.2. Розподіл магнітного потоку вздовж рейки (*а*) та напруга на виході індукційної котушки (*б*).

швидкості сканування (руху) може виявити структурні періодичності (характерні відхилення магнітного потоку від постійного значення) накладені дискретною довжиною стандартних рейок та міжшпальними інтервалами, а також електромагнітними властивостями елементів шпальних та рейкових скріплень.

2. ПКВП, як модель сигналу-відгуку від електромагнітного збурення рейки

2.1 Сигнал, отриманий електромагнітним збуренням бездефектної рейки, поданий на рис. 3. Він має чітку коливну структуру, яка дещо спотворюється випадковими завадами. Найбільш адекватною моделлю сигналу в цьому випадку буде модель у вигляді періодично корельованого випадкового процесу (ПКВП). Модель у вигляді ПКВП здатна описувати як регулярність коливань, так і їж стохастичність і є узагальненням детерміністичного і стаціонарного підходів.

ПКВП - це періодично нестаціонарний в широкому змісті випадковий процес з інваріантним відносно зсувів на T математичним сподіванням і кореляційною функцією

$$m(t) = E\xi(t) = m(t+T),$$

$$b(t,u) = E\,\overset{\circ}{\xi}(t)\,\overset{\circ}{\xi}(t+u) = b(t+T,u),$$

$\overset{\circ}{\xi}(t) = \xi(t) - m(t)$. T - період корельованості процесу [3].

При виконанні умов

$$\int_0^T |m(t)| < \infty, \qquad \int_0^T |b(t,u)| < \infty, \qquad \int_0^T |f(\omega,t)| < \infty,$$

математичне сподівання, кореляційна функція і спектральна густина можуть бути представлені у вигляді рядів Фур'є

$$m(t) = \sum_{k\in Z} m_k e^{ik\frac{2\pi}{T}t}, \quad b(t,u) = \sum_{k\in Z} B_k(u) e^{ik\frac{2\pi}{T}t},$$

$$f(\omega,t) = \sum_{k\in Z} f_k(\omega) e^{ik\frac{2\pi}{T}t},$$

де

$$m_k = \frac{1}{T}\int_0^T m(t) e^{-ik\frac{2\pi}{T}t}\,dt,$$

$$B_k(u) = \frac{1}{T}\int_0^T b(t,u) e^{-ik\frac{2\pi}{T}t}\,dt,$$

$$f_k(\omega) = \frac{1}{T}\int_0^T f(\omega,t) e^{-ik\frac{2\pi}{T}t}\,dt,$$

$m_k \to \infty$, $B_k(u) \to \infty$, $f_k(\omega) \to \infty$ при $k \to \infty$. Функції $B_k(u)$ і $f_k(\omega)$ мають назву кореляційних і спектральних компонентів. Коефіцієнти Фур'є m_k і $B_k(u)$ є кількісними характеристиками форми періодичних кривих відповідно математичного сподівання і кореляційної функції.

3. Результати обробки

Видається цікавим дослідити статистичні відмінності двох ПКВП: сигналу від завідомо бездефектної рейки та сигналу від рейки з характерним дефектом методами стаціонарної і нестаціонарної обробки на рівні математичного сподівання та дисперсії.

При цьому модель ідеальної рейки будуємо в межах максимальної довжини такої рейки (не більше 25 м), а наявність дефекта оцінюємо в межах міжшпального інтервалу (зрозуміло, що при додатковому вимірюванні відстані точність локалізації дефекта можна піднімати відповідно).

3.1. Дослідимо методами стаціонарної обробки сигнал-відгук від бездефектної рейки (рис. 3).

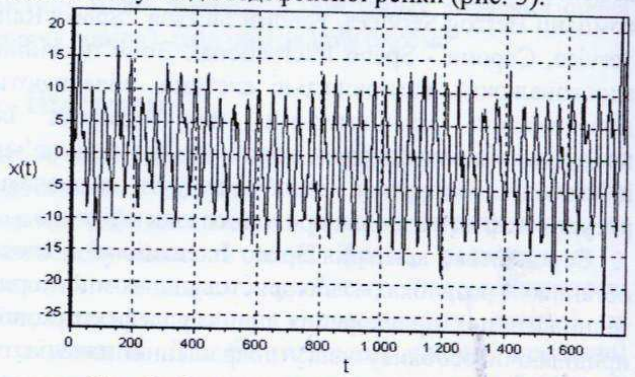


Рис. 3. Сигнал від бездефектної рейки

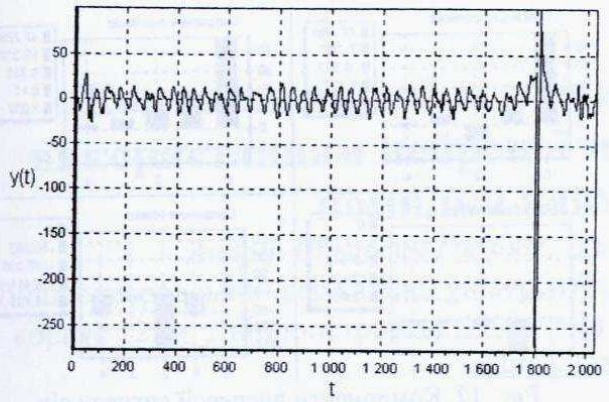


Рис. 4. Сигнал від рейки з дефектом

На рис. 5, 6 подані кореляційна функція і спектральна густина сигналу в стаціонарному наближенні. Кореляційна функція є заникаючою і має коливний характер з одним переважаючим періодом, що відповідає відстані між шпалами (~49.7 см). Рівень шуму в сигналі складає близько 15% по амплітуді. На графіку спектральної густини окрім несучої частоти, що відповідає відстані між шпалами, присутня також бокова пелюстка на частоті, кратній несучій, що можна пояснити технічними характеристиками приймача, що приймає сигнал від двох шпал.

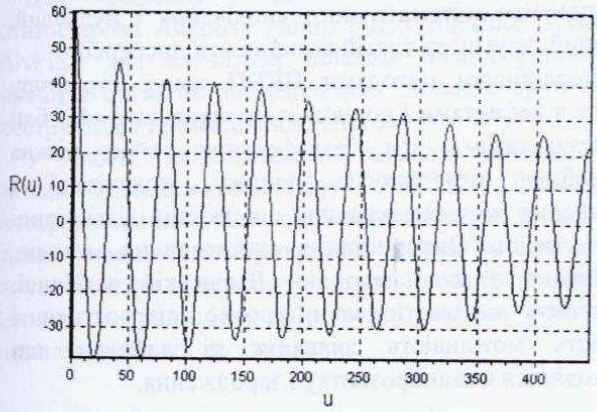


Рис. 5. Кореляційна функція сигналу від бездефектної рейки

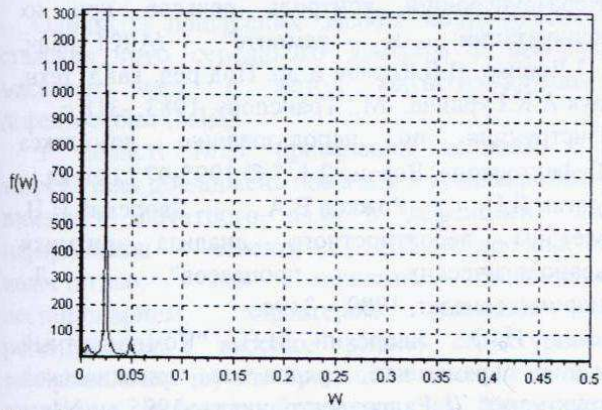


Рис. 6. Спектральна густина сигналу від бездефектної рейки

3.2. Розглянемо статистичні властивості сигналу від рейки з дефектом. Кореляційна функція (рис. 7) має вже значний рівень шуму і не таку правильну коливну форму, як у випадку бездефектної рейки.

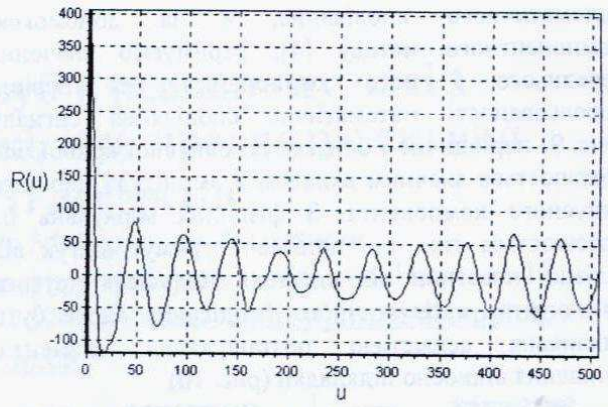


Рис. 7. Кореляційна функція сигналу від рейки з дефектом

На графіку спектральної густини (рис. 8) присутній переважаючий пік, що відповідає міжшпальному проміжку, але з'явилися нові частоти, які забезпечують підняття цього піку на «плато», а також підняття високочастотних областей. Для визначення більш тонких відмінностей і вибору інформативних параметрів сигналів від залізничних рейок дослідимо статистичні характеристики даних тестових сигналів з точки зору нестаціонарності в часі.

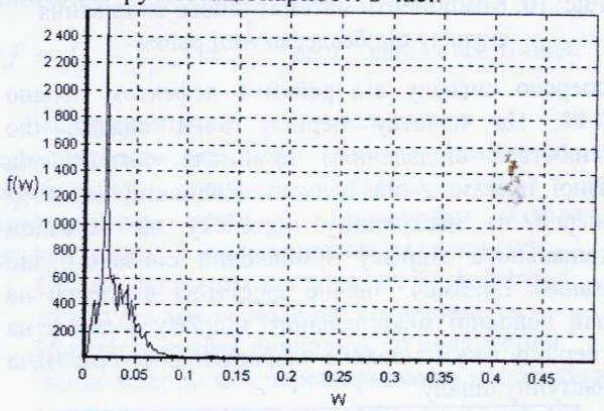


Рис. 8. Спектральна густина сигналу від рейки з дефектом

3.3. Розглянемо методами ПКВП математичне сподівання і дисперсію сигналу від бездефектної рейки.

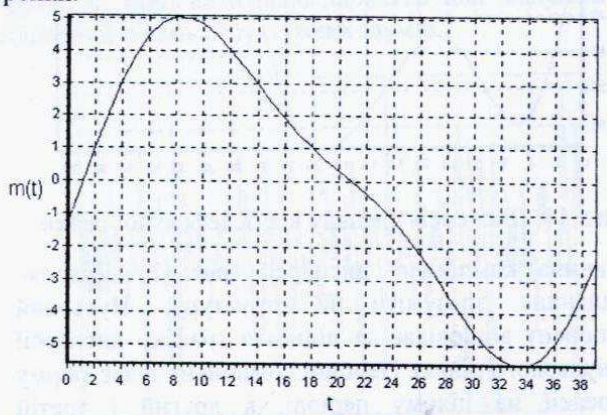


Рис. 9. Математичне сподівання сигналу від бездефектної рейки

Використовуючи когерентний метод оцінювання ПКВП, уточнюємо кількість компонентів

математичного сподівання, а за допомогою компонентного методу [4], отримуємо значення середнього і його компонентів. На періоді корельованості математичне сподівання сигналу (рис. 9) періодичне і близьке до синусної кривої, що пояснюється значним вкладом в амплітуду першого синусного компонента. З фізичних міркувань це пояснюється тим, що приймач отримує відгук від залізної пластини під шпалою. Наявність другого компонента математичного сподівання може бути викликана асиметією розташування елементів кріплення відносно підкладки (рис. 10).

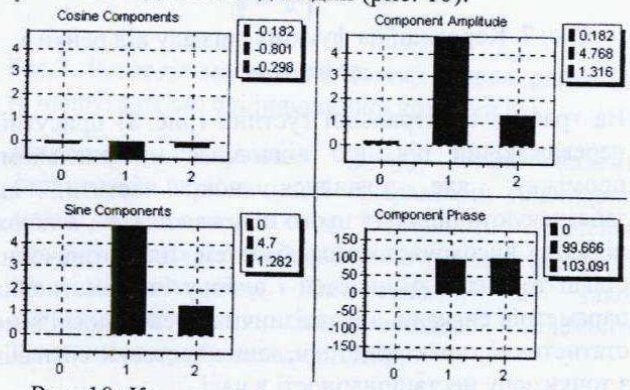


Рис. 10. Компоненти математичного сподівання сигналу від бездефектної рейки

Дисперсію сигналу від рейки з дефектом подано рис. 11. На початку періоду вона спадає, що пояснюється віддаленням приймача сигналу від залізної пластини під рейкою. Коливний характер дисперсії на наступному проміжку до середини міжшпального відрізку зумовлений сигналом, що «доганяє» приймач. Значне зростання дисперсії на другій половині міжшпального відрізку з піком на 3/4 періоду вже зумовлене «наїжджанням» приймача на наступну шпалу.

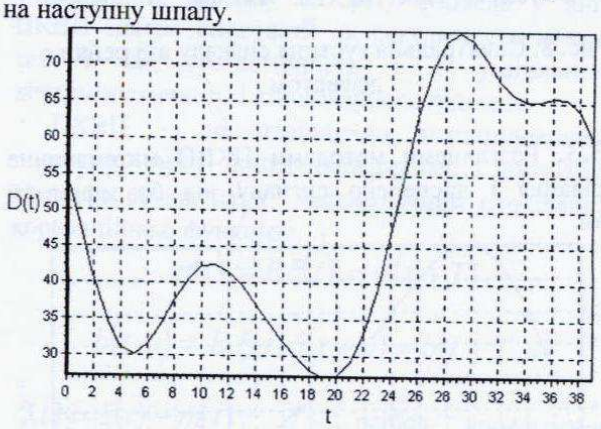


Рис. 11. Дисперсія сигналу від бедефектної рейки

Величина компонент дисперсії (рис. 12) дозволяє детальніше зрозуміти її структуру. Нульовий компонент відповідає за підняття графіку дисперсії від нульового рівню. Перший компонент задає форму дисперсії на цілому періоді, а другий і третій компоненти відповідають за зміни на пів- і чверть-періоді. Зважаючи на незначний рівень четвертого компоненту дисперсії, його можна не брати до уваги.

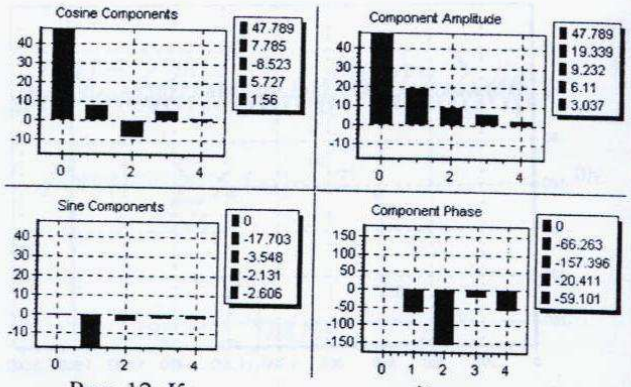


Рис. 12. Компоненти дисперсії сигналу від бездефектної рейки

ВИСНОВКИ

Статистичний аналіз сигналів, отриманих від електромагнітного збурення залізничних рейок, дозволяє проводити швидкісну діагностику стану рейок під час руху вагона-дефектоскопа. З використанням методів теорії періодично корельованих випадкових процесів, можна вибрати інформативними параметрами діагностики кількість компонентів математичного сподівання і дисперсії. Для бездефектної рейки присутні перший і другий компоненти математичного сподівання і нульовий, перший, другий та третій компоненти дисперсії.

Дослідивши методами ПКВП тонку структуру рейок з дефектами і порівнюючи результати обробки з результатами для бездефектних рейок можна розробити методологію пошуку, локалізації і виявлення типу дефектів при швидкісній діагностиці стану рейок. Використання статистичних методів, порівняно з існуючими на Львівській залізниці, дозволить автоматизувати процес діагностики і надасть можливість виявляти ті дефекти, що знаходяться в стані розвитку і зародження.

ЛІТЕРАТУРА

1. "Неразрушающий контроль рельсов при их эксплуатации и ремонте" /А.К.Гурвич, Б.П.Довнар, В.Б.Козлов и др. Под ред. канд. техн. наук А.К.Гурвича, -М.: Транспорт, 1983, -318 с.
2. "Инструкция по использованию комплекса «Дефектоскоп». Версия 2.1. " © 1995-97 Logi+ka.
3. Драган Я.П., Рожков В.А., Яворский И.Н. "Методы вероятностного анализа ритмики океанологических процессов".- Л.: Гидрометеоиздат, 1989. - 319с.
4. Исаев И.Ю., Яворский И.Н. "Компонентный анализ временных рядов с ритмической структурой" // Радиоэлектроника.- 1995. - N1. - с.34-45

Контроль якості сталевих труб в реальному часі з використанням нейроної мережі, яка прогнозує сигнал помилки наявності дефектів

Андрій Павлович Гвоздак, Олександр Мубаркович Ахметшин
Дніпропетровський державний університет, Радіо-фізичний факультет, вул. Наукова 13, корпус 12, м. Дніпропетровськ 49050. тел. 056-7769092. e-mail: andrey@mail.dsu.dp.ua, akhm@mail.dsu.dp.ua.

Магнітно-індукційний метод контролю якості безшовних труб у реальному масштабі часу характеризується високим рівнем структурних шумів, що мають складний закон розподілу щільності ймовірностей які змінюються від партії до партії виробів.

Традиційний метод виявлення дефектів труб пов'язаний із використанням еталонних дефектів. Проте форма реальних дефектів є випадкової, що не дозволяє використовувати методи оптимальної фільтрації для їхнього виявлення.

Для рішення задачі була використана багатошарова Adaptive Neuro Fuzzy Interface System (ANFIS), яка навчалася шляхом подачі усіляких варіантів сигналів бездефектних ділянок труб що реєструються приймальною системою.

У якості аналізованого сигналу розглядався сигнал помилки прогнозу ANFIS. Проведені експерименти показали, що метод дозволяє виділити сигнал випадкових протяжних дефектів навіть у ситуаціях коли відношення сигнал / шум було менше одиниця і традиційні амплітудні методи селекції сигналів дефектів не виявляли.

1. Вступ

Одним із поширених засобів контролю якості сталевих труб середнього діаметра в реальному масштабі часу є метод магнітно-індукційної дефектоскопії (МІД).

У області МІД проведений великий обсяг теоретичних досліджень пов'язаних із моделюванням взаємодії магнітного поля з дефектами, проте інформаційні аспекти МІД досліджувані недостатньо, що можливо пояснюється нестаціонарним характером сигналів що реєструються і високим рівнем структурних і вимірювальних шумів. Традиційні підходи до задачі виявлення дефектів базуються на використанні методів смугової фільтрації сигналів що реєструються [1, 2]. Проте детальний аналіз ситуації показує що, подібний підхід не узгоджується з фізичною структурою сигналів що реєструються, що як слідство веде до зниження реальної чутливості контролю. Це вимагає необхідність пошуку нових підходів до аналізу даних МІД на базі нових інформаційних технологій.

2. Модель сигналу що реєструється

Важливою прикладною особливістю більшості методів МІД є використання систем намагнічування які обертаються навколо труби зі зніманням вихідної інформації за допомогою безконтактних індуктивних перетворювачів. Останнє означає, що спектральна характеристика сигналів що реєструються є лінійчатої з відстанню між гармоніками $\Delta f = f_k - f_{k-1} = \frac{1}{T_0}$, де T_0 період обертання.

Експериментальні дослідження також підтверджують той факт, що, у відмінності від ідеалізованої ситуації, реальна труба має багатогранник, причому з невідомим заздалегідь числом граней. З цього випливає два висновки:

1. Сигнал що реєструється навіть бездефектної структури має періодичну послідовність.
2. Наявність граней приводить до періодичної зміни відстані між перетворювачем і поверхнею труби.

Слідством останнього ефекту є наявність і паразитної амплітудної модуляції сигналу що реєструється. Це призводить до зниження реального порога чутливості системи контролю. На рис. 1 показаний типовий сигнал бездефектної ділянки труби, де явно визначено характер цієї модуляції і відносний рівень структурних шумів.

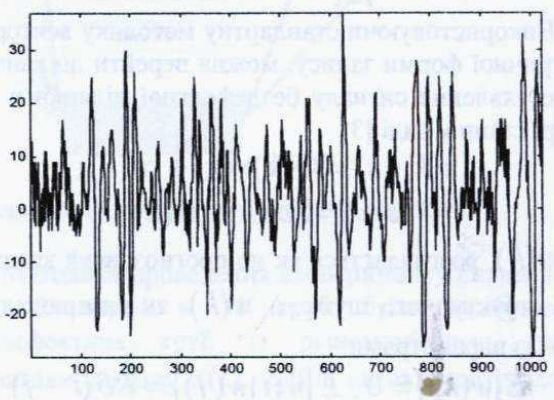


Рисунок 1. Типовий сигнал бездефектного ділянки труби $s(t)$ при магнітно-індукційному методі контролю.

Загальний аналіз характеру сигналів $s(t)$ бездефектних ділянок труб дозволяє зробити два додаткових висновки:

-використання систем відбракування (контролю) із фіксованим порогом прийняття рішення можливо лише в тому випадку, коли амплітуда сигналу дефекту перевершує максимальну амплітуду бездефектної ділянки. Проте дуже часто спостерігалася ситуація, коли поява періодичного сигналу дефекту збігалося з мінімумами сигналу $s(t)$, що не дозволяло ідентифікувати наявність сигналу дефекту при використанні фіксованого порога прийняття рішення.

-використання смугової фільтрації ефективно лише у випадках, суттєвої відмінності спектральних характеристик апріорно невідомого сигналу дефекту і сигналу окремої грані. Це можливо не завжди і в усіх випадках використання смугової фільтрації буде приводити до зниження реальної чутливості контролю оскільки як сигнал дефекту так і сигнал грані буде мати приблизно однакову ширину спектра (але різну інтенсивність спектральних компонент).

Відзначені особливості зажадали пошуку альтернативних підходів до задачі магнітно-індукційного контролю труб, по своїй сутності еквівалентних використанню ковзного порога прийняття рішення для підвищення реальної чутливості контролю.

У рамках поставленої задачі були досліджувані два підходи: метод адаптивної фільтрації Калмана і Adaptive-Network-based Fuzzy interface System (ANFIS) основана що до прийняття рішення про дефектність труби на основі аналізу помилки прогнозу сигналу дефектоскопу в часовій області.

3. Метод адаптивної фільтрації Калмана.

Ідея методу заснована на розгляді сигналу бездефектної ділянки у виді стаціонарної стохастичної динамічної послідовності, що описується в загальному випадку рівнянням авторегресії порядку p, тобто

$$s(t_k) = \sum_{i=1}^{p} a(i)s(t_{k-i}) + u(t_k). \quad (1)$$

Використовуючи стандартну методику векторної-матричної форми запису, можна перейти до канонічного уявлення сигналу бездефектної ділянки в просторі стани у виді [3]

$$x(k+s) = \Phi(k)x(k) + ru(k);$$
$$z(k) = \mathrm{H}x(k) + w(k), \quad (2)$$

де $u(k)$ розглядається як не прогнозована компонента структурного шуму, а $w(k)$ як вимірювальний шум із параметрами

$$E\{w(k)\} = 0;\ E\{w(i)w(j)\} = R\delta(i-j)$$
$$E\{w(k)\} = 0;\ E\{w(i)w(j)\} = Q\delta(i-j). \quad (3)$$

Поява дефекту можна характеризувати моделлю:

$$x(k+1) = [\Phi(k) + \Delta\Phi(k)]x(k) + \Gamma u(k), \quad (4)$$

де $\Delta\Phi(k)$ - випадкова фінітна функція яка характеризує дефект.

Тому що вихідний фільтр розраховується в припущенні без дефектності труби, то поява дефекту веде до невизначеності в розрахунковій моделі, що може привести до великих помилок аж до розходження оцінок фільтрації. Саме цей факт і визначає необхідність використання адаптивної фільтрації. Для локалізації випадкового дефекту необхідно зіставлення дефектограми що фільтрується з деяким еталонним сигналом. У якості якого ми розглядали відфільтрований динамічний сигнал апріорно бездефектної труби

$$s_{et}(k) = s(k) + s_N(k), \quad (5)$$

де $s(k)$ розглядається як квазідетермінована компонента, а $s_N(k)$ - як неув'язка фільтрації з параметрами

$$E\{s_N(h)\} = 0;\ E\{s_N^2(h)\} = P_{et}(k,h), \quad (6)$$

при цьому задача виявлення і локалізації дефекту зводиться до обчислення і порівняння з заданим порогом виразом:

$$y = \frac{[s(k) - s(k,k)]^2}{P_{et}(k,k) + P(k,k)} \rightarrow \begin{cases} < \alpha : \mathrm{H}_0 \\ \geq \alpha : \mathrm{H}_1 \end{cases}, \quad (7)$$

де α - поріг прийняття рішення; H_0 - відсутність дефекту; H_1 - наявність дефекту.

У процесі проведених експериментів були перевірені різноманітні варіанти адаптивних алгоритмів фільтра Калмана [5] і в багатьох ситуаціях були отримані цілком задовільні результати. Проте в цілому проявилася одна проблема пов'язана зі структурною невизначеністю характеру сигналів $s_{et}(k)$, оскільки, як було відзначено вище число граней труби заздалегідь не відомо. Адаптивний фільтр Калмана відслідковував зміни параметрів моделі, але це завжди приводило до появи сигналу розладу на фактично бездефектних ділянках у початковій стадії контролю. Ця задача може бути вирішена на основі використання сімейства паралельних фільтрів Калмана, але для їх виконання в реальному часі, буде потрібно достатньо складна багатопроцесорна система яка складається з гребінки сигнальних процесорів.

4. Контроль якості на основі аналізу помилки прогнозування вихідного сигналу навченої нейронної мережі.

Оскільки штучна нейрона мережа може бути подана як паралельно розподілена обчислювальна мережа, то використання нейроної мережі може бути розглянуто і з позиції альтернативи використання

гребінки фільтрів Калмана. Крім цього, у реальних ситуаціях часто виникає проблема представлення сигналу $s(t)$ до канонічної форми (2), оскільки теоретичні варіанти рішення задачі припускають, що як $w(t)$ так і $u(t)$ представляють гаусовскі послідовності.

При виборі методу рішення задачі визначену роль зіграли результати роботи [6] по прогнозуванню Mackye-Glass часових послідовностей, оскільки зіставлення даних на рис.22 [6] і бездефектного сигналу $s(t)$ на рис. 1 показує визначення їх подібність (хоча структура сигналу $s(t)$ звісно набагато складніше).

У рамках подібного підходу сигнал $s(t)$ варто розглядати як сигнал що генерується нелінійною системою, а не як стаціонарний часовий ряд при розрахунку фільтра Калмана. Такий погляд на природу сигналу $s(t)$ дає додатковий аргумент на користь використання нейроної мережі, тому що структура нейроної мережі може адаптуватися до сильно не лінійних послідовностей, що забезпечує її істотну перевагу в порівнянні з лінійними методами при опрацюванні нелінійних часових послідовностей.

При проведенні експериментів була обрана 5 шарова гібридна нейрона мережа з ANFIS архітектурою на основі обчислювального еквівалента Thikomato методу представлення (рис.2). Для прогнозування сигналу дефектоскопа $s*(t)$ використовувалися 4 входи $s(t_{k-6})$, $s(t_{k-12})$, $s(t_{k-18})$ і $s(t_{k-24})$.

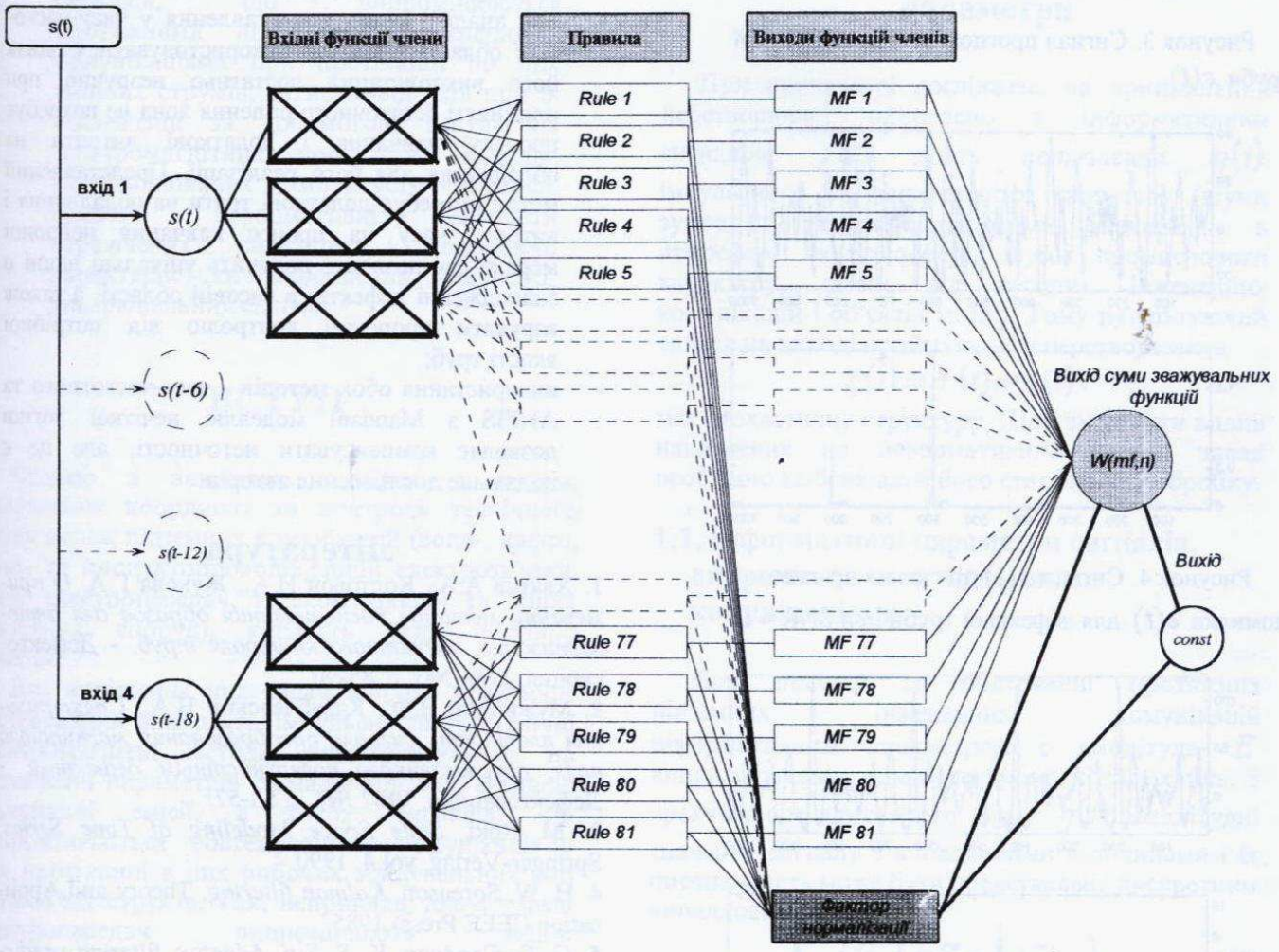


Рисунок 2. Структура ANFIS моделі використаної для процесу прогнозу сигналу дефектів.

Кожному входові ставилося у відповідність три лінгвістичні перемінні (low, med-ium, high). Кожному входові при фузіфікації ставилася у відповідність 3 трикутні функції приналежності що характеризуються трьома параметрами. Загальне число параметрів побудованої моделі ANFIS склало 429 із 81 числом правил фузіфікації відповідно.

Методика проведення експерименту складалася в так. З партії в 50 труб були відібрані 10 бездефектних труб із різноманітними типами динаміки сигналу $s(t)$. Всі ці сигнали розглядалися в якості навчальних для ANFIS.

Типовий сигнал помилки-прогнозу $e(t)$ для бездефектної труби показаний на рис.3. Амплітуда

$e(t)$ знаходилася в межах [-20 ÷ 20] відносних одиниць. Зіставлення сигналів $s(t)$ і $e(t)$ для дефектних труб при SNR ≈ 1 показано на рис. 4, а для SNR ≈ 2 на рис. 5.

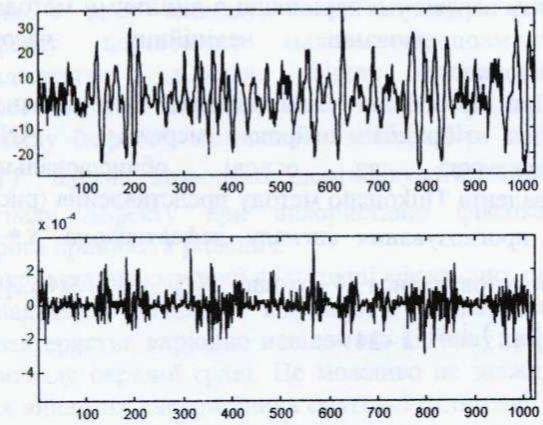


Рисунок 3. Сигнал прогнозу для бездефектної труби $e(t)$.

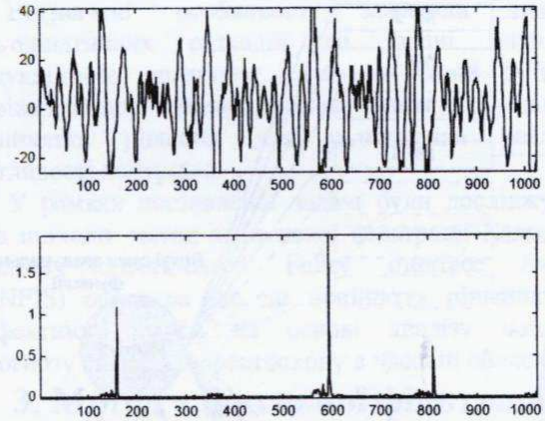


Рисунок 4. Сигнал $s(t)$ та сигнал прогнозу помилки $e(t)$ для дефектної труби при SNR ≈ 1.

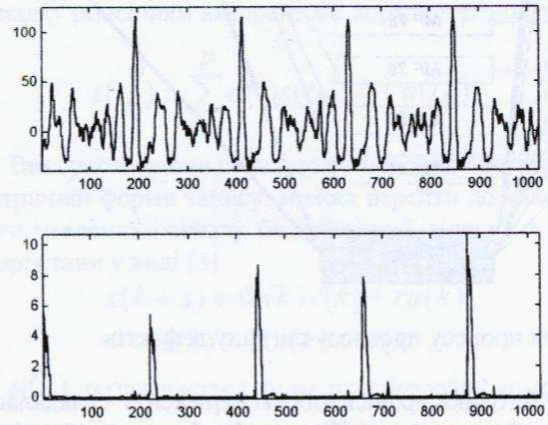


Рисунок 5. Сигнал $s(t)$ та сигнал прогнозу помилки $e(t)$ для дефектної труби при SNR ≈ 2.

Зіставлення результатів на цих малюнках показує, що застосування ANFIS дозволило підвищити відношення сигнал/шум у залежності $e(t)$ в 10 разів у порівнянні з вихідною залежністю $s(t)$. При SNR<1 результати були більш суперечливі і часто призводили до помилкових виявлень дефектів. У таких випадках для мінімізації можливості помилкової інформації про дефект доцільно використання декількох алгоритмів обробки інформації, що засновуються на різноманітних фізичних припущеннях про природу дефектів[7].

5. Висновки.

Використання моделі ANFIS дозволило:

- Поліпшити відношення сигнал шум вихідних сигналів завдяки використання лінійних функцій підлеглості для процесу дефузіфікації;
- дозволив обмежити присутність дефектів в часовій області при порівнянні видимих амплітуд сигналів дефектів;
- при аналізу метод представлення у часу-частотної області [7] може використовуватися однак його використання достатньо незручно при прийнятті остаточного рішення хоча не потребує процесу навчання і додаткові витрати на обладнання для його реалізації. Представлений метод потребує додаткові трати на обладнання і витрату часу на процес навчання нейроної мережі але дозволяє получить унікальні данні о знаходженні дефектів в часовій області, а також керувати процесом контролю від потрібної якості труб;
- використання обох методів - часу-частотного та ANFIS з Mandani моделлю нечіткої логіки дозволяє компенсувати неточності, але це є подальше дослідження авторів.

Література

1. Хватов А.А., Колтовой И.А., Жукова Г.А. *О применении методов распознавания образов для автоматизации магнитного контроля труб.* - Дефектоскопия, 1986 №3, с. 85-90.
2. Мужицкий В.Ф., Карабчевский В.А.. *Спектральная плотность сигнала преобразования магнитного поля, обусловленного поверхностным дефектом.* - Дефектоскопия, 1987 №7, с. 53-57.
3. M. Roki, *State Space Modeling of Time Series*, Springer-Verlag, vol.4, 1990.
4. H. W. Soreuson, *Kalman filtering*, Theory and Application (IEEE Press),1985.
5. G. C. Goodwin, K. S. Sin, *Adaptive filtering prediction and control*, Prentice-Hall, Englewood Cliffs, N.Y., 1984.
6. Y. S. R. Yang, ANFIS: *Adaptive-Network-Based Fuzzy Inference System*, IEEE Transaction on Systems, Man and Cybernetics, vol. 23, N3, pp. 665-685, 1993.
7. A.P. Gvozdak, V. M. Dolgov, *Time-frequency NDT in real time mode based on DSP*, Signal/Image Processing and Pattern Recognition, Proceeding, The Fourth All-Ukrainian International Conference, 19-23 October, Kiev, pp. 113-116, 1998.

Пошук та обстеження захованих об'єктів за їх сигналами

Петро Драбич, Олег Драбич

Фізико-механічний інститут ім. Г.В.Карпенка НАН України
вул. Наукова, 5, 79601, МСП, Львів, Україна,
тел. 8 (0322) 63 33 55, факс 8 (0322) 64-94-27
e-mail: drabich@ipm.lviv.ua

Описана структура стохастичних сигналів, що випромінюються протяжними підземними інженерними комунікаціями при протіканні по них змінних струмів, сигналів-відгуків при їх обстеженні за допомогою імпульсних електромагнітних полів та сигналів, випромінюваних ними акустичної емісії. Вказані інформативні елементи перелічених сигналів і приведені характеристики розроблених пошуково-вимірювальних систем.

ВСТУП

Однією з важливих проблем є пошук, визначення координат та контроль технічного стану мереж підземних комунікацій (водо-, нафто, газо- та продуктопроводів, ліній електрозв'язку, ліній електропередач, телекомунікацій та інших технічних споруд), кількість яких постійно зростає.

Для вирішення цих задач широке практичне застосування знаходять імпульсні та гармонічні електромагнітні методи, які ґрунтуються на визначенні параметрів сигналів-відгуків, сигналів акустичної емісії, а також сигналів, що випромінюються обстежуваними комунікаціями при протіканні в них робочих, зондувальних або блукаючих струмів. Так, наприклад, діючі кабелі електропередач випромінюють змінні електромагнітні коливання з частотою 50 Гц, а магістральні нафто- та газопроводи - коливання з частотою 100 Гц, обумовлені протіканням в них струму катодного захисту, а також сигнали акустичної емісії, обумовлені навантаженням трубопроводів.

При відсутності в комунікаціях робочих струмів електромагнітні коливання можуть збуджуватись приєднанням до них спеціальних генераторів або безконтактним індуктивним методом.

1. Структура випромінювальних сигналів та їх інформативні параметри

При проведенні досліджень на приймальний перетворювач одночасно з інформативним сигналом $f(t)$ діють поля-завади $\eta(t)$ імпульсного та неперервного характеру (шуми зумовлені електромагнітними явищами в атмосфері та іоносфері, шуми промислового характеру, шуми від сусідніх інженерних комунікацій і об'єктів та ін.). Тому результуючий сигнал на виході приймального перетворювача

$$\xi(t) = f(t) + \eta(t)$$

має стохастичну структуру. Щоб зменшити вплив накладених на інформативний сигнал завад необхідно здійснювати його статистичну обробку.

1.1. Інформативні параметри сигналів, випромінюваних підземними комунікаціями

При пошуку та обстеженні протяжних підземних інженерних комунікацій інформативним параметром є амплітуда E випромінюваних ними електромагнітних хвиль. З врахуванням адитивного шуму $\eta(t)$ амплітудні значення сигналу є випадковими величинами і їх послідовність може бути представлена дискретним випадковим процесом

$$\{E_i, i \in Z = 1,2,3,\ldots\}. \qquad (1)$$

В цьому випадку алгоритм обробки сигналу полягає в його фільтрації та оцінці значення амплітуди E шляхом усереднення випадкових значень E_i.

При обробці сигналу використовується теж фізична закономірність, яка полягає в тому, що напруженість магнітної складової електромагнітного поля, що випромінюється протяжним струмопровідним об'єктом, прямо пропорційна до величини змінного струму $I(t)$,

© UkrObraz'2000

що протікає вздовж нього і обернено пропорційна до відстані h між точкою спостереження і віссю об'єкта, тобто магнітне поле розглядають як поле лінійного струму, суміщеного з віссю об'єкта [1,2]. Визначивши на двох відстанях h та $h+a\,(a=const)$ амплітуди E_1 і E_2, наведених в приймальних індукційних перетворювачах (котушках) електрорушійних сил, обчислюють глибину залягання h та амплітуду I_m струму, що протікає в комунікації, за формулами [3].

$$h = a\frac{E_2}{E_1 - E_2},$$

$$I_m = h\,E_1.$$

За характером зміни величини струму та електричного потенціалу на поверхні ґрунту, обумовленого витіканням в ґрунт струму в місцях пошкодження ізоляції комунікації оцінюють технічний стан останньої та характер корозійних процесів, що у ній протікають [4].

1.2. Інформативні параметри розміщених над комунікаціями електромагнітних маркерів.

Для пошуку і швидкої локалізації функціонально важливих елементів чи відрізків комунікацій, а також для точної прив'язки вздовж осі даних внутрішньо-трубного обстеження чи обстежень з поверхні землі нафто- та газопроводів використовують електромагнітні маркерні системи, що складаються з розміщених в ґрунті над віссю комунікації електромагнітних маркерів у вигляді коливних LC - контурів та пошукових пристроїв для їх виявлення. Пошукові пристрої налаштовані на резонансну частоту маркерів. Після опромінення з поверхні ґрунту електромагнітними імпульсами маркер починає генерувати в паузах між збуджуючими імпульсами заникаючі коливання, які фіксуються приймальним перетворювачем і є інформацією про наявність в ґрунті певного маркера [5].
Інформативними параметрами випромінюваних маркерами коливань є:

- початкова амплітуда E_0;
- амплітуди заникаючих коливань E_i;
- сума амплітуд у відеоімпульсі

$$E_\Sigma = E_0 + \sum_{j=1}^{m} E_j;$$

- характеристики, що визначаються функціями інформативних елементів

$$f\left(E_0, E_1, \ldots, E_n\right) = {E_0}^2 + \sum_{j=1}^{n} {E_j}^2;$$

- енергетичні характеристики (енергія або потужність відеоімпульса) та ін.

Очевидно, що на зареєстровану приймальним перетворювачем послідовність відеоімпульсів накладаються теж шуми природного і промислового характеру, а також шуми підсилювального тракту індикаторного пристрою. Оскільки зареєстрований сигнал має стохастичний характер, його доцільно представляти дискретними випадковими процесами типу (1).

При вивченні структури сигналів маркерів виявлено, що найбільш завадостійким параметром є сума амплітуд E_Σ, що використано при побудові електромагнітної маркерної системи МС-3 [6].

1.3. Інформативні ознаки сигналів акустичної емісії.

Серед широкого спектру методів технічної діагностики зварних трубопроводів є метод акустичної емісії [7], який полягає у випромінюванні в процесі експлуатації трубопроводу пружних хвиль.

Стохастичність сигналів акустичної емісії обумовлена випадковістю появи та розміщення дефектів та наявність в трубопроводі та ґрунті акустичних шумів.

Інформативними елементами сигналу акустичної емісії можуть бути інтервали часу між моментами випромінювання, амплітуда, тривалість відеоімпульсу, його енергія, число коливань і ряд інших.

При представленні акустичного сигналу у вигляді сукупності послідовностей випадкових значень вказаних елементів, тобто процесами типу (1), алгоритм його обробки полягає в тому, що в першу чергу виділяють і формують вказані елементи, а потім здійснюють статистичну обробку сформованих послідовностей.

2. Оцінка інформативних параметрів стохастичних сигналів.

Оцінка інформативних параметрів описаних і подібних до них неперервних та імпульсних стохастичних сигналів, як правило, здійснюється за однією реалізацією, що при дослідженні комунікацій в русі вимагає виконання вимоги локальної стаціонарності дискретних випадкових процесів.

При виконанні умови локальної стаціонарності параметри сигналу оцінюються методами простого (лінійного) або вагового усереднення.

Вагове усереднення на основі рекурентного співвідношення, що описується в загальному випадку неоднорідним різницевим рівнянням першого порядку [8] одночасно забезпечує

необхідне у багатьох практичних випадках функціональне перетворення оцінок, наприклад піднесення до степеня, ділення, перемноження, логарифмування і т.п. Тому при побудові портативних електромагнітних пошуково-вимірювальних систем часто використовується вагове усереднення, зокрема воно використане в таких розробках Фізико-механічного інституту ім. Г.В.Карпенка НАН України як прилад для пошуку підземних кабелів і трубопроводів та обстеження їх ізоляції ІМК-4 [6] та електромагнітна маркерна система МС-3.

3. Технічні характеристики електромагнітних пошуково-вимірювальних систем.

3.1. Прилад ІМК-4

Забезпечує визначення осі, глибини залягання підземних магістральних нафто-, газо- та продуктопроводів, кабелів зв'язку, електропередач, телекомунікацій, тощо та місць пошкодження їх ізоляції.

Принцип дії приладу заснований на вимірюванні випромінюваних комунікацією електромагнітних полів, а також електричних полів, зумовлених витіканням в ґрунт електричного струму з комунікації в місцях пошкодження її ізоляції.

Визначення координат комунікації в приладі здійснюється за різницевим сигналом двох розміщених на одній штанзі магнітоприймачів, що дає йому переваги перед вітчизняними та зарубіжними аналогами за завадостійкістю і точністю визначення осі і глибини при роботі в умовах значного рівня завад від електромережі та промислових об'єктів, а також при наявності поруч з обстежуваною інших комунікацій.

Місця пошкодження ізоляції комунікації визначаються шляхом реєстрації величини електричного потенціалу або градієнта потенціалу на поверхні ґрунту над віссю.

Прилад дозволяє знаходити підземні (підводні) комунікації на глибині до 10м з похибкою не більше ±5%. Передбачена можливість роботи на трьох частотах: 50; 100 і 222 Гц. Діапазон вимірювання постійного або змінного потенціалів від 10 до 2000 мВ.

Частота 222 Гц використовується в складних умовах, коли значна кількість комунікацій сконцентрована поруч або при відсутності в комунікації струмів з частотами 50 або 100 Гц.

В приладі є звукова індикація для пошуку осі і цифрова - для реєстрації глибини та градієнта постійного і змінного потенціалів на поверхні ґрунту.

Живлення приладу здійснюється від одного акумулятора типу 7Д-0,115-VI УН.

Струм живлення - 4,0 мА.

Для роботи на частоті 222 Гц прилад комплектується генератором змінного струму ГС-1 [9].

Основні технічні характеристики генератора:

- потужність не більше 100 ВА;
- частота вихідного сигналу - 100, 222, 300 і 1000 Гц;
- похибка установки частоти - не більше ±5%;
- нестабільність частоти - не більше 0,1%;
- діапазон зміни вихідної напруги - від 10 до 60 В з дискретність 10 В.

3.2. Маркерна система МС-3.

Призначена для пошуку трас силових, світловодових і телефонних кабелів, їх підземних з'єднань та консерваційних ділянок, розгалужень, приєднань, водопровідних, каналізаційних труб та нафто-, газо- і продуктопроводів, їх трас прокладки, ревізійних люків, переходів під вулицями міст, головних розгалужень, муфт, трійників і т.п.

Принцип дії системи ґрунтується на збудженні маркера імпульсним електромагнітним полем з наступною реєстрацією випромінюваних маркером згасаючих електромагнітних хвиль. Для пошуку осі комунікації маркерна система містить трасошукач, який реєструє горизонтальну складову магнітного поля комунікації. Ефективність і надійність в роботі цієї системи полягає в тім, що в індикаторному пристрої здійснюється подвійна фільтрація інформативного сигналу за допомогою послідовно з'єднаних вузькосмугового підсилювача змінної напруги і перетворювача усереднення за рахунок чого забезпечується суттєве придушення зовнішніх завад та шумів підсилювального тракту.

Технічні характеристики системи МС-3:

- глибина пошуку осі комунікації до 10 м;
- глибина виявлення маркера до 1,2 м при радіусі чутливості на земній поверхні відносно його осі не більше 0,5 м.

ВИСНОВКИ

В доповіді розглянуті вимірювальні сигнали, які випромінюються підземними (підводними) протяжними струмопровідними комунікаціями при протіканні в них змінних електричних струмів, сигналів-відгуків при їх обстеженні за допомогою імпульсних електромагнітних полів, а також сигналів випромінюваної ними акустичної емісії. Показано, що ці сигнали є стохастичними за рахунок адитивних завад природного та промислового характеру і їх доцільно представляти сукупностями дискретних випадкових процесів.

Стохастична модель сигналу дозволяє сформулювати критерії його ефективної обробки, які суттєво підвищують завадостійкість електромагнітних пошуково-вимірювальних систем.

Запропоновано використовувати для оцінки інформативних параметрів вимірювальних сигналів вагове усереднення, що описується неоднорідним різницевим рівнянням першого порядку, яке одночасно забезпечує необхідне функціональне перетворення оцінок.

Підтвердженням ефективності запропонованого підходу є широке практичне використання розроблених в ФМІ ім. Г.В. Карпенка на цій основі електромагнітних пошуково-вимірювальних приладів і систем типу ІМК-1÷ІМК-4 та МС-1÷МС-3.

ЛІТЕРАТУРА

1. Шимони К. *Теоретическая техника*: Пер. с англ. / Под. ред. проф. К.П.Поливанова. - М.: Мир, 1964. - 773с.
2. Patent 2070783A (GB) INTCL G01R19/00. *Measuring current in a conductor* / M.I.Howell, A.Douglas, S.Charmers - Published 9 Sep. 1981.
3. Григорович К.К., Ягола Г.К. *Методы и средства оценки состояния изоляционных покрытий подземных газопроводов* // Сер. «Коррозия и защита в нефтегазовой промышленности». - М.: ВНИИОЭНГ, - 1979. - №9. - С.17-20.
4. Улич Г.Г., Реви Р.У. *Коррозия и борьба с ней. Введение в коррозионную науку и технику: Пер. с англ.* / Под. ред. А.М.Сухотина. - Л.: Химия, 1989. - Пер. изд., США, 1985. - 456с.
5. Драбич П.П., Зубик Й.Л., Кичма А.А., Стецько И.Г. *Электромагнитные маркеры и устройства для их обнаружения* // Метрология. - М.: Изд-во стандартов. -1998. - №8. - С.23-30.
6. Драбич П.П. *Електромагнітні вимірювальні прилади і системи для дослідження середовищ та пошуку і обстеження захованих об'єктів* / Сб. науч. тр. «Системы контроля окружающей среды» - Севастополь: МГИ НАН Украины, 1990. С. 37-41.
7. В.В.Шип, Г.Б.Муравин, В.Ф.Чабуркин. *Вопросы применения метода акустической эмиссии при диагностике сварных трубопроводов* // Дефектоскопия. - 1993. - №8. - С.17-19.
8. *Алгоритмы функционального преобразования и оценки параметров случайных сигналов* // Электронное моделирование. - Киев: Наукова думка. - 1990. - т.12, №4. - С.28-33
9. Патент 10884А (Україна). *Генератор для пошуку і обстеження підземних комунікацій* / П.П.Драбич - Опубл. 25.12. 1996, Бюл.№4.

NONDESTRUCTIVE MATURITY DETERMINATION OF DURIAN BY FORCE VIBRATION

*W. Kongrattanaprasert, S. Arunrungrusmi, B. Pungsiri, K. Chamnongthai and M. Okuda**

King Mongkut's University of Technology Thonburi
Tung-kru, Bangkok, Thailand 10140
Phone: +66-2-470-9064, Fax: +66-2-427-9636
Email: sanchai_k@hotmail.com

*Nippon Institute of Technology
Japan

ABSTRACT

This paper proposes a method for determining maturity levels of durian by using nondestructive vibration. The vibration is transferred directly through durian at a region between prickles that located at the middle of durian. The measurement of frequency responds from vibration of durian is done by using laser doppler. The signal is processed for finding voltages, frequency, time interval and velocity of vibration. These data relate to % dry weight and elastic constant. For the experimental results, the maturity of durian can be classified into two groups, immature durian and mature durian, related to threshold value of velocity of vibration at 500 cm/sec. and elastic constant at 300 $(dyne/cm^2)^{1/2}$ and it can be classified with 95% of accuracy.

1. INTRODUCTION

The durian is one of the high economical fruits that Thailand has exported for several years. To guarantee the qualification of them is an important process. There are many methods to determine their qualities, both destructive and non-destructive. The non-destructive approaches are preferred. Unfortunately, the autonomous quality control in non-destructive way is very difficult, since there are varied factors needed to consider, for example, the thickness of its peel, non-uniformity, roughness of its prickle skin, and so on.

In quality determination of fruits, Ketsa and Daengkanit [1] presented the relation between the maturity of durians and the internal chemical organization. This approach is a destructive method. [2] and [3], the authors presented quality measurement of fruits by considering their external characteristics such as color, size, and shape. Slaughter [4] utilized the near infrared to determine the chemical structure inside tomatos and peachs. However, the infrared is not appropriate for the pericarp fruits such as durians. Schatzki et. al [5] and Clark et.al[6] proposed the film x-ray analysis and MRI analysis, respectively. However, both methods need the expert system. For [7]-[10], these researchers used the ultrasonic for the maturity determination of avocado and watermelon. By using ultrasonic , the transducers need to touch the object not less than 96% of its region. Therefore this method is not applied for durian. [11]-[12] presented maturity determination of durian by knocking, and then took its signals to analyze for determination of frequency response. But the results showed low accuracy because of non-linearization of knocking.

Therefore, this paper presents the nondestructive maturity determination of durian by using the linear vibration for study in the relation of voltages, frequency response, and velocity of vibration between input and ouput vibration. Our proposed method uses transfer of vibration directly through durian at a region between prickles located at the middle of durian. Because middle region of durian have the least thickness of skin, and the space between prickles and the internal air cavity are very high. Laser doppler is used to measure output vibration, and then its signal is classified by determination of relation of % dry-weight. Peerapong [12] presented method for determining %dry-weight related to immature durian and mature durian at threshold value 30% dry-weight. By the pulp of durian 50 g. is fumed at 65°C for 48 hours, then % dry-weight can be obtained from:

$$\%dryweight = \frac{W_{behind}}{W_{before}} * 100 \qquad (1)$$

Where w_{behind} is weight behind fuming and W_{before} is weight before fuming

2. SYSTEM

The prototype system architecture for determination maturity of durian is shown in Fig. 1. After durian is carried by clamp and sets on the soft rubber of the prototype. First, the vibration set (1) which can vibrate with frequency 0-300 Hz is adjusted by setting the T-needle in perpendicular location between the prickles

© UkrObraz'2000

located at the middle of durian. Second, at the opposite side of first setting, the laser doppler set (2) is adjusted by setting the focus in coincident location between the prickles too. The focus is far from target 2.5 cm. The incident angle is equal to the reflect angle. After that the signals from this two sets are sent to amplifier (3) and (4), respectively. We can observe these signals on a monitor of oscilloscope (5) and process them by computer (6) for automatic classification of maturity of durian.

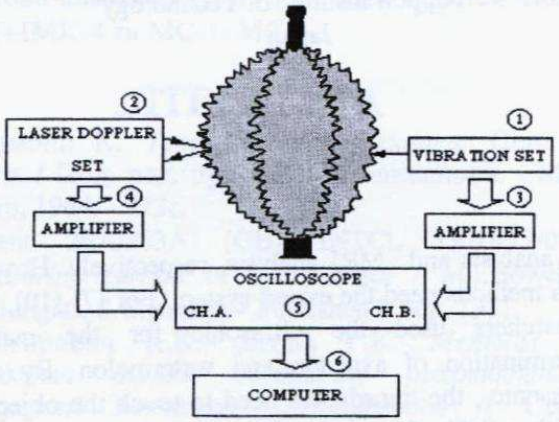


Fig. 1. System Architecture for maturity determination of durian

3. MATURITY DETERMINATION

3.1 Input and Output Vibration

The characteristic of mechanical vibration in medium with mass , elastic and viscostic components, H.L.Oestreicher in [13] show that 1) most parts of the energy radiated from the surface propagates as a transversal wave in the medium when the frequency of low –frequency vibration is less than about 1 kHz. and 2) the velocity of transversal wave in medium related to characteristic of medium are given by

$$V_t = \left(\frac{2(\mu_1^2 + \omega_b^2 \mu_2^2)}{\rho(\mu_1 + \sqrt{\mu_1^2 + \omega_b^2 \mu_2^2})} \right)^{1/2} \quad (2)$$

The frequency of input vibration selected at linear output signal and amplitude response is the highest. From the experiments at frequency 30 Hz. is the best. Because can to observed difference of amplitude and frequency of output vibration is the highest. shown in Figure 3. The ripe durian have velocity of vibration less than the young durian where V_t is velocity of vibration , ρ is the density of the medium , ω_b is the angular frequency of vibration ,and μ_1 and μ_2 are the coefficients of shear elasticity and shear viscosity , respectively.

Velocity of vibration related to frequency and wave propagation of vibration. So, property of shear elasticity can be found by velocity measurement of vibration and wave propagation at low frequency.

In (1) if the shear elasticity is dominant compared with the shear viscosity so that $\mu_1 >> \omega_b \mu_2$ is satisfied , the velocity is written as

$$V_t = (\mu_1 / \rho)^{1/2} \quad (3)$$

The amplitude and frequency of vibration determining have many methods such as ultrasonic and accelerometer. But laser doppler method can fix location of target to higher than accuracy and response of vibration. So, the laser doppler method is used in the vibration measuring of durian.

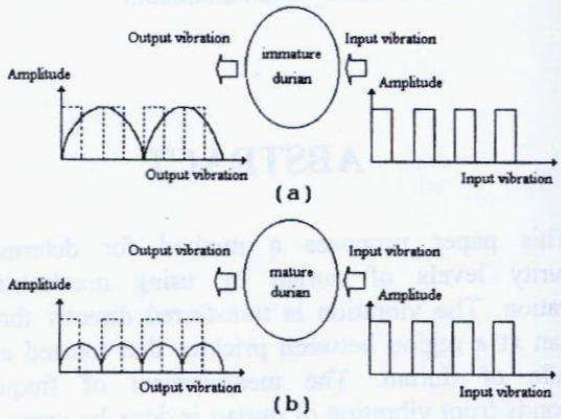


Fig. 2 Relationship between signal of input and output vibration in (a) immature durian (b) mature durian

For the durian with difference maturity , when transfer input vibration at linear amplitude and frequency. The signal of output vibration will have difference amplitude and frequency related to internal air cavity , mass and firmness follow to maturity level of durian. In Fig.2(a) the immature durian have amplitude and frequency of the signal output vibration higher than the mature durian, shown in Figure 2(b). Because the mature durian have more internal air cavity and elasticity than the immature durian , but have less density and firmness. So, the signal of output vibration have less amplitude and frequency than the immature durian.

3.2 Signal

The frequency of input vibration selected at linear output signal and highest amplitude response of the signal of the output vibration. From the result in Fig. 3. the error bar shows the deviation of the values measured at ten different points. We can see at the frequency 30 Hz. is the best. Because it can observe difference of highest velocity of vibration. The mature durian have less the velocity of vibration than the immature durian. And we can not observe the difference of the velocity of vibration when frequency have more than 45 Hz.

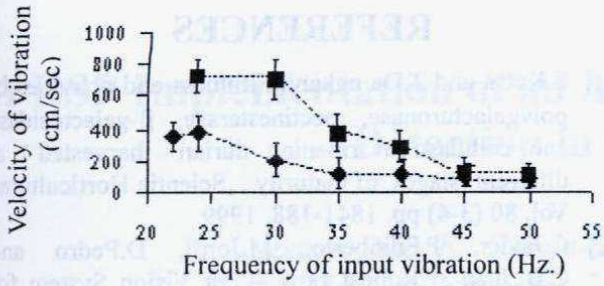


Fig. 3 Relationship between frequency of vibration and velocity of output vibration in
◆ mature durian , ■ immature durian

4. EXPERIMENTAL RESULTS

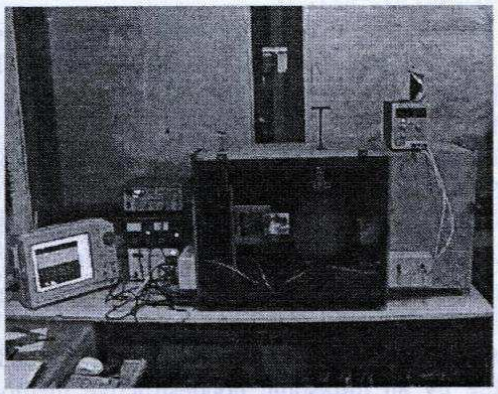


Fig. 4 The prototype system for determination maturity of durian by vibration

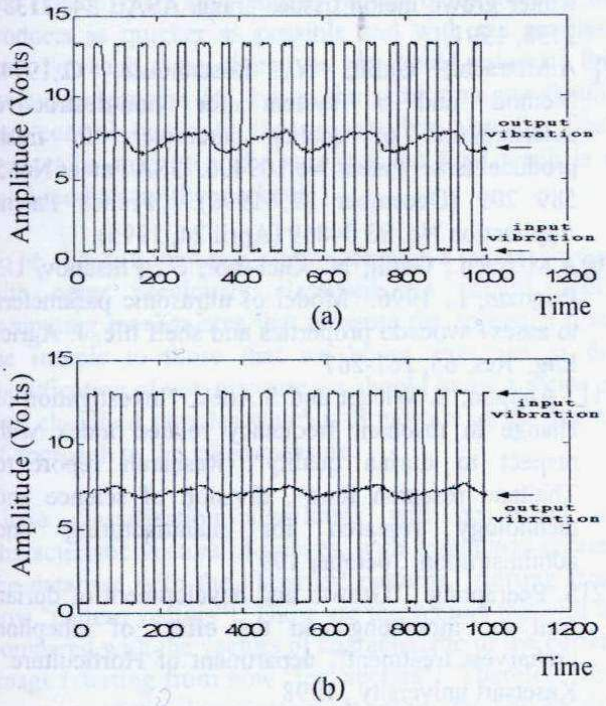


Fig. 5 Relationship between input and output vibration in (a) mature durian
(b) immature durian

The prototype system architecture for determination *maturity of durian by vibration is shown in Fig. 4. The* experimental results, 30 Hz input square wave and output sine wave, are shown in Fig. 5. We can observe that signals of output vibration in the different maturity of durian have different amplitudes, frequency, and velocity of vibration. Fig. 5(a) is signal of mature durian which has less amplitude and frequency than signal of immature durian in Fig. 5(b).

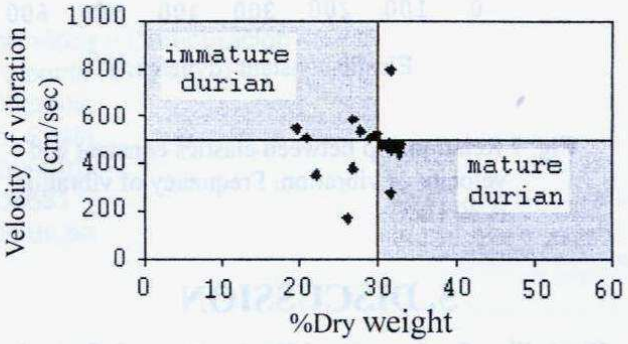


Fig. 6 Relationship between velocity of vibration and %dry weight in immature durian and mature durian

In experiment uses monthong durian 20 fruits, 3/2 shape, special class, weight 4-4.5 kg. The signal of output vibration is processed by computer program, when reference amplitude of signal of output vibration from the prototype is 4.6 V/cm. for determine amplitude, frequency, time interval , distance of vibration and velocity of vibration. Results of velocity of vibration are compared to %dry-weight for relation determination between velocity of vibration and % dry-weight shown in Fig. 6. This relation is compared to the maturity determination by using an Effegi firmness tester with a 5 mm. plunger tip which was inserted to a depth of 0.5 cm. and the force neccesary to do this recorded in newtons for threshold value determination of velocity of vibration. We can set threshold lines at velocity of vibration is 500 cm/sec and 30% dry-weight for results classification of the experiment is best three groups , immature durian and mature durian and error. And it can be classified with 80% of accuracy

Durian having different maturity will also have different characteristics of structure, hard or soft. Due to incresing of quantity of water depends on maturity of durian, so in (3) density and elastic constant of durian will be important signs to classify the maturity. From Fig. 7. the threshold line for classification of three results groups, as immature durian, mature durian, and two error groups, is obtained. The threshold line of elastic constant and velocity of vibration are 300 $(dyne/cm)^{1/2}$ and 500 cm/sec. Respectively. The efficiency of proposed classification method is 95%.

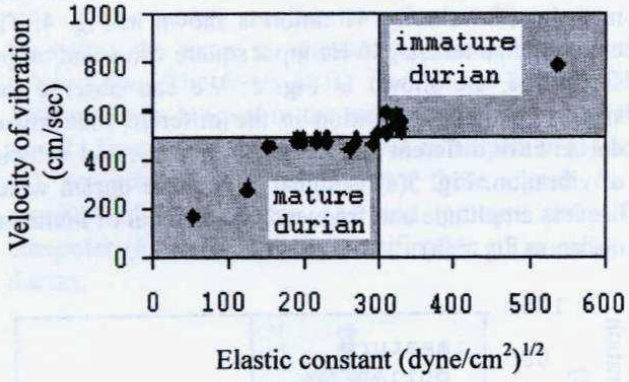


Fig. 7 Relationship between elastics constant and velocity of vibration. Frequency of vibration is 30 Hz.

5. DISCUSSION

From Fig. 6., the error 20% is occurred from the reason that the tip of durian has broken. This makes the experimental results are in a region of high %dry-weight and high velocity of vibration. Not only that, some durian having narrow region between prickles and uncomplete petals makes the results are in the region of high %dry-weight and low velocity of vibration.

To reduce this error, therefore, the experiment of durian having narrow region between prickles is done by cutting their prickles to make sufficient region for measurement process with laser doppler. And uncomplete petal durian will be vibrated at a region as near the middle of petal as we can. For the case of durian with broken tips, we must hold the below of tip of durian with the prototype strongly. The end of T-needle must touch tightly to a region between prickles located at the middle of petal, which has the largest region. The below of durian must touch to the rubber base at least but it must touch enough to control rotating of durian.

From Fig. 7., the error comes from external light sources which make the signals are error. The cover of the prototype is used to solve this problem. The receiver set must be set in the line of the vibration set.

6. CONCLUSIONS

Durian is one of the high economical fruits which requires nondestructive guarantee of their quality. This paper presented nondestructive maturity determination of durian by vibration at low frequency and measure output vibration by laser doppler. The signal of output vibration are taken for automatic maturity classification of relationship between velocity of vibration and elastic constant. And it can be classified into two groups, immature durian and mature durian, with 95% of accuracy rate.

REFERENCES

[1] S.Ketsa and T.Da ngkanit,Firmness and activities of polygalacturonase, pectinesterase, β-galactosidase and cellulase in ripening durian harvested at different stages of maturity . ,Scientia Horticulturae Vol. 80 (3-4) pp. 1841-188, 1999

[2] C.Javier, P.Filioberto, M.Jordi, D.Pedro and C.B..Juan , "Robust Low –Cost Vision System for Fruit Grading", IEEE conference on Electrotechnical, vol .3, pp. 1710–1713 , 1996

[3] Q.Yang, "apple stem and Calyx Identification with Machine Vision", J.agric. EngngRes, vol.63, pp.229-236, 1996

[4] D.C.Slaughter, D.Barrett, M.Boersig, 1996. Nondestructive determination of soluble solids in tomatoes using near-infrared spectroscopy. J. food Sci. 61, 695-697.

[5] T.F.Schatzki, R.P.Haff, R. young, I.Can, Le, L.c., x-ray imaging. Sensors for nondestructive Testing : Proceeding Sensors for Nondestructive Testing International Conference, Orlando, FL, 18-21 February 1997. NRAES (Northeast Reg Agric Eng Svc) Coop Extn, Ithaca, NY , pp. 161-171

[6] C.J.Clark, P.D.Hockings, D.C. Joyce, R.A.Mazucco, Application of magnetic resonance imaging to pre- and post-harvest studies of fruit and vegetables. Postharvest Biol. Technol. 11, 1-21, 1997

[7] A.Mizrach, U.Flitsanov, Predicting avocado shelf life by an ultrasonic nondestructive method. Acta Hortic., no. 421, 1996.

[8] A.Mizrach, N.Galili, G.Rosenhouse, D.C.Teitel, Acoustical, mechanical and quality parameters of winter grown melon tissue. Trans. ASAE 34, 2135-2138, 1991.

[9] A.Mizrach., Galili, N., Rosenhouse, G.,1994, Method and a system for non-destructive determination of quality parameters in fresh produce. Israel Patent No.109406. USA Patent No. 5 589 209 (December 31, 1996). French Patent Application No. 95 04869 (April 24, 1995).

[10] A.Mizrach., Galili, N., Gan-mor, S., Flitsanov, U., Prigozin, I., 1996. Model of ultrasonic parameters to assess avocado properties and shelf life. J. Agric. Eng. Res. 65, 261-267.

[11] T.Anupun, S. Wichet and S.Anek, "Investigation of change in resonant frequency related index with respect to durian quality", Research report to Thailand research fund , division of science and technology research for manufacturing and administration. , october 1997

[12] S. Peerapong , "Growth and development of durian fruit cv. monthong and the effect of ethephon preharvest treatment", department of Horticulture , Kasetsart university , 1998

[13] H.L. Oestreicher, "Field and impedance of an oscillating sphere in a viscoelastic meduim with an application to biophysics", J.Acoust .Soc. Am., vol .23, pp.707-714, June 1951 .

A DSP Implementation of an AOM and its Application to Defects Detection in Textile Material

Ibarra-Picó, F.; Garcia-Crespi, F.; Cuenca-Asensi, S.A.; Morales-Benavente, J.L.; Lorenzo-Quintanilla J.J.

Departamento de Tecnología Informática y Computación
Campus de San Vicente
Universidad de Alicante
E03080 - Alicante, Spain
Phone: +34-965909387
Fax: +34-965903681
Email: ibarra@dtic.ua.es

Abstract. This paper explains a method of defects detection in textile material using a DSP. This **Supervised Learning** method will allow the detection of defects in anyone of the phases of production. An algorithm of pattern classification based on minimum distance is used to carry out this method. Scalar distance in an **AOM network** is used to provide a measure of the angle which form the 2 compared vectors too. Finally, some experiments have been carried out.

1 INTRODUCTION

The objective of any industry is the fabrication of its products as quicker as possible and with the smallest cost. To reach these objectives we should automate the production process **[4]**, but at the same time we should also control not only the process but also the final product in order to make sure that the final result is a valid product for the market.**[1]**.

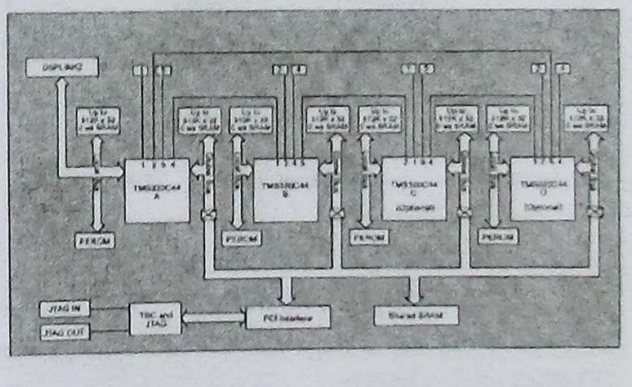
To be able to compare a sample taken in a certain instant with other previously classified we should apply computing mechanisms that measure the approach from the sample to those that we know. Previous to the classification of any measure we should know a series of well classified samples which will indicate us the degree of approach to the taken measure.

These classification algorithms will be based on characteristic vectors of images, with which was created the database with the classified patterns, (starting from now "pattern vectors"). These are the vectors that will be compared with the vectors of characteristic of the current image (starting from now "test vectors"). Therefore, and by means of the classification algorithm, we will be able to know to which pattern vector approaches more each test vector, and, consequently, if we know that the pattern vector contains defects, we could, in the same way, say that the test vector contains the same or similar defects.

2 DEVELOPMENT SYSTEM

To carrying out the experiments of defects detection in textile stuff based on AOM distance, we have worked with a card of digital signal processing (DSP). It is a Texas Instruments's card named PCI/C44S **[2]** with two processors C44 (that works at 50MHz) which works in parallel to carry out the high speed classification of the captured images **[5]**.

The diagram of blocks of the PCI/C44S card is the following one:



3 AOM CLASSIFIER

To compare what they look like each other two vectors whose values are normalised, this algorithm obtains a AOM distance among them, as the value of the scalar product of the vectors. That is to say, the degree of similarity among the vectors will come given by the angle that they form between them. The distance obtained is always a value between 0 and 1, since this value is the cosine of the angle that the two vectors form. As the values of the vectors has been previously

© UkrObraz'2000

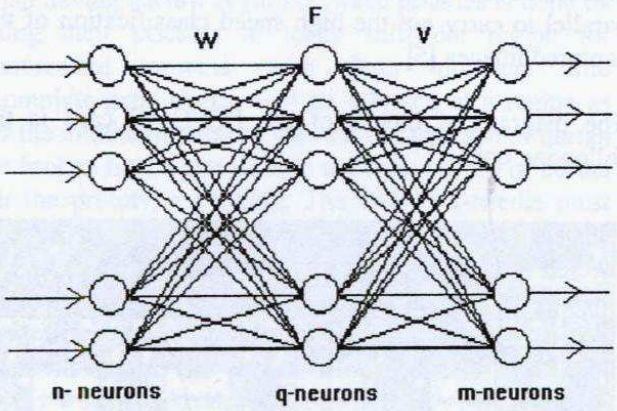
normalised, the modules of the vectors are rejected simplifying vastly the calculations to carry out.

Let a set of **q** patterns pairs $(\mathbf{a_i}, \mathbf{b_i})$ for $\mathbf{i=1, .., q}$ belonging to the vector spaces $\mathbf{R^n}$ and $\mathbf{R^m}$, where the patterns $\mathbf{a_i}$ are normalised in same way so $|\mathbf{a_i}| = \mathbf{k}$ where $\mathbf{k} \in \mathbf{R}$ and $\mathbf{i = 1..q}$. We will build two learning matrices in the following way:

$$A = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{q1} \\ a_{12} & a_{22} & \cdots & a_{q2} \\ & \cdots\cdots & & \\ a_{1n} & a_{2n} & \cdots & a_{qn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{21} & \cdots & b_{q1} \\ b_{12} & b_{22} & \cdots & b_{q2} \\ & \cdots\cdots & & \\ b_{1m} & b_{2m} & \cdots & b_{qm} \end{bmatrix}$$

where **n** dimension patterns are the columns of matrix **A** while **m** dimension patterns are the columns of matrix **B**.

The memory is built as a neural network (see figure below) with two synaptic matrices (containing hebbian correlations) **W** and **V**, which are computed $\mathbf{W = AQ^t}$ and $\mathbf{V = QB^t}$. Where **Q** is an intermediate orthogonal matrix (Walsh, Householder, and so on) of dimensions **qxq**. The $\mathbf{q_i}$ vectors of matrix **Q** are an orthogonal base of the vector space $\mathbf{R^q}$.



If **Q** is a Householder matrix (built as $\mathbf{Q = I - 2uu^t}$ where $\mathbf{u^t = [2/q\ 2/q\ ...\ 2/q]}$) then the intermediate filter (**F**) in the hidden layer is

$$F(x_i) = \begin{cases} 1 - 2 / q & x_i \geq 0 \\ -2 / q & x_i < 0 \end{cases}$$

where only the winner neurone $\mathbf{x_i}$ which is associated to the $\mathbf{a_i}$ pattern rises to **1 - 2/q**.

As a result, the neural network has an entrance level with **n** neurones, a hidden level with **q** neurones and an exit level with **m** neurones. Each processing element of a layer is connected with all the processing elements of the following layer. The AOM network can expand by backtracking with $\mathbf{W^t}$ and $\mathbf{V^t}$, what allows its bi-directional use.

4 EXPERIMENTS

Considering the appropriate experiments to carry out in order to measure the power of the programs and the card with which it works, it will be formed 3 series of experiments, two of which will be commented later on.

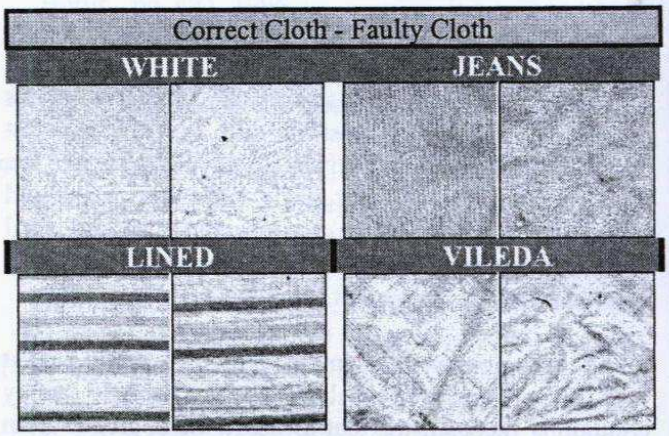
With regard to the AOM classifier, and in order to reduce not only the necessary operations but also the processing time, we have eliminated the hidden layer of the AOM network so that the calculation of the AOM distance is reduced to:

$$V_A = (a_1, a_2, a_3, \ldots, a_n) \qquad \text{AOM distance} = \sum a_i \cdot b_i$$
$$V_B = (b_1, b_2, b_3, \ldots, b_n) \qquad \forall i \in 1,2,\ldots n$$

To finish, we are going to show the parameters that have been used in the AOM classifier:

Parameters of AOM classifier	
Number of patterns	4 x 10
Distance method	Scalar Product
Filter f1	Scalar Product
Filter f2	Normalisation

4.1 Preparation Of Experiments



For the experiments carried out in this project, a battery of 40 images it has been created by scanning 4 different types of cloths (white cloth, jeans cloth, lined cloth, and "vileda" cloth). All the images have a size of 512x512 pixels, 256 grey levels and a 100 dpi resolution.

When creating the series of experiments, and because the group of samples is reduced (10 images for each cloth), we have considered the method **Leaving One Out**

This method consists on the use of a sample as test, and the other ones as training repeating the experiments, leaving every time a different sample, that's why its name.

The error rate will be the proportion between the number of fault experiments and the number of total experiments.

Image size	n x n
Window size	m x m
Number of Images	K
Number of Classifications x Image	$(n/m)^2$

$$T_c = \frac{\sum_{i}^{k} \frac{n^o\ errors}{(n/m)^2}}{k}$$

4.2 Series Of Experiments

Next, the results of each one of the carried out series of experiments are shown. For all the carried out series, we have measured the time of execution and the error rate.

4.2.1 Series 1

In the first series, it has been carried out experiments on the four different types of cloths and applied the following parameters for the MC calculation:

Windows size:	64 x 64					
Planes:	5					
Limits:	-1	32	64	96	128	255
Pieces:		2x2	3x3	4x4		

The previous parameters have been chosen to comparing them with the results of the following series that it will have less number of piece sizes. The obtained error rate has been:

Cloth Type				
Image	Jeans	White	Lined	Vileda
0	0	0	1	0
1	4	1	9	0
2	6	0	8	0
3	6	0	6	0
4	3	0	16	0
5	7	1	12	1
6	2	0	20	1
7	0	0	14	2
8	1	0	19	1
9	0	3	16	2
Total Errors	29	5	121	7
Error Rate	4.53%	0.78%	18.91%	1.09%
Execution Time	4.801s	4.833s	4.725s	4.831s

We have obtained disparate results depending on the cloth type we are analysing, having the highest error rate for the most complex cloth type, that is to say, the lined cloth.

As far as the execution times is concerned, they are practically equal because they won't depend on the complexity of the cloth to analyse but on the MC parameters that will be considered.

4.2.2 Series 2

For the second series, we have changed the MC parameters with regard to the first series and we have considered different piece sizes to process for each image, since we wanted to observe the influence that this change would suppose in the error rate as well as in the execution times.

Windows size:	64 x 64					
Planes:	5					
Limits:	-1	32	64	96	128	255
Pieces:		2x2				

We have only considered a 2x2 piece size for each window of each image and we have obtained the following results:

Cloth Type				
Image	Jeans	White	Lined	Vileda
0	0	0	0	0
1	2	1	10	0
2	3	0	17	0
3	6	0	9	0
4	4	0	17	0
5	5	0	17	0
6	2	0	15	1
7	0	0	19	2
8	1	0	19	1
9	1	3	17	3
Total Errors	24	4	140	7
Error Rate	3.75%	0.63%	21.88%	1.09%
Execution Time	2.859s	2.850s	2.826s	2.853s

If we compare the error rate of this series with those of the first series, we can see that the results are pretty equal. The only difference is that these results have increased because we have lost certain information in the MC calculation.

Now, we are going to study the behaviour of the execution times of the programs with these parameters. As we can observe the times have decreased notably due to the decrease of the complexity in the MC calculation.

4.2.3 Series 3

For the series 3 we have wanted to verify the importance that suppose either the number of planes and the limits selection. These limits have been selected in a homogeneous way, that is to say all the planes have the same size (they embrace the same number of grey tones). It has been decreased the number of planes with regard to the previous series from 5 to 4 planes.

Windows size:	64 x 64					
Planes:	4					
Limits:	-1	64	128	192	255	
Pieces:		2x2	3x3	4x4		

The error rate and the execution times obtained are the following:

Cloth Type				
Image	Jeans	White	Lined	Vileda
0	0	0	0	0
1	9	10	11	14
2	9	3	19	11
3	9	9	9	9
4	12	8	18	12
5	23	1	22	11
6	15	2	20	13
7	14	1	26	12
8	15	3	25	13
9	21	2	18	15
Total Errors	127	39	168	110
Error Rate	19.84%	6.09%	26.25%	17.19%
Execution Time	3.557s	4.408s	4.108s	4.218s

When decreasing the number of planes to analyse, the error rate has suffered a notable increment in all the analysed types of cloths. Therefore the number of planes and the disposition of the same ones supposes one of the main factors to reach acceptable results.

The times of execution have also suffered a slight variation, decreasing in relation to the first series because the number of planes is lower. However, as we can see, the results follow the same rules that the previous series.

4.3 Final Results

To summarize, we will expose the results obtained in the three exposed series of experiments.

To obtain the answer time necessary to evaluate a window, we have to divide the total execution time of the complete image by the number of processed windows. In our case, the number of windows is always the same and it is equal to 64 (an 512 x 512 image can be divided in 64 windows of 64 x 64).

	Success Rate	Total Execution Time	Answer Time
Series 1	93.67%	4.797 s	0.075 s
Series 2	93.16%	2.847 s	0.044 s
Series 3	82.66%	4.072 s	0.064 s

5 CONCLUSIONS

Once commented the experiments, we can extract some conclusions that guide us in later tests. In general, in anyone of the cases we can appreciate that the method doesn't require an excessive processing time; producing rates of success around 94%. These results are quite encouraging if we keep in mind that it has been analysed some complex cloth types (such as lined cloth).

To finish, and since the results obtained both in error rate and in execution times have been quite good, the application of this method can be very advantageous, moreover knowing that the development environment used is relatively simple.

6 REFERENCES

[1] Badía Jiménez, Albert (1998) "*Calidad: Enfoque ISO 9000 Normalización, homologación, certificación, acreditación, aseguramiento (ISO 9000) y auditoria*" Ediciones DEUSTO.

[2] LSI plc – "*PCI/C44S Board User's Manual*"

[3] LSI plc – "*NETAPI C4x – User's Manual*"

[4] Bodson, Marc (November 1994). "*Emerging Technologies in Control Engineering*" IEEE Control Systems, p. 10-12.

[5] Neubauer, C. (1992). "*Segmentation of defects in textile fabric*", Proc. IEEE, p. 688-691.

[6] Pal N.R. and Pal S.K. (1993). "*A review on image segmentation techniques*", Pattern Recognition, Vol. 26, No.9, p. 1277-1294.

[7] Ibarra-Picó, F. and García-Chamizo, J.M. (1993) "*Bi-directional Associative Orthonormalized Memories*". Actas AEPIA, vol 1, p 20-30.

[8] Ibarra-Picó, F., García-Chamizo, J.M. and Satorre-Cuerda, Rosana (1995) "*An Orthogonal Neural Network with Guaranted Recall by Iterative Filters and its Application to Texture Discrimination*". Artificial Neural Nets and Genetic Algorithms, p 186-189.

ДО ПРОБЛЕМИ ТЕХНІЧНОЇ ДІАГНОСТИКИ ТОВСТОЛИСТОВИХ ВИРОБІВ МЕТОДОМ УЛЬТРАЗВУКОВОЇ ТОМОГРАФІЇ

В.В.Кошовий, Є.В.Кривін, В.П.Курій, О.Є.Левицький, І.М.Романишин,
Фізико-механічний інститут ім. Г.В. Карпенка НАН України
79601, м. Львів, МСП, вул. Наукова, 5
тел: (0322) 637-218, електронна пошта: Koshovy@ah.ipm.Lviv.ua

Розглядається принцип і метод ультразвукової реконструктивної томографії для відновлення просторового розподілу фізико-механічних характеристик матеріалу (ПР ФМХ) в товстолистових виробах при односторонньому їх прозвучуванні дзеркально-тіньовим методом стосовно задачі технічної діагностики якості зварного з'єднання та матеріалу в його околі.

ВСТУП

Підвищення рівня технічної діагностики (ТД) вимагає не тільки визначення дефектності виробу, але і оцінки переддефектного стану. Для оцінки переддефектного стану виробу необхідна інформація про просторовий розподіл (ПР) фізико-механічних характеристик (ФМХ) і параметрів напружено-деформованого стану (НДС). Один з ефективних методів експериментального визначення ПР ФМХ і параметрів НДС полягає в застосуванні ультразвукової томографії [1].

Аналіз стану неруйнівного контролю (НК) і ТД виробів дозволив зупинитись на необхідності контролю товстолистових виробів при односторонньому доступі, які на даний час є найбільш актуальними в багатьох галузях науки і техніки. Приведені дальше результати теоретичних досліджень, чисельного моделювання та експериментальних досліджень свідчать про те, що найбільш підходящими об'єктами для НК запропонованими томографічними методами є зварні з'єднання, які мають достатньо високочастотний характер розподілу ФМХ матеріалу в площині виробу (XOY). Задача експериментального визначення ПР ФМХ матеріалу та параметрів НДС у виробах із зварними з'єднаннями на даний час далека від задовільного рішення і є актуальною, особливо враховуючи суттєвий неоднорідний розподіл параметрів ФМХ у зварному з'єднанні.

В роботі задача експериментального визначення ПР ФМХ матеріалу розв'зується шляхом побудови томографічних зображень ПР акустичних характеристик - швидкостей поширення та загасаня різних типів УЗ хвиль (повздовжніх, поперечних, поверхневих) та обчислення на основі цього ФМХ матеріалу (пружних, міцністних, структурних, технологічних, а також параметрів НДС). ПР акустичних характеристик матеріалу в переддефектному стані є слабоконтрастним, тому можна обмежитись наближенням геометричної акустики при визначенні шляху поширення сигналу.

Особливості УЗ контролю, пов'язані з труднощами забезпечення акустичного контакту, не дозволяють використати ефективні алгоритми томографічного відновлення, розроблені в рентгенівській томографії. Для забезпечення високої продуктивності контролю та універсальності щодо форми контрольованого виробу розроблені принципи побудови томографічних методів і систем в ультразвуковому НК, які полягають в наступному:

- збір проекційних даних проводиться променевими проекціями;
- томографічне відтворення проводиться методом оберненого проектування і побудови сумарного зображення (СЗ), яке будується в процесі збору даних і служить пошуковим для швидкого визначення аномальних областей, що будуть підлягати більш детальному дослідженні на наступних етапах;
- побудова шуканого ПР здійснюється шляхом прямого і оберненого проектування СЗ та фільтрації проекцій СЗ.

Одним з методів ультразвукового контролю листових виробів при односторонньому доступі є дзеркально-тіньовий метод (ДТМ). Для дзеркально-тіньового методу характерна обмеженість ракурсу збору даних. В літературі раніше не розглядались можливості ДТМ в задачах томографічного відновлення. В даній роботі такий розгляд приведений.

1.ТЕОРЕТИЧНІ РЕЗУЛЬТАТИ.

Ефективність роботи томографічної системи та алгоритмів відновлення визначаються схемою збору даних.

При односторонньому доступі збір проекційних даних (прозвучування) в кожній точці скануючої поверхні листового виробу можна здійснювати за

© UkrObraz'2000

різними схемами: із зміною напрямів зондування по азимутальному куту (φ) та ракурсу (θ), а також при їх комбінації. Подальший теоретичний аналіз, чисельне моделювання та експериментальна апробація підтвердили оптимальність одноракурсної схеми збору даних в конусі.

На основі проекційної теореми проведено аналіз можливостей одноракурсної схеми збору даних в конусі. Отримано, що при такому зборі даних відсутні просторові спектральні складові, що належать конусу з просторовим кутом при вершині $\pi/2-\theta$.

З проекційної теореми слідує, що найбільше інформації міститься при найбільшому ракурсі збору даних. Отже, можна обмежитися одним ракурсом збору даних. Після вибору схеми збору даних головною задачею стає отримання співвідношення між шуканим ПР та відновлюваним згідно вибраної схеми збору дпних.

Сумарне зображення, яке в загальному вигляді записується як

$$b(\vec{r}) = \hat{K}_3^{\theta_0}\{g(\vec{r})\} = \int_{-\infty}^{\infty}\int_{(2\pi)} g(\vec{r}+\vec{m}k)\delta(\vec{m}\cdot\vec{n}_{\theta_0})d\Omega dk,$$

можна представити у вигляді:

$$b(x,y,z) = g(x,y,z)***\frac{\delta(z\sin\theta_0 - \cos\theta_0\sqrt{x^2+y^2})}{\sqrt{x^2+y^2+z^2}},. \quad (1)$$

де $K_3^{\theta_0}$ - оператор оберненого проектування,

$\vec{n}_{\theta_0} = (-\cos\varphi\cos\theta_0,\ -\sin\varphi\cos\theta_0,\ \sin\theta_0)^T$,

$g(\vec{r})$ - ПР неоднорідності.

Отримати формули інверсії можна досить легко, якщо перейти в спектральну область. Згідно співвідношення (1) отримуємо:

$$B(u,v,w) = G(u,v,w)x\begin{cases}\dfrac{4\pi}{\sqrt{u^2+v^2-w^2tg^2\gamma_0}}, & |wtg\gamma_0| < \sqrt{u^2+v^2},\\ 0, & |wtg\gamma_0| \ge \sqrt{u^2+v^2}\end{cases} \quad (2)$$

Із (2) випливає, що можна відновити тільки функцію G^*, яка із спотвореннями відображає шукану функцію:

$$G^*(u,v,w) = \begin{cases}G(u,v,w), & |wtg\gamma_0| < \sqrt{u^2+v^2},\\ 0, & |wtg\gamma_0| \ge \sqrt{u^2+v^2}\end{cases}.$$

Звідси можна отримати:

$$g^*(x,y,z) = \frac{1}{4\pi}\Im_3^{-1}\left\{\Im_3\{b(x,y,z)\}\sqrt{u^2+v^2-w^2ctg^2\theta_0}\right\}. \quad (3)$$

Як видно із отриманого співвідношення (3), для відновлення шуканого зображення розподілу необхідно використовувати операції тримірного перетворення Фур'є, що є досить трудоємкою процедурою для прикладної реалізації. Тому був розроблений більш ефективний алгоритм відновлення шуканого розподілу на основі СЗ.

Більш ефективним є застосування алгоритму відновлення ПР на основі операцій оберненого проектування проекцій СЗ:

$$g^*(\vec{r}) = -\frac{1}{(4\pi)^2}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}-ctg^2\theta_0\frac{\partial^2}{\partial z^2}\right)\hat{K}_3^{\theta_0}\{b(\vec{r})\}$$

Приведений вище алгоритм вимагає наступних обчислювальних процедур:

1. проектування СЗ у вигляді паралельних проекцій;
2. фільтрація отриманих проекцій шляхом подвійного диференціювання;
3. обернене проектування по фільтрованим проекціям.

2.ЧИСЕЛЬНЕ МОДЕЛЮВАННЯ

З метою вибору оптимального ракурсу прозвучення та дослідження можливостей розроблюваної системи томографічного відновлення проводилося чисельне моделювання. Досліджувався вплив на якість відновлюваного зображення величини кута ракурсу, неповноти спектрального складу та збільшення числа ракурсів прозвучування.

Чисельне моделювання підтвердило можливість обмеження проозвучування одним ракурсом. В результаті чисельного моделювання отримано, що для якісного відновлення ПР ФМХ матеріалу по товщині виробу необхідно мати достатньо високочастотний розподіл ФМХ матеріалу в площині виробу, що характерно для виробів із зварним з'єднанням. Чисельні експерименти показали, що при відновленні ФМХ матеріалу у зварних з'єднаннях можна обмежитися ракурсом прозвучування, рівним 60^0.

3.РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТАЛЬНОЇ АПРОБАЦІЇ.

На рис.1 приведена блок-схема УЗ томографічної системи UST-2000 збору іобробки даних.

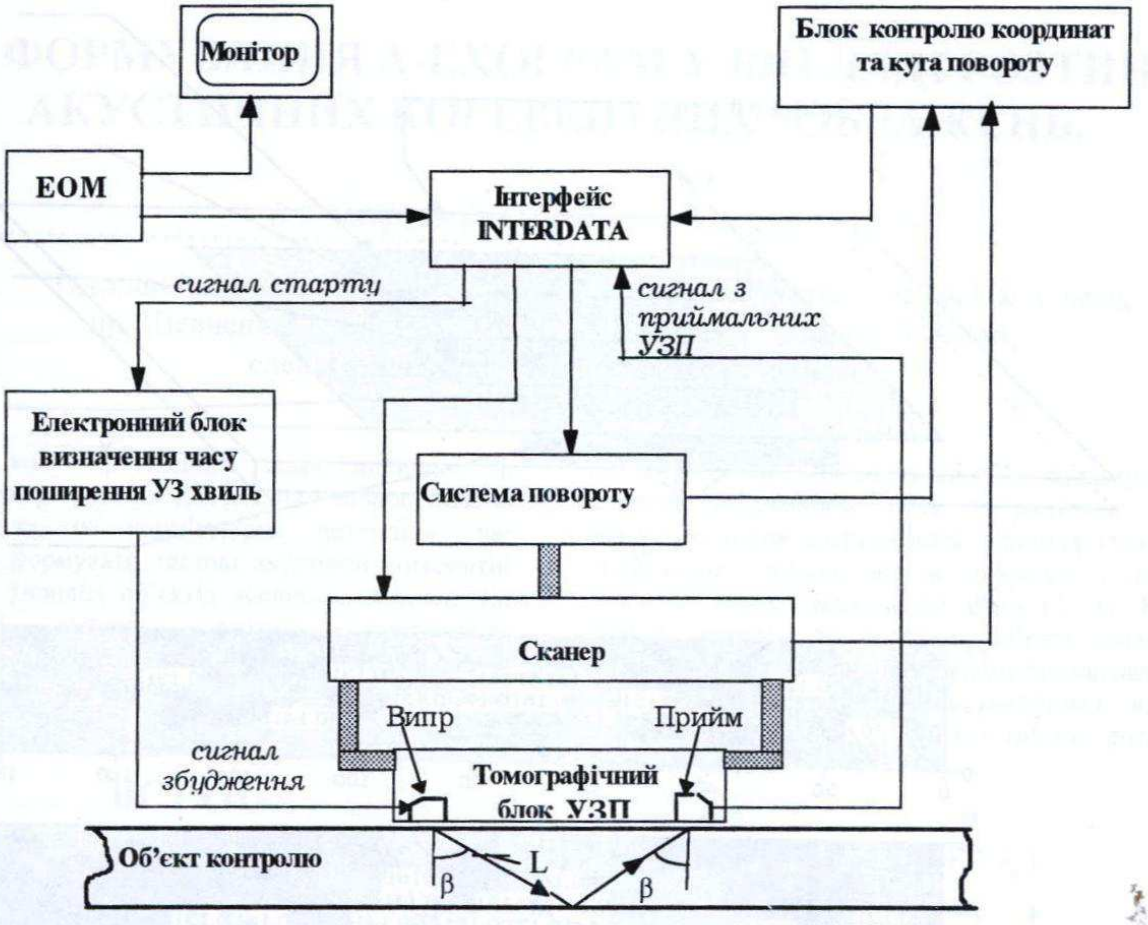


Рис.1. Структурна схема томографічної системи UST-2000

Томографічна система включає в себе наступні основні блоки: ЕОМ з монітором , інтерфейс обміну даних та сигналів, електронний блок визначення часу поширення УЗ хвиль в об'єкті контролю, багатоканальний томографічний блок УЗ перетворювачів (УЗП), скануючий пристрій 3D-SCAN, що забеспечує сканування по поверхні об'єкту прозвучування томографічним блоком УЗП і поворот останнього на заданий кут, оптоелектронний блок контролю координат геметричного центру і кута повороту томографічного блоку УЗП в полі контролю.

На рис.2б, як іллюстрація томографічного відновлення ПР ФМХ згідно розроблених алгоритмів відновлення, представлено томографічне зображення ПР розміру зерен D (мм) вздовж шва зварного з'єднання в п'яти характерних вертикальних січеннях згідно рис.2а. Отримані зображення свідчать про високі можливості розробленої томографічної системи.

Розроблені у ФМІ НАН України томографічна система UST-2000 та прикладний пакет програм томографічної реконструкції дозволяє визначити неруйнівним методом просторовий розподіл пружних (модулі пружності), міцністних, технологічних, структурних характеристик матеріалу та параметрів неоднорідного напружено-деформованого стану (тензор напружень та його інваріанти).

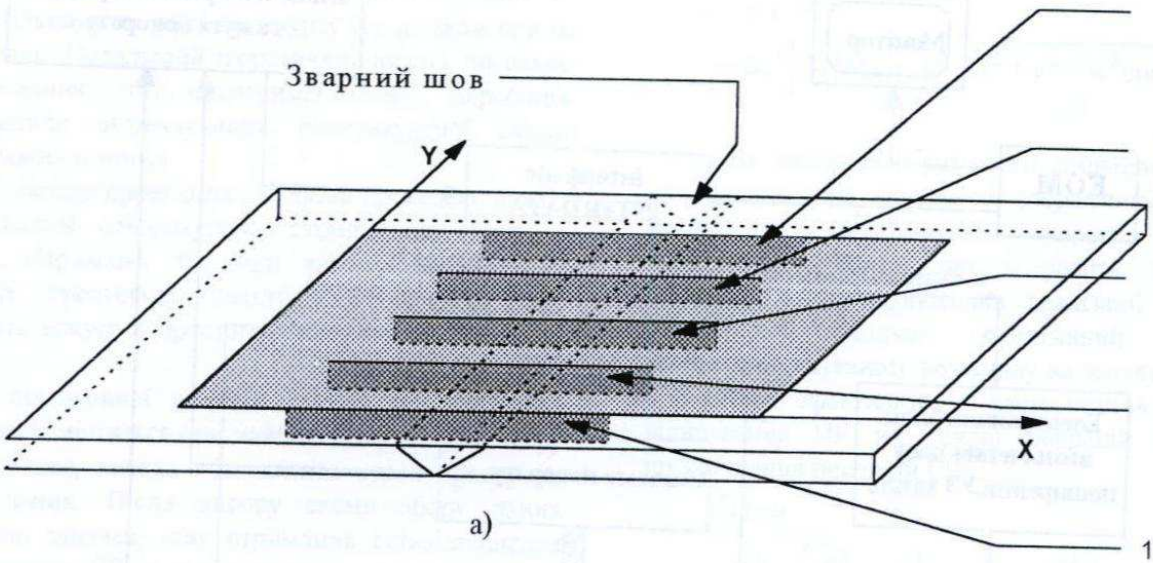
Розробка проводилась в рамках проекту УНТЦ #612/

ТЕХНІЧНІ ХАРАКТЕРИСТИКИ ТОМОГРАФІЧНОЇ СИСТЕМИ:

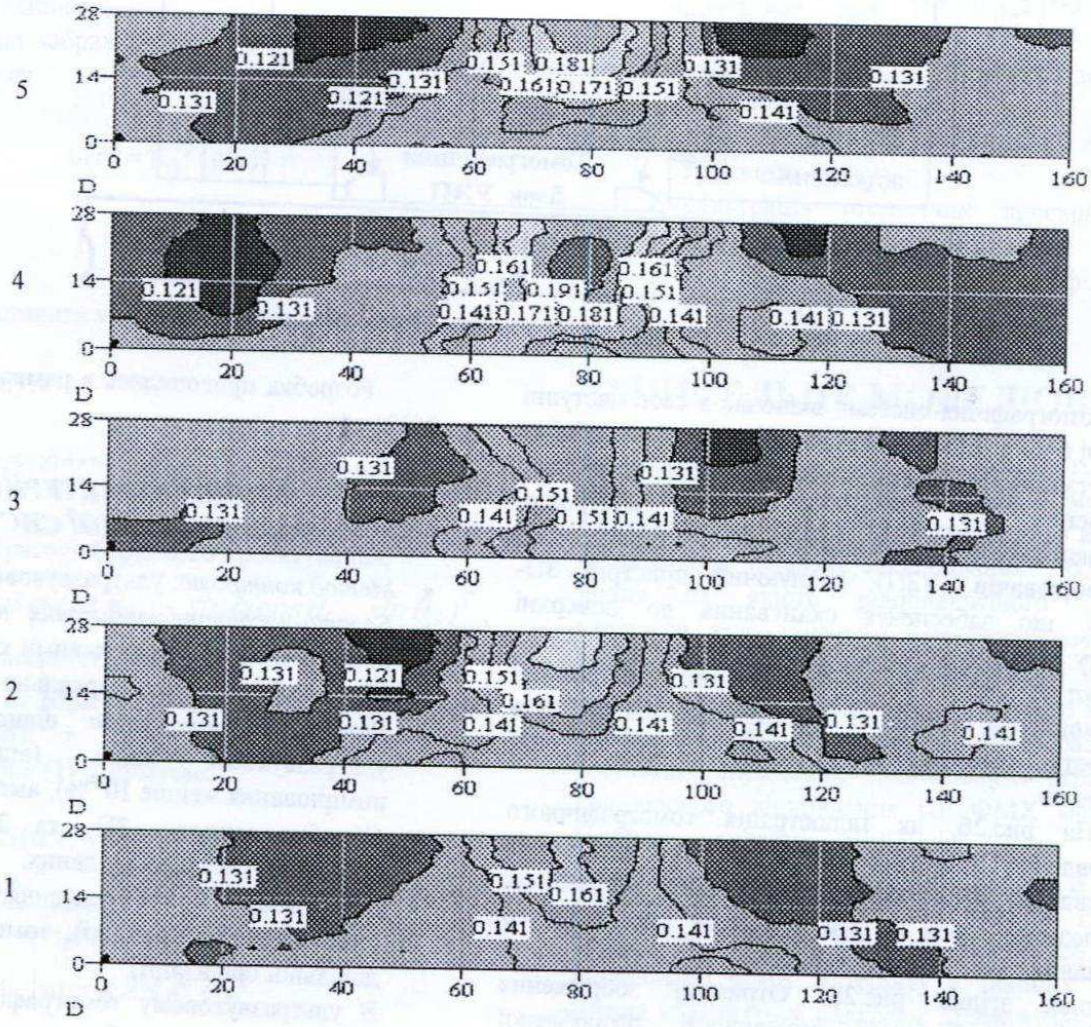
- *Метод контролю*: ультразвуковий.
- *Багато-частотне зондування:* повздовжніми, поперечними та поверхневими хвилями.
- *Метод збору даних:* дзеркально-тіньовий.
- *Вимірювані параметри:* швидкість поширення ультразвукової хвилі (відносна похибка вимірювання менше 10^{-4}%), амплітуда.
- *Обробка сигналу:* 2D- та 3D- томографічна променева схема збору даних.
- *Відновлювані зображення:* ортографічне (усереднене по товщині); томографічне (січення довільної орієнтації).

В ультразвуковому томографі використовується універсальна система збору даних, яка забезпечує швидку переналадку на контроль виробів різноманітної форми.

Можливі області застосування: фундаментальні дослідження матеріалів, атомна енергетика, ракетно-космічна і авіаційна техніка, суднобудівництво, обладнання хімічного, нафтохімічного та газового виробництва, контроль адгезійної міцності покрить неруйнівними методами.



а)



б)

Рис.2. Відновлене томографічне зображення ПР розмірної структури зерен в різних січеннях вздовж зварного шва у виробі листової форми (крок січень 60мм)

ЛІТЕРАТУРА

1. Кошевой В.В., Кривин Е.В., Романишин И.Н Томографія листовых изделий // *Дефектоскопия.-1998, №9.*-С.-20-27.

ФОРМУВАННЯ А-ЕХОГРАМ У ВИГЛЯДІ РОЗТИНУ АКУСТИЧНИХ КОГЕРЕНТНИХ ЗОБРАЖЕНЬ.

В.Ф. Миргород, С.М. Ковешніков, А.Г. Кисель

Одеський державний політехнічний університет, Україна, 65044, м. Одеса,
пр. Шевченка, 1; тел.: +380 482 250-104, факс: +380 482 343-848;
електронна пошта: acrobat@ospu.odessa.ua

Для кола прикладних задач медичної та технічної ультразвукової діагностики запропоновано алгоритми та їх комп'ютерна реалізація, що дозволяють формувати тестові акустичні когерентні зображення типових об'єктів дослідження у вигляді А-ехограмм: кутового розтину акустичного зображення. Виконано комп'ютерне моделювання таких ехограм в порівнянні з експериментально одержаними.

ВСТУП

Ультразвукова медична та технічна діагностика є визнаним методом виявлення аномальних об'єктів у досліджуваному середовищі. Прогрес у розвитку технічних засобів ультразвукової діагностики дозволяє одержувати високоякісні акустичні зображення об'єктів у вигляді кутових розтинів (А-ехограм), або у двомірному вигляді [1]. Подальша їх класифікація становить певну проблему у зв'язку з невідповідністю акустичних когерентних зображень об'єктів дослідження їх оптичним аналогам [2]. Така невідповідність зумовлена когерентним характером зондуючих коливань та специфікою їх розсіювання на поверхні об'єктів складної форми. Саме тому акустичні когерентні зображення об'єктів дослідження досить часто мають мало загального з їх відповідними оптичними зображеннями. Для розробки автоматичних класифікаторів та оцінки їх ефективності необхідно мати можливість формувати навчаючі тестові послідовності, наприклад у вигляді А-ехограм, що вірогідно відтворюють особливості створення вторинного когерентного акустичного поля поверхнею об'єкта класифікації. Такі ехограми пропонується формувати у вигляді кутових розтинів відповідних акустичних когерентних зображень як результат просторово-часової звіртки зондуючого поля $S(t,\vec{r})$ та функції розсіювання (ФР) об'єкта класифікації $h(t,\vec{r})$.

1. МЕТОДИКА ФОРМУВАННЯ А-ЕХОГРАМ

Когерентний характер зондуючих коливань зумовлює локальний характер процесів розсіювання акустичних хвиль на поверхні об'єкта, тому основна енергія вторинного поля формується деякою кількістю точок дзеркального відбиття (так званих "блискучих" точок), що в дійсності є охрестям Френеля точок стаціонарної фази [2, 3]. В таких точках нормаль до поверхні об'єкта співпадає з напрямком на точку прийому відбитого сигналу. Тому для когерентних зображень найбільш вигідною моделлю функції розсіювання є модель дискретних розсіювачів [1].

$$h(\vec{r}) = \sum_{n=1}^{N} a_n \delta(\vec{r} - \vec{r}_n), \qquad (1)$$

де N – кількість розсіювачів;

a_n – амплітудний коефіцієнт відбиття кожного розсіювача;

$\delta(\cdot)$ – функція Дірака;

$\vec{r}_n$ – радіус-вектор, що з'єднує точку прийому з відповідною точкою стаціонарної фази (ТСФ).

Найбільш суттєвою відмінністю когерентних зображень є значний вплив на їх формування та вигляд явища інтерференції парціальних компонент в (1), характер якої залежить від фазових співвідношень вказаних компонент. Для урахування цієї залежності в [3] запропонована модель функції розсіювання когерентних зображень у вигляді:

$$h(\tau,q) = \sum_{n=1}^{N(q)} a_n(q)\exp[j\varphi_n(q)]\delta(\tau - \tau_n(q)), \qquad (2)$$

де q – кут зору;

$a_n(q)$ – нормований амплітудний розподіл ФР;

$\varphi_n(q)$ – відносний фазовий розподіл ФР;

$\tau_n = |\vec{r}_n| / c$ – відносна затримка;

c – швидкість розповсюдження коливань у середовищі.

Дислокація ТСФ зв'язана функціональною залежністю з формою поверхні об'єкта, але оскільки ця залежність невідома, одержання фазового розподілу $\varphi_n(q)$ є досить складною проблемою. Для

© UkrObraz'2000

об'єктів з акустично-жорсткою поверхнею запропонована експериментально розрахункова методика реконструкції відносного фазового розподілу за даними експериментальних вимірювань ефективної поверхні розсіювання (ЕПР) об'єкта в цілому та окремо "блискучих" точок. Для об'єктів вказаного класу ФР з урахуванням фазового розподілу вичерпно характеризує вторинне акустичне поле і може змінюватися лише із зміною типу об'єкта. Тому для відомих типових об'єктів може бути сформована бібліотека їх акустичних ФР (сигнатур), яка не потребує значних об'ємів пам'яті. Дійсно, для кожного кута зору ФР формується з N значень амплітуди, відносної затримки та фази кожної з $3N$ ТСФ. За даними експериментальних досліджень та літературних джерел [4] число N не є значним: $N = 3 \ldots 8$, оскільки при досить значних N вторинне поле втрачає просторову когерентність і модель (1) переходить в модель Делано. Комп'ютерна генерація тестової навчальної А-ехограми по відомій ФР типового об'єкта виконується згідно співвідношенню

$$y_A(t,q) = S[t,\vec{r}(q)] * h[t,\vec{r}(q)], \qquad (3)$$

Відповідно (3) сформована А-ехограма є кутовим розтином для певного заданого кута зору q акустичного когерентного зображення об'єкта дослідження.

2. КОМП'ЮТЕРНЕ МОДЕЛЮВАННЯ ТА ЕКСПЕРЕМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

Комп'ютерне моделювання акустичних когерентних зображень виконано у середовищі "MatLab 6.0". Для експериментальних досліджень використано спеціалізований вимірювальний комплекс.

На рис.1 зображена А-ехограма тестового об'єкта, одержана з високим просторо-часовим розподілом на частоті зондування 2 МГц з довжиною хвилі λ =0,75 мм. По цій реалізації реконструйовано елементи функції розсіювання.

Для підвищення вірогідності експериментальних вимірювань поряд з об'єктом додатково розміщувались калібрований приймач зондуючого сигналу та еталонна металева сфера діаметром 30 мм. За допомогою каліброваного приймача фіксувалась дійсна форма падаючого на поверхню об'єкта зондуючого поля, що дозволило врахувати його зміну середовищем розповсюдження.

При розповсюдженні прямокутного зондуючого імпульса у реальному середовищі його фронти набувають гаусівської форми, що враховано при компьютерному моделюванні.

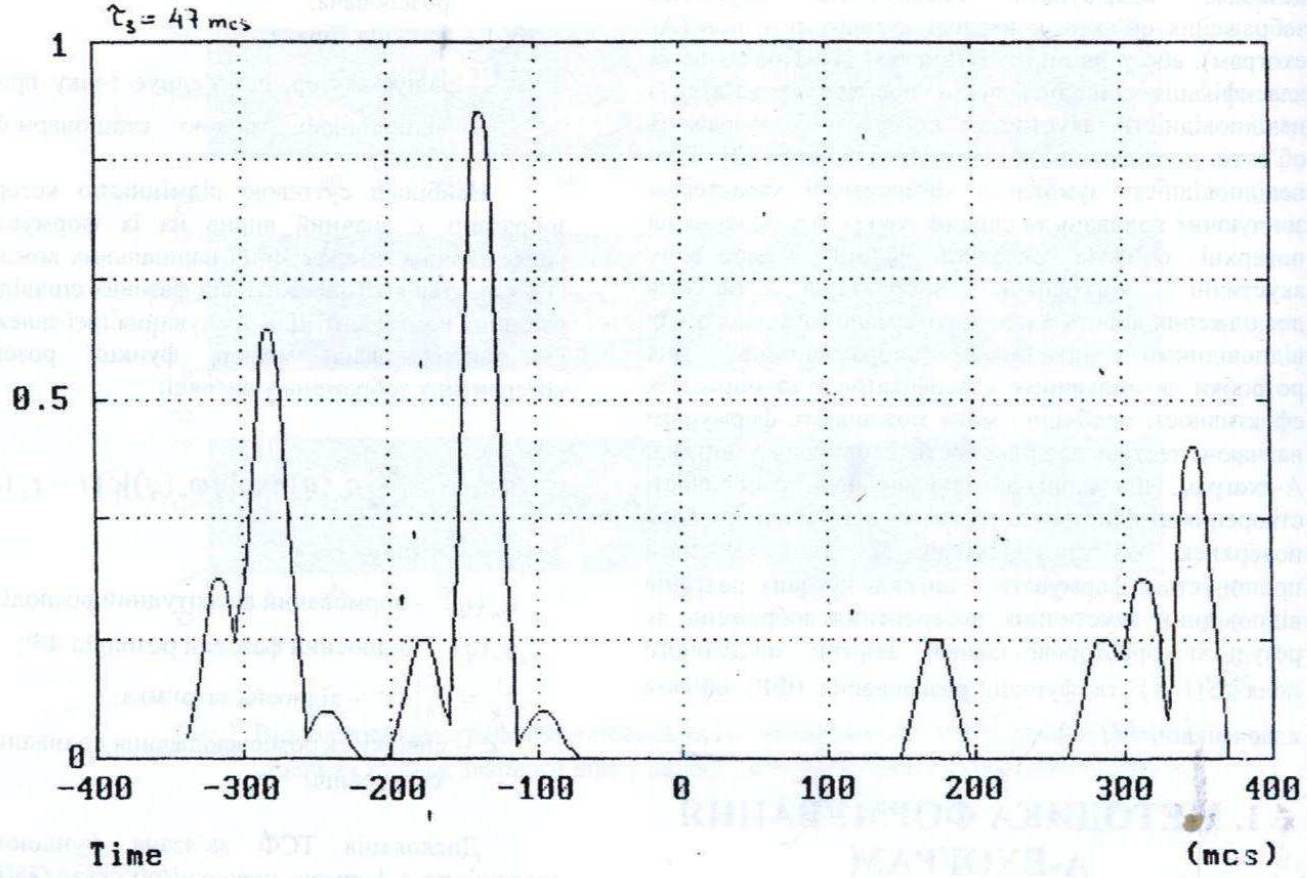


Рис. 1. А-эхограма тестового об'єкта на частоті f = 2МГц.

Еталонна сфера має відому ЕПР та коефіцієнт відбиття для вказанної частоти зондування і видбитий нею сигнал використовувався для амплітудного масштабування рівня експериментально одержаних ехограм тестового об'єкта.

Частота повторення зондуючих імпульсів встановлювалась в діапазоні 50…60 Гц, що дозволило виконати статистичну обробку одержаних ехограм по декількам сотням реалізацій для кожного кута зору, який мав змогу замінюватись у двох площинах. Експеріментально досліджувались сталість елементів ФР від коливань кута зору в діапазоні $\pm 5^{\circ}$ у двох площинах, та від зміни дістанції зондування ±20% .

На рис.2 запропоновано для порівняння згенерована та експериментальна ехограми для довільного імпульсного зондуючого сигналу, які ілюструють явище інтерференції вторинного акустичного поля. Порівняння ехограм виконувалось за допомогою кореляційного J_1 та середньоквадратичного J_2 критеріїв, значення яких також приведені на рисунку. Комп'ютерне моделювання А-ехограм виконано при використанні зондуючих сигналів змінної акустичної довжини та різної форми (гаусівських та прямокутних імпульсів), а також для різних кутів зору в широкому діапазоні у двох площинах.

середньоквадратичне відхилення не перевищує 5%. Найбільш важливим є одержане високе значення коефіцієнта кореляції, що підтверджує близькість форми зформованого просторово-часового розподілу вторинного акустичного поля його дійсному розподілу з урахуванням інтерференції. Найкращі результати одержано при гаусівській апроксимації фронтів та в діапазонах кутів зору, де значення N не перевищує 10, що пояснюється частковою втратою когерентності при $N > 10$. Але і в такому випадку немає принципових перешкод для моделювання акустичних зображень з частковою чи повною втратою когерентності: в моделі ФР (2) слід передбачити можливість завдання фазового розподілу стохастичним з деякою функцією щільності ймовірності, конкретний вигляд якої визначається механізмом руйнування когерентності вторинного поля.

Тому запропонована модель ФР акустичних когерентних зображень у вигляді (2) має досить загальний характер. Але найбільш важливою її перевагою є експеременталько встановлена відповідність реальним фізичним процесам формування вторинного акустичного поля складною поверхнею об'єктів дослідження, і тому відповідність задачам їх класіфікації.

Модель "блискучих" точок Делано,

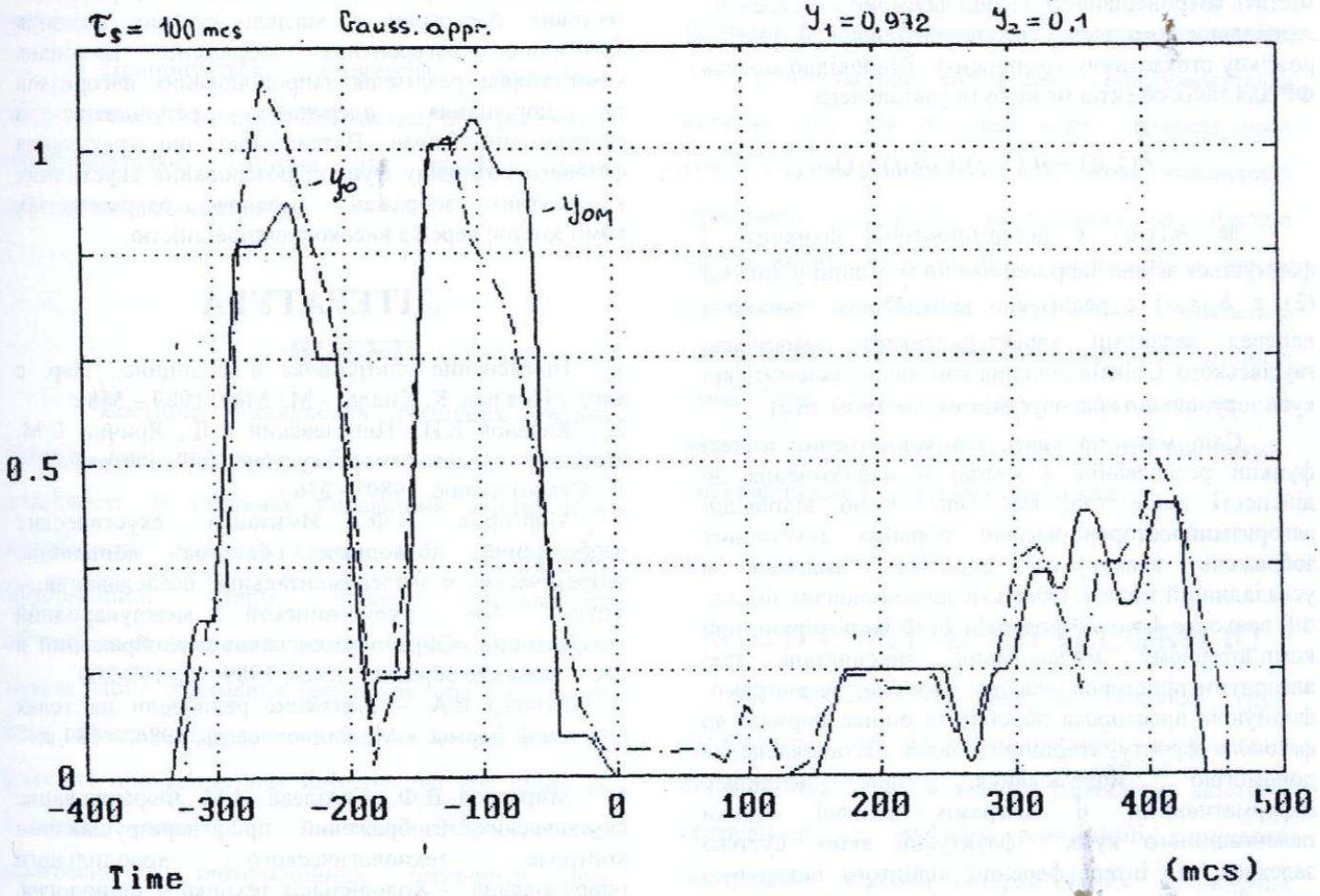


Рис. 2. Згенерована та експериментальна ехограми.

Одержана задовільна відповідність згенерованих ехограм експериментально одержаними: коефіцієнт кореляції близько 0,9;

узагальнена для акустики Ольшевським, пропонує стохастичне об'єднання їх полів у гаусівське вторинне поле, відповідно тому має інтегральний

характер і не несе в собі структурної інформації про характеристики поверхні об'єкта зондування. Відома модель "локальних джерел" Штагера [4] також грунтується на стохастичному об'єднанні полів, відбитих локальними дільницями на поверхні об'єкта, але з урахуванням певної функції щільності ймовірності негаусівського характеру. Тому відомі моделі ФР мало пристосовані до вирішення класифікаційних задач, оскільки не містять параметрів, що мають безпосередньо залежність від особливостей форми поверхні відбиття об'єкта. На відміну від вказаних відомих моделей, запропонована модель ФР таку залежність враховує. Дійсно, амплітудні коефіцієнти відбиття залежать від радіусів кривизни поверхні в точці стаціонарної фази, відносні затримки разом з зафіксованим кутом зору вказують на їх дислокацію на поверхні, а тип фазового розподілу враховує різний характер розсіювання на гладких поверхнях та ребрах, зламах, кутах та інше.

Загальним обмеженням запропонованої моделі є наближення Кірхгофа, тобто радіуси кривизни поверхні відбиття повинні багаторазово перевищувати довжину хвилі зондуючих коливань. Для більшості задач медичної та технічної діагностики таке обмеження не є суттєвим.

Поверхня багатьох об'єктів дослідження не є цілком гладкою, і крім структурних формуючих нерівностей, що мають детермінований характер, містить мікронерівності. Перші формують когерентну дзеркальну компоненту вторинного поля, а другі – розсіяну стохастичну компоненту. Відповідно модель ФР для таких об'єктів може бути узагальнена:

$$h(\tau,q) = h_1(\tau,q) + \alpha(q)h_2(\tau,q),$$

де $h_1(\tau,q)$ є детермінованою функцією і формується згідно запропонованій методиці у вигляді (2), а $h_2(\tau,q)$ є реалізацію випадкового процеса з наперед заданими характеристиками, наприклад гаусівського. Співвідношення компонент залежить від кута зору, що що враховується параметром $\alpha(q)$.

Слід мати на увазі, що ускладнення моделі функції розсіювання з метою її наближчення до дійсності лише тоді має сенс, якщо відповідні алгоритмипросторово-часової обробки акустичних зображень враховують додаткові чинники в ускаладниній моделі. Оскільки запропонована модель ФР враховує фазовий розподіл ТСФ на поверхні, при комп'ютерному моделюванні використані такі алгоритми просторово-часової обробки, як діаграмо-формуюча просторова обробка та оцінка нормалі до фазового фронту вторинного поля. Встановлено за допомогою моделювання, що найбільш інформативними є діаграми фазової оцінки пеленгаційного кута, флуктуації якого суттєво залежить від інтеренференції відбитого поверхнею поля. Таким чином додаткова структурна інформація, врахована в ФР, млже бути виявлена відповідними алгоритмами обробки вторинного акустичного поля.

Зауважимо, що одержана функція розсіювання деякого типового об'єкта з урахуванням фазового розподілу, вичерпним чином характерізує його вторинне акустичне поле у наближенні Кіргофа і є невід'ємною характеристикою відповідного акустичного зображення. Складність одержання такої ФР компенсується повнотою структурної інформації про його поверхню відбиття, яка може бути реконструйована по відповідній ФР, розрахованій для різних кутів зору.

Окрім вирішення задачі формування навчаючих послідовностей для забезпечення автоматичних класифікаторів, запропонована модель ФР може бути використана для оптимального вибору відповідних зондуючих сигналів, узгоджених з особливостями поверні відбиття. Мірою такої узгодженності може бути максимальне значення функції взаємокорелляції чи двомірної функції невезначеності [4,5].

ВИСНОВКИ

Для збудування навчальних послідовностей автоматичних класифікаторів аномальних об'єктів стосовно медичної та технічної ультразвукової діагностики запропоновано алгоритм формування тестових А-ехограм в вигляді кутових розгинів акустичних когерентних зображень. Виконана комп'ютерна реалізація запропонованих алгоритмів та порівняння одержаних результатів з експериментальними. Встановлено, що урахування фазового розподілу функції розсіювання акустичних когерентних зображень дозволяє одержати їх комп'ютерні версії з високою відповідністю.

ЛІТЕРАТУРА

1. Применение ультразвука в медицине.: Пер. с англ. / Под ред. К. Хилла. – М.: Мир, 1989 – 568 с
2. Качанов Е.И., Пигулевский Е.Д., Яричин Е.М. Методы и средства гидроакустической голографии. – Л.: Судостроение, 1989. – 256 с.
3. Миргород В.Ф. Имитация акустических изображений подводных объектов: концепция, теоретические и экспериментальные исследования. – Труды 4-й Всеукраинской международной конференции «Обработка сигналов и изображений и распознавание образов», Киев, 1989 – С. 199-200
4. Штагер Е.А. – Рассеяние радиоволн на телах сплошной формы. – М.: Радио и связь, 1986. – 184 с.
5. Миргород В.Ф., Гвоздева И.М. Формирование акустических изображений при неразрушающем контроле технологического холодильного оборудования. – Холодильная техника и технология, вып. 64, 1999. – С. 126-129.

ІНТЕГРОВАНА ІНСТРУМЕНТАЛЬНА СИСТЕМА ДЛЯ РІШЕННЯ ЗАДАЧ ДІАГНОСТИЧНОГО КОНТРОЛЮ

Канд. техн. наук В.Д. Павленко, О.О. Фомін, В.О. Шибаєв

Одеський державний політехнічний університет

65044 пр. Шевченка 1, Одеса, Україна

тел. (0482) 288-170, e-mail: pavlenko_vitalij@mail.ru

Розглядається комплексний підхід до проблеми діагностичного контролю, що включає спільне використання програмних засобів і алгоритмів статистичної класифікації. Аналізується склад алгоритмів для рішення широкого кола задач діагностичного контролю: навчання, класифікації, дослідження інформативності діагностичних ознак об'єкта.

Пропонується структура організації алгоритмічної і інтерфейсної частин, інтегрованих в інструментальну систему для рішення задач діагностичного контролю.

ВСТУП

Постійне ускладнення технічних систем ставить нові проблеми забезпечення їх якості і надійності. У системах управління якістю при виробництві складних систем і їх сертифікації інформаційні етапи процесу управління об'єднуються операцією діагностичного контролю, задача якої - отримання інформації про фактичний стан об'єкта контролю і прийняття на цій підставі рішення про віднесення його до тієї або іншої із заданих категорій якості. Впровадження прогресивних інформаційних технологій при автоматизації операцій контролю в системах управлінні якістю і сертифікація складних систем дозволяє використати для рішення задач діагностичного контролю ефективні процедури інтелектуальної обробки інформації, засновані на методах статистичної класифікації (розпізнавання образів) [1].

У роботі пропонується комплексний підхід до проблеми підвищення ефективності систем управління якістю і сертифікації, заснований на розвитку як алгоритмічних, так і програмних засобів. Тут розглядається інтегрована інструментальна система (ІС) для рішення задач діагностичного контролю, що поєднує в собі нові можливості організації інтерфейсу користувача на підставі прогресивних інформаційних технологій, а також ефективні статистичні алгоритми навчання, класифікації і дослідження діагностичної цінності ознак при багатокласовому розпізнаванні на підставі даних, поданих як звичайними вибірками, так і вибірками з нечітким описом класів.

1. АЛГОРИТМІЧНІ ЗАСОБИ ІС

Розроблена ІС містить декілька груп алгоритмів, для багатостороннього дослідження статистичних даних в задачах управління якістю і сертифікації. У програмі реалізовані алгоритми попередньої обробки вибірки даних: заповнення прогалин в даних статистичними методами [2]. Ця проблема набуває особливе значення при роботі з реальними даними, коли внаслідок цілого ряду

© UkrObraz'2000

об'єктивних і суб'єктивних причин в описі деяких об'єктів статистичної вибірки відсутні окремі компоненти (діагностичні ознаки), а збір нової інформації про ці не повністю описанні об'єкти часто зв'язаний з додатковими витратами ресурсів (часових, матеріальних і т.д.) або ж взагалі не є можливим.

Друга група алгоритмів призначена для побудови вирішальних правил статистичної класифікації. В якості вирішальних правил статистичної класифікації використовуються поліноми першого, другого і третього ступеня. Для побудови вирішальних правил статистичної класифікації використовуються параметричні і непараметричні методи навчання розпізнаванню образів: метод статистичних рішень і метод стохастичної апроксимації [1, 2]. Розвиток цих методів дозволив отримати більш ефективні процедури побудови вирішального правила на підставі методу статистичних рішень і оптимізації вибору порога розпізнавання:

$$d(\mathbf{x}) = -\frac{1}{2}(\mathbf{x}-\mathbf{m}_1)^{\prime}\mathbf{S}_1^{-1}(\mathbf{x}-\mathbf{m}_1)+ \\ +\frac{1}{2}(\mathbf{x}-\mathbf{m}_2)^{\prime}\mathbf{S}_2^{-1}(\mathbf{x}-\mathbf{m}_2)+\frac{1}{2}ln\frac{|\mathbf{S}_2|}{|\mathbf{S}_1|}+\lambda \quad (1)$$

де **x** - вектор параметрів стану об'єкта розмірністю n; $\mathbf{m}_i$ - вектор середніх значень вимірювань параметрів стану об'єктів для і-го класу (i=1,2); Si - коваріаційна матриця параметрів стану, що вимірюються для і-го класу; $|Si|$ - визначник матриці Si; $\mathbf{S}_i^{-1}$ - матриця, зворотна матриці Si, λ - поріг розпізнавання.

Другий метод для ефективного пошуку вирішального правила полягає у використанні в якості початкових значенньь коефіцієнтів квадратичного вирішального правила в алгоритмі методу стохастичної апроксимації значень, попередньо отриманих на етапі навчання за допомогою методу статистичних рішень, що дозволяє значно скоротити кількість ітерацій в рекурентній процедурі методу стохастичної апроксимації і дає можливість отримати вирішальне правило з більш високим процентом правильного розпізнавання. При цьому перехід між розглянутими методами здійснюється по формулах:

$$\mathbf{A} = \frac{1}{2}(\mathbf{S}_2^{-1} - \mathbf{S}_1^{-1})$$
$$\mathbf{b} = \mathbf{S}_1^{-1}\mathbf{m}_1 - \mathbf{S}^{-1}{}_2\,\mathbf{m}_2 \quad (2)$$
$$c = -\frac{1}{2}(\mathbf{m}_1^{\prime}\mathbf{S}_1^{-1}\mathbf{m}_1 - \mathbf{m}_2^{\prime}\mathbf{S}_2^{-1}\mathbf{m}_2) + \frac{1}{2}ln\frac{|\mathbf{S}_2|}{|\mathbf{S}_1|}$$

де **A** - матриця коефіцієнтів при квадратичних членах вирішального полінома; **b** - вектор коефіцієнтів при лінійних членах вирішального полінома; c - вільний коефіцієнт.

Важливим етапом всякого процесу класифікації є вибір набору діагностичних параметрів. При цьому звичайно прагнуть обмежитися як можна меншою їх кількістю. Скорочення опису об'єктів спрощує збір додаткових даних, робить матеріал більш оглядним, а пристрій, що розпізнає - більш простоєм і швидкодіючим. Зменшення параметрів опису об'єктів, як правило, поліпшує якість прогнозу [3]. Крім того, скорочення числа діагностичних ознак дозволяє знизити мінімально необхідну кількість об'єктів статистичної вибірки, зберігаючи її репрезентативність. Тому після збору статистичних даних необхідно розглянути можливість виключення частини ознак так, щоб помилки розпізнавання були менш допустимих.

Очевидно, що кількість ознак, необхідна для успішного рішення деякої задачі розпізнавання, залежить від розділяючих якостей вибраних ознак. Задача вибору ознак ускладнюється звичайно тією обставиною, що найбільш важливі ознаки не завжди легко виміряти або, як виявляється в багатьох випадках, відповідні можливості вимірювання стримуються економічними факторами. У ІС включена група алгоритмів, спрямованих на дослідження інформативності параметрів статистичної вибірки і відбір найбільш

інформативних наборів з них. Ці алгоритми представляють різні методи перебору параметрів статистичної вибірки: повний, скорочений і випадковий перебір (метод Монте-Карло) [3].

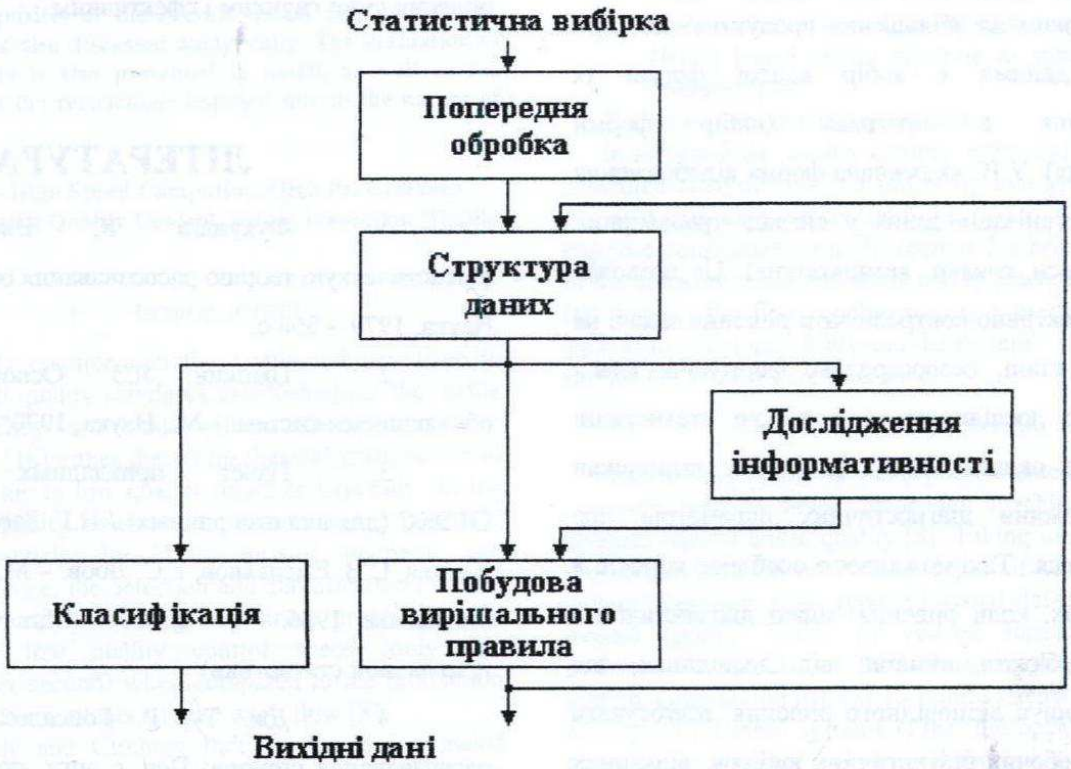
При роботі з багатоклассовыми статистичними вибірками для пошуку найкращого рішення в задачах дослідження інформативності пропонується використати критерії *F(P)*, що враховують різницю в якості розпізнавання різних класів на кожному наборі діагностичних параметрів:

$$F(P) = \frac{\sum_{i=1}^{m} P_i}{\alpha \cdot (P_{\max} - P_{\min}) + 1} \qquad (3)$$

$$F(P) = \frac{\sum_{i=1}^{m} P_i}{e^{\beta \cdot (P_{\max} - P_{\min})}} \qquad (4)$$

де P_i - відсоток правильного розпізнавання i-го класу, P_{max} і P_{min} (максимальне і мінімальне значення відсотка правильного розпізнавання серед заданих класів на наборі діагностичних параметрів, що розглядається; m - кількість класів в задачі; α, β - вагові коефіцієнти (величина зворотна сумі $\sum_{i=1}^{m} P_i$).

Використання множини алгоритмів обробки статистичної вибірки викликає необхідність розв'язання проблеми організації обчислювального процесу в ІС. Для забезпечення ефективної взаємодії методів обробки даних в ІС, що пропонується, реалізований оригінальний засіб інтеграції перерахованих алгоритмів в єдиний обчислювальний процес. На мал. 1 представлена структурна схема організації програмного інтерфейсу, що реалізовує обчислювальні алгоритми статистичного розпізнавання і дослідження інформативності діагностичних об'єктів.



Мал. 1. Структурна схема організації програмного інтерфейсу ІС

2. ІНТЕРФЕЙСНІ ЗАСОБИ ІС

З метою скорочення втрат часу на непродуктивні операції підготовки і узгодження даних між різними алгоритмами і забезпечення ефективної взаємодії методів обробки даних, а також зручності роботи зі статистичними вибірками розроблений сучасний графічний інтерфейс користувача ІС, відповідний стандартам операційної системи Windows 95/98, що дозволяє гранично спростити, стандартизувати і візуалізувати процеси маніпулювання з даними, алгоритмами дослідження початкових даних, дозволяючи створювати, змінювати статистичні вибірки, переглядати результати досліджень і здійснювати вибір переважних рішень.

Основні функції інтерфейсу користувача полягають в наданні простого і інтуїтивно зрозумілого засобу відображення початкових даних і формування повчальних статистичних вибірок. Першим кроком до збільшення продуктивності при роботі з даними є вибір вдалої форми їх представлення в програмі (вибір форми відображення). У ІС як основна форма відображення прийнята організація даних у вигляді трьохмірних таблиць (класи, ознаки, вимірювання). Це дозволяє швидко і ефективно контролювати рішення задачі на будь-якому етапі, безпосередньо формуючи класи об'єктів, що досліджуються, і робочі статистичні вибірки, що складаються з довільних підмножин повної множини діагностичних параметрів, що досліджуються. Такі можливості особливо корисні в тих випадках, коли рішення задачі діагностичного контролю об'єктів вимагає від дослідника, що здійснює пошук відповідного рішення, підготувати цілий ряд робочих статистичних вибірок, відмінних одна від одної лише набором включених до їх складу діагностичних параметрів об'єктів, що досліджуються, або кількістю і складом класів об'єктів, що розглядаються, тоді як непродуктивний час, затрачений на безпосереднє формування робочих вибірок перевищує у багато разів час, що затрачується на рішення задачі.

ВИСНОВКИ

Використання розробленої ІС для рішення практичних задач діагностичного контролю в системі управління якістю і сертифікація, розглянутих в роботі, демонструє істотне скорочення часу, що відводиться на пошук задовільних рішень задачі, скорочення кількості діагностичних параметрів в описі об'єкта досліджень без зниження якості класифікації його можливих станів, зниження об'єму статистичної вибірки і отримання вирішальних правил, що забезпечують надійне розпізнавання.

Реалізація різноманітних методів в ІС дозволяє зробити процес пошуку задовільного рішення дуже гнучким і ефективним.

ЛІТЕРАТУРА

1. Фукунага К. Введение в статистическую теорию распознавания образов. - М.: Наука, 1970. - 368 с.

2. Цыпкин Я.З. Основы теории обучающихся систем. – М.: Наука, 1970. - 399 с.

3. Пакет прикладных программ ОТЭКС (для анализа данных) / Н.Г. Загоруйко, В.Н. Ёлкина, С.В. Емельянов, Г.С. Лбов. – М.: Финансы и статистика, 1986. – 160 с.: ил. – (Мат. Обеспечение прикладной статистики).

4. Дж. Ту, Р. Гонсалес. Принципы распознавания образов: Пер. с англ. под ред. Ю.И. Журавлёва.. – М.: Мир, 1978 - 411 с.

Requirements and Specification for High Speed Computing unit for flaw detection in fabrics

Anagnostopoulos C.[1], Hasan S.[4], Kayafas E.[1], Knaak U.[2], Loumos V.[1], Nassar S.[3], Saqer Abdel-Rahim[4], Stassinopoulos G.[1] and Vergados D.[1]

[1] Computer Science Division, Department of Electrical & Computers Engineering,
National Technical University of Athens, GR 15773 Zographou, Athens, Greece
Tel: +30 1 772 2538 Fax: +30 1 7722534, E-mail: canag@central.ntua.gr, vergados@telecom.ntua.gr
[2] Parsytec AG, Aachen, Germany
[3] Electronics Research Institute, Cairo, Egypt
[4] Royal Scientific Society, Amman, Jordan

Abstract --

Industrial High Speed units (or High Performance Computing – HPC)are of strategic interest for the textile industry as they could form the basis of a system achieving a high degree of accuracy on textile inspection.

Due to the specific nature of textile, the defects encountered within textile production must be detected and corrected at early stages of the production process. Thus, the visual defect detection is of utmost importance for the product's overall quality and cost.

This paper presents the requirements of the textile companies, in regards to the quality control process by assessing the feasibility of using a HPC based machine vision to substitute human operators.

The prerequisites or the overall system and its software components are also discussed analytically. The evaluation of the first results is also presented in detail, as well as the limitations and the restrictions imposed due to the nature of the problem.

Keywords -- **High Speed Computing, High Performance Computing,Textile Quality Control, Fabric Inspection, Textile Industry.**

I. INTRODUCTION

The quality engineers in the textile industry have to retain the high quality standards established for the textile industry [5]. They have to deal with an extensive variety of fabric defects [16] either due to mechanical malfunction of the loom, or due to low quality fibers and spreads. At the present, the quality assessment procedures are generally performed manually by expert quality engineers and technicians. Hence, the detection and classification of these defects is a time consuming and tiring procedure. In addition, the low quality control speed (only few centimeters per second) when compared to the production speed reveals the bottleneck in the work flow [8].

The Textile and Clothing Industry has an estimated turnover of 150 billions ECU's and employs about 2,5 million persons in EUROPE [1]. The majority of the companies in this sector are SME's, which are facing an ever increasing competition by low price imports. Automation and technological development are suggested as key factors for the survival of this industrial sector. Another key factor is the production of new top-of-the-range products, which are less sensitive to price competition [12]. Although the clothing industry has benefited from technological innovations, particularly in CAD, it is still a labor-intensive industry.

- This paper studies in depth the quality control requirements in the vision systems of the textile clothing industries, by analyzing the quality control processes, evaluation of the performance of the vision system and establishing the feasibility analysis of using High Performance Computing (HPC) based vision machine to substitute human operators [14].

The paper is organized as follows:

In section 2 the quality control in the textile industry is presented analytically. The following two section describe the system's architecture and the specifications for the required constituent parts. In section 5 a brief introduction to the software requirements is taking place. Finally, in the last section, the first conclusions are presented and issues such as the interoperability and the system's performance in the Textile Industry are evaluated.

II. QUALITY CONTROL

Fabric inspection is mainly effectuated in a textile company through fabric inspection machines, where an operator checks fabric quality [8]. Taking into account that the textile industry uses many types of yarns and wave patters, there are many types of visual defects, that affect overall quality. There are several formal systems to evaluate the quality of the fabric, namely the 4-point system, Graniteville [6], [2] 10-point systems etc. The main concept of all these systems is that the operator calculates the numbers of major and minor defects as point values per m^2 and then considers the quality of the fabric as "first" or

© UkrObraz'2000

"second" quality. Each fabric, either woven or knitted, presents specific defects.

III. Overall System Architecture

The HPC-surface inspection system is typically made up of the following main components:

- The inspection bridge (top and bottom side) with cameras and lighting
- The defect analysis computer system including a main computer (server) and a workstation per camera.
- The operator console, and the quality control console for production purposes.
- The archive subsystem for the strip and defect databases.
- The cabling of the system including the interface to the operational database and the rotational encoder to determine the speed.

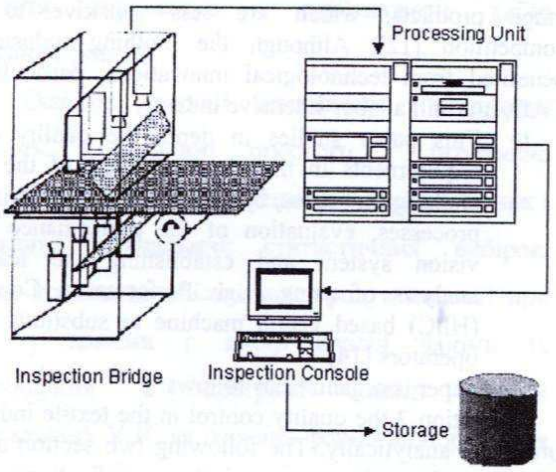


Fig.1 Overall System Architecture

IV. System Specifications

The HPC system specification addresses the overall architecture and detailed data for the required constituent parts like, cameras (number, resolution, spectral sensitivity, etc), illumination and flashing sources and their triggering, mechanical construction, number of processing nodes, configuration and memory, operating and programming software environment, functional specification of algorithms developed, non functional requirements based on production environment and ergonomics. The HPC system based on a prototype developed for the textile industry includes an inspection bridge, illumination modules, image acquisition technology [12] (cameras, objectives, filters, frame grabbers, cables etc) with corresponding peripherals, as well as a high performance computer, which enables the scalability concerning speed and width of material. The system also includes the necessary development environment such as the parallel framework and the software modules [14].

The HPC inspection system consists of three main components [15]:

- the **inspection bridge** with cameras and lighting,
- the **inspection computer system**, consisting of networked industrial PC's, and
- the **inspection console** with the **System Management Interface**.

An extensive requirement analysis took place for all of the mentioned components of the HPC System.

A. Inspection bridge installation requirements

The inspection bridge is installed above and below the strip. The images acquired with the cameras have to be transmitted via cable to the analysis computer.

The bridge can be installed anywhere in the production process where a straight strip is available. If a specific application requires the concurrent use of two different image capturing and lighting methods, two inspection bridges may be installed one after the other.

Due to the fact that the quality and resolution of the taken images of the strip surface have a big influence on the quality of the inspection, the determination of the acquisition system (camera and illumination) is the first important step. Generally one needs several pixels in order to detect a defect. On non-structured surfaces (or pre-processed images), the minimal number of pixels needed for a safe detection of a defect is 4 pixels. For a classification of various defect types more pixels are necessary. A minimum for each classification is 12 pixels and for a more precise classification about 30 pixels is necessary.

B. Inspection computer system requirements

The computer unit for the evaluation of surface images can be located remotely. The connection to the consoles can be done via standard Ethernet. Access to the inspection results can be provided throughout the mill over an already existing local area network.

C. Online inspection console requirements

The online inspection console is usually placed next to the production line at some distance after the inspection bridge. This enables inspection personnel to intervene in the production process when necessary.

D. System Management - Graphical user Interface

A Graphical User Interface has been developed in order to control, monitor and manage all the components of the HPC System.

This GUI (System Management Console) enables the operator to perform the following actions:

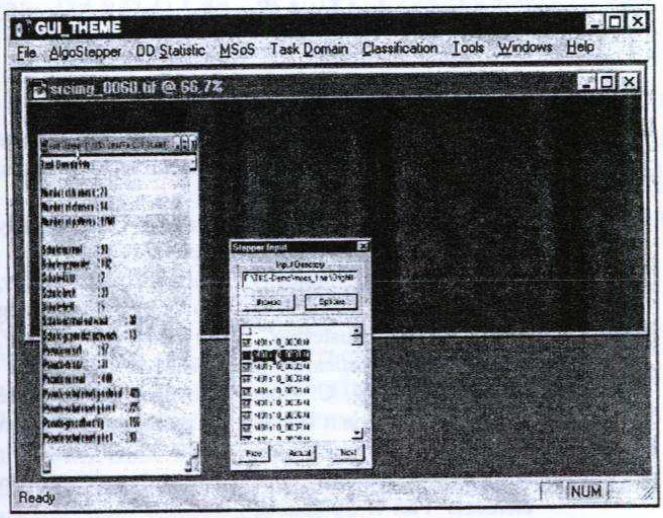


Fig.2 The GUI of the HPC System

- **Load the database**
- **Select the feature space**
- **Select and calibrate cameras**
- **Select the classifier type**
- **Training the classifier**
- **Fabric selection**

E. Illumination and Camera Technology

The dark field illumination has a great significance for surface damage detection, due to the fact that it reacts very sensitively to any changes of the surface smoothness. For this purpose infrared LED's (as used e.g. for remote control for various audio and video systems) are combined into a condensed packaged light deal. Through methods that have been patented the angle of the beam has been so reduced that the total light beam is close to parallel. The overlapping of the beam area is desired and allows the lighting module to be used even if some LED's are not functioning. A pulsed light is used for the lighting of the surface rather than a regular light, allowing a dense packaging of the LED's and a high light intensity without dissipating too much power (i.e. heat). These flashes are extremely short (4 µs up to 100 µs, in the standard mode 38 µs). The use of infrared technologies (wave length 800 to 950 nm) has the additional advantage that the employees, who work in the area of the inspection system, are not bothered by the flashing and through the use of infrared filters in front of the cameras, interfering light can be suppressed. Defects that don't have an edgy kind of surface damage can be detected with the extra use of a diffused bright field light. This lighting is based on fluorescent lights and is implemented in the flash (photo) technology. The bright field images are taken by separate cameras. It is also possible if desired, for these defects to have a different resolution than in the dark field. Color cameras can be used for special recognition projects.

In order to capture the entire width without gaps the HPC-System has several cameras positioned in a row. For the complete configuration the following parameters are relevant: image sequence frequency (50 Hz per row of cameras) and brightness area of the lighting. The uses of a second row of cameras can double the web speed. On special request, systems can be provided with more rows of cameras. The cameras are delivered in a closed housing and are being securely mounted.

F. Mechanical Configuration of the cameras

The camera row is installed in a vertical position over the web with the determined spacing (based on the chosen resolution). The LED infrared dark field lighting can be set at angles between 0 and 70 degrees depending on the amount of reflection of the inspected material. Camera/lighting unit is delivered in a stable frame with covering. The housing also contains modules for the synchronization of the cameras and lighting as well as video signal adapters for the fiber optic transfer.

TABLE I

STRIP SPEED AND CAMERA RESOLUTION

CCD Camera Resolution (square mm)	**Max. Strip Speed** (m/sec)
0.10 x 0.20	2.4
0.15 x 0.30	3.7
0.20 x 0.40	4.9
0.25 x 0.50	6.2

V. SOFTWARE REQUIREMENTS

The software core for an automatic Inspection system should be robust and flexible in order to detect the defected region on the inspected fabric. Due to the complexity of the problem and the extensive variety of the possible defects, the software modules must incorporate the most advanced image processing techniques[14]. In short terms, the software requirements includes the following operation:

- **Correction of Non-Uniform Illumination**

The goal of this step is the correction of local varying mean grey values and variances [11]. This way the illumination irregularities within the image are corrected.

- **Detection of possible defect**

On the basis of this observation a simple operation was developed in order to describe the "local" regularity of the underlying fabric. This operation was motivated by the fact that the standard deviation value of a processing block should not vary with the ones in the neighboring blocks. Hence, the standard deviation difference between neighbor blocks should be near zero or at least should vary smoothly

across the texture plane. The scope of this step is the detection of possible defects in the fabric. Defects manifest themselves by irregularities in the texture of the image [15]. The algorithm should be able to identify possible irregularities performing appropriate measurements in the inspected image [10].

Threshold values are chosen according to the specific type of the fabric. Pixels, whose measurements exceed the above thresholds, are considered to be candidates for a possible location of a defect in the fabric [13].

- **Regions of Interest (ROI)**

During this step, the candidate pixels are clustered, in order to form rectangular Regions Of Interest (ROI), roughly describing the shape of detected defects [4]. As additional information, further statistics of the features in the respective regions should be computed and evaluated [7].

G. Defect Classification

The last step classifies the detected defects according their shape or other criteria into macro classes. The classification into micro classes requires further specific knowledge provided by texture features and 2nd order statistics [3].

The complete inspection results can be archived in a database. The data of all defects can also be made available (archived) for statistical evaluation. This is useful in order to establish long term trends or occurrences of individual time periods [9].

VI. CONCLUSIONS

The objective of this paper is to present the requirements of an automatic system for the textile companies, in regards to the quality control process by assessing the feasibility of using machine vision to substitute human operators. The requirements and specification of such as a system, are studied in depth, and a technical approach is given very briefly (due to words limitation).

The emerging technologies in the field of Information Technologies integrated with available relatively low cost high quality vision machines, can be the most appropriate solution for the Textile Industry. It is anticipated that IT technology can render feasible advanced quality solutions, thus providing an important competitive advantage.

ACKNOWLEDGEMENTS

The work described in this paper is partially supported by the European Commission (Project THEME, project framework INCO, project number: 972572).

The participants in THEME are the National Technical University of Athens (Greece), Parsytec AG (Germany), Electronics Research Institute (Egypt), Royal Scientific Society (Jordan) and National Research Centre (Egypt).

REFERENCES

[1] The effect and Cost of Fabric Faults in Garment Manifacture, Wool Industries Research Association, London, 1971.

[2] Experience with a Demerit System for Cloth Quality Grading, Shirley Institute, Manchester, 1973.

[3] J.G. Campbell, C. Fraley, F. Murtagh, A.E. Raftery, "Linear Flaw Detection in Woven Textiles using Model-Based Clustering", University of Ulster, Faculty of Informatics, Preprint INFM-97-001, 1997.

[4] C. Lu, P. Chung, C. Chen, "Unsupervised Texture Segmentation via Wavelet Transform", Pattern Recognition, 30(5): 729-742,1997.

[5] "ISO 9004-1: Quality Management and quality system elements- Part 1: Guidelines", Berlin 1994.

[6] Manual of Standard Fabric Defects in the Textile Industry, Graniteville Co., South Carolina, 1975.

[7] Ch. Jacquelin, A. Aurengo, G. Hejblum. "Evolving Descriptors for Texture Segmentation", Pattern Recognition **30** (7): 1069-1079,1997.

[8] Mahall K., Quality assessment of textiles: Damage detection by microscopy, Springer- Verlag, Berlin, 1993.

[9] F. De Natale, "Rank-Order Functions for the Fast Detection of Texture faults", International Journal of Pattern Recognition and Artificial Intelligence, **10** (8): 971-984,1996.

[10] J. Canny, " A computational approach to edge detection", IEEE Trans. Pattern Analysis and Machine Intelligence, **8** (6): 678-698, November 1986.

[11] Ch. Daul, R. Rösch, B Claus, J. Grotepaß, U. Knaak, R. Föhr; "A fast image processing algorithm for quality control of woven textiles", Mustererkennung, 20. DAGM Symposium 1998, Springer, Heidelberg 1998.

[12] Parker J.R., "Algorithms for Image Processing and Computer Vision", Wiley Computer Publishing, New York, 1997; pp 155 – 160.

[13] C. Neubauer, "Segmentation of Defects in Textile Fabrics, International Conference on Pattern Recognition, pp. 688-691, 1992.

[14] Anagnostopoulos C. et al., "High Performance Computing Application for the Textile Quality Control", to be appeared in the International Conference on Intelligent Information Processing (IIP2000) proceedings, Federated Conference of the World Computer Congress (WCC2000), 21-25 August 2000, Beijing, China.

[15] Jain R., Kasturi R., Schunck B.G., Machine Vision, McGraw - Hill Inc., New York, 1995, pp. 122 –123.

[16] ISO 8498:1990 Woven fabrics—description of defects—Vocabulary.

МЕТОД ВИБОРУ ТЕСТІВ ДІАГНОСТУВАННЯ МІКРОСХЕМ ПАМ'ЯТІ ПРИ НЕЧІТКИХ АПРІОРНИХ ДАНИХ

В.Г. Рябцев, Н.Л. Кириленко, М.Х. Кхан
Черкаський інженерно-технологічний інститут, 18006, м. Черкаси, бул. Шевченка, 460,
т.: (0472) 43-56-28; факс: (0472) 42-21-65;
e-mail:masum@bd.nensi.cherkassy.ua

Для розпізнавання діагностичних властивостей тестів поряд з визначеними часовими параметрами пропонується використовувати лінгвістичну змінну – вірогідність діагностування. Розглядається отримана таким чином нечітка модель діагностичних властивостей тестів, яка може бути застосована при розробці програм ефективних випробувань мікросхем пам'яті.

ВСТУП

Задача вибору послідовності тестів для ефективного діагностування мікросхем та модулів пам'яті є недокінця формалізованою, тому що при її вирішенні ми оперуємо нечіткими апріорними відомостями про діагностичні властивості застосованих тестів. Дана ситуація часто зустрічається при освоєнні виробництва дослідної партії виробів або під час зміни технології їх виготовлення.

На вірогідність винаходження відмов мікросхем та модулів пам'яті впливають особливості схемної та топологічної реалізації елементів, які запам'ятовують (ЗЕ), дешифраторів адреси та вузлів вводу-виводу даних. Під час діагностування в умовах, які не є адекватними режимам експлуатації, може виникнути парадоксальна ситуація, коли мікросхеми пам'яті, які були признані придатними після випробувань за допомогою ефективних тестів, побудованих на основі складних алгоритмів, виявляються недієспроможними у складі обчислювальних систем [1]. Наприклад, дефекти вузлів управління, які призводять до прямої передачі вхідних даних на вихідні контакти корпусу мікросхеми пам'яті, не виявляються деякими пристроями тестового діагностування [2].

Для виявлення таких масових відмов необхідна модифікація алгоритмів тестів, завдяки якій перевірка зчитування одиниці або нуля проводиться при значенні даних, які на входах мікросхем пам'яті в першому циклі дорівнюють інверсному коду зчитаних даних, а в другому – прямому коду зчитаних даних. Даний метод дозволяє зафіксувати відмови в вузлах управління мікросхем пам'яті, в яких зберігається прямий код вхідних даних, а також у вузлах, які інвертують ці дані [3]. Але для більшості залишених відмов неможливо точно визначити тести, які будуть забезпечувати їх вірогідну локалізацію.

1. МОДЕЛЬ ВІДМОВ МІКРОСХЕМ ПАМ'ЯТІ

Якщо уявити множину усіх можливих відмов мікросхем пам'яті у вигляді універсальної множини σ, то її підмножина η - це множина тих відмов, які виявляються за допомогою дослідних тестів. Множина η утворюється шляхом об'єднання підмножин η_i.

$$\eta = \bigcup_{i=1}^{\kappa} \eta_i , \qquad (1)$$

$$\overline{\eta} = \sigma \setminus \eta; \ \overline{\eta} \notin \bigcup_{i=1}^{\kappa} \eta_i , \qquad (2)$$

де $\{\eta_1, \eta_2, ..., \eta_k\}$ - підмножини відмов, які були виявлені тестами $\pi_1, \pi_2, ..., \pi_\kappa$ відповідно;

$\overline{\eta}$ - множина невиявлених відмов.

Множину всіх можливих видів відмов можна представити у вигляді діаграми Ейлера-Венна, яка приведена на рисунку 1.

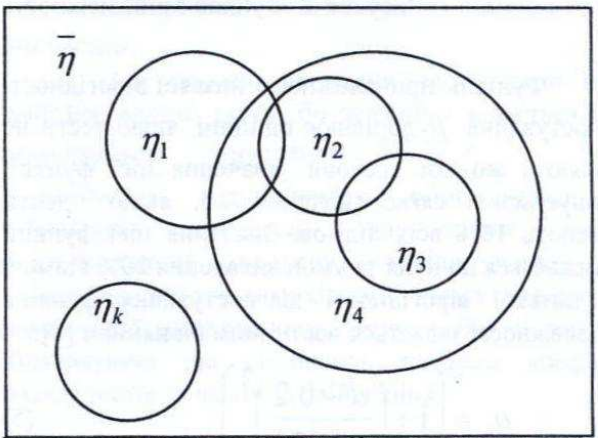


Рисунок 1. Діаграма Ейлера-Венна виявлення відмов мікросхем пам'яті діагностичними тестами

Якщо позначити вірогідність діагностування тестом π_i через

$$d_i = \frac{\eta_i}{\bigcup_{i=1}^{\kappa} \eta_i}, \qquad (3)$$

то діагностичні властивості тесту π_i можна представити у вигляді кортежу

$$j_i = (d_i, t_i), \quad (4)$$

де t_i – тривалість виконання і-го тесту.

2. МЕТОД ВИБОРУ ЕФЕКТИВНОЇ ПОСЛІДОВНОСТІ ТЕСТІВ

Час виконання тесту залежить від складності застосованого алгоритму, має функціональну залежність від обсягу виробу, який діагностується та легко обчислюється.

Значно складніше, а іноді й зовсім неможливо апріорно визначити точне значення вірогідності діагностування різних тестів, тому дану змінну можна вважати нечіткою лінгвістичною змінною, яка може приймати наступні значення:

l - низька;
lm – нижче середньої;
m - середня;
hm – вище середньої;
h - висока.

Графіки функції приналежності для наведених вище значень лінгвістичної змінної вірогідності діагностування зображені на рисунку 2, де d – вірогідність діагностування.

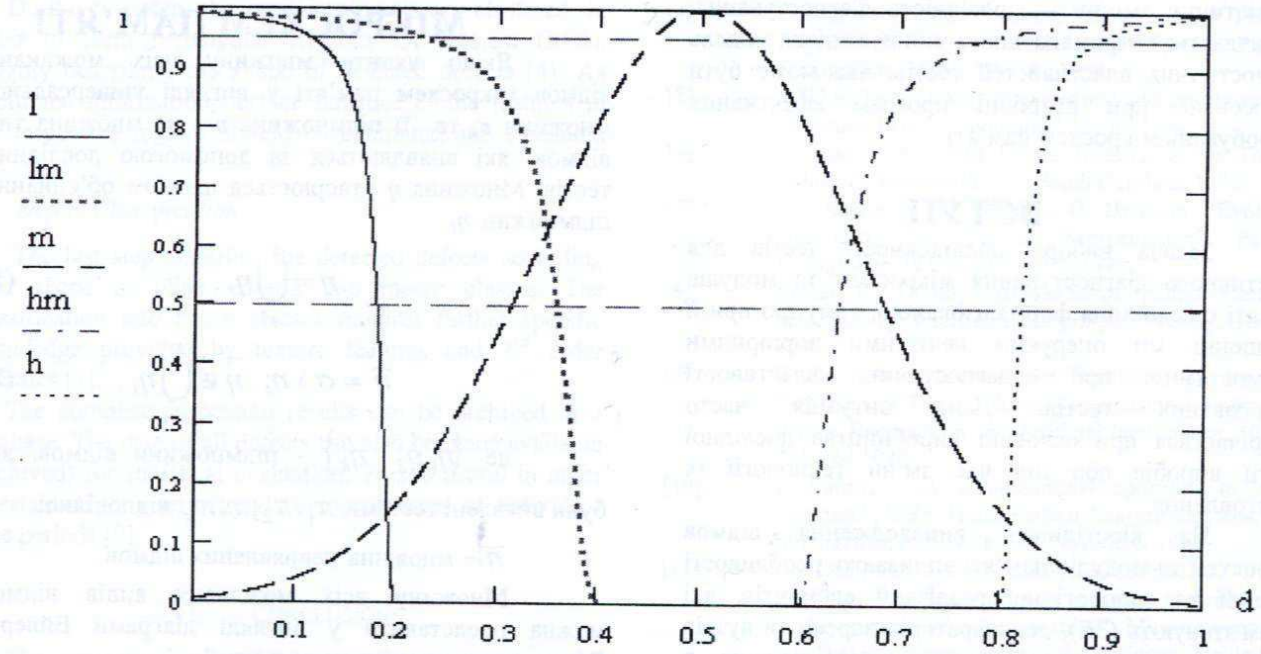


Рисунок 2. Функції приналежності лінгвістичної змінної вірогідності діагностування

Функції приналежності низької вірогідності діагностування μ_l дорівнює одиниці, якщо тести не виявляють жодної відмови. Значення цієї функції зменшується і сягає значення 0,5, якщо тести виявляють 18% всіх відмов. Значення цієї функції наближається до нуля за умов виявлення 20% відмов. Для низької вірогідності діагностування функція приналежності задається наступним рівнянням [4]:

$$\mu_l = \left(1 + \left(\frac{d - 0.2}{0.02}\right)^2\right)^{-1}; \quad (5)$$

Вірогідність діагностування нижче середньої мають тести, функція приналежності яких також дорівнює одиниці при невиявлені жодної відмови, спадає та сягає значення 0,5, якщо вони виявляють 36% всіх відмов. Функція приналежності в цьому випадку має вигляд :

$$\mu_{lm} = \left(1 + \left(\frac{d - 0.4}{0.04}\right)^2\right)^{-1}; \quad (6)$$

Тести, функція приналежності яких дорівнює одиниці при виявленні 50% відмов, мають середню вірогідність діагностування. При зростанні або зменшуванні вірогідності діагностування значення функція приналежності таких тестів зменшується і з достатньо високою точністю може бути задана модифікованою кривою Гауса (7).

$$\mu_m = e^{-20(d-0.5)^2}; \quad (7)$$

Функція приналежності лінгвістичної змінної вірогідність діагностування, значення якої є вище середньої, починає зростати при виявленні 60% відмов і досягає 0.5 при виявленні 64% виявлених відмов та одиниці за умови виявленні всіх відмов. Дана функція може бути задана рівнянням (8):

$$\mu_{hm} = \left(1 + \left(\frac{d - 0{,}6}{0.04}\right)^{-2}\right)^{-1}; \quad (8)$$

Функція приналежності значення висока лінгвістичної змінної вірогідність діагностування зростає починаючи з 80% виявлених відмов і також досягає значення 1 при виявленні всіх відмов. Ця функція визначається рівнянням (9):

$$\mu_{h} = \left(1 + \left(\frac{d - 0{,}8}{0.02}\right)^{-2}\right)^{-1}; \quad (9)$$

Як бачимо з рисунку 2, значення лінгвістичної змінної мають нечіткі межі. Тести з однаковими величинами функцій приналежності можуть мати різні значення цієї змінної: нижче середньої або середня, середня або вище середньої. Крім того, практично неможливо апріорно встановити діагностичні властивості тестів по виявленню всіх можливих відмов, так як кожний тест, як правило, був розроблений з метою локалізації окремих видів відмов основних вузлів мікросхем пам'яті. Лінгвістичні оцінки діагностичних властивостей найбільш розповсюджених тестів по виявленню окремих відмов мікросхем пам'яті наведені в таблиці1.

Таблиця 1. Апріорні діагностичні властивості тестів

№	Назва тесту	Діагностичні властивості виявлення відмов				
		1	2	3	4	5
1	Запис-зчитування	l	l	lm	l	l
2	Марш	hm	m	lm	lm	l
3	Хрест	hm	lm	lm	lm	l
4	1/0, що біжить	h	hm	l	hm	l
5	Галоп по рядку	h	hm	lm	lm	h
6	Галоп по стовпчику	h	hm	lm	lm	l
7	Метелик	h	hm	m	m	hm
8	Обертання адреси	h	hm	h	l	h

В таблиці 1 наведені наступні позначення окремих видів відмов:

1 – відмова комірок пам'яті;

2 – відмова дешифраторів адреси;

3 – невідповідність часу вибірки адреси;

4 – перевищення часу відновлення підсилювачів нормативного значення;

5 – невідповідність часу регенерації до встановленого значення.

Під час здійснення діагностування мікросхем пам'яті можна визначити імовірність виникнення відмов різних видів (10).

$$\sum_{i=1}^{5} c_i = 1,\ c_i = \overline{0,1}, \quad (10)$$

де c_i - імовірність виникнення відмови і-го виду.

Твердження (10) буде справедливим тільки під час процесу діагностування неполагоджених мікросхем пам'яті, коли вже відомо про існування відмови.

При рівно імовірному розподілу всіх видів відмов $c_1 = c_2 = ... = c_5 = 0.2$.

Однак, у деяких випадках такий розподіл може не існувати на практиці. Наприклад, для мікросхем пам'яті статичного типу $c_5 = 0$. Для даних мікросхем відсутній такий параметр, як час регенерації даних, і тому відмов 5-го виду для них не існує.

Таким чином, задаючи імовірність виникнення відмов різного виду можна враховувати особливості структури виробів, які діагностуються, та визначати середнє значення імовірності виявлення відмов кожним тестом .

$$d_{cp}^{j} = \sum_{i=1}^{5} d_i^j * c_i, \quad (11)$$

j – номер тесту;

i – вид відмови;

c_i - імовірність виникнення відмови і-го виду;

d_i^j - вірогідність виявлення відмов і-го виду j-м тестом.

Але не завжди можна застосувати навіть найефективніші тести, бо звичайно задається ще й максимально допустимий час $t_{\max}$, який визначається виробничими обмеженнями. Ці обмеження зумовлені досягненням високої продуктивності процесу діагностування, яка забезпечує випуск необхідного обсягу продукції за визначений інтервал часу: зміну, добу, місяць і т.д. Враховуючи цю обставину, введемо коефіцієнт якості тестів за часом їх здійснення.

$$k_j = \begin{cases} 0, & при\ t_j > t_{\max}, \\ \varepsilon, & при\ t_j = t_{\max}, \\ \dfrac{t_{\max} - t_j}{t_{\max}}, & при\ t_j < t_{\max}, \end{cases} \quad (12)$$

де ε - наперед задане мале число;

t_j - час здійснення j-го тесту;

$t_{\max}$ - часове обмеження процесу діагностування;

k_j - коефіцієнт якості за часом для j-го тесту.

Комплексний критерій якості Q_j кожного тесту можна визначити за його лінгвістичним та часовим параметрами:

$$Q_j = d_j * k_j, \qquad (13)$$

Потім, виходячи з аксіоматичних визначень, які приведені в [5], можна здійснити розподіл тестів на неперспективні, перспективні та оптимальні за Парето. З упорядкованої послідовності нас цікавлять тести, які мають високі значення комплексного критерію якості та відповідають економічним можливостям виробництва:

$$\sum_{k=\alpha}^{\omega} t_k \le t_{\max}, \qquad (14)$$

де $k \in \{\alpha, \beta, ..., \omega\}$ - номери тих тестів, які були включені у ефективну тестову послідовність;

t_k - час роботи k-го тесту;

$t_{\max}$ - максимальний час, відведений на діагностування.

Крім цих тестів до рекомендованої для діагностування послідовності можуть бути включені й інші тести, які користувач на основі свого досвіду вважатиме ефективними за даними обставинами.

Комплексний критерій якості бажано обчислювати за допомогою (13), якщо вплив часових та лінгвістичних характеристик для користувача рівноважний. Але коли більш важливим є використання найбільшої кількості тестів у заданий інтервал часу, можна замість (13) використовувати наступне рівняння:

$$Q_j = d_j + k_j. \qquad (15)$$

Для кожного і-го виду відмов визначаємо вірогідність його виявлення, яку можна спрогнозувати:

$$d^i_{np} = \max d^i_k, \qquad (16)$$

де $k \in \{\alpha, \beta, ..., \omega\}$, $i = \overline{1,5}$.

Середнє значення вірогідності виявлення всіх розглянутих видів відмов, яку можна спрогнозувати d_{cp} має наступний вигляд:

$$d_{cp} = \frac{1}{5}\sum_{i=1}^{5} d^i_{np}. \qquad (17)$$

Можна виділити песимістичний та оптимістичний види прогнозу. Песимізм заснований на невпевненості у вірогідності апріорних даних і, за звичай, є притаманний спеціалістам з малим досвідом. Оптимісти впевнені, що даним тестам властиві найкращі діагностичні характеристики.

Значення вірогідності діагностування, яку можна спрогнозувати, наведені в таблиці 2.

Таблиця 2. Вірогідність діагностування для різних видів прогнозу

Значення лінгвістичної змінної	Песимістичний прогноз		Оптимістичний прогноз	
	μ_{nec}	d_{nec}	μ_{onm}	d_{onm}
Низька	0.95	0.11	0.5	0.18
Нижче середньої	0.95	0.23	0.5	0.36
Середня	0.5	0.31	0.5	0.69
Вище середньої	0.5	0.64	0.95	0.78
Вища	0.5	0.82	0.95	0.89

Як бачимо з таблиці 2, вірогідність діагностування відмов мікросхем пам'яті, яку можна спрогнозувати, для окремого розглянутого тесту може змінюватись у межах [0.11;0.89], що відповідає практично всьому діапазону змінної, яка нами досліджується.

ВИСНОВКИ

Метод вибору послідовності тестів, який був розглянутий вище, є заснованим на принципі нечітких апріорних даних і дозволяє враховувати об'єктивний досвід та кваліфікацію спеціалістів в області діагностування мікросхем пам'яті. Даний метод можна використовувати в системі підтримки прийняття рішень, де враховуються як точні кількісні параметри, так і інформація, яка представлена в лінгвістичній формі, яка легко сприяється.

ЛІТЕРАТУРА

1. Nissen P. *Some memory testing paradoxes* – Semiconductor test conference: LSI and Boord, Cherry Hill, N.Y.,1978,Did.Pap. Hew York, p.91-95.
2. Данилин Н.П., Попель Л.М. *Установка функционального контроля БИС ОЗУ "Элекон ФЗУ".* – Электронная промышленность, 1981, №2, с.15 – 18.
3. Рябцев В.Г., Торшина В.Е. *Модификация теста "запись-считывание" для диагностирования дефектов устройства управления БИС ОЗУ.* – Электронная техника, серия 8, 1987, вып.1, с.6 – 8.
4. Заде Л. *Понятие лингвистической переменной и ее применение к принятию приближенных решений.* Пер. с анг. – М.: Мир, 1976. – 165с.
5. Кириленко Н.Л., Рябцев В.Г., Кхан М.Х. *Метод принятия решений при разработке программ испытаний памяти.* – Проблемы программирования, 2000, №1-2, с.507 – 512.

Nonlinear model of qualitative classification on base of wave-packet decomposition in space of Volterra kernels.

Akhmetshin A.M., Shutin D.V.

Dnepropetrovsk State University
Email: akhm@mail.dsu.dp.ua, faria@email.dp.ua

ABSTRACT

In the given paper the new approach to the usage of nonlinear models for qualitative classification has been presented. The method is oriented on classification of texture images and stochastic signals. The wave packet decompositions of the initial sequence were interpreted as input and output signals of virtual nonlinear system. The peculiarities of nonlinear transfer characteristics were than considered for performing a qualitative classification. This technique was applied to classification of gastritis on the base of IR spectrograms of the mucous membrane, x-ray CT and x-ray images of the bone trabecular structure, affected with osteoporosis. It was shown, that the proposed method requires much smaller resolution and it is also more sensitive to the changes in the initial sequence.

Keywords: Nonlinear identification, Volterra kernels, wave-packet decomposition.

1. INTRODUCTION

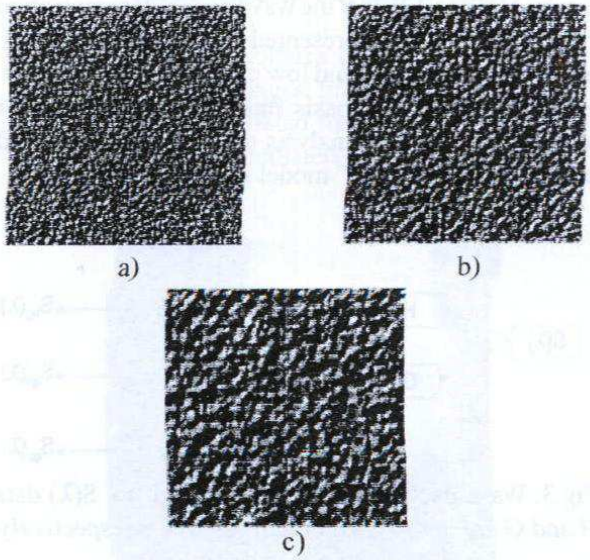
Stochastic processes (signals) find their application in almost every natural system. This fact very often makes them an object of research. In terms of information theory, signal acquired from the system can be interpreted as a message, transmitted to the receiver. The receiver's main role is to "read" the message correctly, in other words, to identify some distinctive characteristics of the message (signal) and classify it. The correct classification of the information received plays one of the most important roles in further processing. In case of medical data processing we would most often speak of correct diagnostics. Therefore, choosing the proper method of signal classification would increase the probity of the results received and quality of the analysis as a whole.

Current methods of data classification are based on linear and nonlinear models. Linear models are most often simple enough. They can be easily simulated and analyzed, but these models are quite rough and unsuitable for the complex processes. On the other hand, the usage of nonlinear models lacks simplicity, but it does provide an exactness and accurateness of the results obtained.

In the given paper we presented a method of qualitative classification of the stochastic signals and images, which is based on nonlinear models and wave-packet decomposition.

2. METHOD

Let's consider three grained texture images (Fig.1.). To increase the sensitivity of the processing procedure, it was proposed to analyze the texture image by the cut. The cut should preserve the spectral information of the image however it will lose information of any objects that are on the texture.



a) b) c)

Fig.1. Three grained texures.

The taken cuts S(λ) of the presented images are shown on Fig.2. In case of stochastic signals, it would be convenient to use AR models[1] for describing them, i.e.

$$S(\lambda) = \sum_{i=1}^{M} P_i S(\lambda_{k-i}) + a(\lambda_{k-i}) \; ; \; k=1,\ldots,N. \quad (1)$$

In terms of the model (1), observable data S(λ) can be interpreted as an output of a linear system excited with a stationary white noise $a(\lambda_k)$. This approach is very efficient when the processes, being simulated, are stationary. However in many cases these signals are non-stationary. It means that the input of our system must be nonstationary as well. In this case we should use another approach to specify the input of the system so that the

© UkrObraz'2000

requirements mentioned above were satisfied. To achieve this, the wave packet decomposition[2] was used. It was

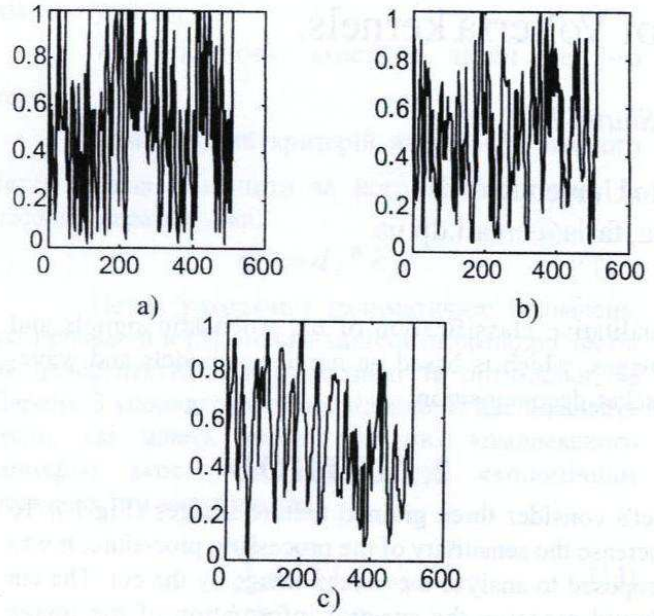


Fig.2. Slices of the grained textures, presented on the Fig.1.

shown by Mallat S.[2], that the wave packet decomposition (till the 2 level)could be presented as it is shown on Fig.3. *H* and *G* are special high and low pass filters respectively. They play a role of the basis functions in wave-packet decomposition. Further analysis of $S(\lambda)$ was conducted in terms of "input-output" model shown on Fig.4.

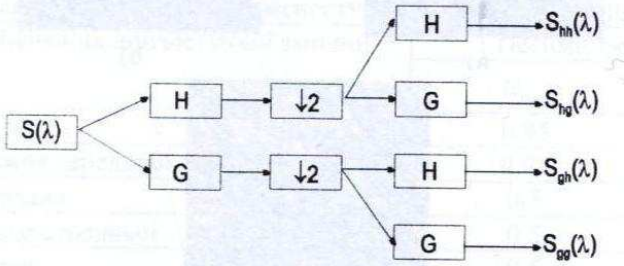


Fig.3. Wave packet decomposition of given $S(\lambda)$ data; H and G are lowpass and highpass filters respectively; ↓2 – dyadic downsampling.

It was assumed, that $S_{hh}(\lambda)$ and $S_{gh}(\lambda)$ are the input and output of the model respectively. Practicability of this approach has following grounds:

1. $S_{hh}(\lambda)$ and $S_{gh}(\lambda)$ are sharing the same "genetic root" (i.e. $S(\lambda)$) and have concerted Fourier spectra. It ensues from the idea of wave packet decomposition.
2. $S_{hh}(\lambda)$ and $S_{gh}(\lambda)$ are non-stationary alternating-sign sequences, which envelope curves are correlated. This fact is the main reason for using model, presented in Fig.4.

The main idea of the analysis is the connection between $S_{hh}(\lambda)$ and $S_{gh}(\lambda)$ through functional second order Volterra series. As long as the second order Volterra kernel $F(\tau_1,\tau_2)$ is an image it should reflect the changes in the signal, making them more obvious for a visual observation. This fact specified the choice of the second order Volterra kernel as a major informative characteristic for the qualitative classification. The kernels of the system with arbitrary input and output sequences can be calculated in the frequency domain by non-parametric method of Powers[3].

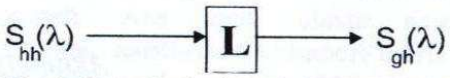


Fig.4. Hypothetical model to form $S_{gh}(\lambda)$; L - nonlinear operator, characterizing the system excited with $S_{hh}(\lambda)$.

Volterra kernels, calculated for the grained textures(Fig.1) are shown on Fig.5.

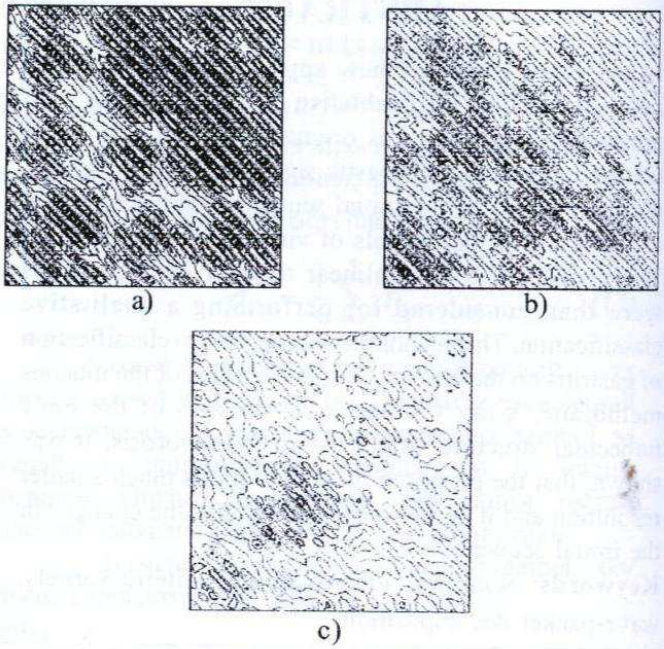


Fig.5. Volterra kernels, calculated for textures, shown on the Fig.1.

3. APPLICATION OF THE PROPOSED METHOD

The proposed method has been applied to qualitative characterization of different stochastic processes, in particular to characterization X-Ray CT and X-Ray images of a coxa, affected with osteoporosis and IR spectrograms of gastric mucos membrane. X-Ray CT images of coxa are presented on Fig.6. The main advantage of using X-Ray CT images is that they give the full information of the bone topology, revealing the local changes of trabeculae. This fact makes the diagnostics of osteoporosis on the base of the tomographic images very effective, especially, on the beginning stages of the disease development. X-Ray CT images, shown on Fig.6. present coxae of a woman at the age of 45. The right coxa (b) is affected more deeply with osteoporosis than the left one (a). This fact has been proved by earlier clinical analysis. White lines show the places where the cuts have been taken from the image for the analysis. It should be noted, that only the oscillating component was used for the analysis because it is responsible for the trabecular bone structure. Therefore, only this component should be analyzed for osteoporosis diagnostics. The Volterra kernels, calculated by the

proposed method form these component are shown on Fig.7. Considering calculated kernels, it can be seen that the peculiarities of the kernel topology reflect the

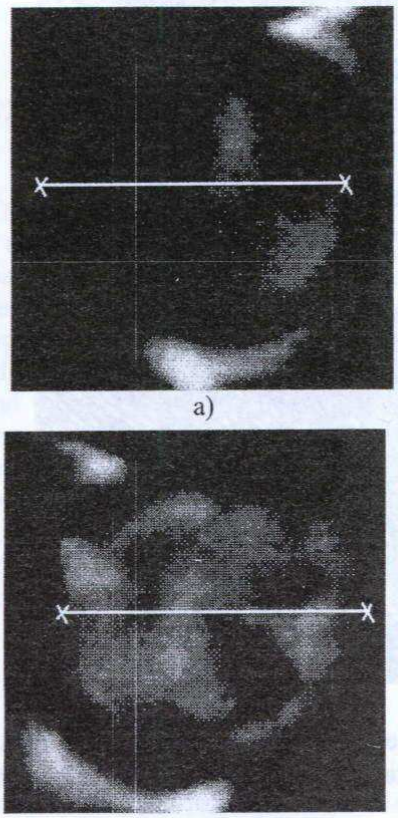


Fig.6. X-Ray CT images of the human coxae, affected with osteoporosis.
a) normal coxa; b) affected with the desease.

changes in bone trabecular structure and therefore they can be used for osteoporosis diagnostics. The proposed method was also applied to the diagnostics of osteoporosis on the base of X-Ray images. One of the classical methods for osteoporosis classification on the base of X-Ray images is an X-Ray densitometry techniques, which are based on dual energy x-ray absorptiometry[4]. This approach gives good results, but the resulting images could not be processed any further, because the brightness of the pixels is connected to the density of the bone and any nonlinear transformation would destroy this brightness/density relationship. Therefore, the method, which could be applied to the X-Ray images acquired with a usual x-ray camera, would be of a great use.On Fig.8. there are two X-Ray images of the left coxa of a person suffering from a quickly developing osteoporosis.The treatment began in 1997 but the drug therapy failed to stop the disease development and in two years the coxa was replaced with prosthesis. It was assumed that a cut of the X-Ray image (the projection of trabecular structure on a plain surface) should contain averaged sizes of trabeculae. Therefore, local changes in trabeculae bone structure should be reflected in local characteristics of oscillating component of a slice. From looking at calculated images of second order Volterra

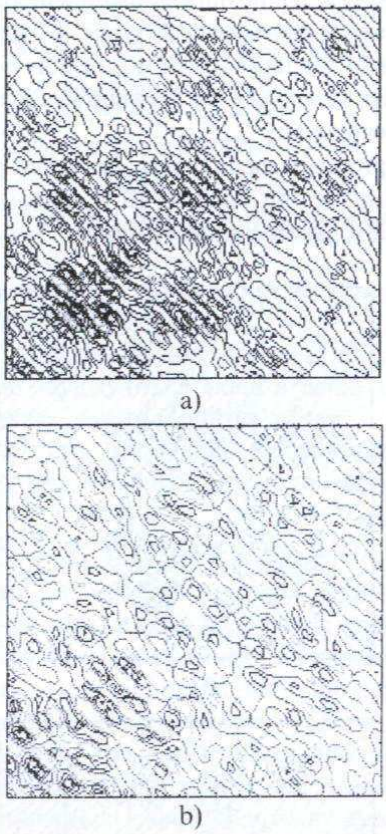


Fig.7. Second order Volterra kernels, calculated for the X-Ray CT images, presented on Fig.5.
a) kernel culculated for the normal coxa;
b) kernel calculated for the ill coxa.

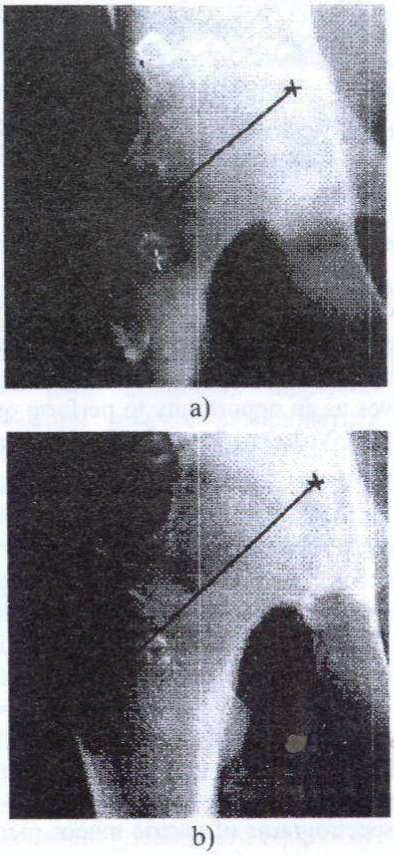


Fig.8. X-Ray images of the coxa,affected with osteoporosis. a) Year 1998;b) Year 1999.

kernels (Fig.9.) it can be noticed that the topology of the image changes as the disease develops. To our mind, the peculiarities of the topology of the kernels cohere with the trabecular structure – the less uniform kernel is, the thicker trabeculae are.

The proposed method was also tested on gastritis classification on the base of IR spectroscopy data (Fig.10.) of the gastric mucos membrane. In comparison

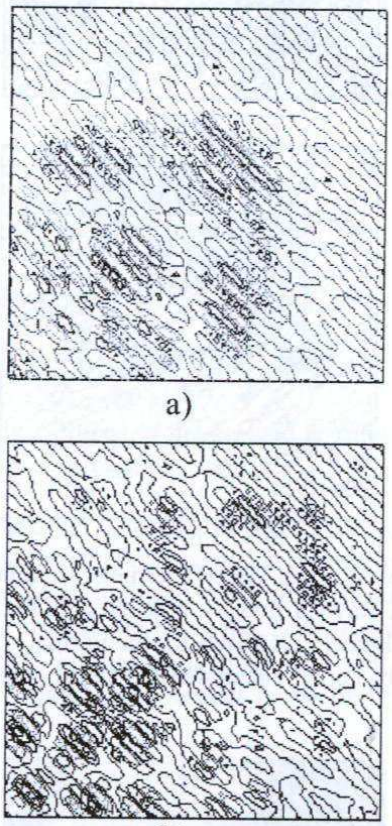


Fig.9. Kernels, calculated for the X-Ray images. presented on Fig.7.
a) Year 1998; b) Year 1999

with other methods of testing for gastroenterological diseases, IR spectroscopy is the most harmless and easy to be conducted. The use of the proposed method gives very good results, especially, when the calculated kernels were further processed (Fig.11.)[5]. Contrasting images on Fig.11 gives us an opportunity to perform qualitative distinctions of Volterra kernels for "normal" and "affected" cases.

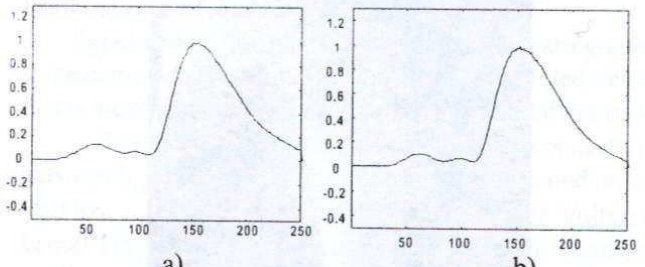


Fig.10. IR spectrograms of gastric mucos membrane;
a) normal; b) gastritis;

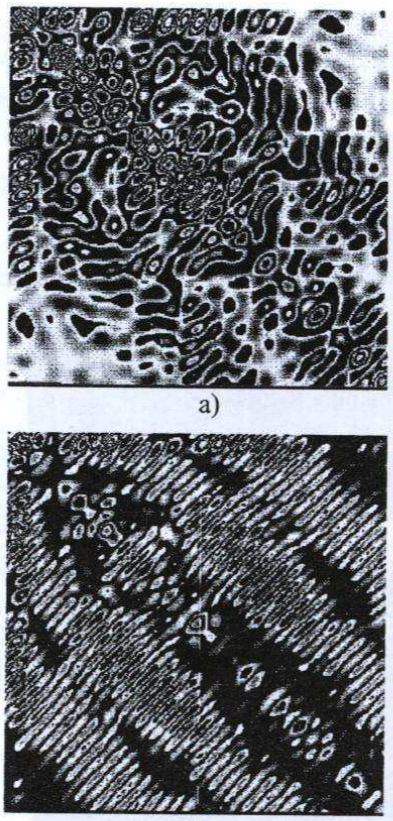


Fig.10. Kernels, calculated for the IR spectrograms presented on Fig.9. a) normal case. b) ill case.

4. CONCLUSION.

It has been shown that the proposed method proved to be sensitive enough to be successfully applied to the analysis of trabecular bone structure on the base of both X-Ray CT and X-Ray images. The results of IR spectroscopy data analysis have also proved that this method can be effectively used for gastritis diagnostics. Nonlinear approach and wave-packet decomposition, implemented in the method, make it effective in analysis of complex stochastic processes, especially, when the physico-mathematical model of the processes being studied is very complex.

REFERENCES.

[1]. Box G.E.P., Jenkins G.M., *Time series forecasting and control*, San Francisco, 1970.

[2]. Mallat S., "A theory for multidimensional signal decomposition: the wavelet representation", IEEE Trans. On Pattern Analysis, v.11, pp.674-693, 1989.

[3]. Powers, E.J., "Applications of digital polyspectral analysis to nonlinear systems modeling and nonlinear wave phenomena," *Proc. Workshop on Higher-Order Spectral Analysis*, pp. 73-77, Vail, Colorado, June 1989.

[4]. Lafferty F.W. "Correlation of dual-energy X-ray absorptiometry, quantitative computed tomography, and single photon absorptiometry with spinal and non-spinal fractures". *Osteoporosis Int* 6:407-415, 1996.

[5]. Akhmetshin A., "Pseudocoherent phase-contrast imaging method with varying focus depth: an effective tool of low-contrast radiological image analysis", SPIE Proc., Medical Image Processing, v.3338, pp.904-915, 1998.

СТРУКТУРНИЙ МЕТОД ОПИСУ І ВИДІЛЕННЯ ВІДРІЗКІВ ПРЯМИХ ЛІНІЙ У КОНТУРІ.

Володимир Калмиков, Дмитро Калмиков, Тетяна Романенко
Інститут проблем математичних машин і систем НАН України
02187 Київ 187 просп. Академіка Глушкова 42, тел. 266 1369, E-mail: kvg@immsp.kiev.ua

The structural method of the description and allocation of pieces of direct lines in a contour is offered which is based on a hypothesis about allocation of pieces of any direct lines in primary visual cerebral cortex. The offered hypothesis demonstrates one of probable mechanisms of display and recognition of pieces of direct any lengths and directions by the neurophysiology means, and can be used as a working hypothesis at experimental researches of visual perception. The received way of the description of image contours is not economical, but the algorithm, received on its basis, allows to restore the contour information as a sequence of pieces of straight lines, which was lost while the image was discretized.

1. ВСТУП.

Запропонований структурний метод опису і виділення відрізків прямих ліній у контурі бінарного об'єкту має в основі гіпотезу про виділення відрізків довільних прямих ліній у первинній зоровій корі мозку людини, що розроблена на основі наступних відомих даних про механізм зорового сприйняття [1].

1. Зображення у полі зору ока проектується на сітку, що утворена зоровими рецепторами, яку можна розглядати як якийсь двомірний дискретний простір;

2. Кожний нейрон поля зору ока (*гангліозна* клітина) збирає інформацію від деякої множини *зорових рецепторів*, що її називатимемо *рецептивним полем*. Середня відстань τ між гангліозними клітинами в центральній частині поля зору ока визначає дискретність поля зору. Кожній гангліозній клітині відповідає нейрон первинної зорової (*стріарної*) кори мозку. Будемо називати такі нейрони *первинними* нейронами.

3. У стріарній корі виявлені нейрони, що збуджуються при пред'явленні у визначеному місці поля зору відрізка визначеної довжини і визначеного напрямку – рис. 1а (зірочками позначені нейрони, дугами – зв'язки їх зі стимулами, що ті нейрони збуджують). Розмір відрізків, на які реагують такі нейрони, складає менше 2° поля зору. Будемо називати такі нейрони *лінійними*. Кількість відмінних один від одного напрямків відрізків, що перетинають одну точку, обмежена і залежить, очевидно, від кількості первинних нейронів, що утворюють рецептивні поля лінійних нейронів, конфігурації цих рецептивних полів і не перевищує 2 - 3 десятків. Природно припустити, що кожній гангліозній клітині в центральній частині поля зору відповідає лінійний нейрон.

4. У стриарній корі виявлені також *кінцеві* нейрони, що збуджуються якщо в даній точці поля зору має місце кінець відрізка прямої. Природно також припустити, що кожній гангліозній клітині в центральній частині поля зору відповідає кінцевий нейрон.

2. ГІПОТЕЗА.

Отже, по сигналах лінійних нейронів можна судити про те, чи є в поле зору ока відрізки прямих деяких фіксованих напрямків, наприклад, як на рис.1а. Чим довше відрізок прямої, що розглядається, тим більше лінійних нейронів даного напрямку, відповідних відрізкам, що взаємно продовжуються, видають сигнали – 1b. У той же час загальновідомо, що зір людини виділяє відрізки прямих дуже великої, значно більшої кількості напрямків серед множини

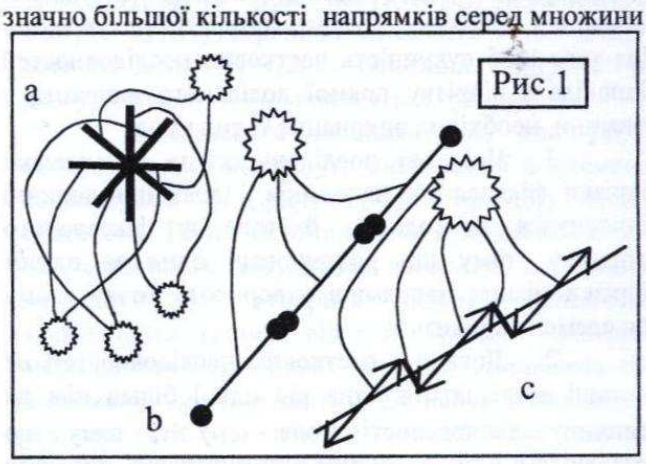


Рис.1

довільних ліній, ніж тієї, що виявлена експериментально. Дійсно, із звичайних повсякденних спостережень з очевидністю випливає, що відмінності між напрямками двох довільних відрізків достатньо великої довжини можуть вимірюватися кутовими хвилинами і, навіть, секундами. Для цього потрібно значно більша дозвільна спроможність, ніж можуть забезпечити самі по собі сигнали лінійних нейронів. У такому випадку повинен існувати спеціальний зоровий механізм, що реалізований у вигляді зв'язків між нейронами, який забезпечує виділення і розпізнавання відрізків прямих ліній довільної довжини і довільного напрямку при розумних витратах кількості нейронів.

Природно припустити, що сигнали лінійних нейронів про відрізки фіксованих напрямків використовуються для побудови образа відрізка прямої довільної довжини і довільного напрямку.

© UkrObraz'2000

Якщо напрямок деякого відрізка прямої довільної довжини в поле зору не збігається цілком із жодним із тих фіксованих напрямків, що виділяються за допомогою лінійних нейронів, тобто цей напрямок є проміжним між двома сусідніми фіксованими напрямками, то такий відрізок не може бути відображений сигналами лінійних нейронів, що відповідають продовжуючим один одного відрізкам обмеженої довжини та одного фіксованого напрямку. У той же час окремі лінійні нейрони, що мають напрямок найближчий до напрямку відрізка, що розглядається у полі зору, видають відповідні сигнали. Проте відрізки, що відповідають цим нейронам, уже не завжди утворюють зв'язану послідовність, тобто кінець попереднього відрізка не завжди збігається з початком наступного – мал 1с. У залежності від того, наскільки напрямок відрізка у полі зору не збігається з фіксованими напрямками, послідовність продовжуючих один одного відрізків фіксованої довжини переривається, "спрацьовують" кінцеві нейрони і починається нова послідовність продовжуючих один одного відрізків фіксованої довжини. Ця нова послідовність зміщена щодо попередньої в напрямку, перпендикулярному фіксованому напрямку відрізків. Такі послідовності, що складаються з одного або деякої кількості продовжуючих один одного відрізків фіксованої довжини і напрямку, будемо називати *частковими.* Для того щоб сукупність часткових послідовностей відповідала відрізку прямої довільного напрямку і довжини, необхідно виконання таких умов:

1. Часткові послідовності та їх складові відрізки фіксованих напрямків і довжини повинні відноситися до одного й того ж фіксованого напрямку, тому що розташовані один за одним відрізки різних напрямків утворюють точку зламу або елемент кривизни;
2. Довжини часткових послідовностей не повинні відрізнятися одна від одної більш ніж на величину дискретності поля зору τ, тому що розташовані один за одним відрізки різної довжини утворюють точку зламу або елемент кривизни;
3. Відстані між кінцями і початками розташованих одна за одною часткових послідовностей не повинні перевищувати величини дискретності поля зору τ тому що в іншому випадку між ними може бути поміщений відрізок, що відрізняється по довжині або напрямку.
4. Напрямки зміщення наступних часткових послідовностей відносно попередніх не змінюються.

Умови 1,3,4 можуть бути легко виконані шляхом утворення рецептивних полів нейронів більш високих порядків із лінійних і кінцевих нейронів у якості рецепторів. Умова 2 може бути здійснена шляхом порівняння електричних зарядів, отриманих від нейронів сусідніх часткових послідовностей, оскільки величина передаваного заряду залежить від довжини відрізка у полі зору.

Запропонована гіпотеза демонструє один із можливих механізмів відображення і розпізнавання відрізків прямих довільних довжин і напрямків засобами, характерними для нейрофізіології і може бути використана в якості робочої гіпотези при експериментальних дослідженнях зорового сприйняття.

3. ВІДНОВЛЕННЯ ВІДРІЗКІВ ПРЯМИХ У ОПИСІ КОНТУРУ.

Далі розглянемо ту ж задачу виділення відрізків прямих у процесі структурного аналізу контурів об'єктів бінарного зображення з використанням основних положень розглянутої гіпотези.

Зображення задане в двомірному просторі на прямокутній гратці $M*N$, як деяка функція $V(m,n) = (0,1)$, $m = 1,2,\ldots ,M$, $n = 1,2,\ldots ,N$. Зображення включає об'єкти - сукупності компактно розташованих пікселів, таких що $v(m,n) = 1$, у той час як $V(m,n) = 0$ для фонових пікселів. Об'єкти звичайно описуються контурами - замкнутими ламаними лініями, що відокремлюють піксели, що належать об'єкту, від пікселів, що належать фону.

Як результат дискретизації вихідного зображення відрізки прямих і кривих довільного напрямку, що спочатку утворювали контур, перетворилися у відрізки вертикальних і горизонтальних ліній - *граней*, що чергуються. Спільні точки сусідніх граней будемо називати *точками зламу*. Представлення контурів у вигляді граней незручно з ряду причин, головної з них є втрата представлення контуру як послідовності відрізків прямих (або кривих) ліній довільного напрямку. Звичайно для відновлення відрізків прямої по окремих точках використовують методи, алгоритми згладжування, що дозволяють одержати приблизне (а не точне) рішення навіть при відсутності поміх.

У роботі [2] був запропонований опис контурів бінарних об'єктів послідовностями L-елементів. Розглянемо виділення відрізків прямих у послідовностях L-елементів. Складові елементарні відрізки - L-елементи утворюються парами, що складаються з розташованих один за одним горизонтального і вертикального (або навпаки, вертикального і горизонтального) відрізків. Перший відрізок може мати довільну довжину: $l = 1,2,\ldots ,L$, де L - максимально можлива довжина відрізка в даній реалізації, у той час як довжина другого відрізка в парі дорівнює 1 (у деяких випадках вона дорівнює 0), що цілком очевидно випливає з визначення алгоритму дискретизації. L-елементи відповідають відрізкам обмеженої довжини у викладеній вище гіпотезі про виділення відрізків прямих довільного напрямки.

Параметрами L-елемента є:

- l - довжина L-елемента, у якості якої приймається довжина першого відрізка; як уже

згадувалося, l приймає значення з ряду натуральних чисел;

- g - напрямок L-елемента, у якості якого приймається напрямок першого відрізка (для певності приймемо g = 0 - "нагору", g = 1 - "управо", g = 2 - "униз", g = 3 - "уліво");

- q = 0,1 - напрямок другого відрізка (0 - поворот управо щодо напрямку першого відрізка, 1 - поворот уліво щодо напрямку першого відрізка).

Розглянемо послідовності L-елементів, що утворять відрізки прямих, для випадку $\Delta y > (x (0$, вважаючи, що для $\Delta x > (y (0$, результат буде *симетричним.* Такі послідовності звичайно утворюються в результаті роботи алгоритму дискретизації відрізків по методу Брезенхама.

Як і для випадку зорового сприйняття, L-елементи, що складають відрізок прямої, співпадають за напрямками першого і другого відрізків, відрізняються по довжині не більш, ніж на величину дискретності простору - один піксел. .

Насамперед відзначимо, що $n=\Delta x$ - кількість L-елементів, що утворюють відрізок. Ціла частина від ділення $\Delta y/\Delta x$ - l=(ЦЧ)$\Delta y/\Delta x$ - це довжина деякої частини L-елементів, що складають відрізок. Довжина інших L-елементів складає l+1. Кількість таких L-елементів - n^* дорівнює залишку від ділення $\Delta y/\Delta x$, тобто n^*=(О)$y/(x$, у той час як кількість L-елементів довжини l складає n-n^*. L-елементи довжин l, l+1, що утворюють відрізок прямої, рівномірно розподілені в межах цього відрізка. Якщо при відображенні відрізка прямої L-елементами $n^*=n-n^*$, то L-елементи довжини l і l+1 чергуються. Якщо $n^*>n-n^*$, то L-елементи довжини l рівномірно розподіляються між L-елементами довжини l+1, і навпаки, у протилежному випадку. Послідовність з однакових L-елементів довжини l, що закінчується L-елементом довжини l+1 або l-1, будемо називати *комплексним L-елементом* або K-елементом. Послідовність з однакових L-елементів довжини l, що починається L-елементом довжини l+1 або l-1, а далі продовжується L-елементами довжини l також є K-елементом. Такий K-елемент будемо називати зворотним. Кількість K-елементів, що утворюють відрізок прямої визначається як $\min\{(n-n^*), n^*\}$. Параметрами K-елементу, таким чином, є:

- параметри першого вхідного в нього L-елемента (для зворотних K-елементів - другого);

- k - кількість L-елементів, що утворять K-елемент; k=(ЦЧ)n/j будемо називати "довжиною" K-елементу; зворотні K-елементи будемо позначати знаком "-" при величині k.

- r = -1, 0, +1; значення r = -1 або r = +1 визначають відповідно зменшення або збільшення довжини останнього в порівнянні з попереднім L-елементом на 1. Якщо r=0, то всі L-елементи даного K-елементу однакові.

Через несумісність величинь n і j деякі K-елементи містять j+1 L-елементів. Їхня кількість визначається як j^* = (О) n/j. K-елементи, що містять різні кількості L-елементів, рівномірно розподіляються в межах відрізка.

Послідовність із k_1 однакових K-елементів довжини k, що починається або закінчується K-елементом довжини k+1 або k-1, будемо називати K1-елементом. Кількість K1-елементів, що утворюють відрізок прямої j_1 =min$\{(j-j^*), j^*\}$. Параметрами K-елементу, таким чином, є:

- параметри першого K-элементу, що входить в нього;

- k_1 - кількість K-елементів, що утворюють K1-елемент; k_1=(ЦЧ)j/j_1 будемо називати "довжиною" K1-елемента; зворотні K-елементи будемо позначати знаком "-" при величині k_1;

- r_1 = -1, 0, +1; значення r_1 = -1 або r_1 = +1 визначають відповідно зменшення або збільшення довжини першого або останнього в послідовності K-елемента в порівнянні з іншими K-елементами даного K1-елемента на 1. Якщо r_1=0, то всі K-елементи даного K1-елемента однакові.

При несумісності величин j і j_1 деякі K1-елементи містять j_1+1 K-елементів. Їхня кількість визначається як j_1^* = (О) j/j_1. K1-елементи, що містять різні кількості K-елементів рівномірно розподіляються в межах відрізка. У цьому випадку можуть бути утворені K2-елементи в кількості j_2 =min$\{(j_1-j_1^*), j_1^*\}$. При несумісності величин j_2 і j_1 можуть бути утворені K3-елементи і так далі, аж доти для деякого i=1,2,... – показника сумірності (або рангу) значення j_i і j_{i-1} виявляться сумірні .

Із сказаного випливає, що для усякого відрізка прямої може бути визначений K-елемент із параметрами $\{l,g,q,k,r,...k_i,r_i,i\}$, причому єдиним способом. Такий засіб опису не є ощадливим, проте дає можливість побудувати алгоритм, що дозволяє визначити, чи утворить задана послідовність чергуючихся граней відрізок прямої. Алгоритм визначається перевіркою умов, чи утворюють грані L-елемент, чи утворюють L-елементи K-елемент, чи утворюють K-елементи K1-елемент і так далі доти чергова пара граней перестане разом задовольняти умовам приналежності до відрізка прямої, вже утвореному раніше переглянутими гранями. Тоді відрізок прямої вважається розпізнаним, а наступна пара граней, певне, уже не є продовженням відрізка прямої і є початком іншого відрізка.

Стосовно до зорового сприйняття сказане вище може позначати наступне. Якщо часткові послідовності, що складають довільний відрізок прямої відрізняються одна від одної на величину τ по довжині, то вони повинні бути розподілені рівномірно уздовж відрізка прямої, подібно тому як K-елементи неоднакової довжини рівномірно розподіляються в межах відрізка прямої. Якби в процесі зорового сприйняття не існувало механізму такого рівномірного розподілу, то відрізки прямых, складені з часткових послідовностей довжини тільки довжини l або тільки довжини l+τ не розрізнялися б один від одного, чого насправді не відбувається. У такому випадку запропонована гіпотеза повинна

бути доповнена ще одною умовою перевірки рівномірного розподілу часткових послідовностей різноманітної довжини в межах відрізка прямої.

Подібно L-елементам часткові послідовності мають утворювати структури, подібні K-елементам, що можуть бути реалізовані у вигляді нейронів часткових послідовностей старших порядків, рецептивні поля яких утворені нейронами часткових послідовностей молодших порядків.

Алгоритм виділення послідовності K-елементів у послідовності L-елементів поданий у вигляді таблиці, де t = 0,1,2,...– номер L-елементу у послідовності.

k	l_t, l_{t-1}	$l_t = l_{t-1}$	$l_t = l_{t-1} + 1$	$l_t = l_{t-1} - 1$	$\|l_t - l_{t-1}\| > 1$
k=1		k_i++;t++;	k_i++; t++; k_i=(-1)* k_i;	k_i++; t++; k_i=(-1)* k_i;	k_i=1; i++;
k>1		k_i++;t++;	k_i++; i++; t++; r_i=+1; k_i=1;	k_i++; i++; t++; r_i=-1; k_i=1;	
k<-1		k_i--;t++;	k_i=1; i++;		

Розглянемо приклади роботи алгоритму. На рис. 2,3 наведені дві різні частини контуру у вигляді відрізків прямих ліній та послідовностей L-елементів, не пошкоджені завадами. Послідовність L-елементів на мал.2 – *de* утворена при дискретизації одного відрізка *DE* розмірами |Δx|=62; |Δy|=45. Послідовність L-елементів на мал.3 – *abc* – утворена при дискретизації двох зв'язаних відрізків *AB* та *BC*, що незначно відрізняються нахилом, загальними розмірами |Δx|=53; |Δy|=36. Розміри першого відрізка – *AB* - |Δx|=33; |Δy|=23, другого - *BC* - |Δx|=20; |Δy|=13. Сегментувати конфігурацію на мал.3 на окремі відрізки важко навіть вручну. Але ж коректна сегментація контурів має суттєве значення при обробці зображень . Така сегментація не пошкоджених завадами контурів може бути виконана при синтаксичному аналізі послідовності L-елементів, що утворюють контур. При послідовному застосуванні наведеного алгоритму послідовність *de* утворює один K2-елемент, у той час як послідовність *abc* утворює два K2-елементи, розділяючи послідовність L-елементів саме у точці *b*.

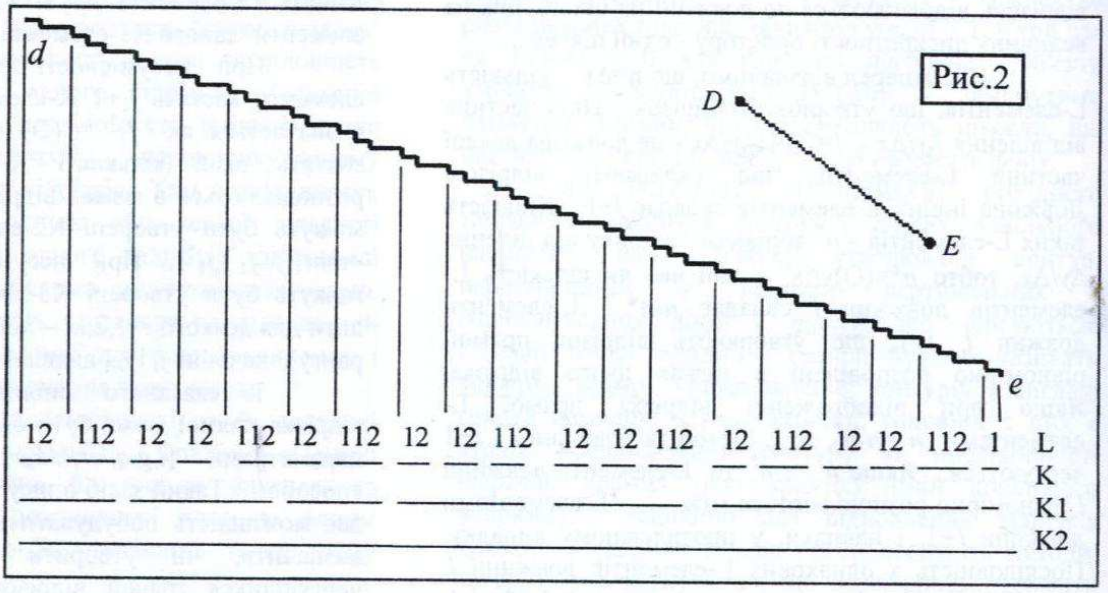


Рис.2

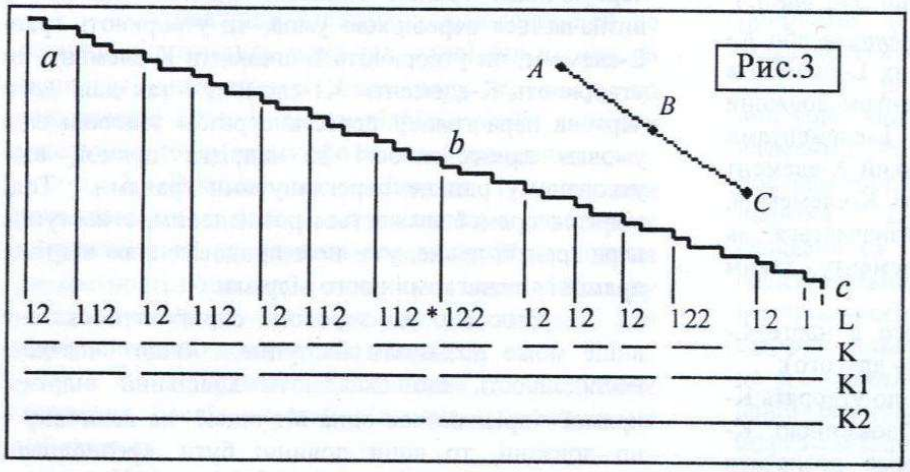


Рис.3

4.ВИСНОВКИ.

Запропоновано метод та алгоритм структурного аналізу відрізків прямих, що дозволяє відновити опис контуру у дискретному просторі, як послідовність відрізків прямих. Проведено зіставлення з механізмами зорового сприйняття, показано їх подібність відносно операцій над елементами зображення, що дозволило сформулювати гіпотезу, придатну до експериментальної перевірки.

ЛІТЕРАТУРА.

1. Хьюбел Д. "Глаз, мозг, зрение." – Москва: Мир, 1990 - 239 с.

2. Калмыков В.Г. Вишневский В.В. Анализ контуров объектов в бинарных изображениях. // Математические машины и системы - 1997 - №2 - с.68- 71.

IMPROVEMENT OF LOW – CONTRAST MEDICAL IMAGE ANALYSIS SENSITIVENESS AND EXACTNESS USING MARKOV RANDOM FIELD MULTIDIMENSIONAL SEGMENTATION METHOD

A.M. Akhmetshin, A.E. Fedorenko

Dnepropetrovsk State University

Abstract

In this work algorithm of medical images visual properties improvement for image analysis simplification was realized. The exactness and sensitiveness increasing of different image objects detection is achieved via gradient correction and segmentation. Pixels classification proceeds using Markov random field (MRF) conception. In order to increase the input segmentation data the multiresolution Wavelet decomposition method was proposed to take in consideration the information about more coarse region forms and borders. Using a number of "approximation" images generated by Wavelet transform together with the initial image allows us to make the resulting clusterisation more exact.

1. INTRODUCTION

In the area of medical images processing the resulting inaccuracy is usually conditioned by both the absence of sufficient input data amount and weak photos contrast. Accurate definition of weak contrast regions is more convenient to be accomplished with any image segmentation (clusterisation) methods. There are two different segmentation ways: "unsupervised" and "supervised". The obvious weakness of supervised algorithms such as K – means method or fuzzy C – means method [1],[2],[4] often lies in the absence of knowledge about true number of image objects, which must be detected. Due to this problem we used unsupervised clusterisation method. Moreover, during the segmentation process local spatial image characteristics of MRF model were taken into account what allows us to consider the input image not only as a simple set of different brightness pixels but also as a complex textured object.

Lets consider some gray – scaled image $Y=\{y_s\}$, where s is a pixel characterized by its coordinates: $s=(i,j)$, $i\in\{1..N\}$, $j\in\{1..M\}$, y_S is a gray level of s. The segmented version of Y is $X=\{x_s\}$; $x_s=k$ means that pixel s can belong to the segmentation class k.

The most known pixels classification method for the segmentation process is the maximization of a posteriori probability (MAP), which is usually defined by Bayesian relationship:

$$P(X \mid Y) = P(Y \mid X) \cdot P(X) \qquad (1)$$

$P(X)\sim p(x_s=k)$ is a priori probability of pixel s belonging to segmentation class k, $P(Y|X)\sim p(y_s=y|x_s=k)$ is a conditional distribution of pixels intensity within certain class k. Both a priori and a posteriori probabilities are defined with respect to MRF model.

2. GAUSSIAN MARKOV RANDOM FIELD MODEL

Let an image Y be modeled on a finite lattice L by Gaussian Markov random field (GMRF) where $Y=\{y_{ij}: 0\le i\le M,\ 0\le j\le N\}$ and $L=\{(i,j): 0\le i\le M, 0\le j\le N\}$. A neighborhood system η_s is a set of all the pixels which are the nearest neighbors of pixel s. η_s must satisfy the following requires:

- s is not an element of η_s;
- if $k \in \eta_s$, then s should be an element of η_k.

The joint probability distribution of pixel brightness values is:

$$P(y) = \frac{1}{(2\cdot\pi)^{M\cdot N/2}(\det\Sigma)^{1/2}} \cdot e^{-\frac{1}{2}[y]^T\cdot[\Sigma]^{-1}\cdot[y]} \qquad (2)$$

where $y=[y_{11},y_{12},\ldots,y_{MN}]^T$, and Σ is a covariance matrix of y.

In addition given image must satisfy the Markovian property:

$$P(y_{ij} \mid y_{pq},(p,q)\neq(i,j)) = P(y_{ij} \mid y_{kl},(k,l)\in\eta_{ij}) \qquad (3)$$

The nearest pixels spatial interaction is described statistically by Gibbs distribution.

3. GIBBS DISTRIBUTION

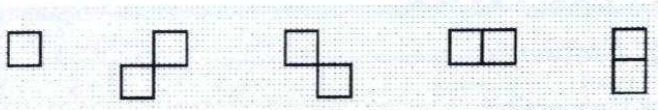
Lets define a clique as a set of one or two neighboring pixels. Given image Y can be modeled as MRF with respect to η when and only when its joint distribution is defined according to Hammersley – Clifford theorem:

$$P(X) = \frac{1}{Z}\cdot\exp\{-U(y)\}, \qquad (4)$$

© UkrObraz'2000

where Z is a normalization constant, U(y) is an energy function value, which defines the generalized interaction of certain pixel with it's neighbors:

$$U(y_S) = \sum_{c \in C} V_C(y_S) \quad (5)$$



a) b) c) d) e)

Fig. 1. Different cliques used in work and pixels mutual spatial location within each kind of clique. a) Single clique; b)– e) Double cliques

V_c is a potential function of type c which defines a degree of clique pixels interaction in a way of fuzzy pixels similarity evaluation. Membership function of the fuzzy relationship " y_s and y_l are similar":

$$\Phi(y_s, y_l) = \frac{|y_S - y_l|}{N} \quad (6)$$

Two – pixel clique potential V_c is defined by value $\Phi(s,l)$, thus it is an increasing function in a closed interval $[0,\beta]$

$$V_c(y_s) = \Phi(y_s, y_l \mid l \in \eta_s) \quad (7)$$

It corresponds to simplified fuzzy inference rule: "if pixels s and l are similar, then V_c is **V**". Based on the given principle the a priori probability as well as the conditional distribution defined in (1) are calculated.
A priori probability:

$$p(x_s) = \exp\{-U(x_s \sim y_t, t \in \eta_s) \quad (8)$$

where

$$U(x_s \sim y_t, t \in \eta_s) = \sum_{l=1}^{7} \frac{|y_t - y_l|}{N}, \quad \forall l \in \eta_s \quad (9)$$

Conditional on certain class probability:

$$p(y_s \mid x_s) = \exp\{-U(y_s \mid x_s \sim y_t, t \in \eta_s)\} \quad (10)$$

where

$$U(y_s \mid x_s \sim y_t, t \in \eta_s) = \frac{|y_s - y_t|}{N} \quad (11)$$

4. METHOD

Proposed method is baseed on neighboring pixels gradient analysis. Image clusterisation algorithm consists of several steps:
1. For extraction of homogeneous regions, the initial image is exposed tos local covariance transform using following equation:

$$\tilde{a}_{ij} = a_{ij} \cdot (1 - c_{ij}) \quad (12)$$

where

$$c_{ij} = \left.(q_{ij} - \mu_{ij}) \cdot (q_{ij} - \mu_{ij})^T \right/ Norm \quad (13)$$

q_{ij} is a row of the nearest eight neigboring to (i,j) pixels brightness values, $\mu_{i,j}$ is mean of q_{ij}. This operation pickes out all image homogenous regions (fig. 2 (b)).
2. Homogeneous regions contrast increasing is achieved by gradient correction and histogram equalization method. Unfortunately, clear histogram equalization method application not always allows us to receive desired exactness of image details. Therefore, before using the histogram equalization algorithm the image gradient correction is accomplished by:

$$b_{ij} = \tilde{a}_{ij} - \Delta a_k \quad (14)$$

where Δa_k is a brightness difference on direction k, κ=1..8:

$$\Delta a_k = a_{nm} - \tilde{a}_{ij}, \quad (m,n) \in \eta_{ij} \quad (15)$$

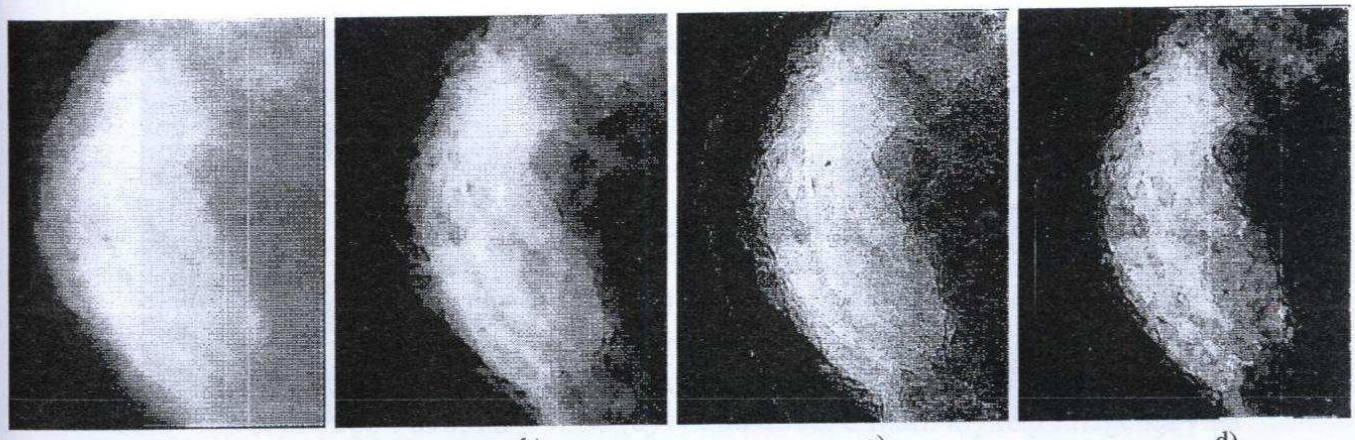
It can be seen that there are eight different values of brightness gradient for each pixel. Therefore it's necessary to define the best gradient meaning according to some criterion. In this work choosing of gradient values is accomplished with famous feature detection algorithm according to entropy minimization principle (fig.2 (c)).
3. Achieved on stage 2 image is then used as initial for segmentation. Afterwards it is corrected with Bayesian segmentation method, based on MRF theory. Under classification of each pixel the following rule is used:

$$x_s = \arg\max_t P_t(X \mid Y) \quad (16)$$

where $P_t(X|Y)$ is calculated for each of eight directions t from pixel s. During clusterisation process together with the input image two "low-pass" images found by "herringbone" algorithm of two–dimensional discrete Wavelet transformation are processed to evaluate common region forms. In this case we have deal with multidimensional version of unsupervised clusterisation based on MRF model (fig.2 (d)).

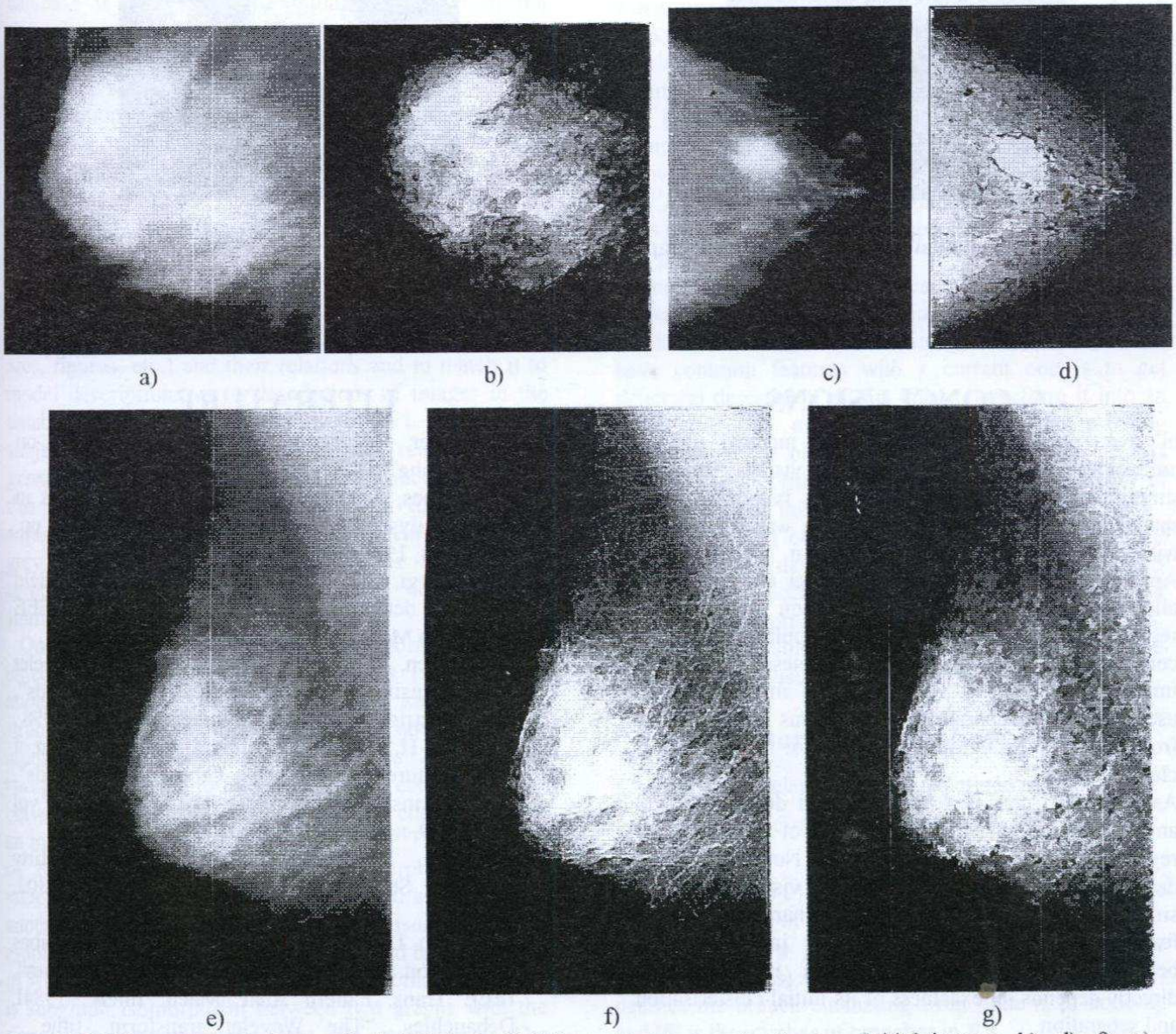
5. RESULTS

Given algorithm was tested on series of digital mammograms and on several tomographic images. Testing results show that the application of the method gives quite good results in the increasing of exactness of different medical images analysis in the sense of more accurate detail extraction and contrast improvement of the researched image object forms and borders (fig. 3). Due to clusterisation described method has a real possibility to be used as a method of suspected regions extraction before their direct recognition.



a) b) c) d)

Fig. 2 Results of method application on different algorithm steps. a) Initial image; b) Image with extracted homogeneous regions; c) Image (b) after gradient correction; d) Multidimensional MAP – segmentation result with application of MRF model and Wavelet decomposition.



a) b) c) d)

e) f) g)

Рис. 3 MAP – segmentation method tested on digital mammograms. a), c), e) – Initial images, b), d), f) ,g) – Segmentation result.

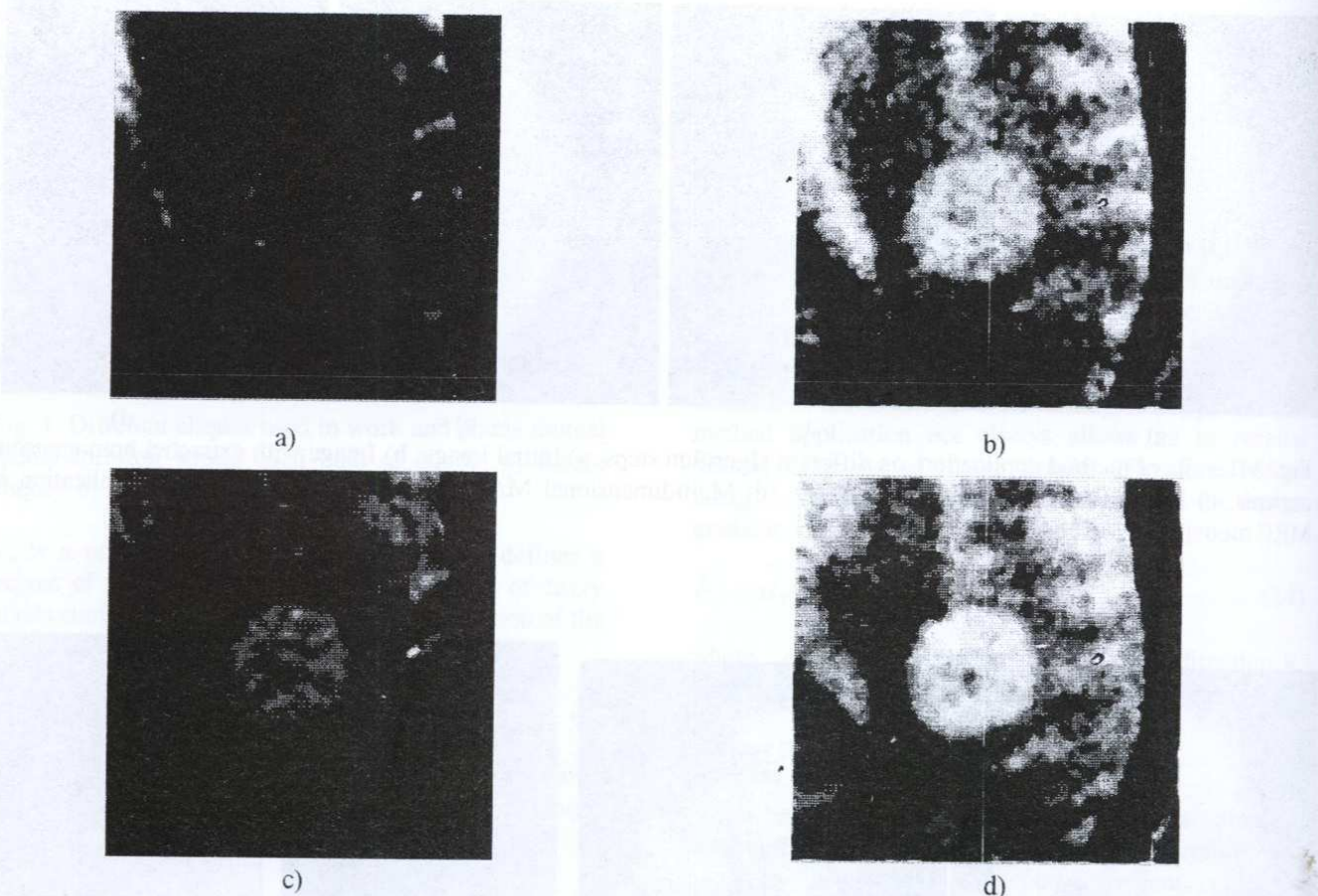


Fig.4. MAP – segmentation method tested on tomographic images. a), c) – Initial images, b), d) - Segmentation result.

6. CONCLUSIONS

In given paper we have developed medical images processing method for image visual analysis simplification. Proposed technique is based on local spatial image features analysis, whish was carried out with the help of Markov random field model. Application of multiresolution Wavelet decomposition allows us to make additional correction of segmented regions. Results of algorithm application show considerable improvement of exactness and visual image characteristics. Thus proposed method can be used in tasks of separating of suspicious tumor regions from normal background tissues for their further classification.

As it can be seen, performed method doesn't contain any classifier for the final recognition of ill and normal regions within certain medical image. Nevertheless, the described problem of image visual analysis simplification in a way of its visual characteristics and features improvement plays a very important role, because the final classification of image objects directly depends on exactness of its initial clusterisation and separation.

LITERATURE

[1] R. Kohler, "A segmentation system based on thresholding", CGIP, vol. 15, pp 319 – 338, 1981

[2] R.C. Dubes, A.K. Jain, "Random fields models in image analysis", J. Applied Stat., vol. 16, no 2, pp. 131 – 164, 1989

[3] M. Callergi, L.P. Clarke, "Markov Random Field for tumor detection in digital mammography", IEEE Trans. On Medical Imaging, vol. 14, no 3, 1995

[4] C.H. Chen, G. G. Lee, "Multiresolution wavelet analysis using gaussian markov random fields", Dep. Electrical Computer Engineering, UMD, 1998

[5] H. Derin, H. Elliott, "Modeling and segmentation of noisy textured images using Gibbs random fields", IEEE Trans. Pattern Anal. Machine Intell., vol PAMI – 9, no 1, pp 39 – 55, 1987

[6] J. Besag, "On the statistical analisys of dirty pictures", Stat. Soc., vol. 48, no 3, pp 259 – 302, 1986

[7] S. Geman, D. Geman, "Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images", IEEE Trans. Pattern Anal. Mach. Intell., 1984I. Debauchies, "The Wavelet transform, time – frequency localization and sygnal analisys", IEEE Trans. Inform. Theory, no 36, pp 961 – 1005, 1981

Attributed relational graphs based description and retrieval of shoeprints in the image database system SHARS

Ch.Feist[1], *V. Kiiko*[2], *R.Knöfel*[1], *G. Stanke*[1]

[1]Gesellschaft zur Förderung angewandter Informatik e. V. (GFaI),
Albert-Einstein-Str. 16, D-12489 Berlin
Tel.:+49 (0)30 6392-1610, Fax: +49 (0)30 6392-1661, e-mail: stanke@gfai.de
[2]Int. Research and Training Centre of Information Technologies and Systems of National Academy of Sciences of Ukraine, Prospect Academica Glushkova 40, 252022 Kiev
Tel.:+380 (0)44 266-3018, Fax: +380 (0)44 266-1570, e-mail: kiyko@image.kiev.ua

Abstract

In this paper attributed relational graphs based approach implemented in the image database system SHARS for archival and retrieval of shoeprints from a database is described. The initial description of the shoeprints is a set of drawn by user figures, which then is transformed to an attributed relational graph (ARG) describing shoeprints in terms of some primitives and their relations. To retrieve shoeprints from the database this ARG is matched with all ARGs in the database using proposed exact matching algorithms.

1. Introduction

So called "geometric" approach for retrieval of images from the database is to get a structural description of the image as a set of extracted geometric features (lines, arcs, figures, etc.) and their relations and to match it to model descriptions or to descriptions of images in the database. The automatic detection of such features is a subject of many researches but still remains on in general unsolved problem. One of the recent works in this field [1] represents an algorithm for search or adjusting figures consisting of line segments and arcs in grey scale images. This algorithm is based on a computationally efficient procedure for defining similarity between two images.

One of the powerful tools for describing structured objects are attributed relational graphs (ARGs). The nodes of ARG represent primitives or subpatterns of structured objects, and branches between nodes represent mutual relations between these primitives [2]. To match two graphs we need to find a one-to-one correspondence between their vertices and between their edges such that an incidence relationship is preserved [3]. This matching of two graphs is called graph isomorphism. If the matching is fulfilled between a graph and a subgraph of another graph, then it is called subgraph isomorphism or graph monomorphism. Terms of exact and best matching of ARGs are also important. Under exact matching there is subgraph isomorphism between two graphs with the same attributes on correspondent vertices and edges. The best matching means an optimal solution of the subgraph isomorphism problem when the distance between two matched graphs is minimal.

A number of algorithms are proposed for matching attributed relational graphs [2,4-10]. It is known that the graph matching problem belongs to the class of *NP* complete problems and computational complexity of any conventional search algorithm may grow exponentially when increasing the size of matched graphs. Some proposed algorithms have polynomial complexity for special types of graphs but still have exponential complexity at the worst case.

In [10] the algorithm for detection subgraph optimal isomorphism between two ARGs is proposed which is known to have an $O(N^4)$ average case complexity and an $O(N^5)$ in the worst-case complexity. In [11] the algorithm for detection subgraph isomorphism between two planar graphs in time $O(c^{w\log w}\, n)$ is described where w and n are numbers of nodes in input and reference graphs and c is some constant value.

Widely used way to retrieve images from database that have common features with a current one is to get structural description of this image, transform it into an ARG, then to match it with ARGs representing structures of the other images in the database. In this paper some algorithms for description and retrieval of shoeprints along this way are considered.

The definition of the attributed relational graph and also steps of the description of the shoeprints in SHARS are shown in Section 2. Section 3 introduces some subgraph isomorphism based retrieval algorithms. Experimental results are presented in Section 4, and finally, the conclusions are given in Section 5.

2. Description of shoeprints

An attributed relational graph is defined as [5]

$$G=(N, B, A, E, G_N, G_B),$$

where

$N=\{n_1, n_2,\ldots, n_{|N|}\}$ is a finite set of nodes; $|N|$ is the number of nodes in N;

$B=\{ b_1, b_2,\ldots, b_{|B|} \}$ is a set of ordered node pairs (or directed branches) i.e., $b=(n_i, n_j)$ for some $1\le i,j \le |N|$ denotes the branch emanating from node n_i to node n_j, and $|B|$ is the number of branches in B;

A is an alphabet of node attributes;

E is an alphabet of branch attributes;

$G_N : N \to 2^{|A|}$ is a function that defines the node attributes;

© UkrObraz'2000

G_B:B→$2^{|E|}$ is a function that defines the branch attributes.

Let us consider now separate steps of description of the shoeprints and representing them by ARGs in image database system SHARS. The initial description of the shoeprint is a set of geometrical figures drawn by user in the image of this shoeprint. User can also define one or more characteristic subsets of this set, which must be on retrieved shoeprints.

Each of currently used figures (line, half of the circle, circle, polyline, polygon, square, triangle, rectangle, parallelogram, rhombus, rings, wave, tooth and position) contains so called control points and to draw or modify any figure in the image only these points must be defined.

On the next step every figure is represented by ARG with some attributes on its branches and nodes. Every node or branch of this graph has one context mark or attribute (centre of the circle, centre of half of the circle, end of line segment, line segment) and none or some of the following metric attributes: coordinates, radius, length and direction of the line segment, cross angle and relation of lengths of two neighbour line segments. The number of used metric attributes depends on some options of the program and also of the possibility for the user to define the position and orientation of the shoeprint in the image as well as the resolution of this image.

As an example let us consider attributed relational graph G_p of the figure „polyline" that consists of two line segments. This graph consists of three nodes and two branches. The nodes of the graph G_p have the context mark „end of line segment" and the branches have the context mark „line segment". The nodes of the graph have metric attributes "coordinates" if: 1) the resolution of the image is known, 2) the user has estimated the position and direction of the shoeprint in the image by drawing the figure „position" and 3) the option „Coordinates" is switched on. The middle node of the graph in any case has also metric attributes „lengths relation of neighbour line segments" and „angle between neighbour line segments". The branches of the graph have the metric attribute „length", if the resolution of the image is known, and the metric attribute „direction", if the user has defined the orientation of the shoeprint in the image and the option „Angle" is switched on.

It follows that

$$G_p=(N, B, A, E, G_N, G_B),$$

where

$N=\{ n_1, n_2, n_3 \}$;

$B=\{ b_1, b_2 \}$;

$A=\{$ end of the line segment, coordinates, lengths relation of neighbour line segments, angle between neighbour line segments $\}$.

$E=\{$length, direction$\}$.

Attributed relational graphs that define characteristic groups of figures in the shoeprint are similar to ARGs of the figure "polyline" but instead of the node attribute "end of line segment" they have the node attribute "label of figure".

3. Retrieval algorithms

Retrieval algorithms are based on matching (subgraph isomorphism) of the input ARG with ARGs of shoeprints in the database and also on the distance measure between these graphs. For calculating the distance between two ARGs the concept of error-correcting transformation [5,7,10] is adopted. The cost of matching two ARGs (or distance between these graphs) is defined as the cost of the sequence of transformations with minimal total cost that must be performed on one of the two ARGs in order to produce another ARG. These transformation operations are: *node insertion, node deletion, branch deletion, branch insertion, node label substitution and branch label substitution.* The costs corresponding to each operation are: w_{ni}, w_{nd}, w_{bi}, w_{bd}, w_{ns} and w_{bs} respectively. It is reasonable in our case to assume that $w_{nd}=w_{bd}=0$, $w_{ni}>0$, $w_{bi}>0$, $w_{ns}=w_{bs}=1$ and the distance $dist(x,y)$ between two numeric attributes x and y is equal to v=abs(x-y) if v exceeds some threshold value and to 0 otherwise. It can be defined that we have the exact matching if the distance between two graphs is equal to 0 and the best matching otherwise. In both cases the part of one graph must be defined that has the minimal distance with another graph but for exact matching this distance is equal to 0 and less computational time is usually required.

ARGs, corresponding to currently use in SHARS figures, are either chains or cycles. That is why the way of checking exact matching between two ARGs during moving lesser graph along greater one can be used. The computational complexity of this algorithm is $O(M*N)$.

There are some criteria for matching of input and reference descriptions of shoeprints implemented in the system SHARS. By default two descriptions are matched if at least one figure in input description is exact matched with one or more figures in the reference shoeprint. The user may also define the number of pairs of figures or group of figures in the input shoeprint that must be matched with ARGs of the reference shoeprints.

To match the group of figures we need to find subgraph isomorphism between two graphs. One of them is the complete (fully connected) graph such that nodes of this graph represent different figures in the shoeprint from the database and each of these nodes is connected to all other nodes of the graph. The other graph G represents the chain of figures in the input shoeprint. Each node of this graph has a label of the corresponding figure in the group and number of nodes in G is equal to number of figures in the group. The following algorithm checks some necessary conditions of subgraph isomorphism between these two graphs. During the first part of the algorithm for each node of the graph G the list of exact matched with this node figures from the reference description is constructed. During the next step all possible excludings of those figures from these lists that do not satisfy to local relations between the neighbour figures in the group are carried out. It means that each figure F must be excluded from the list if there are no in neighbour lists two other figures that satisfy with F to local relations of neighbour relations of figures

in the group. The process of excluding figures is stopped when additionally none of the figures can be excluded. Therefore the input and reference graphs are considered as matched if the final lists of these figures are not empty in spite of the fact that there is exactly subgraph isomorphism between these graphs only when each of lists has only one label or figure and all these labels are different.

For matching ARGs thar represent figures of other types than currently used in SHARS one of the algorithms [2,4-10] can be implemented. For this aim new algorithm for exact matching ARGs was also developed which is described below.

The main problem using the error-correction distance measure is the computational complexity of the matching that may grow up exponentially when the graph size increases. To reduce the complexity original graphs can be decomposed into smaller subgraphs (basic attributed relational graphs or BARGs) and the matching of the original graphs to fulfil through the matching of these decomposed subgraphs [10]. The BARG represents a one-level tree, which consists of a node (root), and every node (descendant) connected to that node by one emanating or incoming branch. The cost of matching two BARGs G_k and G_p can be defined as

$$\text{Distance}\,(G_k, G_p) = w_{ns} * \text{dist}(r_k, r_p) + \text{dist}(b's, e's),$$

where

$\text{dist}(r_k, r_p)$ is the distance between root node labels and attributes of two BARGs that is computed depending on their data types,

k and p are the number of branches connected to the root node BARGs being matched, and

dist(b's, e's) is calculated as the minimum of a weighted bipartite graph constructed from branches of BARGs as their nodes and its structure is as shown in Fig. 1.

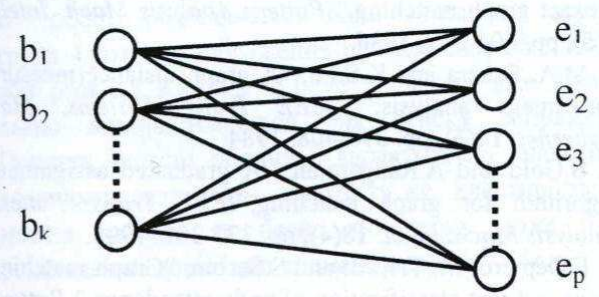


Fig. 1. Weighted bipartite graph corresponding to branches distance.

The minimum of a weighted bipartite graph is defined as the complete matching with the minimum sum of the matching edges weights. There are some algorithms for such matching with an $O(N^2)$ average case complexity and an $O(N^3)$ worst-case complexity [12].

Exact matching between branches of two BARGs means that dist(b's, e's) = 0, i.e., there is one-to-one correspondence between branches of these BARGs and correspondent branches have the same attributes. To define such exact matching it is necessary to exclude from bipartite graph in Fig. 1 those branches that have non-zero weights and then to use one of the algorithms f.e., [13] to find assignment for remain branches.

Let G_i and G_r be input and a reference ARGs. The first step of the algorithm for checking exact matching between these two graphs consists in their decomposition into BARGs. Then the distances between decomposed BARGs are defined, it means that a distance matrix D between BARGs of reference and input graphs is constructed. The $D_{i,j}$ element of this matrix represents the distance between the ith BARG in the input graph and the jth BARG in the reference graph, and $D_{i,j}=0$ if there is exact matching between these two BARGs and $D_{i,j}>0$ otherwise. On the base of distance matrix for every node of lesser graph (having less nodes than another one) a list of BARGs of greater graph that have exact matching with this node and not lesser number of connected branches is made out.

During the next step of the algorithm we scan constructed lists of BARGs and exclude every BARG B in the current list if the list of BARGs in the neighbour node is either empty or contains only BARGs that are not neighbours to B in the reference graph. During this step can be also helpful some additional tests, f.e. : a) if some list L contains only one BARG this BARG can be excluded from all other lists or b) if the number of different BARGs in all lists less than the number of nodes in input graph the input and reference graphs can not be matched.

This process is stopped when additionally none of the BARGs can be excluded. Therefore the input and reference graphs are not matched if the final lists of BARGs are empty and are exact matched if each of these lists contains only one BARG and all these final BARGs are different. In the case when final lists contain more than one BARG additional testing is needed.

Let us estimate computational complexity of the algorithm. Suppose that the input and reference graphs have M and N nodes relatively. To decompose both G_i and G_r into BARGs we need quadratic time for each graph, i.e., this step has a time complexity of $O(M^2+N^2)$. The computational complexity of exact matching between two BARGs is $O(M*N)$, assuming that any node is connected to all other nodes in the graph, and the computational complexity of calculating the distances between all BARGs of both input and reference graphs is $O(M^2*N^2)$. The computational complexity of the second (BARGs excluding) step of the algorithm is $O(M^2*N^3)$, $M<N$ in the worst case. In summary, the computational complexity of the proposed algorithm for checking the exact matching between two ARGs is also $O(M^2*N^3)$ in the worst case.

4. Experiments

During experiments 150 shoeprints were encoded and saved in the database. After that the testing group of the 25 different shoeprints was formed and shoeprints of this group were then encoded and retrieved by some users.

Each shoeprint from this group has one or more shoeprints of the same type in the database, which differ from each other by size, position and many other features, for example some of these shoeprints are only parts of other ones (Fig. 2).

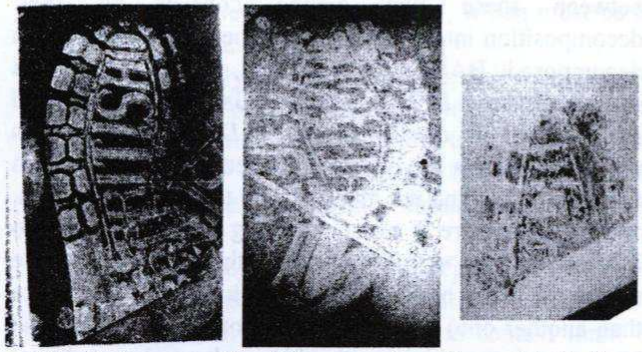


Fig. 2. An example of the same type shoeprints in the database.

Many shoeprints in the database are not supplied by metric information because of the absence of data about resolution of the image or the position of the shoeprint in this image. The number of retrieved shoeprints by one figure or by one pair of figures for such shoeprints is not so small. As follows from experimental data the most effective possibility for the retrieval of most shoeprints from a database consists in defining of a group of figures in the input shoeprint. This possibility provides a sharp reduction of number of retrieved shoeprints and preserves the majority of the same type retrieved shoeprints from the database.

It is important that retrieval of the shoeprints from database is fulfilled using not the names and parameters of figures but using attributed graphs that represent these figures and also on the base of subgraph isomorphism of these graphs. This fact reduces dependence of the retrieval results on subjective actions of the user during describing or coding input shoeprints because in many cases the same object in the shoeprint can be described by figures that have different names and/or different numbers of its components but attributed graphs that represent these figures can be matched. Introduced options and some possibilities for the definition of an inquiry to a database (retrieval by figure, by pairs or by groups of figures) can be also useful in this direction.

5. Conclusions

In this paper ARG based approach and algorithms for the retrieval of shoeprints from a database are proposed. The proposed algorithms are based on the exact matching of ARGs that describe input and reference shoeprints. The distance is calculated based on the concept of error-correcting transformations with taking in account some peculiarities of the shoeprints description and retrieval problem. From experimental results the following remarks can be concluded.

The developed tools can be used for the description and retrieval of shoeprints from a database. The proposed algorithms for detection subgraph isomorphism between input and reference ARGs and also the introduced different possibilities for the definition of an inquiry to a database reduce dependence of the retrieval results on subjective actions of the user during describing or coding input shoeprints. On the next step the algorithms for optimal subgraph isomorphism will be tested and added to the system for the further reduction of this dependency.

The implemented ARGs based tools and algorithms are an addition but not a substitution for the other method implemented in SHARS that is based on coding shoeprints by rather global geometrical features, that describe shoeprint in its different parts.

References

1. I.Aksak, Ch. Feist and others, "Detection of the Objects with the Given Shape on the Grey-Valued Pictures," CAIP'97 "Computer Analysis of Images and Patterns", 1997, Kiel, p.551-558
2. A. Sanfelio and K.S. Fu, "A distance between attributed relational graphs for pattern recognition," *IEEE Trans. Systems, Man Cybernet.* **13(3),** pp. 353-362, 1983.
3. Narsingh Deo, "Graph Theory with Applications to Engineering and Computer Science," Printice-Hall, Englewood Cliffs, New Jersey, 1995.
4. A.K.C. Wong, M. You and S.C. Chan, "An algorithm for graph optimal monomorphism, *IEEE Trans. Systems, Man Cybernet.* **20(3),** pp. 628-636, 1990.
5. W.Tsai and K.S.Fu, "Subgraph error-correcting isomorphism for syntactic pattern recognition," *IEEE Trans. Systems, Man Cybernet.* **13(1),** pp. 48-62, 1983.
6. M.A. Eshera and K.S.Fu, "An image understanding system using attributed symbolic representation and inexact graph matching," *Pattern Analysis Mach. Intell.* **8(5),** pp. 604-617, 1986.
7. M.A. Eshera and K.S.Fu, "A graph distance measure for image analysis," *IEEE Trans. Systems, Man Cybernet.* **14(3),** pp. 398-408, 1984.
8. S.Gold and A.Rangarajan, "A graduated assignment algorithm for graph matching,"*IEEE Trans. Pattern Analysis Mach. Intell.* **18(4),** pp. 377-388, 1996.
9. F.Depiero, M. Trivedi and S.Serbin, "Graph matching using a direct classification of node attendance," *Pattern Recognition,* **29(6),** pp. 1031-1048, 1996.
10. Yasser El-Sonbaty and M.A.Ismail, "A new algorithm for subgraph optimal isomorphism," *Pattern Recognition,* **31(2),** pp. 205-218, 1998.
11. D.Eppstein, "Subgraph isomorpism in planar graphs and related problems," Department of information and computer science, University of California, Irvine, CA 92717, Tech. Report 94-25, p. 27, 1994.
12. S.Yamada and T.Hasai, "An efficient algorithm for the linear assignment problem," *Elect. Comm. Japan,* Part 3, **73(12),** pp. 28-36, 1990.
13. О.Оре. Теория графов. Москва, 352 с, 1968.

ВИКОРИСТАННЯ "ВЛАСНИХ ОБЛИЧ" ДЛЯ РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ

Левко О. В., Нюнькін К. М.
Інститут проблем штучного інтелекту
83050, м.Донецьк, проспект Богдана Хмельницького, 84
Тел.: (0622)926082, факс: (0622)926082
E-mail: kmn@iai.donetsk.ua

ABSTRACT

In this paper usage of a method of image decomposition on eigenfaces for human face recognition is considered. The experiments permitting are conducted to define limitations of this approach.

ВСТУП

Комп'ютерне розпізнавання людських облич привертає в останній час все більше уваги. З одного боку це пов'язано з численними задачами, в яких може застосовуватися ця техніка, з іншого - навально зростаючі потужності сучасних комп'ютерів роблять можливим вирішення таких задач.

Всю безліч підходів до розпізнавання облич можна поділити на дві великі групи: методи, що використовують окремі деталі зображення, й алгоритми, що розглядають все зображення цілком. В роботі розглядається один з підходів другий групи - розклад зображення по "власним обличчам" або розклад Karhunen-Loeve [1.2], що вже досить давно використовується для стиску зображень. Поняття "власні обличчя" вводяться по аналогії з власними векторами і мають ту же властивість: їх можна вибрати як базис, по якому може бути розкладений довільний елемент простору облич.

1. ПРОСТІР ОБЛИЧ

Зображення в комп'ютері представляється матрицею окремих пікселів, яка має V рядків по H пікселів у кожному. У випадку сірого напівтонового зображення значення пікселів визначаються яскравістю відповідної точки зображення. Таке зображення можна уявити у вигляді вектору. Для цього необхідно з'єднати послідовно всі рядки зображення, отримавши один довгий рядок. Значення кожного пікселя зображення буде відповідати одній компоненті вектору. Отриманий вектор належить простору усіх можливих зображень, що мають розмір H*V пікселів. Цей вектор має вимірність H*V. Ясно, що працювати з векторами такої великої вимірності дуже складно.

При вирішенні задачі розпізнавання людських облич нас цікавлять не будь-які зображення, а тільки зображення облич. Всі обличчя схожі друг на друга: всі вони мають ніс, очі, губи й інші характерні деталі, розташовані на певних місцях. Тому зображення облич в просторі зображень компактно згруповані і утворюють порівняно невеликий кластер. Отже, повний простір зображень є надмірним для опису облич і можна побудувати спеціальний простір, що буде описувати обличчя оптимальним чином, маючи при цьому значно меншу вимірність, ніж вихідний простір зображень. Природно цей простір назвати простором облич [3].

Нехай I=H*V - кількість пікселів у зображенні та число компонентів вектору простору зображень; K - кількість зображень облич в навчальній вибірці; X - матриця вимірності K*I, що визначається навчальною вибіркою. k-ий стовбчик матриці відповідає k-му зображенню обличчя навчальної множини.

Надмірність простору зображень є наслідком того, що кожний піксель обличчя корелює з іншими пікселями і коваріаційна матриця для множини облич не є діагональною.

Для оптімального подання облич необхідно знайти такий простір, в якому коваріяційна матриця для навчальної множини облич буде діагональною. Ця умова буде виконуватися, якщо за базисні вектори простору взяти власні вектори. Базисні вектори простору облич називають основними компонентами. Введемо матрицю Y, що містить основні компоненти, k-ий стовбчик матриці - подання k-го зображення обличчя в просторі облич. Її коваріяційна матриця $\Sigma_Y=Y*Y^T$ буде діагональною.

Для отримання основних компонентів використовуються лінійні перетворення. Нехай P -

© UkrObraz'2000

матриця перетворення, тоді: $Y=P^T*X$, $X=P*Y$. В [3] запропоновано за матрицю P взяти власні вектори матриці Σ_X: $\Sigma_Y*P=P*\Lambda$, де Λ - діагональна матриця, що містить власні значення матриці Σ_X. Тоді $\Sigma_X=P^T*P*\Lambda=\Lambda$, тобто, Σ_Y - діагональна матриця, що містить власні вектори Σ_X, чого і вимагалося досягнути. Для знаходження власних векторів використовується метод обертань [4].

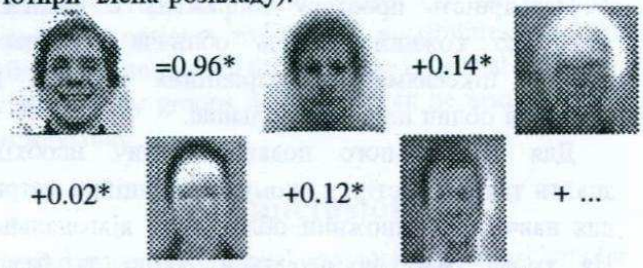
Величина власного значення визначає важливість або вагу відповідного йому власного вектору для подання простору. Власні значення можна відсортувати за убуванням, та відкинути вектори, що відповідають найменшим власним значенням, не вносячи цим істотної помилки в подальше розпізнавання [3]. При цьому вимірність простору ще зменшиться в декілька раз.

Кожний із знайдених власних векторів визначає точку в просторі зображень; тобто являє собою деяке зображення. Більше того, отримані зображення схожі на обличча. Тому їх аналогічно до векторів називають "власними обличчами" (ВО), підкреслюючи цим їхній зв'язок зі зображеннями облич. Приклад зображень перших ВО, упорядкованих за убуванням відповідних їм власних значень, наведений на малюнку 1.



Мал. 1. Приклад перших ВО

Зазначимо, що перше ВО є середнім всіх облич, поданих в навчальній множині; друге відповідає напрямку найбільших змін; третє - напрямку наступних по величині змін; і т. д. Зображення будь-якого обличча може бути уявлене в вигляді лінійної комбінації ВО. На малюнку 2 показаний приклад розкладу одного з облич (тільки перші чотири члена розкладу).



Мал. 2. Приклад розкладу зображення обличча по базису ВО

Після побудови базису ВО, з'являється можливість подання довільного зображення в знайденому базисі. Причому коефіцієнт розкладу при першому власному векторі може служити детектором наявності людського обличча на зображенні. Експериментально встановлено, що коефіцієнт розкладу при першому векторі для зображень, що містять обличча, складає в середньому 0.96; а для інших зображень - 0.83. Проведені дослідження показали можливість успішного застосування цього методу. Відносна кількість зображень, що були невірно ідентифіковані (обличчя / не обличчя), не перевищувала 2%.

Коефіцієнти розкладу являють собою вектор, вимірність якого дорівнює кількості ВО і звичайно складає декілька десятків. Маючи коефіцієнти розкладу зображення довільного обличча і знаючи набір ВО, по якому цей розклад виконувався, можна відновити вхідне зображення. На малюнку 3 наведені вхідне і відновлені зображення одного і того ж обличча при різній кількості коефіцієнтів розкладу. Точність відновлення тим більша, чим більше ВО використовується при розкладі. Тим не менше вже при 50 членах розкладу відновлене зображення мало відрізняється від оригіналу. Таким чином, замість зображень облич можна розглядати вектори в n-мерном просторі.



a) b) c) d) e)

Мал. 3. Відновлення зображення обличча людини по різній кількості ВО: a) вхідне зображення; b) - e) відновлене по 100, 50, 25 і 10 ВО відповідно

Для відновлення зображень на мал.3 ВО обчислювалися по всім 212 изображениям (приблизно по 20 зображень облич кожної з 11 людин) і їх було отримано також 212. Після цього отримані ВО ранжувалися в порядку убування відповідних їм власних значень, і при відновленні використовувалося тільки 100, 50, 25 або 10 перших з них.

Слід підкреслити, що для успішного застосування цього методу зображення облич повинні бути справді "схожі" друг на друга, тобто обличча повинні мати однаковий масштаб, ракурс і розташуватися в одному і тому ж місці на зображенні. Бажано, щоб обличчя займало все зображення. Це дозволить виключити вплив фону.

2. РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ

Подання зображень людських облич як векторів в просторі ВО дає можливість вирішувати задачу розпізнавання облич. Для цього спочатку за декількама десятками зображень довільних облич

формуються ВО. Після цього декілька зображень кожного з "відомих" людей розкладаються по цим ВО і вектори коефіцієнтів розкладу зберігаються в *базі даних як еталони для цих людей. Зазначимо*, що зображення, що розкладаються, могли і не використовуватися при формуванні ВО. База даних еталонних векторів в будь-який момент може бути доповнена новими обличчями. Набір ВО при цьому залишається незмінним. Кожне нове введене зображення обличча людини також розкладається по тим же самим ВО, що й еталонні зображення.

Еталонні вектори, що отримані розкладом по ВО зображень облич кожної людини, утворюють класи. Тепер можна ставити і вирішувати задачу приналежності вектору, отриманого розкладом по ВО зображення довільного обличча, одному з класів.

2.1. БАЗИ ДАНИХ

В експериментах по розпізнаванню облич використовувались зображення з декількох баз даних. Оскільки розміри облич на зображеннях різних баз даних істотно розрізняються, експерименти з зображеннями кожної бази даних проводилися незалежно від інших. У базах були представлені обличча людей різної статі та расової приналежності, з вусами, бородами, окулярами та без них. Варіювався також вираз обличча. Нижче наведені більш детальні параметри кожної з баз даних.

Манчестерська база даних: 214 зображень розміром 64x64 (20 зображень облич кожного з 9 чоловіків і 3 жінок). Мають місце слабкі повороти голови, трохи розрізняється масштаб, на окремих кадрах наявні окуляри.

База даних дослідної лабораторії Оліветті: 400 зображень розміром 92x112 (по 10 зображень облич кожного з 35 чоловіків і 5 жінок). На зображеннях однієї й тієї же людини розрізняються вираз обличча та ракурс. Масштаби практично однакові.

База даних Йельського університету: 165 зображень розміром 320x243 (по 11 зображень облич кожного з 14 чоловіків та 1 жінки). На зображеннях однієї й тієї ж людини розрізняються вираз обличча, умови освітлення і наявність окулярів. Масштаб та ракурс практично однакові.

База даних університету Берну: 150 зображень розміром 512x342 (по 10 зображень облич кожного з 15 чоловіків). На зображеннях однієї й тієї ж людини розрізняються вираз обличча і ракурс. Масштаб практично однаковий.

База даних MIT: 450 зображень розміром 128x120 (по 27 зображень облич кожного з 16 чоловіків). На зображеннях однієї й тієї ж людини розрізняються вираз обличча, ракурс, масштаб та умови освітлення. Тільки серед зображень цієї бази є бічніі нахили голови.

2.2. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Перед виконанням розпізнавання зображення всіх баз даних були заздалегідь оброблені: фон був обрізаний і залишено тільки обличчя. Для порівняння манчестерська база даних, на зображеннях якої фон слабкий, була оброблена також і без попереднього виділення облич.

Експерименти проводилися слідуючим чином. З зображень бази даних формувалася навчальна вибірка, розмір якої варіювався від всієї бази до одного зображення на людину. По зображенням навчальної вибірки обчислювалися ВО. Після цього зображення кожного з наведених в базі людей ділилися на дві рівні частини. По першій половині обчислювалися еталонні вектори коефіцієнтів розкладу, а зображення другої розглядалися як невідомі і підлягали розпізнаванню. Після цього половини мінялися місцями. Еталонні вектори коефіцієнтів розкладу, стосовні до однієї людини, осереднялися для отримання середнього вектору.

Кожне нове зображення обличча невідомої людини розкладалося по тим же самим ВО, що й еталонні зображення. Вважалася, що обличчя на введеному зображенні належить тій людині, чий середній еталонний вектор виявлявся найближчим.

Частки невірно розпізнаних облич при різних розмірах навчальної вибірки і кількості ВО наведені в таблиці 1.

Результати експериментів показують, що даний метод розпізнавання людських облич надто чутливий до фону на зображенні. В першій серії експериментів зображення манчестерской бази даних були оброблені разом з фоном (перший рядок в табл.1). Частка помилок при розпізнаванні в цьому випадку була дуже висока і досягала половини. Це, звичайно, менше, ніж імовірність помилок при випадковому виборі, але значно більше, ніж можна допустити в реально працюючій системі. Усунення фону дозволило в 4-5 раз знизити частку помилок (третій рядок в табл.1).

Ще один шлях зниження помилок - використання більш досконалої процедури розпізнавання. Так пошук найближчого з всіх еталонних векторів (а не середнього) дозволяє зменшити імовірність помилки в півтора рази (другий рядок в табл. 1). Це означає, що кластери, утворені еталонними векторами окремих людей, не мають сферичної форми і погано рознесені в просторі.

Найкращі результати були отримані для зображень бази даних Йельского університету, незважючи на те, що зображення обличча однієї й тієї ж людини розрізнялися виразом, умовами освітлення та наявністю окулярів. Отже ці умови не такі важливі як масштаб обличча і ракурс голови, які на цих зображеннях практичні однакові.

Більш за все помилок було зареєстровано при розпізнаванні облич на зображеннях бази даних MIT, серед яких були наявні бічні нахили голови та істотна зміна розміру обличча. Справді, у випадку бічного нахилу голови зображення обличча сильно викривляється і мало схоже на зображення, де голова не нахилена. Зображення, на яких обличча мають різні розміри також значно розрізняються.

При обробці зображень з баз даних дослідної лабораторії Оливетти й університету Берну, що містять зображення облич при поворотах і вертикальних нахилах, отриман невеликий відсоток помилок, отже такі зміни ракурса припустимі для систем розпізнавання, що основани на поданні людських облич через власні обличча.

Табл.1 Результати експериментів

N	База Даних	Кількість зображень в навчальній вибірці	Кількість ВО	Розмір зображення	Еталони/Невідомі		
					Всі/всі	1/2	2/1
1	Manchester*	212	64	64x64	4.7%	48.1%	50.0%
		66	64		12.3%	49.0%	40.7%
		33	32		13.2%	51.0%	46.3%
		11	11		20.8%	56.7%	50.0%
2	Manchester**	212	64	64x64	0%	35.6%	35.2%
3	Manchester	212	32	32x48	3.3%	15.4%	10.2%
		33	32		3.3%	17.3%	9.3%
		11	11		9.4%	26.9%	13.9%
4	ORL	400	32	92x112	4.0%	21.0%	16.0%
		40			5.5%	18.0%	20.0%
5	Yale	135	32	112x128	2.22%	5.0%	12.0%
6	Bern	150	32	96x112	6.7%	16.0%	13.3%
7	MIT	430	32	64x64	29.1%	34.4%	44.7%

* Манчестерська база даних оброблена без виділення облич.

** При розпізнаванні використовувався найближчий еталонний вектор.

ВИСНОВКИ

Таким чином, розклад зображень людських облич на ВО може бути використаний при побудові системи розпізнавання людей за їхніми обличчами.

Методи, що основані на ВО дозволяють распізнавати людей за їхніми обличчами незалежно від вираза обличча та умов освітлення.

Цей метод має високу чутливість до навколишнього фону, розміру обличча на зображенні, а також до бічних нахилів голови. При створенні реально працюючої системи, що розпізнає, необхідно передбачити автоматичне визначення положення, розміру і нахилу обличча на зображенні з тим, щоб компенсувати викривлення, які вносяться, і усувати з розгляду навколишній фон.

Слід також приділити особливу увагу застосуванню ефективних засобів вирішення задачі розпізнавання образів, що здатні істотно підвищити якість розпізнавання.

ЛІТЕРАТУРА

1. Turk M. A., Pentland A. P. Face Recognition Using Eigenfaces.//Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1991.-p. 586-591.
2. Pentland A., Moghaddam B., Starner T., Turk M. View-based and Modular Eigenfaces for Face Recognition.//Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1994.
3. Romdhani S. Face Recognition Using Principal Component Analisis./The Msc Thesis, 1996.-57p.
4. Березин И.С., Жидков Н.П. Методы вычислений. - М., Наука, 1965.

CLASSIFICATION OF TEXTURE IMAGES WITH THE USAGE OF ARTIFICIAL NEURAL NETS AND HIGHER-ORDER STATISTICS

A.M. Akhmetshin, D.A. Mikhalev

Dniepropetrovsk state university

Abstract

The object of the research is photos of metal-graphic microstructures.

The aim of the work is the choice of algorithmic and elaboration of software for effective classification of texture images with the help of neural nets, and also classification the process of exactness rise researches of questions and diminution of general processing time of texture images.

The method of the research is comparative analysis of two classifiers, the first one is built on Hamming neural net, the second one is built on Kohonen neural net.

Results of the researches were employ for classification of the photos of metal-graphic microstructures, which were obtain through digital photoscanner.

1. Introduction

The concept of the texture is more often used in different areas of a science and engineering, such as hylology and research of mining and metallurgical processes, the control of the quality of industrial production, image processing of the metal-structures. It follows from the fact, that the surfaces of a great number of real objects in these fields are textures and under condition of their homogeneity, they can be described with the help of a comparatively little number of parameters. By virtue of this, fact, the solution to this problems of the texture classification is actual, in particular, in geology and geodesy, in a snap analysis of micro- and of macrostructures in metallurgy.

The general scheme of texture classification is shown in the fig. 1.

As well as in the system of pattern recognition, two main tasks can be determined here: the first one is feature extraction and the second one is classification of textures to the evaluated features.

Two different approach to extraction of texture features is considered in the work. The first one is connected with higher-order statistics and is based on calculation of third-order cumulant. The second one is based on the statistics of the second order and consists in calculation 10 texture features from spatial gray-level dependence (SGLD) matrices.

Structurally the classifier is represented by two principally different schemes of artificial neural net. The first classifier, which is used during calculation of third-order cumulant, consists in application of Hamming neural net, the second one, which is used during calculation of statistics of second order consists in application Kohonen's self-organizing maps.

There are 5 test images of microstructures, each of which represents the separate class. The size of each image for a simplicity of a illustration is selected 128×128. All these images are shown on fig. 2.

One of the main problems during the classification of textures consists in uncertainty of texture images orientation. In this connection each texture, presented on the fig. 2, in addition can be rotated on 90^0, 180^0 and 270^0. Thus, 20 images from 5 classes are presented.

It is required to interprete the input image correctly and refer it to one of 5 classes.

2. Feature extraction

2.1. Higher-order statistics

For the selection of informative features of textures the apparatus of higher-order statistics can be used, and as features the estimation of 3-rd order cumulants can be used [6].

Cumulant is one of the numerical characteristics of random variables, it is closed to concept of the moment of the high order. In practice, for the finite sampling of data $\{x(n)\}_{n=0}^{N-1}$ the cumulantes are calculated according the following formulas:
second-order cumulant:

$$\hat{C}_{xy}(k)=\frac{1}{N_3}\sum_{n=N_1}^{N_2}x^*(n)y(n+k), \qquad (2.1)$$

third-order cumulant:

$$\hat{C}_{xyz}(k,l)=\frac{1}{N_3}\sum_{n=N_1}^{N_2}x^*(n)y(n+k)z(n+l), \qquad (2.2)$$

where $N_1, N_2 \in [0, N]$. N_3 are usually set equal to N and calculated their estimations are asymptotically unbiased.

The algorithm for application of the higher-order statistics apparatus (algorithm 1) for obtaining the matrix of informative features of textural images consists of the following steps [8]:

1. Texture image I(x, y) is transformed in one-dimensional array y(k) by joining of successive rows in a sequence.

© UkrObraz'2000

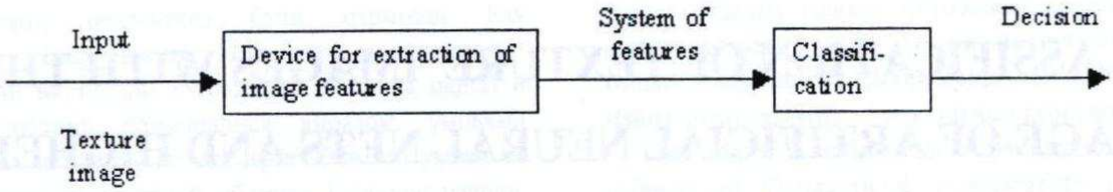


Fig. 1. The generalized scheme of texture classification and recognition.

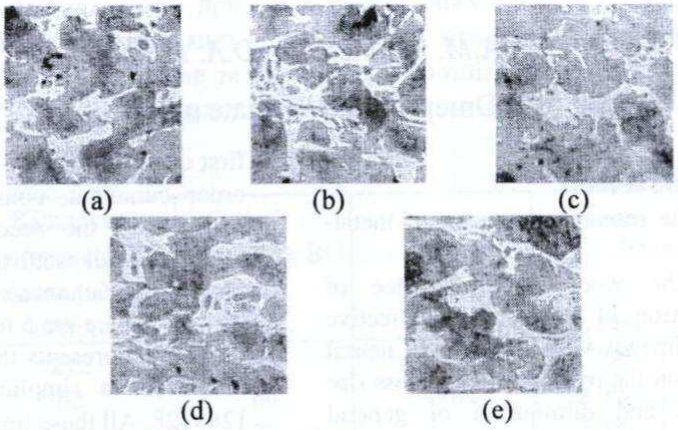


Fig. 2 Textures of different classes

2. The dimension of matrix of estimated cumulants is 2*n+1.
3. Time-series y(k) is segmented into records of m samples each, with no overlap; the biased estimates of the third-order cumulants are obtained from each segment on the formula (2) and then averaged.
4. As result the matrix C(x, y) is calculated where (i, j) element of matrix is the estimate of C_{3y}(i-n-1, j-n-1) for i, j = 1, ..., 2*n+1.

2.2. Spatial gray-level dependence (SGLD) matrices

The essence of one of most widely used methods of description of texture consists in calculation of texture features from spatial gray-level dependence (SGLD) matrices.

The *(i, j)*th element of the SGLD matrix is the joint probability that gray levels *i* and *j* occur in direction at a distance of θ pixels apart in an image. Ten texture features, including correlation, entropy, angular second moment, mean, difference entropy, contrast, inverse difference moment, deviation, different angular second moment, different mean were used. The definition of the texture measures are given in the literature [1, 2, 4].

These measures were extracted from each SGLD matrix at pair distance d = 5 and four directions (0°, 90°, 180°, and 270°). These features contain information about image characteristics such as homogenity, contrst, and the complexity of the image.

3. Classification

3.1. Hamming neural network

Among the different configurations of artificial neural networks there are the ones, in which classification according the principle of training, strictly speaking, doesn't approach either training with the teacher, or training without the teacher [4]. In such networks the weight coefficients of synapses set up only once before the beginning of work of the network on the basis of the information about processed data, and all training of the network is amounted to this calculation. The Hamming network falls into the class of networks with similar logic of work which is usually used for organization of associative memory. This network is used in case of the binary input images.

Neural net solution of this task can be obtained on the basis of the Hamming architecture. The network has one layer of identical neurons, number of which is equal to quantity of classes. Each neuron is coupled to each the input, which number is equal to the dimension of considered library of images.

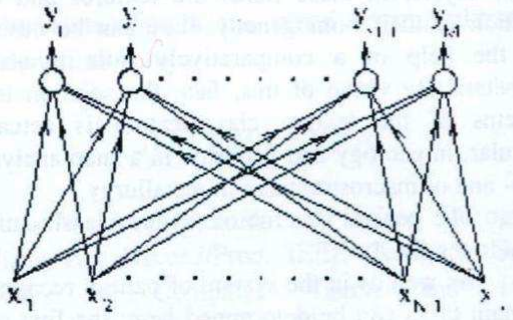


Fig. 3. Hamming neural network

The network choose the sample with minimum of Hamming distance to unknown input signal, therefore as a result one output of the networks appropriating to this sample will be activated.

It is reached by introduction of additional feedbacks between neurons arranged by a principle "lateral braking". Such mechanism is called "Winner Take All" (WTA) [4, 7].

3.2. Self-organizing maps

Self-organizing map (SOM) is neural net with straight ties, in which the teaching algorithm without teacher is used. By means of process, which named is self-organize, from outgoing elements SOM will make the topologic conception of initial data.

Algorithm SOM is based on competitive teaching without teacher. It provides reflection, preserving the topology, from space of big dimension on elements of map. Elements of the map, or neurons, usually make a two-dimensional grid. Thus this reflection is the reflection from space of big dimension on flatness. Property of topology preservation of signifies, that SOM distributes the similar vectors of input data on neurons: points, which is near to each other in space of inputs, on near situated elements of map in SOM.

Kohonen's net consist of r neurons, according to number of classes: to each class corresponds its neuron, and inside out. Input vector X is given on net input, and on input of each neuron. Its dimension n depends on task. Each neuron calculates function f from input vector I, and as answer of net the number of that neuron is given out (this vector get in class with this number), significance of function f in which receive maximum (or minimum, as more convenient). So the principle "winner take all" (WTA) is used.

4. Realization

4.1. Classifier, built on the base of the third-order cumulant

At the first stage of the calculation of texture of informative features the algorithm 1 was used. The size of segments, into which initial array is divided, is selected equal to m = 128. The size of the matrix of 3-rd order cumulant C is equated 2*n+1 = 129, i.e. n = 64. The example of such matrix for the class, presented on fig. 2.a, is shown on fig. 4.

As it has been already pointed, that the input signals for Hamming network are binary. Therefore during the transition to the stage of the classification of initial images it is necessary to make them binary. The threshold for binarization was set equal to 0. Thus, if C (i, j)≥0, then that B (i, j) = 1, if C (i, j) < 0 then B (i, j) = 0, where i, j = 1, ..., 2*n+1, B is binary matrix of informative features of texture.

The example of the matrix B for the class, presented in the fig. 2.a, is shown in the fig. 5.

The fig. 5 demonstrates, that binary feature matrixes for different orientations of textures differ considerably. It is connected, first of all, to the following: for obtaining of the rating matrix of cumulants the initial input was transformed in to the one-dimensional array, and as the image rotates the origin displaces it causes the change of.

It was empirically proved, that the best vector-samples for the Hamming network would be 3-d order cumulants averaged on all four rotations.

Thus, algorithm of the work of the system of texture images recognition (algorithm 2) consists of following steps:

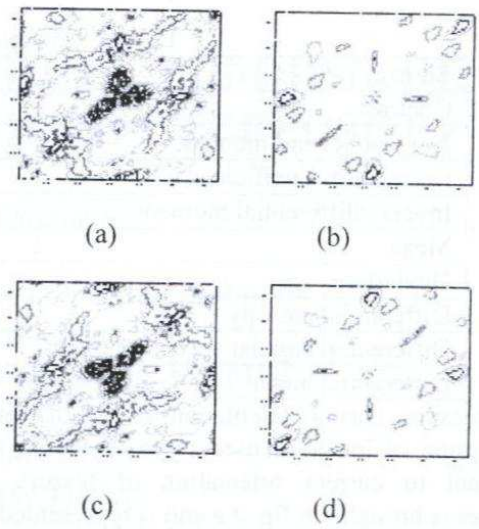


Fig. 4. Third-order cumulants for the textures of the class presented on fig. 2.a in level lines images:
(a) orientation 0^0; (b) orientation 90^0;
(c) orientation 180^0; (d) orientation 270^0;

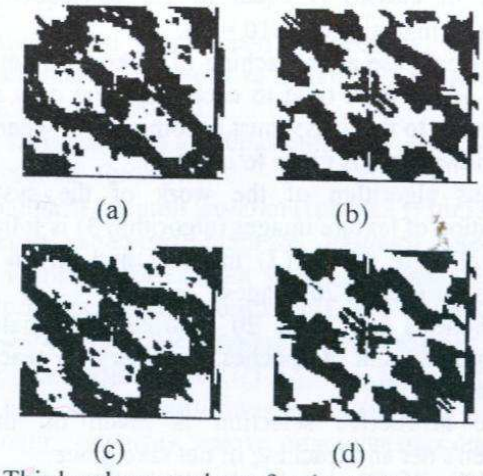


Fig. 5. Third-order cumulants for the textures of the class presented on fig. 2.a in level lines images after binary procedure:
(a) orientation 0^0; (b) orientation 90^0;
(c) orientation 180^0; (d) orientation 270^0;

1. For each of 20 image the cumulants of 3-rd order are calculated.
2. For each of 5 classes averaged 3-rd order cumulants of all four images of the class is calculated, and then it is binaried with threshold 0.
3. All 5 binary matrixes are processing through the Hamming network as vector-samples.
4. On input of the system of recognition unknown test image (one of the described 20 images) is given, and the system should refer this image to the initial class with minimum error of the classification.

4.2. Classifier, built on the base of SGLD features

In elementary phase of practical realization of given classifier is necessary to count the informing features, which in given case are 10 texture features from SGLD matrices.

TABL. 4.1

Texture features from SGLD matrices, shown on fig. 2.a

Entropy	930.2033
Contrast	3.2671e+06
Angular second moment	2.9109e+03
Correlation Coefficient	8.0294e-015
Inverse differential moment	95.8760
Mean	2.4399e+03
Deviation	1.1409e+07
Differential entropy	-1.5082e+04
Differential angular second moment	2.5254e+05
Differential mean	8.9553e+04

Because during calculation of SGLD brightness description of image is used, type of SGLD matrix is invariant to current orientation of texture. Texture features is brought on fig. 2.a and is represented on table 4.1.

For correcting classification the volume of topographic Kohonen's map exemplarily must excel amount of possible classes in twice. In our case the number of classes is equal to 10, consequently, the Kohonen's map size of $2 \times 10 = 20$.

In ideal case after teaching of Kohonen's net only 1 neuron of net must treat to each class and only neuron, conforming to this class must become active when image of unknown class is given to input.

Thus algorithm of the work of the system of recognition of texture images (algorithm 3) is following:

1. Is calculate the SGLD matrices and on 10 texture features for each of 20 images.
2. Kohonen's net from 20 neurons is initialed and maximum amount of epoches, necessary for teaching of net is set.
3. The instructive selection is given on input of Kohonen's net and teaching of net take place.
4. Unknown testing image is given on input of recognition system (one from 20 described above), this testing images must be refered by the system to one or another initial class with minimum mistake of classification by dint of calculation of statistic features of second order and further usage of Kohonen's net.

5. CONCLUSION

The accuracy classification of both classifiers is equal to 95-100 %. The operating time of the system was 10 minutes for evaluation of one matrix of features and 2-3 seconds for the classification with help of PC: P5-366 MHz/64M. Such a high speed of work of the classifier is reached by the simplicity of the principle of work of the Hamming network. For increase of processing speed of work of the extractor of features (see fig. 1) the size of the matrix of features decreased down to 64×64, however the accuracy of cumulants classification decreased (down to 70 %) as the speed of calculation increases.

Advantages of the method are the high accuracy of the classification and little operating time of the classifier for the class of the PC.

Disadvantages are insufficient speed of work of the system during feature extraction, and also restriction of the number of outputs of the classifier which are equal to number of classes. In this situation, when on input of the system is exited with texture image of metal-structure from the 6-th class, the system will still refer it to one of the initial 5.

Thanks to its work principle, the problem of "6-th class" in Kohonen's classifier will be absent. Besides this device of texture classification, built on base of Kohonen's net, works quicker than Hamming's classifier. First of all this reaches at the expense increase of speed of work feature's extractor. In addition to that invariance of SGLD matrix to orientation of texture allows to count up the informing features for each class only onces, where as matrix of estimations of cumulants must be counted 4 times (for each orientation).

The main disadvantage of this algorithm is little number of informing features. The dimension of Kohonen's topographic map must be increased for correcting classification attached to augmentation of amount of classes, but that is gives considerable classifier work time augmentation.

Further researches are directed to removal of disadvantages which were brought above, that can be reached by use, for example, adaptive resonance theory network (ART2).

REFERENCES

1. Andreev G.A., Bazarskiy O.V. "Analysis and synthesis of random spatial textures" // Foreign radioelectronics. – 1984. – №2. – P. 3 - 34
2. Dhawan A., Chitre Y., "Analysis of mammographic microcalcifications using gray-level image structure features" // IEEE Trans. Med. Imag. – 1996. – №15.
3. Wosserman F. "Neurocomputing". – Moscow: World, 1992.
4. Haralick R.M., Shanmugam K., and Dinstein I. "Textural features for image classification" // IEEE Trans. Syst., Man, Cybern. – 1973. – SMC-3. – P. 610-621.
5. Kosko B. "Neural networks for signal processing". – Prentice Hall, Englewood Cliffs, NJ, 1992.
6. Mendel J.M., Nikias C.L. "Signal processing with higher-order spectra" // IEEE Signal Processing Magazine. – 1993. – №10.
7. Terehov S.A., "Lecture under the theory and applications of artificial neural networks". – Snezhinsk: VNIIT, 1994.
8. Nikias C.L., Raguver M.L. "Bispectral estimation employing to digital signals processing" // IEEE Signal Processing Magazine. – 1987. – №7.

ОПТИМІЗАЦІЯ КУСКОВО-ЛІНІЙНОЇ ДЕФОРМАЦІЇ ЗНІМКІВ У ЗАДАЧАХ МАШИННОГО СТЕРЕОЗОРУ

М.І.Шлезінгер[1], Б. Флах[2], І.В.Онищенко[3]

[2]Institute of Artificial Intelligence, Dresden University of Technology
Email: bflach@ics.inf.tu-dresden.de
[1,3]Міжнародний науково-навчальний центр ЮНЕСКО/МПІ
інформаційних технологій та систем.
пр.Академіка Глушкова 40, Київ 03680
т.266-62-08 Email: schles@image.kiev.ua

АНОТАЦІЯ

Розглянуто метод знаходження оптимального монотонного кусково-лінійного відображення відрізка на відрізок. При цьому враховується як ступінь відповідності точок відрізків, так і довжина ламаної. Наведено результати застосування описаного методу при відновленні рельєфу людського обличчя за стереопарою.

ВСТУП

Дана стаття відноситься до класу робіт з відновлення рельєфу поверхні за стереопарою, що грунтуються на використанні паралаксів. Цей напрям було започатковано Г.Л.Гімельфарбом. У його роботі [1] побудовано модель оптичної поверхні, що відтворюється за своїми зображеннями, та модель процесу отримання цих зображень; обгрунтовано зведення задачі стереозору до оптимізаційних задач, зокрема, до задачі динамічного програмування. У роботі [2] запропоновано деякі евристики для звуження класу шуканих поверхонь.

Розв'язування задачі стереозору полягає у знаходженні монотонного відображення скінченної дискретної впорядкованої множини на себе. Єдиним таким відображенням є тотожне відображення. При переході від дискретних множин до неперервних матимемо безліч допустимих відображень, навіть якщо шукатимемо тільки ті з них, що є неперервними та кусково-лінійними. Якщо додатково вимагати, щоб вершини ламаної мали лише цілі координати, то отримаємо скінченну кількість допустимих ламаних.

1 ПОСТАНОВКА ЗАДАЧІ

Для заданого натурального числа n *полем зору* будемо називати відрізок T=[0; n] дійсної осі. Відрізок p_i вигляду [i-1; i] $\subset$ T, i=1,..,n назвемо i-м *пікселом*. *Знімком* будемо називати кусково-постійну функцію r : T → C, де C – деяка скінченна множина *кольорів*. r приймає однакові значення в усіх точках піксела.

Вихідні дані для задачі стереозору задаються двома знімками : лівим $r_Л$ та правим $r_П$, визначеними над одним полем зору T.

Ми шукаємо таке відображення f : T → T, що :
1) f – кусково-лінійне;
2) координати вершин ламаної (графіка f) цілі числа;
3) довжина кожного відрізка ламаної обмежена зверху і знизу; обмеження знизу дозволяє керувати "гладкістю" рельєфа, а обмеження зверху прискорює роботу алгоритму, зменшуючи перебір варіантів.

Відрізок ламаної $e(i_1,j_1,i_2,j_2)$, $i_1,j_1,i_2,j_2 \in \{0,1..n\}$ задається координатами початкової (i_1,j_1) та кінцевої (i_2,j_2) точок, причому мають виконуватися наступні обмеження : значення i_1-j_1 та i_2-j_2 мають належати деякій множині H ; обмеження на довжини відрізків ламаної $dimin \le i_2-i_1 \le dimax$, $djmin \le j_2-j_1 \le djmax$. Через E позначимо множину всіх допустимих відрізків. Позначимо через S множину всіх ламаних s таких, що :
1) кожен відрізок ламаної належить E;
2) ламана починається у точці вигляду (0,k) або (k,0) і закінчується у точці вигляду (n,n-k) або (n-k,n), де $0 \le k \le kmax$.

Під *довжиною* l(e) відрізка e розумітимемо його геометричну довжину. Нехай $\alpha(j,k) : T \times T \to R$ функція, яка кількісно характеризує несхожість кольорів пікселів j лівого та k правого знімків. *Якістю* q(e) відрізка $e(i_1,j_1,i_2,j_2)$ назвемо число де [x] означає цілу частину числа x, $di=i_2-i_1$, $dj=j_2-j_1$,

$$\frac{1}{di \cdot dj} \sum_{n=0}^{di \cdot dj - 1} \alpha\left(\left[i_1 + n/di\right]\left[j_1 + n/dj\right]\right) + \beta \cdot l(e), \quad (1)$$

β - деяка константа. Якість ламаної дорівнює сумі якостей її відрізків. Ми шукаємо таку ламану $s^* \in S$, що

$$s^* = \arg\min_{s \in S} \sum_{e \in s} q(e), \quad (2)$$

де сума береться по всіх дужках ламаної s.

© UkrObraz'2000

2 РОЗВ'ЯЗУВАННЯ ЗАДАЧІ

Позначимо через F(i,j) якість найкращої ламаної з кінцем в точці (i,j) і початком у точці вигляду (0,k) або (k,0), де 0≤k≤kmax. Покладемо F(0,k)=F(k,0)=0 для 0≤k≤kmax. Розв'язком задачі буде ламана з кінцем у точці вигляду (n,n-k) або (n-k,n), де 0≤k≤kmax, що має мінімальне значення F. При розв'язуванні задачі використовується рекурентне співвідношення

$$F(i_2, j_2)= \min_{e(i_1,j_1,i_2,j_2)\in E} (F(i_1, j_1)+q(e)). \quad (3)$$

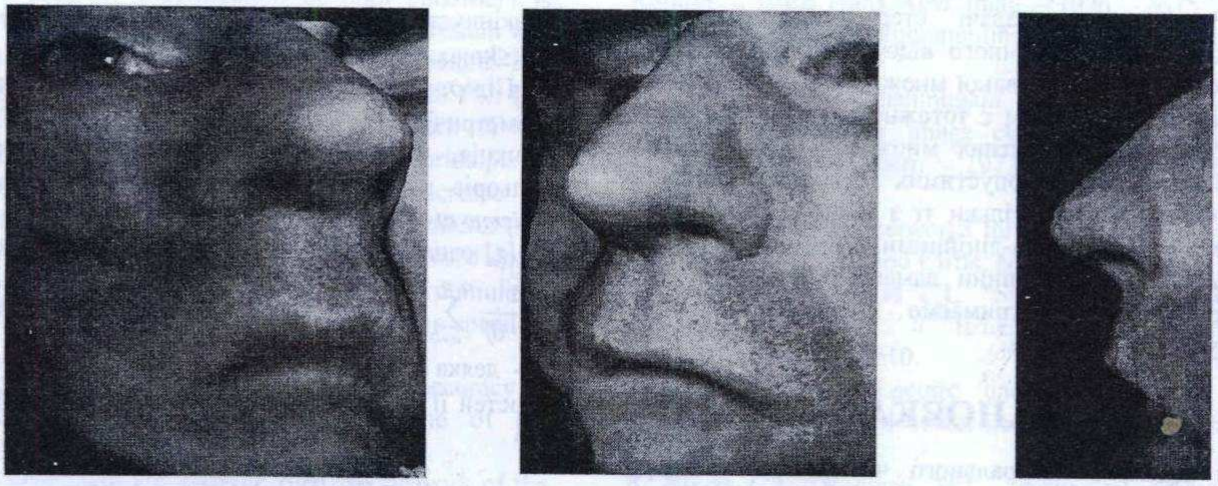
3 ОТРИМАНІ РЕЗУЛЬТАТИ

Описаний алгоритм було реалізовано мовою С. На мал.1 показано вихідну стереопару, а на мал.2 отримані результати для таких значень параметрів : H = {-30,-29,..,29,30}, n=300, dimin=djmin=3, β=0.05, kmax=20, dimax=djmax=5; колір точки задається трійкою (r,g,b), а функція α така, що

$$\alpha^2((r_1,g_1,b_1),(r_2,g_2,b_2)) = (r_1-r_2)^2+(g_1-g_2)^2+(b_1-b_2)^2$$



Мал.1 Вихідна стереопара



Мал.2 Відновлений рельєф обличчя

ВИСНОВКИ

Розглянуто метод оптимізації кусково-лінійної функції однієї змінної, який дозволяє знаходити відповідність між точками пари рядків знімків стереопари. Але при цьому не враховується залежність між сусідніми рядками. Врахування такої залежності приводить до задачі оптимізації кусково-лінійної функції двох змінних.

ЛІТЕРАТУРА

[1] Гимельфарб Г.Л. "Симметричный подход к задаче автоматических стереоскопических измерений в фотограмметрии" // *Кибернетика.* —1979.— №2.— С.73-82.

[2] Гимельфарб Г.Л. "Симметризованное би- и тринокулярное стереозрение: взаимосвязи между теоретическими основами и эвристическими решениями" // *Теоретические и прикладные вопросы распознавания изображений.*— Киев, 1995.— С.4-25.

КАЛІБРОВКА ЗОБРАЖЕНЬ ДЛЯ ПОРЯДКОВИХ МЕТОДІВ РОЗПІЗНАВАННЯ РЕЛЬЄФУ

Д.І. Рябоконь
Міжнародний Науково-Навчальний Центр Інформаційних Технологій та Систем
Проспект Академіка Глушкова 40, Київ-03680. Тел 266-62-08

АНОТАЦІЯ

В роботі розглядаються методи перетворення стереопари зображень до такого вигляду, щоб пари рядків розгортки могли бути використані як вихідні дані для порядкових алгоритмів розпізнавання рельєфу об'єкта. Описується комплекс, що дозволяє створювати ідеальні зображення для стереоскопа та розпізнавати рельєф на цих зображеннях.

ВСТУП

Досить легко створити стереопару для побутових потреб. Для цього треба лише фотографувати об'єкт із двох різних позицій і поглянути на пару знімків через стереоскоп. Так можна визначити, які точки знаходяться ближче, а які далі. Для машинної обробки стереопара повинна бути виконана значно ретельніше. Підвищені вимоги до стереопари зумовлені тим, що більшість методів розпізнавання рельєфу [1, 2] оперують із парами рядків розгортки на зображеннях незалежно. Весь рельєф потім складається зі шматочків, що були відновлені на кожному рядку.

Ідеальною стереопарою будемо називати два таких зображення, що для кожної точки на рядку j першого зображення відповідна точка на другому зображенні знаходиться також у рядку j. Саме для порядкових методів розпізнавання рельєфу і потрібне перетворення двох зображень до ідеальної стереопари (калібровка зображень).

ОБРОБКА НЕКАЛІБРОВАНИХ ЗОБРАЖЕНЬ

1. Епіполярні лінії

Розглянемо випадок, коли при обробці рядку j лівого та правого зображення взагалі неможливо знайти правильну відповідність між точками, бо, скажімо, на одному із зображень використовується інший масштаб. Для застосування порядкових алгоритмів необхідно спочатку знайти сімейства ліній на обох зображеннях, котрі будуть використані в процесі розпізнавання.



Рис. 1. Епіполярні лінії

Кожній точці на лівому зображенні ставиться у відповідність лінія на правому зображенні та навпаки. Епіполярна лінія на лівому зображенні – це проекція променя, що прямує від точки поверхні об'єкта до фокуса правої камери на проекційну площину лівої камери (Рис. 1). Точки на зображенні – це перетин променів із проекційною площиною камери. Оскільки промені, що утворюють точки зображення, проходять через фокус камери, всі епіполярні лінії на кожному із зображень перетинаються в одній точці (ця точка може знаходитись і на безмежності).

Епіполярні лінії на парі зображень зручно подати в однорідних координатах, де кожний елемент зображення індексується трьохвимірним вектором $U=(i,j,1)^T$. Сімейства епіполярних ліній можна тоді записати у вигляді

$$(U_L)^T \; F \; U_R=0,$$

де F – матриця 3x3, що має назву *фундаментальна матриця*.

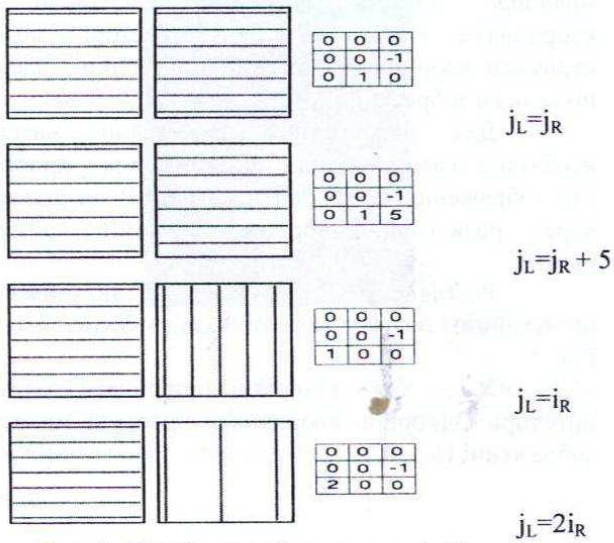


Рис. 2. Сімейства епіполярних ліній

© UkrObraz'2000

Оскільки всі епіполярні лінії перетинаються в одній точці, ранг матриці F повинен дорівнювати двом.

Якщо знайти епіполярні лінії на зображеннях, перетворити їх так, щоб матриця F мала вигляд {(0,0,0), (0,0,-1),(0,1,0)}, то можна буде застосувати порядкові методи розпізнавання рельєфу, обробляючи на стереозображеннях рядки розгортки з однаковими індексами.

2. 8-точковий алгоритм

Якщо на зображеннях можна вказати відповідність між якимись точками, то фундаментальна матриця знаходиться за допомогою *8-точкового алгоритму* [3]. Оскільки для кожної пари відповідних точок виконується

$$(U_L)^T \; F \; U_R = 0 \,,$$

то можна записати наступну рівність.

$$\begin{aligned} & i_L i_R F_{11} + i_L j_R F_{12} + i_L F_{13} + \\ + \; & j_L i_R F_{21} + j_L j_R F_{22} + j_L F_{23} + \\ + \; & i_R F_{31} + j_R F_{32} + F_{33} = 0, \\ & \text{rank}\,(F) = 2. \end{aligned} \qquad (1)$$

Якщо маємо сукупність пар відповідних точок, то можна скласти однорідну систему рівнянь для коефіцієнтів матриці F. Для знаходження всіх коефіцієнтів матриці F необхідно 8 рівнянь, оскільки матриця знаходиться з точністю до постійного множника.

3. Проекційні матриці

Нехай $X=(x,y,z,1)^T$ однорідні координати точки у просторі, $U=(i,j,1)^T$ однорідні координати проекції точки на зображенні. Тоді X та U пов'язані відношенням

$$(i,j,1)^T = P(x,y,z,1)^T / (xP_{31} + yP_{32} + zP_{33} + P_{34}), \qquad (2)$$

де P – матриця 3x4, що має назву *проекційна матриця.* Знаючи проекційну матрицю та координати якоїсь точки в просторі, можна отримати координати проекції цієї точки (індекси пікселя на зображенні).

Для знаходження проекційної матриці необхідно знати координати шести точок у просторі і на зображенні. Коефіцієнти матриці P знаходяться через розв'язок однорідної системи лінійних рівнянь.

Розглянемо приклад знаходження проекційних матриць для об'єкта, зображеного на Рис. 3.

X_0 ... X_8 – однорідні координати точок у просторі. Однорідні координати проекцій точок на зображенні U_0 ... U_8.

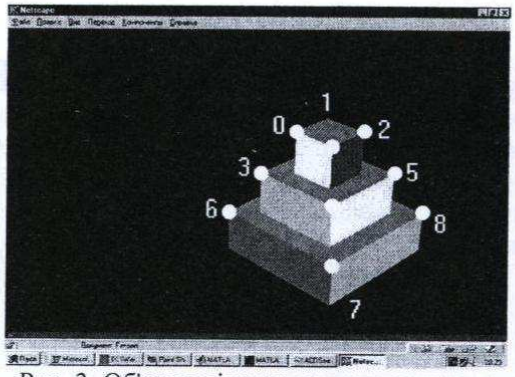


Рис. 3. Об'єкт з відомими просторовими координатами

U_0 =(461, 207, 1) X_0=(-1, 1, 6, 1),
U_1 =(520, 233, 1) X_1=(-1, -1, 6, 1),
U_2 =(569, 209, 1) X_2=(1, -1, 6, 1),
U_3 =(406, 276, 1) X_3=(-2, 2, 4, 1),
U_4 =(521, 331, 1) X_4=(-2, -2, 4, 1),
U_5 =(618, 276, 1) X_5=(2, -2, 4, 1),
U_6 =(355, 340, 1) X_6=(-3, 3, 2, 1),
U_7 =(521, 430, 1) X_7=(-3, -3, 2, 1),
U_8 =(664, 342, 1) X_8=(3, -3, 2, 1).

Для точки 0 рівняння (2) дає наступні два рівняння:

$$461(-P_{31} + P_{32} + 6P_{33} + P_{34}) = (-P_{11} + P_{12} + 6P_{13} + P_{14}),$$
$$207(-P_{31} + P_{32} + 6P_{33} + P_{34}) = (-P_{21} + P_{22} + 6P_{23} + P_{24}).$$

Для точки 1:

$$520(-P_{31} - P_{32} + 6P_{33} + P_{34}) = (-P_{11} - P_{12} + 6P_{13} + P_{14}),$$
$$233(-P_{31} - P_{32} + 6P_{33} + P_{34}) = (-P_{21} - P_{22} + 6P_{23} + P_{24}).$$

І, нарешті, для точки 8:

$$664(3P_{31} - 3P_{32} + 2P_{33} + P_{34}) = (3P_{11} - 3P_{12} + 2P_{13} + P_{14}),$$
$$342(3P_{31} - 3P_{32} + 2P_{33} + P_{34}) = (3P_{21} - 3P_{22} + 2P_{23} + P_{24}).$$

Матриця системи однорідних лінійних рівнянь для 12 невідомих $P_{11}, P_{12}, P_{13}, P_{21}, \ldots, P_{34}$ буде:

$$A = \begin{pmatrix}
-1 & 1 & 6 & 1 & 0 & 0 & 0 & 0 & 461 & -461 & -2766 & -461 \\
0 & 0 & 0 & 0 & -1 & 1 & 6 & 1 & 207 & -207 & -1242 & -207 \\
-1 & -1 & 6 & 1 & 0 & 0 & 0 & 0 & 520 & 520 & -3120 & -520 \\
0 & 0 & 0 & 0 & -1 & -1 & 6 & 1 & 233 & 233 & -1398 & -233 \\
1 & -1 & 6 & 1 & 0 & 0 & 0 & 0 & -569 & 569 & -3414 & -569 \\
0 & 0 & 0 & 0 & 1 & -1 & 6 & 1 & -209 & 209 & -1254 & -209 \\
-2 & 2 & 4 & 1 & 0 & 0 & 0 & 0 & 812 & -812 & -1624 & -406 \\
0 & 0 & 0 & 0 & -2 & 2 & 4 & 1 & 552 & -552 & -1104 & -276 \\
-2 & -2 & 4 & 1 & 0 & 0 & 0 & 0 & 1042 & 1042 & -2084 & -521 \\
0 & 0 & 0 & 0 & -2 & -2 & 4 & 1 & 662 & 662 & -1324 & -331 \\
2 & -2 & 4 & 1 & 0 & 0 & 0 & 0 & -1236 & 1236 & -2472 & -618 \\
0 & 0 & 0 & 0 & 2 & -2 & 4 & 1 & -552 & 552 & -1104 & -276 \\
-3 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 1065 & -1065 & -710 & -355 \\
0 & 0 & 0 & 0 & -3 & 3 & 2 & 1 & 1020 & -1020 & -680 & -340 \\
-3 & -3 & 2 & 1 & 0 & 0 & 0 & 0 & 1563 & 1563 & -1042 & -521 \\
0 & 0 & 0 & 0 & -3 & -3 & 2 & 1 & 1290 & 1290 & -860 & -430 \\
3 & -3 & 2 & 1 & 0 & 0 & 0 & 0 & -1992 & 1992 & -1328 & -664 \\
0 & 0 & 0 & 0 & 3 & -3 & 2 & 1 & -1026 & 1026 & -684 & -342.
\end{pmatrix}$$

Для існування розв'язку такої системи однорідних лінійних рівнянь, ранг матриці A повинен бути нижчім за 12. Якщо це не так, то його необхідно понизити. Для цього в сингулярному розкладі матриці A найменшому власному значенню приписується 0.

4. Знаходження епіполярних ліній за проекційними матрицями

Якщо P_L – проекційна матриця для першої камери, P_R – проекційна матриця для другої камери, то сімейства епіполярних ліній можна записати у вигляді

$$(U_L)^T \; F \; U_R = X^T (P_L)^T F \, P_R \, X = 0, \qquad (3)$$

причому добуток $(P_L)^T \; F \; P_R$ має бути антисиметричною матрицею.

При знаходженні F за проекційними матрицями P_L, P_R можна перетворити їх так, що

$$P_L := P_L H \quad P_R := P_R H, \qquad (4)$$

де матрицю H необхідно вибрати таким чином, щоб P_L стала

$$P_L = \{(1\ 0\ 0\ 0),(0\ 1\ 0\ 0),(0\ 0\ 1\ 0),(0\ 0\ 0\ 0)\}.$$

Використовуючи (3) та (4), маємо:

$$(U_L)^T \; F \; U_R = X^T (P_L H)^T F \, P_R \, H \, X = \\ = X^T H^T (P_L)^T F \, P_R \, H \, X,$$

внаслідок чого

$$F = \{p1,p2,p3,p4\}^T \{[p4,p1], [p4,p2], [p4,p3], [p4,p4]\}^T,$$

де p_i – вектори-стовпці матриці P_R.

Можна переконатися, що добуток $P_L \; F \; P_R$ буде антисиметричною матрицею [4].

Знаючи сімейства епіполярних ліній, нескладно перетворити зображення так, щоб фундаментальна матриця мала вигляд

$$F = \{(0,0,0), (0,0,-1),(0,1,0)\}.$$

5. Результати обробки стереозображень

Для двох закріплених одна відносно одної фотокамер досить один раз розрахувати проекційні матриці. Всі знімки, що будуть надалі зроблені такою системою, калібруватимуться одним і тим самим перетворенням. Калібровані зображення будуть придатними для обробки порядковими методами розпізнавання рельєфу.

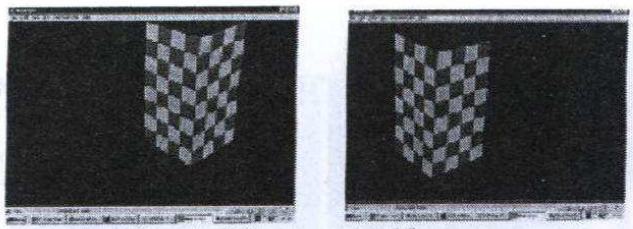


Рис. 4. Об'єкт для калібрування

На Рис. 4 зображена стереопара об'єкта з відомими просторовими координатами. На основі цих зображень знаходяться проекційні матриці та епіполярні лінії для всіх знімків, що будуть зроблені цими ж камерами.

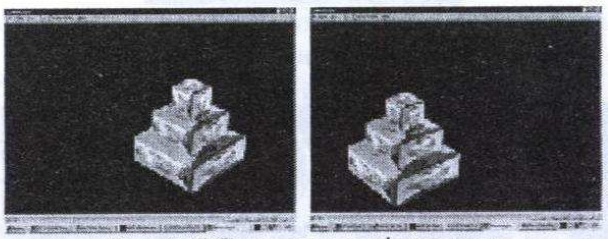


Рис. 5. Об'єкт для розпізнавання

На Рис. 5 зображенна стереопара об'єкта, просторову конфігурацію якого необхідно розпізнати. Рядки розгортки на цій стереопарі не відповідають епіполярним лініям. Зображення слід калібрувати перед розпізнаванням.

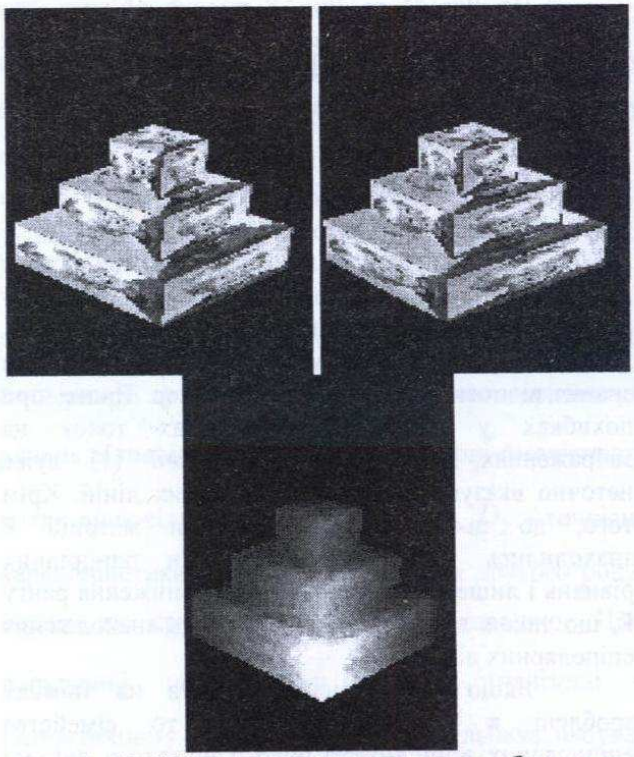


Рис. 6. Калібровані знімки та висотне зображення

На Рис. 6 подані калібровані стереозображення об'єкта та його відновлена просторова конфігурація.

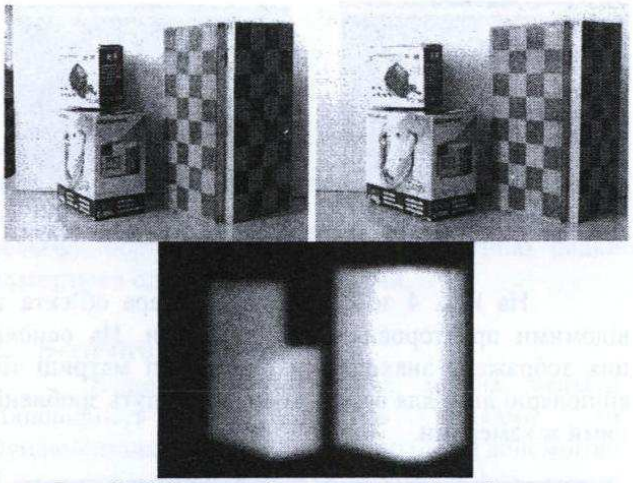


Рис. 7. Об'єкти, сфотографовані цифровою камерою, та відновлений рельєф

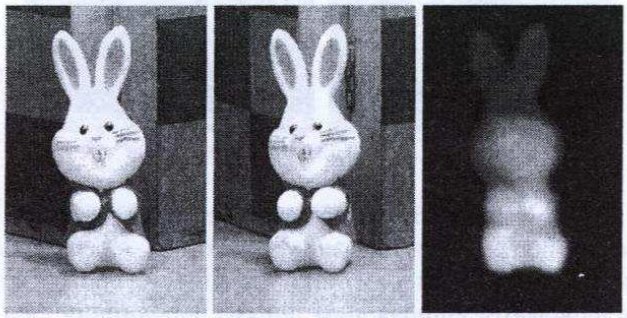


Рис. 8. Об'єкт та його відновлений рельєф

На Рис. 7 та Рис. 8 подано об'єкти, що сфотографовані цифровою камерою, та їх відновлена просторова конфігурація.

ВИСНОВКИ

Для калібровки зображень було запропоновано два методи. Суть першого – у знаходженні епіполярних ліній за вказаними парами відповідних точок на зображеннях, другого – у знаходженні проекційних матриць.

Перший метод можна застосовувати без знання відносного розташування камер. Проте, при похибках у вказаних координатах точок на зображеннях, псевдорозв'язок задачі (1) дуже неточно вказує сімейство епіполярних ліній. Крім того, до цього часу коефіцієнти матриці F знаходились як розв'язок системи однорідних рівнянь і лише потім проводилось пониження рангу F, що також зумовлювало неточність знаходження епіполярних ліній [4].

Якщо всі елементи об'єкта на знімках зроблені в одному масштабі, то сімейства епіполярних ліній можна було б шукати у вигляді паралельних прямих на кожному із знімків. Задача (1) в цьому випадку значно спрощується. Матриця F буде автоматично мати ранг 2 та шукатися у вигляді

$$F = \begin{bmatrix} 0 & 0 & F_{13} \\ 0 & 0 & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix}$$

Матрицю F, що отримали таким чином можна використовувати як перше наближення і для загального випадку калібровки зображень. Елементи матриці F слід знаходити ітеративно, на кожному кроці уточнюючи якийсь стовпець чи рядок матриці при фіксованих інших елементах. Обмеження на ранг в такому випадку буде лінійне по елементам F_{ij} рівняння. Таким чином, на кожному кроці ітерації фундаментальна матриця буде мати ранг 2, а точність знаходження її елементів не погіршуватись.

Калібровка другим способом дозволяє створювати ідеальні стереопари двома фотоапаратами. Для двох камер слід зафіксувати взаємне положення і один раз провести вимір проекційних матриць, сфотографував об'єкт з відомими просторовими координатами. Всі знімки, що будуть отримані надалі, можна буде перетворювати в ідеальні стереопари.

Автором було створено програмно-апаратний комплекс, що дозволяє розпізнавати рельєф на зображеннях з використанням попередньої калібровки. Знімки робились цифровою камерою, що могла переміщатися по спеціальному штативу. Калібровані зображення оброблялись порядковим алгоритмом розпізнавання рельєфу.

ЛІТЕРАТУРА

[1] Г.Л. Гимельфарб. " Симметризованное би- и тринокулярное стереозрение: взаимосвязь между теоретическими основами и эвристическими решениями." Теоретические и прикладные вопросы распознавания изображений. Киев 1995.

[2] R. Sara. "The Class of Stable Matching for Computation Stereo."

[3] R. Hartley. "In defence of the 8-point algorithm." In InternationalConference on Computer Vision, pages 1064–1070, 1995.

[4] T. Werner. "Practice of 3D Reconstruction from Multiple Uncalibrated Unorganized Images. Czech Pattern Recognition Workshop 2000."

[5] Г.Л. Гимельфарб. "Симметричный подход к задаче автоматических стереоскопических измерений в фотометрии." Кибернетика 1979.

АВТОФОКУСУВАННЯ ЕЛЕКТРОННОГО МІКРОСКОПУ ЗА ЗОБРАЖЕННЯМ

А.С. Краснопоясовський, А.М. Скаковська

Сумський державний університет
вул. Римського-Корсакова, 2, м.Суми, 40007
33-03-90, факс (0542) 33-40-58
E-mail: pm@ssu.sumy.ua

Розглядається застосування методу функціонально-статистичних випробувань (МФСВ), основаного на оцінці інформаційної здатності системи розпізнавання образів, для класифікаційної настройки (КН) складного об'єкту контролю та управління.

ВСТУП

Важливою задачею в електронній мікроскопії є автофокусування растрового електронного мікроскопу (РЕМ). Це особливо актуально, наприклад, при автоматичній класифікації об'єктів дослідження, яка вимагає додержання незмінних і однакових умов для прийняття високодостовірних рішень. До цього часу невідомі ефективні алгоритми розв'язання цієї задачі. Розглянемо підхід до розв'язання задачі автофокусування РЕМ за зображенням в рамках МФСВ [1,2].

1 ТЕОРЕТИЧНІ ТА МЕТОДОЛОГІЧНІ ПОЛОЖЕННЯ ПІДХОДУ

Фокусність зображення досліджуваного зразка залежить від багатьох факторів, які впливають на функціонування електронного мікроскопу. Для РЕМ в якості управляючого контрольованого параметру виступає струм об'єктивної лінзи $I_Л$. Ідея автофокусування РЕМ полягає в обчисленні інформаційної міри початкового (базового) розфокусованого зображення X_0 відносно поточного зображення X_S, яке встановлюється на S-ому кроці КН, порівнянні її значення з попереднім і прийнятті рішення про зміну управляючого параметру на задану величину. Процес КН продовжується до тих пір, поки значення струму $I_Л$ не забезпечить прийнятної з практичної точки зору фокусності зображення. В якості критерію функціональної ефективності (КФЕ) КН розглянемо інформаційну міру Кульбака, яка є добутком відношення правдоподібності Λ на міру відхилень розподілів ймовірностей. В загальному випадку відношення правдоподібності подамо у вигляді

$$\Lambda = \log_2 \frac{P_t}{P_f} = \frac{p_1 D_1 + p_2 D_2}{p_1 \alpha + p_2 \beta}, \qquad (1)$$

де $P_t = p_1 D_1 + p_2 D_2$, $P_f = p_1 \alpha + p_2 \beta$ -повні ймовірності правильного і неправильного прийняття рішень відповідно; p_1 , p_2 - апріорні ймовірності появи відповідних подій; α , β , D_1 , D_2 - точнісні характеристики: помилки першого та другого роду, перша та друга достовірності відповідно. При допущенні, що $p_1 = p_2 = 0.5$ (найгірші в статистичному сенсі умови) міра Кульбака набуває вигляду

$$E = 0{,}5 \log_2 \left(\frac{D_1 + D_2}{\alpha + \beta} \right) \left[(D_1 + D_2) - (\alpha + \beta) \right] =$$

$$= \log_2 \left(\frac{2 - (\alpha + \beta)}{\alpha + \beta} \right) \left[1 - (\alpha + \beta) \right]. \qquad (2)$$

© UkrObraz'2000

При введенні оцінок $\alpha = k_1/n^*$, $\beta = k_2/n^*$, де k_1, k_2 - число реалізацій класів X_0 і X_S відповідно, які не знаходяться в області своїх класів; n^* - кількість спостережень, які забезпечують репрезентативність вибіркової послідовності. Тоді формула (2) набуває робочого вигляду

$$E = \frac{1}{n}\log_2\left\{\frac{2n + 10^{-r} - [k_1 + k_2]}{[k_1 + k_2] + 10^{-r}}\right\} * [n - (k_1 + k_2)], \quad (3)$$

де r -число цифр у мантисі значення критерію E, яке вводиться на випадок появи нульових емпіричних частот. В якості розділяючої гіперповерхні (РГП) у першому наближенні береться гіперсфера, центром якої є еталонний вектор (ЕВ) $x_0 \in X_0$.

2 ОБРОБКА ЗОБРАЖЕННЯ

Оскільки сучасні ЕОМ дозволяють розрізняти 256 градацій чорно-білого зображення, то його обробку доцільно здійснювати шляхом сканування рецепторного поля, що є природним для РЕМ, і побудови розподілу яскравості для кожного рядка рецепторів. Квантовані в часі криві розподілів яскравості розглядаються як випадкові реалізації зображення $\{g^{(j)}_{0,n}\}$, де $n = \overline{1, n^*}$; $j = \overline{1, J_{\max}}$ - змінна числа рецепторів у рядку.

З метою формування бітового масиву значень реалізацій $\{x^{(j)}_{0,n}\}$ здійснюється їх квантування на два рівні за формулою

$$x^j_n = \begin{cases} 1, & if \ A_H \le g^{(j)}_n \le A_B; \\ 0, & if \quad else, \end{cases} \quad (4)$$

де A_H, A_B - нижній і верхній контрольний допуски на рецептори. Оптимізація контрольних допусків в інформаційному сенсі здійснюється безпосередньо в рамках алгоритму КН. За початкові значення A_H і A_B приймаються відповідно $A_H = A_0 - \Delta$ і $A_B = A_0 + \Delta$, де A_0 - номінальне (середнє) значення яскравості зображення. Оскільки двобічне поле контрольних допусків вибрано симетричним, то початкове значення Δ дорівнює чверті шкали яскравості, тобто 64. Формування бінарного ЕВ базового класу X_0 здійснюється за формулою

$$x^{(j)}_0 = \begin{cases} 1, & if \ \sum_{n-1}^{n^*} x^{(j)}_{0,n} \ge n^*/2; \\ 0, & if \quad else. \end{cases} \quad (5)$$

Аналогічно обробляється поточне зображення X_S. Отже, в процесі обробки зображень X_0 і X_S здійснюється формування вхідних даних, які за МФСВ необхідні для обчислення інформаційної міри між базовим і поточним зображеннями. Такими вхідними даними є ЕВ x_0, x_s і відповідно випадкові реалізації $\{x^{(j)}_{0,n}\}$, $\{x^{(j)}_{s,n}\}$.

Визначення оптимальної в інформаційному сенсі кодової відстані d^*_0 - радіуса РГП, здійснюється за послідовним алгоритмом

$$d_0(k) = d_0(k-1) + \begin{cases} h, if \ E[d(k-1)] < E[d(k)]; \\ 0, if \ E[d(k-1)] \ge E[d(k)], \end{cases}$$

де k – число прирощень радіуса; h – крок прирощення. Обчислення КФЕ закінчується при знаходженні екстремального значення $E^* = \max_{\{d\}} E$, де $\{d\} = \{1,2,..d(x_0 \oplus x_s)\}$ – множина радіусів РГП класу X_0; $d(x_0 \oplus x_s)$ - міжцентрова кодова відстань.

Оптимізація контрольних допусків здійснюється після вибору d^*_0 для базового класу X_0, за алгоритмом $\{\delta_{k,i}\} = \{\delta^*_{k,i}\}$, якщо $E^*[\{\delta^*_{k,i}\}, d^*_0] = \max_{\{\delta^*_{e,i}\}} E^*[d^*_0]$, де $\{\delta_{e,i}\}$ – система нормованих полів допусків на рецептори.

3 ОПИС АЛГОРИТМУ

Означимо три класи функціонального стану РЕМ: X_H^0, який відповідає прийнятній фокусності, тобто оцінці "НОРМА"; X_M^0, який відповідає значенню параметру $I_Л$ "МЕНШЕ НОРМИ"; $X_Б^0$, який відповідає значенню параметру $I_Л$ "БІЛЬШЕ НОРМИ". Зрозуміло, що класи X_M^0 і $X_Б^0$ мають розфокусовані зображення різного ступеню. Умовою запуску алгоритму є вихід РЕМ зі стану, що відповідає класу X_H^0. За базовий клас X_0 приймаємо стан РЕМ на момент запуску алгоритму КН. На рис.1 наведено структурну схему алгоритму автофокусування РЕМ в рамках МФСВ.



Рисунок 1.

Формування класу X_0 відбувається таким чином:

1 Зчитується масив значень яскравості зображення базового класу X_0.

2 Здійснюється вибір початкових допусків A_H і A_B.

3 Формується бінарний масив $BM0[N,J]$ за процедурою (4).

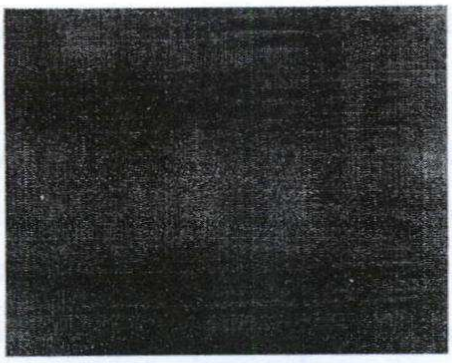
4 Формується ЕВ $EV0[J]$ шляхом статистичного усереднення масиву $BM0[N,J]$ за процедурою (5).

Після визначення напрямку зміни струму $I_Л$ шляхом встановлення приналежності класу X_0 до одного із класів X_M^0 або $X_Б^0$ обирається стратегія зміни $I_Л$ на кожному кроці КН.

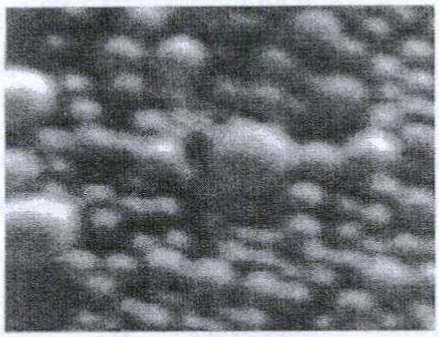
Поточне зображення X_s формується аналогічно X_0 при незмінних контрольних допусках. КФЕ обчислюється за формулою (3) в рамках наведених процедур оптимізації параметрів навчання [2].

4 РЕАЛІЗАЦІЯ АЛГОРИТМУ

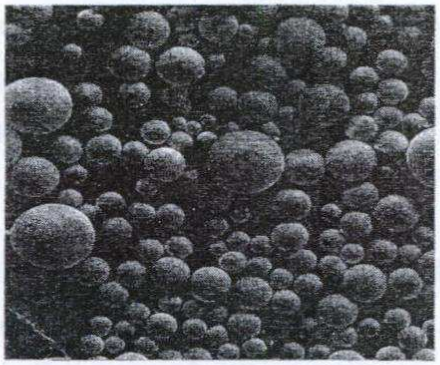
Реалізація розглянутого алгоритму здійснювалася для автофокусування серійного мікроскопу РЕМ-103 виробництва АТ СЕЛМІ (м.Суми, Україна), для якого сталий струм обмотки об'єктивної лінзи змінюється в діапазоні 0.3-2.0 А. При цьому значення струму I_N, яке відповідає класу X_H^0, знаходиться в інтервалі 1.08-1.32 А і залежить від багатьох внутрішніх і зовнішніх випадкових факторів функціонування РЕМ. Приростання струму $I_Л$ на кожному кроці КН здійснювалося шляхом поділу навпіл різниці між розрахованим (номінальним) і поточним його значеннями. На рис.2 наведено зображення досліджуваного зразка: базове розфокусоване зображення (рис. 2,а), зображення після першого кроку настройки (рис. 2,б) і екстремальне сфокусоване зображення (рис. 3,в).



а)



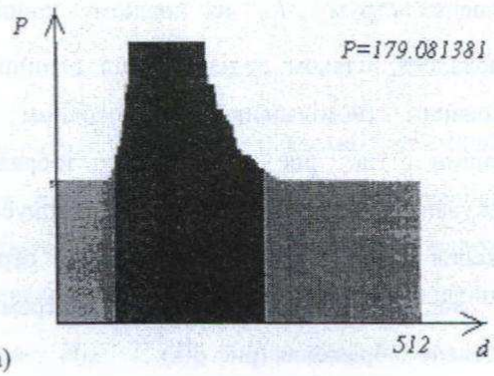
б)



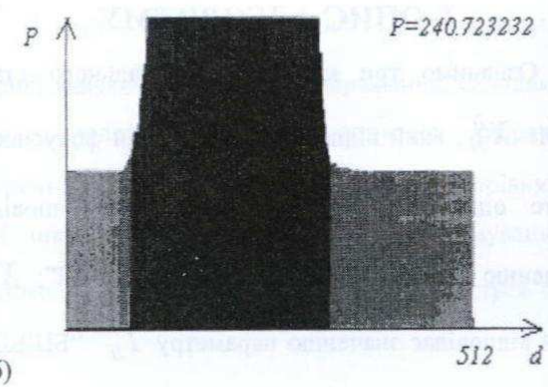
в)

Рисунок 2.

На рис.3,а наведено значення нормованого критерію (3), при $S=1$ і на рис. 3,б - для екстремального кроку настройки. Тут P - площа робочої області значень КФЕ, яка обчислюється на кожному кроці КН за умови, що D_1, $D_2 \geq 0.5$, і на рис.3 визначається затемненою частиною графіку.



а)



б)

Рисунок 3.

Тривалість автофокусування для розглянутого базового зображення становила 5с, що обумовлено інерційністю управляючої обмотки об'єктивної лінзи.

ВИСНОВКИ

1 Автонастройка складної системи розглядається як ітераційна процедура пошуку максимума інформаційної міри різноманітності базового початкового образу настройки і образами, що формуються на наступних кроках настройки.

2 Обчислення КФЕ на поточному кроці КН здійснюється в рамках алгоритму навчання за МФСВ.

3 Автофокусування РЕМ за розглянутим алгоритмом КН є цілеспрямованою ітераційною процедурою, яка здійснюється в межах прийнятної з практичних міркувань оперативності при довільних початкових умовах.

ЛІТЕРАТУРА

1. Краснопоясовський А.С. Технічна діагностика складних систем за методом функціонально-статистичних випробувань//Обробка сигналів та розпізнавання образів: *Праці третьої Всеукраїнської міжнародної конференції*, 1996.-С.174-177.

2. Краснопоясовський А.С., Черниш А.В. Оцінка функціональної ефективності навчання системи розпізнавання. *Вісник СумДУ*, 1997.-№2.-С.112-118.

MULTISTAGE APPROACH TO IMAGE PROCESSING

Y.V. Baybak[1], N.O. Shweiki[1], Y.F. Kutaev[2], A.A. Gertsiy[1], L.I. Timchenko[1].

[1]Vinnitsa State Technical University, Khmelnitskoye Shosse 95, Vinnitsa, 21021, Ukraine,
E-mail: timchen@svitonline.com
[2]State Scientific Enterprise "Astrophisika", Volokolamskoye shosse 95, Moscow, 129820, Russia

ABSTRACT

We present a novel three-dimensional network and its application to pattern analysis. In this paper, a novel approach is proposed, which allows for an efficient reduction of the amount of visual data required for representing structural information in the image. This algorithm is tolerant to minor structural changes and can be used for automatic face recognition.

The approach is based on a multistage architecture, which investigates partial clustering of structural image components. The initial grey-scale representation of the input image is transformed into a structural representation, so that each image component contains information about the spatial structure of its neighbourhood. The output result is represented as a pattern vector, whose components are computed one at a time to allow the quickest possible response. The input pattern is identified as the best match between the output pattern vector and the model vectors from the database. This approach may be employed for the biomedical image processing, such as ultrasonics and X - ray image.

1. INTRODUCTION

Efficient representation of visual information is one the most important issues in the automatic recognition of human faces. In the case of a feature-based representation, each face is stored as a pattern vector [1]. Yet the connectionist models generally operate on image-based representations by using distributed image representations.

Local feature detectors, which are often employed in these models, make direct representations much more effective in terms of the preservation of both texture and shape information [2]. However, they put excessive requirements on the amount of memory to store all possible descriptors [3]. Most of the attempts to reduce the amount of data necessary to store faces are based either on the "winner takes all" rule [4] or on simple statistical calculations [5]. Those models do not attempt to exploit the vast redundancy of the visual representation [6], nor take into account configural information.

2. METHODOLOGY

The aim of the proposed approach is to reduce the amount of visual data whilst preserving structural information in the image. For this task, a novel three-dimensional multistage network architecture is used (Figure 1). The first dimension of the network manifests itself as the parallel channels; the second dimension shows as the ranking of the clustered image components depending on their information content. The third network dimension is specified by a flexible hierarchical structure of further processing and is illustrated in Figure 1 as processing levels. The output of the network is represented as a pattern vector.

The purpose of the processing in the multistage network is to discover the structural regularities and to represent them as a smaller number of more complex units. No data is discarded; instead, data with the higher information content is processed first. The term "information content" is based on the degree of correlation between the data components. This feature allows the computation of the output vector in stages - the first components of the vector contain the larger proportion of the structural output, and the following components contain less correlated information. At each stage, a single output component is computed. Thus, it is possible to analyse the output result after each stage.

Consider a transformation $G(M) = \{a_j \mid j = 1, 2, ..., m\}$ of the input image M, which performs mapping of the initial set of image components into a new set of components. Possible transformations are discussed in Section 3. If the initial image is partitioned into n segments then this transformation G may be applied in each processing channel i, $i = 1, ..., n$, simultaneously. The i-th image segment may be denoted as $M_i^1 = \{a_{ij}^1 \mid j = 1, 2, ..., m^1\}$, where m^1 is the cardinality of a new set. The new components are graded according to their information content.

© UkrObraz'2000

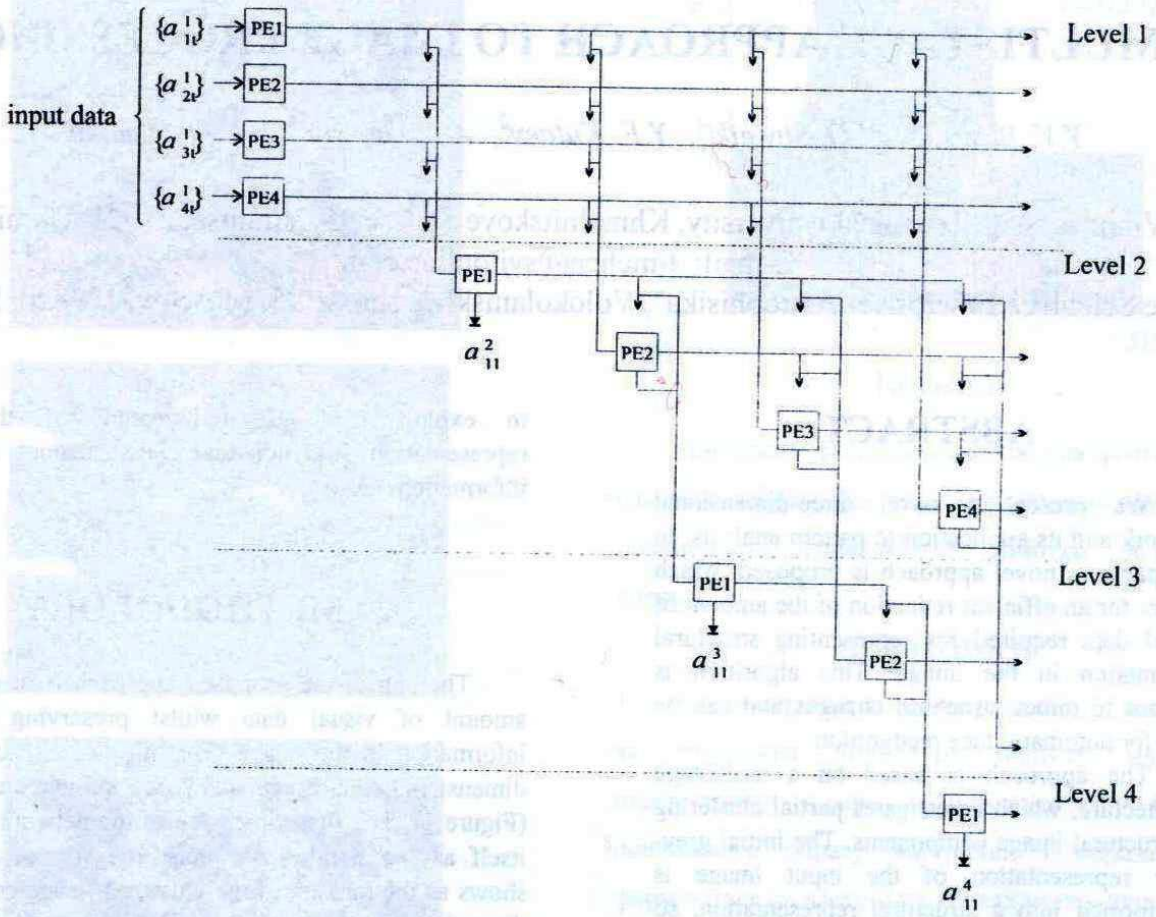


Fig. 1. Topology of the 3-D multistage hierarchical network.

They are extracted to the second processing level with respect to their grades, in order to further investigate similarities between them. At the first level, all channels can be considered as a matrix A_1 of the form:

$$A_1 = \begin{bmatrix} M_1^1 \\ M_2^1 \\ \vdots \\ M_i^1 \\ \vdots \\ M_n^1 \end{bmatrix} = \begin{bmatrix} a_{11}^1 & a_{12}^1 & \cdots & a_{1m}^1 \\ a_{21}^1 & a_{22}^1 & \cdots & a_{2m}^1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1}^1 & a_{i2}^1 & \cdots & a_{im}^1 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1}^1 & a_{n2}^1 & \cdots & a_{nm}^1 \end{bmatrix}.$$

The column index indicates the stage of the component extraction and the row index indicates the channel from which that component is extracted. In algebraic form transmission of the decomposed data A_1 to the second level is represented as transposition T, $A_1^T = T(A_1)$.

On the second level the components are placed into the processing channels according to their grades on the first level. The transformation G is applied again, this time to the rows of A_1^T.

$$A_2 = G(A_1^T) = G(T(G(M))) = \begin{bmatrix} a_{11}^2 & a_{12}^2 & \cdots & a_{1n^2}^2 \\ a_{21}^2 & a_{22}^2 & \cdots & a_{2n^2}^2 \\ \vdots & \vdots & \vdots & \vdots \\ a_{n^2 1}^2 & a_{n^2 2}^2 & \cdots & a_{n^2 n^2}^2 \end{bmatrix}$$

where $n^2 = m^1$ is the number of parallel channels at the second level, and m^2 denotes the number of components (processing stages) at the second level.

The matrix A_2 represents the second processing level after the transformation in the way that each component a_{ij}^2 is extracted at the stage $t=i+j-1$. If the matrix A_2 is aligned so that the column index illustrates the stage of the component extraction, the same processing procedure can be applied on each following level of the hierarchy. This alignment is implemented as horizontal shift of each row of A_2 in order to position the first nonempty element on the leading diagonal. The resultant matrix is denoted A_2', and $A_2' = P(A_2)$, where P is the aligning transformation.

$$A_2' = \begin{bmatrix} a_{11}^2 & a_{12}^2 & \cdots & \cdots & a_{1m^2}^2 & x & x & \cdots & x \\ x & a_{21}^2 & a_{22}^2 & \cdots & a_{2m^2-1}^2 & a_{2m^2}^2 & x & \cdots & x \\ x & x & a_{31}^2 & \cdots & a_{3m^2-2}^2 & a_{3m^2-1}^2 & a_{3m^2}^2 & x & x \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x & \cdots & \cdots & x & x & a_{n^2 1}^2 & a_{n^2 2}^2 & \cdots & a_{n^2 m^2}^2 \end{bmatrix}$$

where x denotes empty positions.

The component a_{11}^2 is the only one to be extracted at the first stage of the processing at the second level. Therefore, it is unrelated with all the other second-level components in time and effectively is an output component. It represents an intermediate result of the processing, and is the only component extracted as a result of the multistage processing at the second level. This component is extracted from the matrix by removing the first matrix column.

The remaining clustered components are stored in the matrix A_2''. This extraction is performed by applying a transformation L, that is, $L(A_2') = a_{11}^2 + A_2''$. The initial matrix of the third level A_2^T is obtained after transposition T of the A_2'', that is $A_2^T = T(L(P(A_2)))$.

The described procedure is appled at each of the following hierarchical levels until the $k_{\max}$-th level, where the matrix $A_{k_{\max}}$ contains a single element. Then $T(L(P(A_{k_{\max}}))) = \varnothing$.

3. SYSTEM PERFORMANCE

Several transformations can be applied in the multistage network. Some possible examples include principal component analysis of local image patches [7], partial clusterisation of the Gabor coefficients [8], and the cosine transform. In this paper, analysis of the facial structure is performed using partial clusterisation of pixel connectivity. A three-level representation [9] of grey-scale facial image is used to specify a three-level pixel connectivity The three connectivity levels are then processed using three multistage networks, each producing a pattern vector (Figure 2).

In each network, the image components are clustered using the ISODATA algorithm [10]. Then, the newly organised clusters are graded on the basis of the uniformity of data in those clusters.

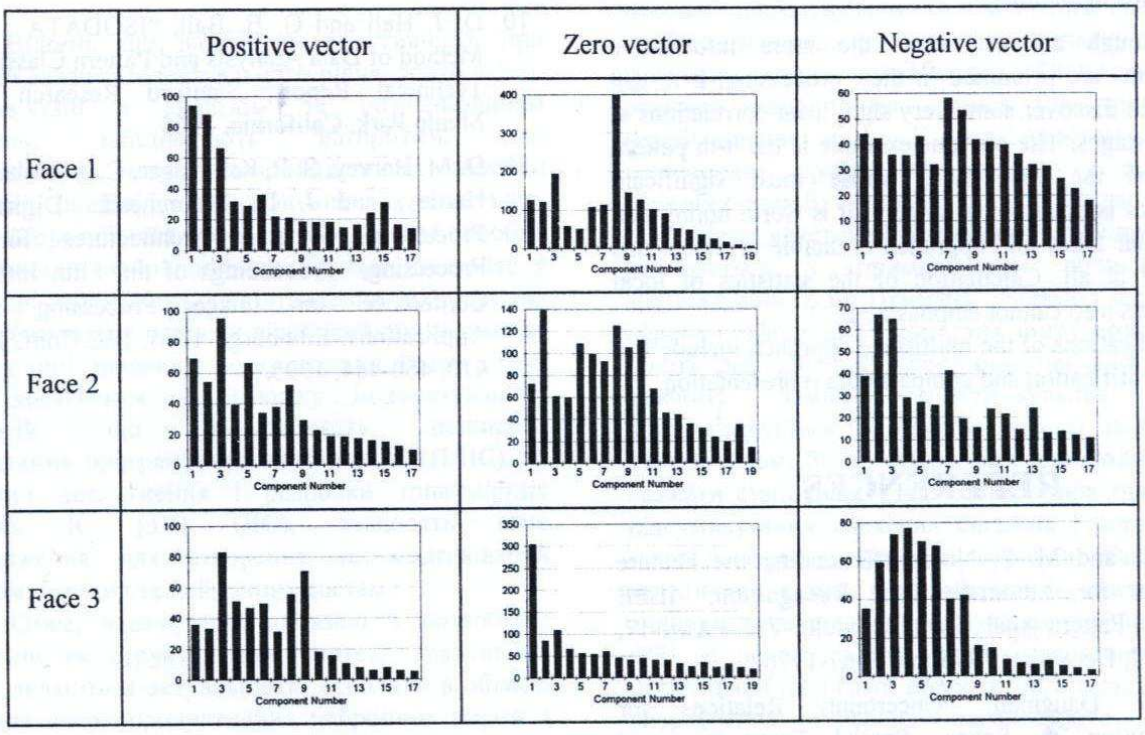


Fig. 2. Pattern vectors generated by the multistage algorithm for three faces.

Components with equal grades are extracted from different channels in order to be processed in the same way at the higher hierarchical levels. At each processing level, starting from the second one, an output component is extracted. Thus, analysis of the output can be performed after the early stages of processing.

The output pattern vectors highlight different structural organisation of each particular face. At the same time, minor structural changes do not affect the first vector components, thus making the algorithm tolerant to slight changes in facial expression.

In the current realisation, the system may be viewed as performing a multistage sorting of the changing data. Even if the components are not correlated at the low levels of the multistage network, they can be found correlated at some higher levels. Any correlations affect the values of the remaining data and influence the output result.

4. DISCUSSION AND CONCLUSIONS

There are several factors, which contribute to the fast speed of this multistage processing:

1. Image data is segmented and fed into a number of parallel channels. Components from different segments may be only processed together at the higher levels, if they are assigned the same rank. Thus, the number of possible combinations is significantly reduced.
2. Calculations in the channels can be performed in parallel, for example, on our custom TMS320C40 image processing computer [11];
3. Components of the output vector are extracted one at a time, with priority given to the more informative ones. Thus, the algorithm is robust to minor structural changes in the image.

Although at each level the more informative components are promoted in their processing, it is still possible to discover some very significant correlations at the later stages. The obvious example is the 0-th pattern vector of the first face, whose most significant component is extracted at stage 8. It is worth noting that the "winner takes all" approach is unable to detect this regularity at all. Calculation of the statistics of local coefficients also cannot emphasize it.

Applications of the multistage approach include fast object identification and compact data representation.

REFERENCES

1. X. Jia and M. S. Nixon, "Extending the Feature Vector for Automatic Face Recognition," IEEE Trans. Pattern Anal. Machine Intell., Vol. PAMI-17, No. 12, December 1995, pp. 1167-1176.
2. J. G. Daugman, "Uncertainty Relations for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters," J. Optical Soc. Amer. A2, Vol. 7, pp. 1160-1169, 1985.
3. M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. v. d. Malsburg, R. P. Wurtz and W. Kohen, "Distortion Invariant Object Recognition in the Dynamic Link Architecture," IEEE Trans. Computers, Vol. 3, pp. 300-311, 1993
4. N. Petkov, "Biologically Motivated Computationally Intensive Approaches to Image Pattern Recognition," Future Generation Computer Systems, Vol. 11, pp. 451-465, 1995.
5. J F. Goudail, E. Lange, T. Iwamoto, K. Kuyma, and N. Otsu, "Fast Face Recognition Using High Order Autocorrelations", Proceedings of 1993 Joint Conference on Neural Networks, pp. 1297-1300, 1993
6. J. M. Tenenbaum and A. P. Witkin, "Perceptual Organization as Building-Blocks for Vision," Journal of the Optical Society of America, A-Optics and Image Science, Vol. 1, No. 12, pp. 1216, 1984.
7. R. P. N. Rao and D. H. Ballard, "Natural Basis Functions and Topographic Memory for Face Recognition," Proc. Intl. Joint Conf. on Artif. Intell. (IJCAI), pp. 10-17, 1995.
8. J. Daugman, "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by 2D Visual Cortical Filters," J. Opt. Soc. Amer. (A), Vol. 2, No. 7, pp. 1160-1169, 1985.
9. J. M. White and G. D. Rohrer, "Image Thresholding for Optical Character Recognition and Other Applications Requiring Character Image Extraction," IBM J. Res. Develop., Vol. 27, No. 4, pp. 400-411, 1983.
10. D. J. Hall and G. B. Ball, "ISODATA : A Novel Method of Data Analysis and Pattern Classification," Technical Report, Stanford Research Institute, Menlo Park, California, 1965.
11. D. M. Harvey, S. P. Kshirsagar, C. A. Hobson, D. A. Hartley, and J. D. Moorehead, "Digital Signal Processing Systems Architectures for Image Processing," Proceedings of the Fifth International Conference on Image Processing and Its Applications, Edinburgh 1995, IEE Conf. Publ. No. 410, 460-464, 1995.

УДК 681.3

Математична модель нейрона на принципах паралельної порогової обробки інформації

Тетяна Мартинюк, Антоніна Буда, Андрій Кожем'яко, Тетяна Васильєва, Віра Козлова

Вінницький державний технічний університет
21021, Вінниця, Хмельницьке шосе, 95
Тел.: (0432) 325718. Факс: (0432) 465772
E-mail: laser@lili.vstu.vinnica.ua

New way to organization of function principle of neuron as basis element of multilayer neural networks is suggested. The principles of multioperand summation of number array with using difference slices are laid to the basis of mathematical model of neuron. The possibility of combined execution two operation-multioperand summation and comparison partial sums with given threshold of processing is shown.

Вступ

Відомо, що найбільшу ефективність при реалізації задач інтелектуального рівня, тобто задач, що пов'язані з аналізом та розпізнаванням зображень, забезпечують алгоритми, що використовують методи сприйняття та обробки інформації, подібні до нейробіологічних, властивих людині. Моделювання процесів паралельної обробки у нейронних і нейроподібних мережах дозволяє у деякій мірі наблизитись до відтворення тих підходів, на яких базуються процеси відображення та аналізу образів у корі головного мозку людини. Разом з тим, значні досягнення у розвитку інтелектуальних технологій, що обумовлюють поширене використання програмуємих логічних ІС (ПЛІС), та інтенсивні дослідження і розробки тривимірних оптичних ІС (3-D ОІС), знаходять своє запровадження для створення та моделювання нейрокомп'ютерів та нейронних систем.

Отже, задачі, що пов'язані з розробкою алгоритмів та структур для синтезу адаптивних систем, являються актуальними, особливо в області створення високопродуктивних нейронних мереж з можливостями навчання і самонавчання, які спрямовані на розв'язання задач "інтелектуального" рівня.

Особливості моделі нейрона

Розглянемо структуру та функціональні можливості нейрона, як елемента шара, що забезпечує інформаційний прийом у багатошаровій нейронній мережі. У загальному випадку він реалізує порогову функцію виду:

$$S = \sum_{i=1}^{m} x_i w_i \geq p, \qquad (1)$$

де x_i – вхідні величини, w_i – ваги, p – поріг обробки.

Двовходова структура, що моделює нейрон із найпростішою пороговою функцією (1), показана на рис.1а. Аналіз цієї структури свідчить про те, що найбільш витратними за часом та складністю є операції підсумовування та порівняння. Разом з тим, треба відзначити, що у даному випадку зовсім не важливим є визначення абсолютної величини суми S, оскільки значення має лише факт рівності або перебільшення суми S над порогом p. Отже послідовний процес, що полягає у формуванні спочатку суми S, а потім у порівнянні двох величин S і p, більш властивий формалізованому процесу, що реалізується у примітивному обчислювачі з архітектурою фон Неймана. Відомо, що процеси обробки інформації у коркових зонах мозку людини мають розподілений просторово-часовий характер, причому замість підсумовування сигналів використовується їх накопичення із порівнянням. Таким чином, більш природним для людини можна вважати суміщення у часі таких двох операцій, як підсумовування зважених сигналів і порівняння їх суми із заданим порогом обробки. Такий підхід можливий, якщо підсумовування масиву чисел виконувати за методом формування різницевих зрізів (РЗ). В основу цього методу покладено принцип паралельної порогової обробки інформації, який був розроблений стосовно паралельного додавання тривалостей групи часових інтервалів [1]. Суть методу полягає в тому, що в якості внутрішнього порогу q обробки у кожному циклі визначається величина, що дорівнює найменшому значущому доданку у поточному масиві чисел, тому значення порогу q у процесі обчислення змінюється. Процес формування суми масиву вхідних чисел закінчується тоді, коли остаточний масив доданків не містить додатних чисел. Сума формується за виразом

© UkrObraz'2000

$$S = \sum_{j=1}^{N} S_j = \sum_{j=1}^{N} q_j \cdot b_j, \qquad (2)$$

де S_j - часткова сума, що формується у j-му циклі; q_j – поріг обробки у j-му циклі; b_j – кількість доданків у масиві j-го циклу, що містять поріг q_j; N – кількість циклів.

виході схеми нейрона свідчить про виконання умови (1). На рис. 1б показана схема, що моделює нейрон за виразом (3). Так само, як і алгоритми паралельного підсумовування з використанням РЗ, запропонований алгоритм порогової обробки масиву чисел (3) може бути реалізований на лінійному систолічному масиві [4].

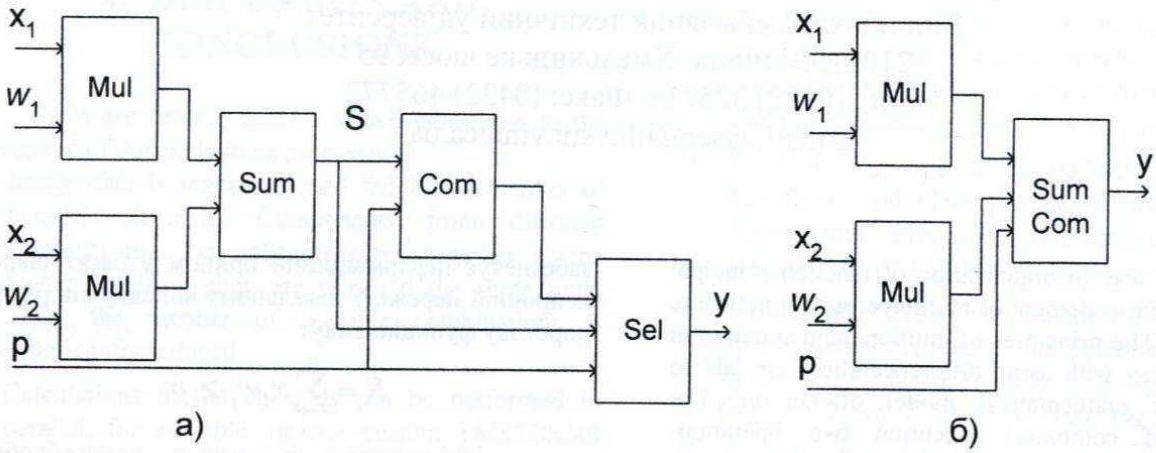


Рис. 1

Слушність виразу (2) доведена рядом теорем [2]. Серед особливостей цього методу необхідно відзначити його універсальність (як для аналогових, так і для цифрових даних), багатооперандність, а також перспективність реалізації на оптоелектронній елементній базі з використанням специфічного логіко-часового кодування інформації

Перевагою даного методу паралельної обробки інформації можна також вважати перспективність його реалізації на систолічній структурі.

Можливість представлення остаточного результату S як суми чаткових сум по всіх циклах обробки дозволяє виконати такі перетворення

$$p - (S_1 + S_2 + \ldots + S_N) = \\ = (\ldots((p - S_1) - S_2) - \ldots - S_N) \le 0. \qquad (3)$$

Таким чином, у кожному циклі обробки необхідно сформувати рзіницю вигляду

$$\Delta_j = \Delta_{j-1} - S_j, \qquad j = \overline{1,N} \qquad (4)$$

де $\Delta_0 = p$, і визначити виконання умови

$$\Delta_j \le 0.$$

Отже, можливе суміщення у часі виконання двох операцій протягом кожного циклу обробки, а саме, порівняння часткової суми S_j із порогом p (4) і формування часткової суми S_{j+1} у наступнму (j+1)-цилі. Процес обробки у такому випадку набуває вигляду конвеєрного обчислення як часткових сум S_j, так і величин Δ_j. Поява одиничного сигналу у на

Висновок

Особливості способу паралельного підсумовування масиву чисел на принципах РЗ дозволяють організувати для реалізації у нероних структурах порогову обробку даних, яка в деякій мірі використовує принципи, властиві нейрологічній обробці інформації.

Література

1. *Свечников С.В., Кожемяко В.П., Тимченко Л.И. Квазиимпульсно-потенциальные оптоэлектронные элементы и устройства логико-временного типа.- К.: Наук. Думка, 1987.-256 с.*
2. *Кожемяко В.П. и др. Параллельная обработка изображений.-Ужгород: Изд-во Ужгород. гос ун-та, 1993. – 89 с.*
3. *Parallel transformation / L. Timchenko, M. Grudin, T. Martynyuk, A. Kozhemyako // УсиМ. – 1999. - №5. С. 93-95.*
4. *Тимченко Л. И., Мартынюк Т.Б. , Загоруйко Л. В. Подход к организации многоуровневой схемы систолических вычислений // Электрон. моделирование. – 1988. №5. – С. 33 – 42.*

INCREASING THE RESOLUTION of SPACE FILTERS

I.M. Gvozdeva, A.S. Paevskiy

The Odessa Hhydrometeorology Institute, Ukraine, 270015,. Odessa,

Str. Lvovskaya, 15; fax(0482)63-63-08, 42-77-67

E-mail auto@ogmi.farlep.odessa.ua

The method of increasing the resolution of space filters with the multiplicative signals processing of receiving elements is considered. The synthesis algorithm of characteristics of space resolution of considered systems, based on their approximation with Chebyshev's polynomials, is developed. In the paper the processor structure of multiplicative processing of receiving elements signals of space filters is represented. The numerical example of synthesis of 8-element multiplicative array according to its linear prototipe is fulfilled.

INTRODUKCTION

The space filters are intended for obtaining the objects portraits in the different scientific fields: hydroacoustics, radiolocation, optics etc.

The modern space filters must possess the high space resolution with the purpose of detalization of investigated objects portraits and their classification. The usage of untraditional (more complex) methods of signals processing receiving elements permits to improve the characteristics of space resolution.

1. THE STRUCTURE SYNTHESIS of MULTIPLICATIVE SPACE FILTERS

The resolution of space filters is determined with resolution of their antennas, which can be realized in some cases as arrays [1,2].

The increase of array resolution with traditional linear processing of signals of receiving elements is possible only at the expense of increasing the physical dimension of array. The way of increasing the array resolution without increase of their physical dimensions, based on the usage of multiplicative processing of their elements signals, is known [3]. However the application of such method is difficult in a view of complexity of synthesizing the structure of processing channel according to the previous stated requirements to the function of space resolution of array. In the given paper the method of structural synthesis of multiplicative array, based on the describing of their output signals with a superposition of Chebyshev's polynomials and approximation in the same base of required function of space resolution.

In a general case the output signal of array has the following form[3,4]:

$$u(\theta) = \int_0^T x^T(t) A x(t) dt, \qquad (1)$$

where

$$x_i = X_{\max} \cdot \cos\left(\omega t + \frac{2\pi i d}{\lambda} \cdot \sin\theta\right), \qquad (2)$$

$$i = \overline{1, N}$$

N - number of receiving elements of multiplicative array,

$X_{\max}$ and ω - accordingly amplitude and angular frequency of output signals of receiving elements,

λ - wavelength of a received signal,

d – the distance between receiving elements,

© UkrObraz'2000

A – the coefficients matrix, defined with structure of processor of space filter,

T - time of an average, $T >> 2\pi/\omega$.

Let's consider the processor of space-time signals processing of array with variable transmission coefficients V_i on the each channel - fig. 1

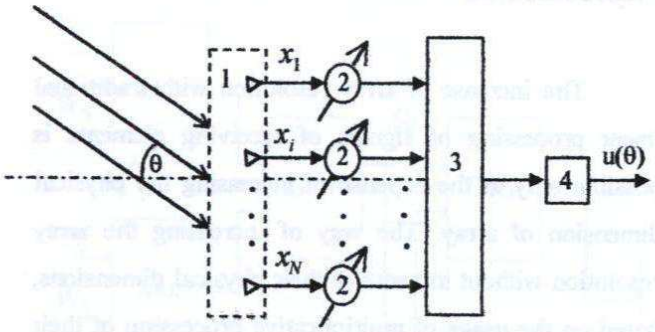


Fig. 1. The structure scheme of space filter with variable transmission coefficients of receiving channels.

In a figure: 1 - unit of receiving elements of array; 2 - units of transmission coefficients V_i, 3 - processor of multiplicative processing, 4 - integrator.

For selected structure of space-time signals processing the matrix A in (1) looks like; $A = \vec{V}\vec{V}^T$. Substituting (2) in (1) and fulfilled integration, we obtaine:

$$u(\theta) = X_{\max}^2 \sum_{m=0}^{N-1} \vec{V}^T E_{m+1} \vec{V} \cos 2\varphi_m = x_0^2 \sum_{m=0}^{N-1} b_m \cos 2\varphi_m, \quad (3)$$
$$m = \overline{0, N-1},$$

where E_k - matrix with 1 on k - lateral diagonal,

$\varphi_m = \frac{\pi m d}{\lambda} \sin\theta$ - relative phase shift,

$b_m = V^T E_{m+1} V$ - the weight coefficients at Chebyshev's polinomials.

Let's enter an intermediate variable: $y = \cos\pi \sin\theta$, and select the interelement distance $d = \lambda/2$. In this case the expression (3) will be transformed to the following form [4]:

$$u(\theta) = X_0^2 \sum_{m=0}^{N-1} b_m T_{2m}[y(\theta)], \quad (4)$$

where T_{2m} - Chebyshev 's polynomials of 2m order.

The expression (4) determines the describing of output signal $u(\theta)$ of multiplicative array with a superposition of Chebyshev's polynomials, that allows at the given transmission coefficients V_i to fulfill its analysis. For the solution of the problem of array synthesis the stated function of space resolution is approximated with piece of row of Chebyshev's polynomials, having the kind (4) with the given degree of approximation, and then the coefficients of amplitude distribution are determined [4]. For their obtaining it is necessary to solve the system of non-linear equations:

$$\left.\begin{array}{l} \vec{V}^T E_{m+1} \vec{V} = b_m \\ m = \overline{0, N-1} \end{array}\right\} \quad (5)$$

where the unknown coefficients are V_i. The solution of the system of non-linear algebraic equations of kind (5) represents itself the independent and enough difficult problem. For enough broads class of array the condition of symmetry of the characteristics of space resolution and, accordingly, of amplitude distribution of transfer coefficients concerning the phase center of array is satisfied:

$$V_1 = V_N; V_2 = V_{N-1}; \ldots V_k = V_{N-k-1} \cdots$$

The condition of symmetry allows to create the recurrent scheme of the solution of a system (5) in the following form:

$$
\left.
\begin{aligned}
&V_1 = \sqrt{b_{N-1}};\\
&V_2 = b_{N-2}\cdot(2V_1)^{-1};\\
&V_3 = (b_{N-3} - V_2^2)\cdot(2V_1)^{-1};\\
&V_4 = (b_{N-4} - 2V_2V_3)\cdot(2V_1)^{-1};\\
&\cdots\cdots\cdots\cdots\cdots\\
&-\textit{for even } k:\\
&V_k = \left(b_{N-k} - 2\sum_{i=1}^{\frac{k-2}{2}} V_{k-i}V_{i+1}\right)(2V_1)^{-1};\\
&-\textit{for odd } k:\\
&V_k = \left(b_{N-k} - 2\sum_{i=1}^{\frac{k-3}{2}} V_{k-i}V_{i+1} + V_{\frac{k+1}{2}}^{\,2}\right)(2V_1)^{-1};\\
&k = \overline{1, N/2}
\end{aligned}
\right\}, \quad (6)
$$

The further standardization of the solution of problem of space filters synthesis can be reached if the technique of synthesis of multiplicative array with according to the linear prototype will be used. At selected interelement distance $d = \lambda/2$ the elements number of the linear prototype accordingly is equal 2N+1, and the physical dimension of array is twice more, that determines the advantage of multiplicative signal processing.

2. NUMERICAL EXPERIMENT

For a numerical example we consider the synthesis of multiplicative 8-element array according to the linear prototype. As its prototype it is necessary to use 17-element linear array as the order of its polynomial coincides to the order of synthesized array. Let's consider the most complex case of multiplicative array synthesis according to the linear prototype, when the space resolution characteristic of linear array has the narrowest main lobe. For this case the amplitude distribution of the prototype is equal $b_m = 1$. Using the recurrent equation (6), we obtain:

$$V_8 = V_1 = 1$$
$$V_7 = V_2 = 1/2 = 0{,}5$$
$$V_6 = V_3 = 3/8 = 0{,}375$$
$$V_5 = V_4 = 5/16 = 0{,}3125$$

The processor structure of multiplicative space-time signals processing is determined with equation (5) and is shown in fig. 2.

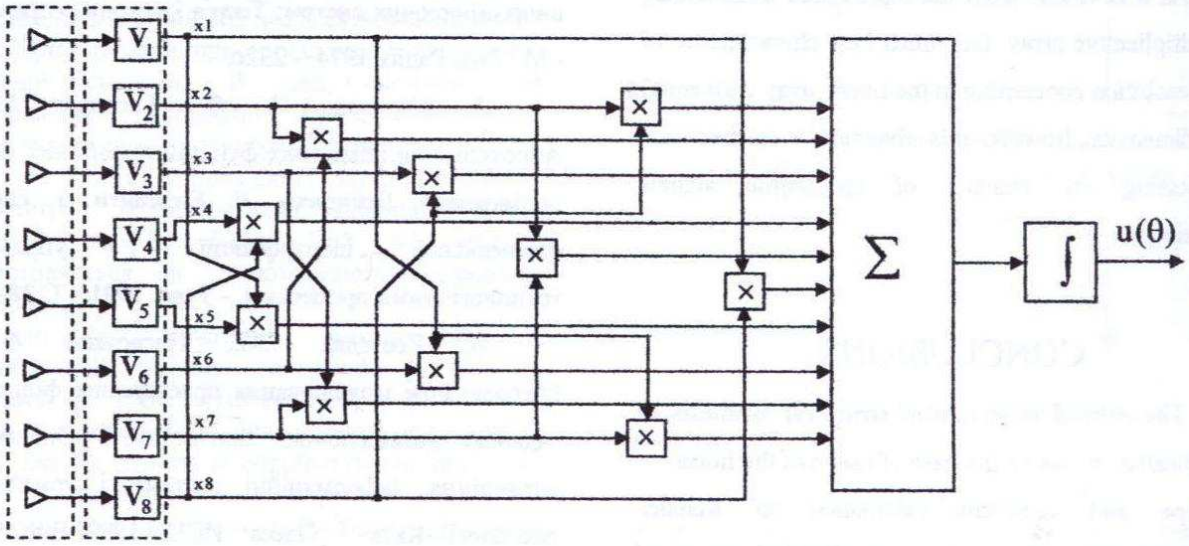


Fig.2. The scheme of the multiplicative processor of space-time signals processing of array.

For comparison in fig. 3 the function of space resolution of synthesized 8-element multiplicative array

according to its linear prototype and 8-element linear array are submitted.

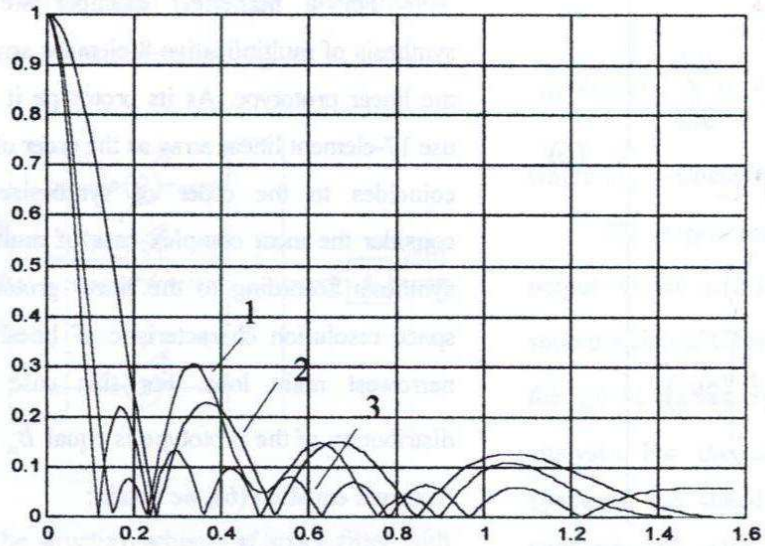


Fig. 3. The functions of space resolution:

1 - 8-element multiplicative array;

2 - 8-element linear array;

3 - 17-element linear array.

As it is visible from the represented illustrations, the multiplicative array, has much best characteristic of space resolution concerning to the linear array with equal linear dimension, however this advantage is reached with complicating the channel of space-time signals processing.

CONCLUSIONS

The offered technique of structural synthesis of multiplicative arrays on the base of usage of the linear prototype and recurrent calculation of transfer coefficients can to be used as an effective means of improvements of space resolution characteristics of space filters at solving the broad class of problems of radiolocation, hydroacoustics, optics, acoustics etc.

REFERENCES

1. Моизінго Р. А., Міллер Т. У. Адаптівні антені решітки: Вступ в теорію: Пер. з англ. - М.: Радіо і зв'язок, 1986. - 448с.

2. Бахрах Л. Д., Кременіцкий С. Д. Синтез випромінюючих систем: Теорія і методи розрахунку. - М.: Род. Радіо, 1974. - 232с.

3. Міргород В. Ф., Гвоздєва І. М. Апроксимація апаратних функцій нелінійних систем поліномами Чебишева // Елементи і системи оптимальної ідентифікації і управління технологічними процесами. - Тула. 1991. - С.28-23.

4. Гвозєдва І.М., Паєвський А. С. Математичне моделювання просторових фільтрів із заданою дозволяючою здатністю // Оптимізація управління, інформаційні системи і комп'ютерні технології.-Київ Одеса: ИСЦ, 1999.-Вип.1.-Ч.1.-256с.

Засоби пам'яті послідовної та паралельної дії на оптоелектронних пристроях.

Михайло Семенюк
Науково-впроваджувальне інноваційне підприємство "СТРУМ",
21001, м.Вінниця, вул. Коцюбинського 11, кв.108.
E-mail: ours_13@aport.ru

Рассматривается оптоэлектронное устройство на базе квантовых запоминающих элементов, которое может быть использовано для построения оптических вычислительных устройств с обработкой сигналов в десятичной системе счисления, и с непосредственной индикацией обрабатываемых данных.

The article deals with the optoelectronic device, which utilizes quantum memory chips. These devices can be used for creation of optical computers based on signal processing in the decimal calculation format, with direct indication of the data being processed.

ВСТУП

Оптимальний діалог людини з засобами обробки даних можливий якщо інформація представлена в звичній для нас десятковій системі.

Перші лічильні пристрої створювались на базі десяткової системи. В цих системах, для представлення інформації та її обчислення, використовувались механічні елементи на базі лічильних коліс (прикладом таких систем є арифмотетри, обчислювальна машиина АСКК, "МАРК-1", та ін. [1]). Цей же принцип використовується в різноманітних механічних лічильниках.

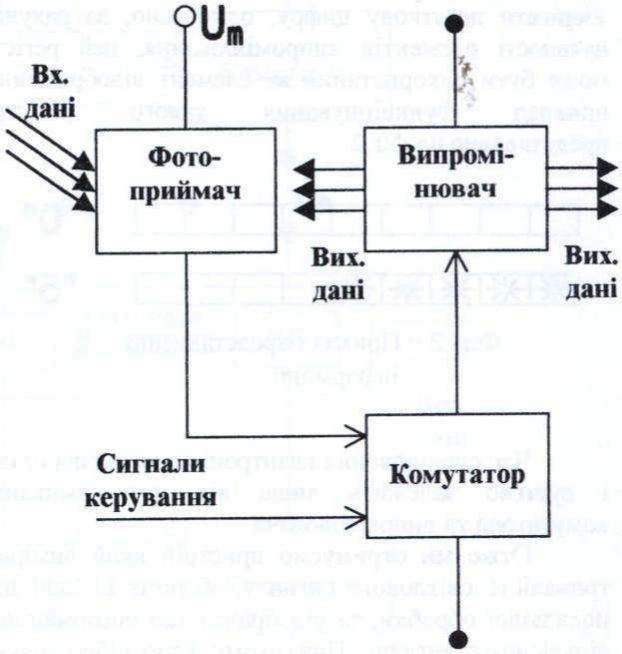
Але з початком використання в обчислювальній техніці електронних засобів від десяткової системи змушені були відмовитись в зв'язку з неможливістю знаходження існуючих на той час засобів пам'яті та обробки більш чим в двох стійких станах. Тому обробка даних в обчислювальних пристроях відбувається в двійковій системі.

Це приводить до необхідності перетворення інформації як на етапі її введення в засоби обробки, так і на етапі виведення. Тому, доцільним є побудова систем які оперують в десятковій системі счислення і не потребують подвійного перетворення даних.

Вирішення проблеми десяткового обчислення можуть дати засоби пам'яті на оптичних пристроях "квантронах" [2]. Побудовані на цих елементах пристрої можуть знаходитись в будь-якій кількості стійких станів, а також використовуватись для сприйняття та відображення інформації.

1. ОПТОЕЛЕКТРОННІ ЗАСОБИ ПАМ'ЯТІ ПОСЛІДОВНОЇ ДІЇ

Квантрон – це одинична комірка (фіг.1) яка складається з оптопари, приймача та випромінювача оптичних сигналів та комутатора електричних сигналів.



Фіг.1 – Квантрон

В якості фотоприймача може буди використаний будь який світлочутливий елемент, випромінювач представляє собою засіб перетворення електричної енергії в світлову, причому діапазон випромінювання (спектр) повинен збігатись з діапазоном сприймання фотоприймача. Комутатор використовується для підключення випромінювача до джерела енергії і може бути збудований на будь якому ключовому елементі.

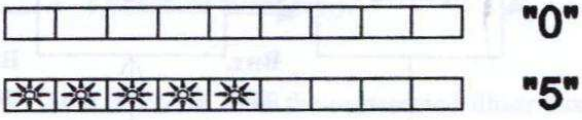
Під час впливу на фотоприймач світлового

сигналу і напруги активізації (збудження) $U_m=U_a$ цей елемент пропускає на комутатор струм достатній для його відкриття та, відповідно, вмикання випромінювача. Світлова енергія випромінювача поширюється в різних напрямках, при цьому відбувається опромінення як власного фотоприймача для підтримання струму необхідного для підтримання комутатора увімкненому стані, так і фотоприймача наступного квантрону для його збудження, а також індикація увімкненого стану.

Зняття вхідного світлового сигналу зменшує струм що надходить на комутатори до значення недостатнього до його відкриття. При цьому відбувається лише підтримання активних квантронів у збудженому стані. Такого ж результату досягають під час дії напруги фіксації $U_m=U_f$.

Таким чином, ті квантрони які були увімкнені, за рахунок зворотного світлового зв'язку залишаються у цьому стані, а нові не вмикаються. Для вимикання квантрону, на відповідний вхід комутатора подається імпульс гасіння. Комутатор закривається, струм через випромінюючий елемент перестає проходити і він гасне.

Якщо послідовно об'єднати дев'ять таких елементів, то отримаємо регістр який здатен зберігати десяткову цифру, одночасно, за рахунок наявності елементів випромінювання, цей регістр може бути використаний як елемент відображення, приклад функціонування такого регістру представлено на фіг.2.



Фіг. 2 – Приклад представлення інформації

Час спрацювання квантронів є величина стала, і суттєво залежить лише від часу вмикання комутатора та випромінювача..

Отже ми отримуємо пристрій який вимірює тривалість світлового сигналу, зберігає ці дані для подальшої обробки, та відображає цю інформацію в кількісному вигляді. При цьому непотрібно ніяких допоміжних засобів для виділення початку і кінця часового інтервалу, аналого-цифрового перетворення вхідних даних.

Відсутність методичних похибок під час обробки інформації досягається за рахунок відсутності не доцільних перетворень і однакового часу спрацювання квантронів [3].

2. ОПТОЕЛЕКТРОННІ ЗАСОБИ ПАМ'ЯТІ ПАРАЛЕЛЬНОЇ ДІЇ

В розглянутому вище методі використовується властивість квантронів до перемикання в точно фіксовані відліки часу. При цьому комутатор використовуються для підключення до джерела енергії, необхідної для переведення випромінювача у збуджений стан.

Але, крім сигналів потужність яких визначається тривалістю, існує також велика кількість сигналів в яких енергетичною характеристикою є величина випромінювання, або спектр частот.

Для обробки таких сигналів також можливе використання квантронів.

У випадку, коли значення інформації залежить від освітленості, в конструкції квантрону доцільнно використовувати два світлочутливих елемента (фіг.3):

– для створення напруги збудження (фотоприймач);

– для отримання напруги підтримування випромінюючого елементи в активному стані (фотоелемент оптопари – ФЕ).

На фіг.3 зображено систему з'єднання лінійки квантронів для створення десяткового регістра паралельної дії.

Вхідний сигнал подається одночасно на фотоприймаючі елементи всіх квантронів. Ці елементи через низку опорів, величина яких однакова і визначається дискретизацією вхідного сигналу, підключається до входу "U_m". Так як, для відкриття комутатора значення струму на його вході повинно досягнути відповідної величини, то, необхідно щоб падіння напруги на опорі "R" було скомпенсоване величиною фотоструму від опромінення фотоприймача вхідним сигналом.

Кількість збуджених елементів (квантронів) визначає величину вхідного сигналу.

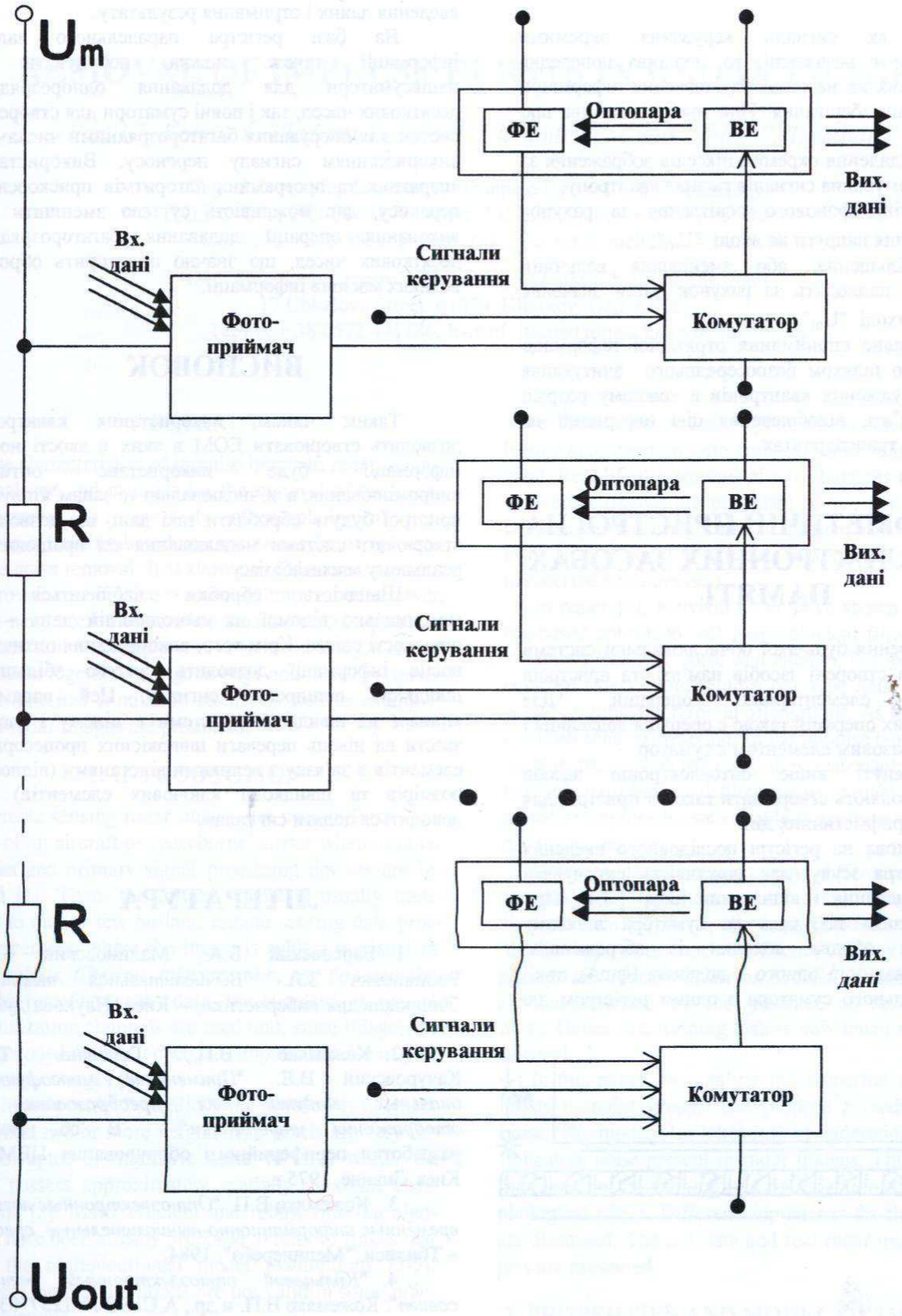
Після припинення дії на фотоприймачі зовнішнього випромінювання, підтримання горіння досягається за рахунок наявності зворотного зв'язку, тобто, опромінення фоточутливих елементів оптопари.

Для нарощування системи (тобто збільшення розрядності представлення інформації) використовується каскадне з'єднання регістрів, при цьому вихід "U_{out}" попереднього розряду з'єднується з входом "U_m" наступного розряду.

Таким чином, отримуємо перетворювач величини сили світла в цифровий код.

Час спрацювання регістрів (фіксації інформації що надходить) не залежить від кількості підключених регістрів, так як вхідний сигнал одночасно надходить до всіх фотоприймачів, а зміна величини сигналу на вході "U_m" відбувається при сталому значенні на інформаційному вході. Цей час дорівнює тривалості спрацювання одного квантрона.

При побудові квантронів (а отже і регістрів) використовуються елементи у яких відсутній "ефект пам'яті", тобто тривалість надходження сигналу не веде до накопичення величини цього сигналу і, отже,



Фіг.3 – Десятковий регістр паралельної дії на квантронах

до хибного сприйняття вхідного сигналу.

Сумарна потужність випромінювання всіх збуджених елементів дорівнює величині вхідного

сигналу, тобто можливе використання значення прийнятих даних для подальшої обробки, відображення, або передачі їх на інші засоби

обробки.

Так як сигнали керування окремими квантронами є незалежні, то можлива попередня обробка даних як на етапі надходження інформації, так і на етапі зберігання. При цьому можливі такі операції:

– видалення окремих пікселів зображення за рахунок формування сигналів гасіння квантрону;

– зміни фонового освітлення за рахунок зміни значення напруги на вході "U_m";

– збільшення, або зменшення величини сигналу що надходить за рахунок зміни значення напруги на вході "U_m".

Візуальне сприймання отриманої інформації можливе або шляхом безпосереднього зчитування кількості збуджених квантронів в кожному розряді масиву пам'яті, відображення цієї інформації на спеціальних транспарантах.

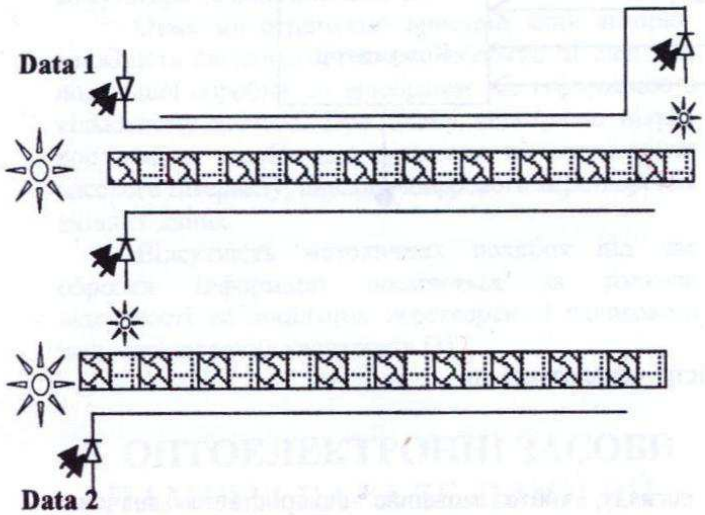
3. АРИФМЕТИЧНІ ПРИСТРОЇ НА ОПТОЕЛЕКТРОННИХ ЗАСОБАХ ПАМЯТІ

Створення будь-якої обчислювальної системи базується на створені засобів пам'яті, та пристроїв виконання елементарних операцій. Для арифметичних операцій такою є операція додавання і відповідно базовим елементом є суматор.

Розглянуті вище оптоелектронні засоби пам'яті дозволяють створювати також і пристрої для виконання арифметичних дій.

Побудова на регістрі послідовного введення даних регістра зсуву, дає можливість виконувати операції додавання і віднімання цифр [3,4]. При цьому можлива побудова як суматора в якому зберігаються обидва доданки із збереженням результату на місці одного з доданків (фіг.4), так і накопичувального суматора з одним регістром для введення даних і отримання результату.

На базі регістра паралельного запису інформації також можна побудувати як напівсуматори для додавання однорозрядних десяткових чисел, так і повні суматори для створення систем для оперування багаторозрядними числами з використанням сигналу переносу. Використання апаратних та програмних алгоритмів прискореного переносу, дає можливість суттєво зменшити час виконання операції додавання багаторозрядних десяткових чисел, що значно прискорить обробку великих масивів інформації.



Фіг.4 – Суматор на регістрі послідовної дії

ВИСНОВОК

Таким чином, використання квантронів дозволить створювати ЕОМ в яких в якості носіїв інформації буде використане оптичне випромінювання, а обчислювальні та запам'ятовуючі пристрої будуть обробляти такі дані, що дозволить створювати системи моделювання які працюють в реальному масштабі часу.

Швидкість обробки наблизиться до максимально відомої на сьогоднішній день – до швидкості світла. Крім того, використання оптичних носіїв інформації дозволить суттєво збільшити швидкість поширення сигналу. Цей параметр впливає на швидкодію системи в цілому і здатен звести на нівець переваги швидкісних процесорних елементів в зв'язку з великими відстанями (відносно розмірів та швидкодії ключових елементів) які доводиться долати сигналам.

ЛІТЕРАТУРА

1. Барковский Б.А., Малиновский Б.Н., Рабиинович З.Л. *"Вычислительная техника"*. Энциклопедия кибернетики, – Киев, Наукова думка, 1974 г.

2. Кожемяко В.П., Головань Т.В., Качуровский В.Е. *"Применение многофункциональных модулей для преобразования и отображения информации"*. В сб. Новые разработки периферийного оборудования ЦВМ, – Киев, Знание, 1975 г.

3. Кожемяко В.П. *"Оптоэлектронные логико-временные информационно-вычислительные среды"*. – Тбилиси, "Мецниереба", 1984.

4. *"Кольцевой оптоэлектронный регистр сдвига"*. Кожемяко В.П. и др., A.C. SU № 1257703 A1

REMOVAL OF IMPULSE BURSTS IN RADAR IMAGES

Pertti T. Koivisto[†], *Vladimir V. Lukin*[‡], *Vladimir P. Melnik*[†], *Oleg V. Tsymbal*[‡]

[†]Signal Processing Laboratory, Tampere University of Technology
P.O.Box-553, FIN-33101, Tampere, FINLAND, Tel. +358 3 365 3860,
Fax +358 3 365 3857, E-mail: peko@cs.tut.fi, melnik@cs.tut.fi

[‡]Dept 504, State Aerospace University (Kharkov Aviation Institute)
17 Chkalova Street, 61070, Kharkov, UKRAINE,
Tel/fax + 38 0572 441186, E-mail: lukin@mmds.kharkov.ua

ABSTRACT

The characteristics of impulse bursts in radar images are analyzed and methods for the removal of the bursts are discussed. As a case study, soft morphological filters utilizing a training-based optimization scheme are used for the noise removal. It is shown that the methods used can provide an effective removal of impulse bursts even if two neighboring rows of an image are corrupted. At the same time the multiplicative noise in images is also suppressed together with good edge and detail preservation. Numerical simulation results as well as examples of real radar images are presented.

1. INTRODUCTION

Remote sensing radar images are usually formed on board of an aircraft or spaceborne carrier where a radar antenna and primary signal processing devices are installed [1]. Then, the formed image is usually transferred to one or few on-land remote sensing data processing centers, where the image is subject to visualization, analysis, filtering, interpretation, etc. For transferring the remote sensing data, the standard or special communication channels are used and, since images are often encoded and then decoded, impulsive noise may be observed in images [2, 3].

In many practical situations the probability of spikes is low and two or more neighboring pixels are very seldom corrupted by impulsive noise. In other words, the spikes possess approximately spatially invariant characteristic [3]. Many efficient and robust filtering algorithms have been already proposed to remove spikes that fulfill the aforementioned model assumptions [3-5]. However, these assumptions are not valid in some practical situations.

In particular, the interference can be long term and so intensive that it corrupts several consecutive image pixels in one or more rows following each other. Such situations may happen, e.g., if the receiver input and circuitry are not well protected against intensive interference or if in the neighborhood of the remote sensing data processing center there are some electromagnetic wave irradiation sources operating in the frequency band which overlaps with communication channel wave band. Real life illustrations of what happens in this case with radar images are presented below. (In this paper, we assume that images are transferred rowwise. Naturally, similar methods can also be applied if images are transferred columnwise.)

In paper [6], Koivisto *et. al.* have shown that training-based optimized soft morphological filters are able to remove line-type noise well enough. For example, the obtained filters could remove line-type noise with horizontal or almost horizontal orientations. As this noise in a certain sense corresponds to the impulse bursts that we have in the considered case, it is reasonable to expect that soft morphological filters being trained for the removal of impulse bursts are able to perform well for radar image recovery.

However, there are also some differences between the task considered here and the design task studied in the reference [6]. First, the impulse bursts differ from the line-type noise since the former one has more complicated and random behavior. Second, the radar images are characterized by the presence of multiplicative noise. Hence, the training task is now much more complicated.

In this paper, we analyze the properties of impulse bursts in radar images and propose a model for this noise. The model also takes into consideration the multiplicative noise present in radar images. This model is then used in the training and optimization of soft morphological filters. Different approaches for the training are discussed. The test data and real radar image examples are presented.

2. PROPERTIES AND MODEL OF IMPULSE BURSTS IN RADAR IMAGES

To get an idea what the impulse bursts are, let us first analyze a radar image that was received by the Center of Earth Remote Sensing (Kharkov, Ukraine) via telecommunication channel from Ukrainian satellite "Sich-1." A 256x219 part of the radar image is presented in Fig. 1a. It can be seen that several fragments

© UkrObraz'2000

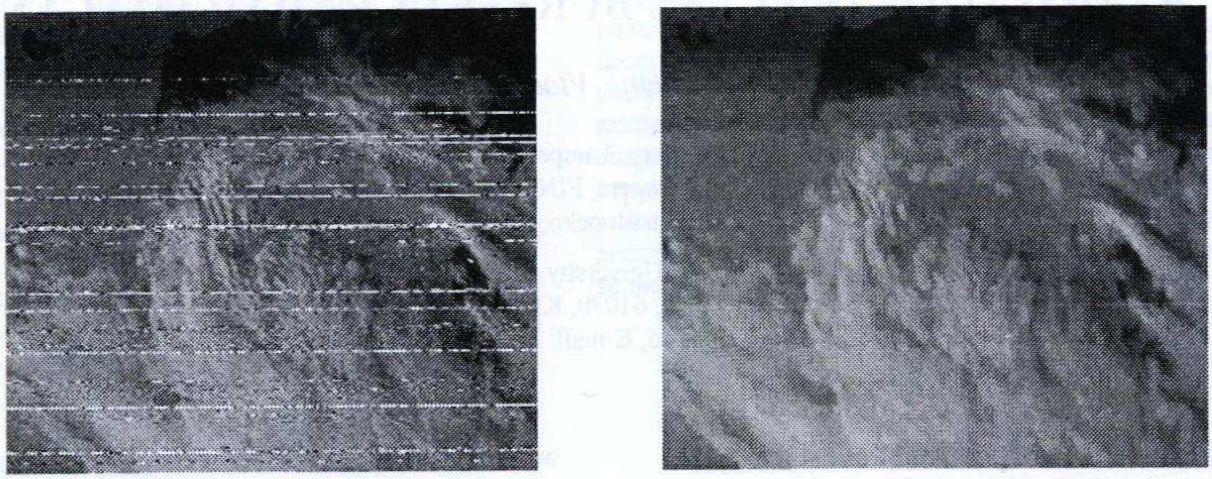


Fig. 1. (a) A part of the original radar image and (b) the part filtered by the proposed soft morphological filter (see text).

in many rows are corrupted by impulse bursts, and the lengths of such fragments are rather different. Moreover, sometimes such fragments occur in two consecutive rows. It can also be observed visually that in some pixels of the considered fragments the values are maximal (i.e., 255 in the 8-bit representation used) while some pixel values differ from 255 but still remain "impulsive" with respect to the values that can be predicted for the radar image from its local analysis.

As we use training-based design method, we need a corrupted and a true (desired) training image. Since it is impossible to have such uncorrupted real radar image, artificial test images must be used.

In order to make an adequate model of a test radar image, we studied the properties of the real image in more details. First, the percentage of pixels corrupted by impulse bursts was calculated. The percentage was approximately 5%. Then, the statistical characteristics of burst lengths were estimated. Finally, the behavior of the bursts was studied row by row for the rows containing bursts. It was found that the mean of the burst fragment values was approximately 180, the bursts contained a quasisinusoidal component and a noisy component. Besides, cut-off effects were observed, that is, there may be several pixels in a row having values equal to 255. The multiplicative noise is also present in the real radar image. The probability density function of this noise is Gaussian with 1 as the mean. The evaluated relative variance is approximately 0.05.

All aforementioned properties of the impulse bursts and the multiplicative noise have been taken into account when generating the noise model for the test image. Since images are transferred as one-dimensional arrays the noise model is also presented for one-dimensional array. More precisely, our noise model is now the following.

First, a Markov-chain with two states is used to determine which samples belong to impulse bursts. The transition probability from "no-burst-state" to "burst-state" is p and the transition probability from "burst-state" to "no-burst-state" is q.

If a sample does not belong to an impulse burst then it is corrupted by the aforementioned multiplicative noise in the usual way. On the other hand, if a sample belongs to an impulse burst then the (corrupted) sample value f_j is obtained using the formula

$$f_j = \text{round}\{\alpha + \beta \sin[(j-k)\,\omega + \varphi_k] + \xi_j\},$$

where k denotes the index of the leftmost sample in the burst in question, α is the average level of impulsive noise, β and ω are the amplitude and frequency of the harmonic component of the burst, respectively, φ_k denotes the phase of the harmonic component of the burst, and ξ_j is the fluctuating noise component of the burst. The phase $\varphi_k \in]-\pi,\pi]$ is a random variable with uniform distribution and the noise component ξ_j is a random variable with Gaussian probability density function with zero mean and standard deviation σ. Hence, α, β, and ω are common to all bursts, φ_k is common to all pixels in some burst, and ξ_j varies from pixel to pixel. Rounding is to the nearest non-negative integer less than or equal to 255.

A test image of size 256x256 corrupted by multiplicative noise and impulse burst is presented in Fig. 2a. The estimated parameter values for the noise model were $p = 0.9993$, $q = 0.989$, $\alpha = 200$, $\beta = 80$, $\omega = 0.7$, and $\sigma = 30$. Then, the percentage of the pixels belonging to the bursts was approximately 5%. The test image contains homogeneous regions, large size objects with different shapes, and small size objects also having different shapes, contrasts and orientations. Besides this noisy test image, we used a noise free test image, a test image corrupted by multiplicative noise only, a test image corrupted by impulse bursts only, and a test image corrupted by impulse burst and negative impulsive noise (i.e., here spikes with amplitude 0). Those test images

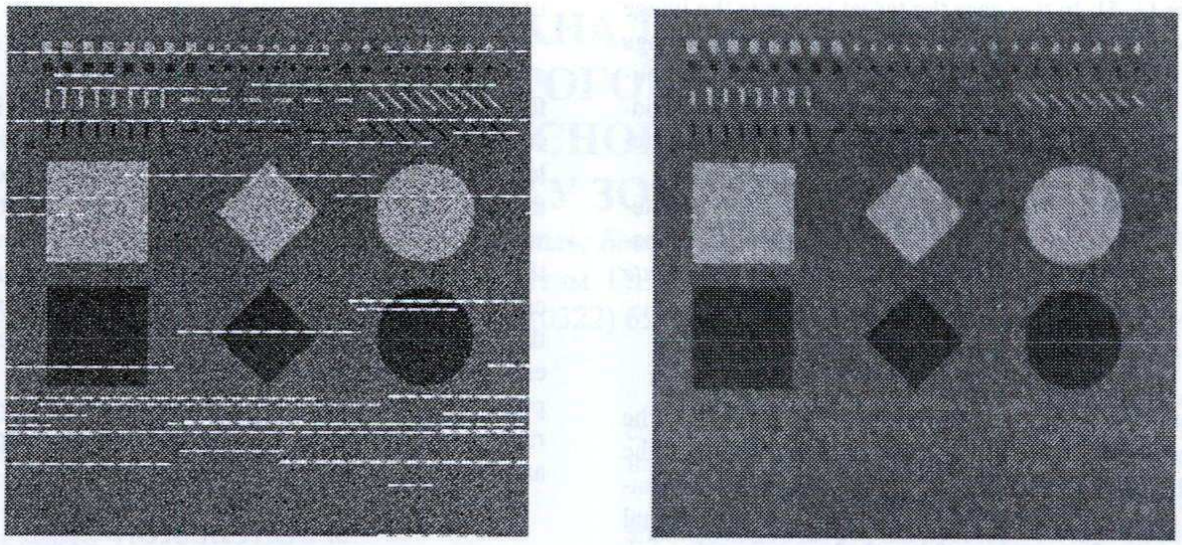


Fig. 2. (a) Original test image with multiplicative noise and impulse bursts and (b) its filtered counterpart (see text).

were needed to study different approaches to the training of the soft morphological filters (see Section 4).

Our desire was also to check whether the soft morphological filters destroy many details while removing the impulse bursts. By comparing the images in Fig. 1a and Fig. 2a, one can see that the structure and general properties of the images are similar enough also for this purpose.

3. SOFT MORPHOLOGICAL FILTERS

Soft morphological filters [7] form a class of stack filters and were introduced to improve the behavior of standard flat morphological filters in noisy conditions. They have many desirable properties, e.g., they can be designed to preserve details well [8]. In addition, they are suitable for impulsive or heavy-tailed noise.

The two basic soft morphological operations *are soft erosion* and *soft dilation*. Based on these operations, two compound operations, *soft opening* and *soft closing*, can be defined in the usual way. As we consider is this paper images only, the definitions are also given in two-dimensional signal space.

Definition 1. The *structuring system* $[B,A,r]$ consists of three parameters, finite sets A and B, $A \subseteq B \neq \emptyset$, in $\mathbf{Z}^2$ and an integer r satisfying $1 \leq r \leq \max\{1, |B \setminus A|\}$. The set B is called the *structuring set*, A its (hard) *center*, $B \setminus A$ its (soft) *boundary*, and r the *order index* of its center or the *repetition parameter*.

The *translated set* T_x, where the set T is translated by x, $x \in \mathbf{Z}^2$, is defined by $T_x = \{x+t : t \in T\}$. A *multiset* is a collection of objects, where the repetition of objects is allowed. For example, $\{1,1,1,2,3,3\} = \{3 \diamond 1, 2, 2 \diamond 3\}$ is a multiset.

Soft morphological operations transform a signal $f: \mathbf{Z}^2 \rightarrow \mathrm{R}$ to another signal by the following rules.

Definition 2. The *soft erosion* (*soft dilation*)} of f by the structuring system $[B,A,r]$ is denoted by $f \ominus [B,A,r]$ ($f \oplus [B,A,r]$), where $f \ominus [B,A,r](t)$ ($f \oplus [B,A,r](t)$) is the r'th smallest (largest) value of the multiset $\{r \diamond f(a) : a \in A_t\} \cup \{f(b) : b \in (B \setminus A)_t\}$.

A *finite composition* of length p of basic soft morphological operations is simply a sequence of p soft erosions and/or dilations, where each operation can have different structuring system. Naturally, each operation is applied to the result of the previous operation in the sequence.

4. OPTIMIZATION

Although there are no analytical criteria for deciding which soft morphological operation (and with which parameters) is the best for some situation, a suitable operation sequence and its parameters can be found using supervised learning methods, e.g., simulated annealing and genetic algorithms [6]. Of course, some training set, for which the desired output is known, is needed.

The optimization methods presented in [6] allow one to handle the impulse bursts in several ways. First, it is possible to search the optimal filter using the noise-free test image as the target image and the image corrupted by impulse bursts and multiplicative (or negative impulsive) noise as the source image.

Another possibility is to use structural constraints, in which case the target image is the noise-free image and the source image is the image corrupted by multiplicative noise. The impulse bursts are presented as constraints and the optimal filter is sought, provided that the impulse bursts are removed (totally or at least to some extent).

A third possibility is to optimize the soft morphological operations to remove impulse bursts only. Then, at the second stage, the multiplicative noise can be sup-

pressed by some conventional technique suited for this purpose [4, 5]. In this case the target image is the image corrupted by multiplicative noise and the source image is the image corrupted by impulse bursts and multiplicative noise. However, this third possibility is not studied in this paper but it is left as a topic for future research.

As the error criterion, it is possible to use any criterion that can be calculated using two images as parameters. In this paper, we have used the *mean absolute error* (MAE) and the *mean square error* (MSE).

5. EXPERIMENTAL RESULTS

The experimental tests in this paper are based on the following test case. The training image pairs are the ones discussed in Sections 2 and 4. The application image is the one shown in Fig. 1a. In each case the optimal composite soft operation of length 2 were sought with overall dimensions 3x3, 3x5 (i.e., 3 columns and 5 rows), and 5x5 (i.e., the structuring sets were inside the 3x3, 3x5, and 5x5 rectangles). The length two was selected since now our noisy image contains both positive and negative impulsive noise and a single soft operation is not able to remove two-sided noise. On the other hand, as the experiments showed, two consecutive soft operations are already powerful enough for our purposes.

When the 3x3 window was used, the optimal filters were not able to remove impulse bursts sufficiently. On the other hand, the filters optimal inside the 3x5 and 5x5 windows were already able to remove almost all of the bursts. Hence, the quality of these filters depended on their ability to remove multiplicative noise and preserve details.

The optimal filter sequence was usually soft erosion followed by soft dilation. This is also natural since the noise was mostly positive. However, the results achieved by the optimal soft openings (i.e., the soft erosion followed by the soft dilation with the symmetric structuring set) were usually almost as good as those achieved with the optimal composite soft operations of length 2. This is important since the optimization of soft openings is much easier than the optimization of composite soft operations of length 2.

The error criterion used (i.e., the MAE or the MSE) did not seem to have crucial effect in the optimization. However, the filters optimized under the MSE produced visually better results although, in general, the differences were small.

When comparing the optimization schemes, we noticed that visually the best results were achieved with the noise-free test image as the target image and the image corrupted by impulse bursts and multiplicative (or negative impulsive) noise as the source image. The results obtained using structural constraints were not as good but, however, of good quality. Moreover, there did not seem to be much difference between the cases in which multiplicative or negative impulsive noise was used together with impulse bursts.

As an example of the results of the method, Fig. 1b and Fig. 2b show two images that are filtered by the found optimal filters (the same filter sequence in both cases). The optimal filter sequence is soft erosion followed by soft dilation and the sequence was found under the MSE and inside the 3x5 window. The target image was the noise-free test image and the source image was the test image corrupted by impulse bursts and multiplicative noise (i.e., the image in Fig. 2a). As can be seen, the image in Fig. 2b is a little blurred and some small, especially horizontal, details are lost. However, all impulse bursts are removed. The same also holds for the radar image in Fig. 1b; impulse bursts have disappeared and small distortion has appeared.

6. CONCLUSION

The characteristics of impulse bursts in radar images were analyzed and a method for the removal of the bursts was discussed. It was shown through experiments that the presented method can remove impulse bursts and multiplicative noise with efficiency and at the same time preserve details well. Although our test case was quite limited the method can easily be applied to other cases as well.

REFERENCES

[1] *Remote Sensing Digital Image Analysis. An Introduction*, Edited by J.A. Richards, Springer Verlag, Berlin, 1994.

[2] K.-H. Szekielda, *Satellite Monitoring of the Earth*, Wiley, New York, 1989.

[3] J. Astola and P. Kuosmanen, *Fundamentals of Nonlinear Digital Filtering*, CRC Press, Boca Raton, Florida, 1997.

[4] I. Pitas and A.N. Venetsanopoulos, *Nonlinear Digital Filters: Principles and Applications*, Kluwer Academic Publishers, Boston, 1990.

[5] V.P. Melnik, Nonlinear Locally Adaptive Techniques for Image Filtering and Restoration in Mixed Noise Environment, Dr. Tech. Thesis, Tampere University of Technology, Tampere, Finland, March 2000.

[6] P. Koivisto, H. Huttunen, and P. Kuosmanen, "Training Based Optimization of Soft Morphological Filters," *Journal of Electronic Imaging*, vol. 5, pp. 300-322, June 1996.

[7] L. Koskinen and J. Astola, "Soft morphological filters: A robust morphological filtering method," *Journal of Electronic Imaging*, vol. 3, pp. 60-70, January 1994.

[8] P. Kuosmanen and J. Astola, "Soft morphological filtering," *Journal of Mathematical Imaging and Vision*, vol. 5, pp. 231-262, September 1995.

ВИБІР ВЕЙВЛЕТУ ТА АНАЛІЗ ВПЛИВУ СТЕПЕНІ РЕГУЛЯРНОСТІ БІОРТОГОНАЛЬНИХ ФІЛЬТРІВ, ПОБУДОВАНИХ НА ОСНОВІ ТЕОРІЇ ПОЛЮСІВ, ДЛЯ СТИСКУ ЗОБРАЖЕНЬ

Олексій Коваль, Богдан Русин

Фізико-механічний інститут ім. Г.В.Карпенка НАН України

290601б м. Львів, вул. Наукова 5, тел. (0322) 65-45-30, e-mail: dep32@ah.ipm.lviv.ua

Обгрунтовується вибір типу вейвлета, пропонується метод оцінки його ефективності для задач компресії зображень та аналізується вплив степені регулярності вейвлетоутворюючих фільтрів на якість стиску .

1. ВСТУП

Прооблема ефективного зберігання візуальної інформації існує в багатьох галузях науки і техніки, починаючи від неруйнівного контролю, баз даних охоронних систем, матеріалознавсва і закінчуючи медичною томографією.

Розв'язок цієї проблеми шукався у наступних напрямках:

- побудові алгоритмів та систем за безвтратним принципом, для яких характерна відсутність втрат якості відтвореного зображення у порівнянні з оригіналом;
- розробці алгоритмів та системи компресії із втратами, яким притаманне внесення спотворень в процесі відновлення стиснутого зображення.
- комбінованому підході, який використовує моделювання зображення та врахування втрат, внесених моделлю у порівнянні з вхідними данними.

Останній є найбільш перспективним, оскільки не тільки дозволяє досягати підвищеної степені якості відтвореного зображення, але й передбачає можливість стиску без втрат ii з ефективністю, яка є для такого типу компресії найвищою [7].

2. ВИБІР ВЕЙВЛЕТА

Очевидно, що ефективність всього процесу стиску при таких методах в значній мірі визначається виборм моделі опису зображення.

Якщо припустити, що воно представляється у вигляді кусково-гладкої функції, то, існує можливість застосування кусково-поліноміальної моделі. Вейвлетне перетворення є реалізацією такого підходу, причому при апроксимації використовується мала кількість базисних елементів.

Вибір вейвлета для застосування в задачах чах стиску зображень є актуальною на сьогодні проблемою [1].

Перш за все варто відмітити, що з точки зору специфіки задачі, що розглядається, бажаною є наявність у вейвлета наступних властивостей [4]:

- ортогональність, яка забезпечує найкращу апроксимацію даних, які перетворюються в сенсі L_2;
- наявність осі симетрії базисної функції, що забезпечить відсутність фазових спотворень, внесених при перетворенні;
- обмежена область визначення вейвлета, що зробить скінченним границі сумування при реалізації дискретної версії вейвлетного перетворення;
- наявність в імпульсній характеристиці фільтрів, що генерує вейвлет, цілочисельних коефіцієнтів, що спрощує реалізацію перетворення.

Оскільки всі властивості, крім першої, можуть бути отримані у випадку біортогонального вейвлетного пертврорення, то, переважно, саме цей тип вейвлетів на сьогодні є таким, що найбільш широко застосовується в області стиску зображень.

Виходячи з того, що існує достатньо велика кількість вейвлетів такого типу, причому постійно зростаюча (зокрема, розроблені згідно [8,9] пари біортогональних фільтрів), то проблема оцінки ефективності застосування того чи іншого з них до стиску зображень є достатньо складною.

В роботі [6] пропонується визначати ефетивність застосування біортогональних вейвлетів для компресії зображень на основі аналізу імпульсної та перехідної характеристики лінійної, неінваріантної до зсуву системи, якою моделюється процес дискретного біортогонального вейвлетного перетво-рення, відновлення вхідних даних при невеликій кількості кроків аналізу-синтезу та неврахуванні детальної інформації.

В якості характеристики використовусться мінімальне співвідношення між центральним та першим боковим пелюстком імпульсної характеристики системи та середнє значення другого піка перехідної характеристики. Мала величина першого парметра та велике значення другого приводить до появи артефактів при відновленні.

Для перевірки адекватності розробленого критерію пропонується використати алгоритм стиску, який включає трирівневе вейвлетне перетво-

рення, адаптивне квантування отриманих коефіцієнтів та кодування результатів квантування методами RLC та методом Хаффмана.

На основі розробленої методики та цього алгоритму стиску авторами [6] вибрано шість найкращих з їхньої точки зору пар біортогональних фільтрів.

Як відзначається в [1], використана в [6] методика квантування є неефективною для випадку біортогонального вейвлетного перетворення.

Також відомі спроби характеристики якості фільтрів для задач компресії, зокрема, на основі степені їх регулярності [2,3].

Існує думка, що однозначного зв'язку між степінню регулярності та якістю відновленого зображення на сьогодні ще не встановлено [5].

В даній роботі буде зроблено спробу проаналізувати вплив регулярності фільтрів, побудованих на основі методики, приведеної в [8,9], на якість відновленого зображення.

3. МЕТОДИКА ПОБУДОВИ ФІЛЬТРІВ

В [8,9] представлено методику та результати побудови біортогональних вейвлетів на основі теорії полюсів [10].

Вона полягає в наступному:

- побудові інтерполяційного виразу на основі теорії полюсів, вагові коефіцієнти якого з точки зору теорії сигналів можуть бути розглянуті як відліки імпульсної характеристики деякого фільтра нижніх частот;

- обрахунку імпульсної характеристики дуального низькочастотного фільтру з використанням відомого з теорії банків фільтрів ідеального відновлення обмеження [6]:

$$H_0(z)H_1(-z) - H_0(-z)H_1(z) = P(z) - P(-z) = c \cdot z^{-(2l-1)}$$

де $H_0(z), H_1(-z)$ - z-образи низькочастотних фільтрів аналізу та синтезу, $l \in Z$.

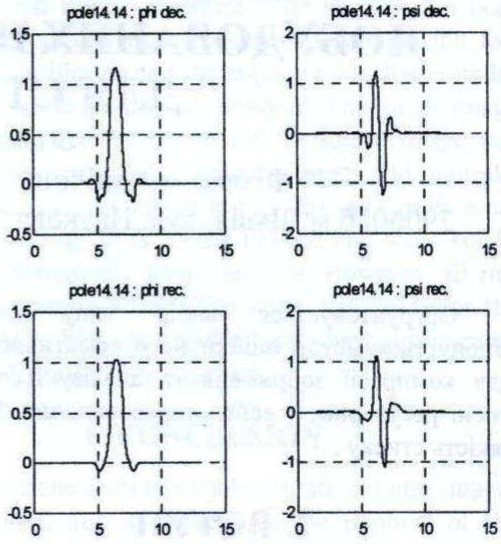
- аналізі границі збіжності імпульсної характеристики еквівалентного фільтру, який утворюється в результаті застосування каскадного алгоритму до одного з низькочастотних фільтрів.

У випадку існування негладкої границі застосовується процедура підвищення регулярності, яка полягає у збільшенні в функції *H(z)* (z-образ імпульсної характеристики того фільтру, який утворює еквівалентний фільтр без гладкої границі) кількості нулів в точці $z = -1$. В цьому випадку випадку виникає повторна необхідність визначення імпульсної характеристики другого фільтра за згаданою вище методикою.

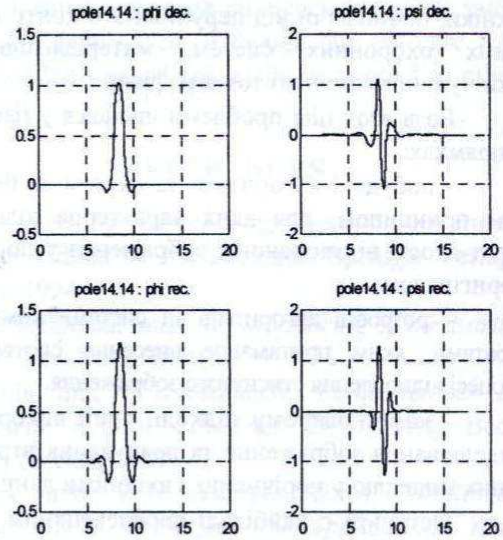
Оцінка степені гладкості отриманих граничних функцій проводилась згідно з розробленим в [2,3] підходом, що грунтується на визначенні регулярності Хелдера.

Тобто, дана методика дозволяє гнучко змінювати степінь регулярності в залежності від умов поставленої задачі.

Результати застосування представленого підходу подані на рис.1,а.



а)



б)

Рис.1. Вигляд $\phi, \psi, \tilde{\varphi}, \tilde{\psi}$ - функцій для отриманої в [9] та скоректованої пари фільтрів (позиції, відповідно, 7 та 8 в табл.1).

4. МЕТОД ОЦІНКИ

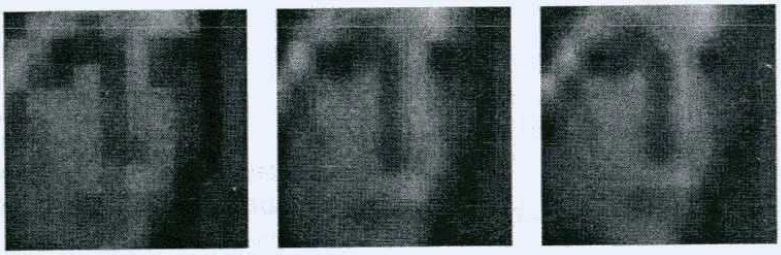
Оскількм результати квантування суттєво впливають на ефективність алгоритма в цілому, то пропонується ефективність застосування біортогонального базису для компресії зображень аналізувати на основі трирівневого вейвлетного перетворення зображення із подальшим його відновленням при умові рівності нулю всіх детальних коефіцієнтів.

Запропонований підхід дозволить вибрати таку пару біортогональних фільтрів, яка б дозволила зосереджувати максимальну частину інформації в коефіцієнтах апроксимації.



а) Lena б) Barbara в) Airport

Рис. 2. Використані тестові зображення .



а) б) в)

Рис.3. Фрагмент відновленого зображення при використанні для стиску пари фільтрів, відповідні біортогональні функція до яких зображені а) на рис. 1, а, б) на рис. 1, б та популярної пари фільтрів 9/7 (таблиця 1(1))

Таблиця 1. Значення коефіцієнтів фільтрів, показники степені регулярності та співвідношення сигнал/шум (С/Ш), dB, для фільтрів, розроблених в [9] та вибраних в [6] як найкращі для компресії зображень

пп	Значення коефіцієнтів		$N+\alpha_N^j$	С/Ш dB
1	2		3	4
1	$H_0(z)$	0.037828 -0.023849 -0.110624 0.377402 0.852699 0.377402 -0.110624 -0.023849 0.037828	1,068	21.434 20.646
	$_1(-z)$	-0.064539 -0.040689 0.418092 0.788486 0.418092 -0.040689 -0.064539	1,701	17.489
2	$H_0(z)$	-0.008473 0.003759 0.047282 -0.033475 -0.068878 0.383269 0.767245 0.383269 -0.068878 -0.033475 0.047282 0.003759 -0.008473	1,899	21.225 20.451
	$_1(-z)$	0.014182 0.006292 -0.108737 -0.069163 0.448109 0.832848 0.448109 -0.069163 -0.108737 0.006292 0.014182	2,324	17.337
3	$_0(z)$	-0.129078 0.047699 0.788486 0.788486 0.047699 -0.129078	0,701	20.966 20.197
	$_1(-z)$	0.018914 0.006989 -0.067237 0.133389 0.615051 0.615051 0.133389 -0.067237 0.006989 0.018914	2,068	17.154
4	$_0(z)$	-0.176777 0.353553 1.060660 0.353553 -0.176777	Розбіг.	21.082 20.281
	$_1(-z)$	0.353553 0.707107 0.353553	1,000	17.249
5	$_0(z)$	0.707107 0.707107	0,000	20.981 20.314
	$_1(-z)$	-0.088388 0.088388 0.707107 0.707107 0.088388 -0.088388	1,000	17.223
6	$_0(z)$	0.033145 -0.066291 -0.176777 0.419845 0.994369 0.419845 -0.176777 -0.066291 0.033145	0,830	21.351 20.562
	$_1(-z)$	0.353553 0.707107 0.353553	1,000	17.447
7	$_0(z)$	[-5 -3593 -40424 73760 523222 523222 73760 -40424 -3593 -5]/1105920	0,897	20.869
	$_1(-z)$	[13262923825 -9530737060645 107244180473480 184089551926368 -1256481605445039 696252040235171 8589582861477144 8589582861477144 696252040235171 -1256481605445039 184089551926368 107244180473480 -9530737060645 13262923825]/16622339109060608	0,479	20.070 17.140
8	$_0(z)$	[-5 -3593 -40424 73760 523222 523222 73760 -40424 -3593 -5]/1105920	0,479	21.519
	$_1(-z)$	[-788464479685 566590575101641 -6366634492961918 -17339362943626002 146734754275933991 29117448535333861 -756498104334163076 506725522884340388 4086421960139065824 4086421960139065824 506725522884340388 -756498104334163076 29117448535333861 146734754275933991 -17339362943626002 -6366634492961918 566590575101641 -788464479685]/ 8976063118892728320;	1.586	20.677 17.553

В його рамках буде проведено аналіз регулярності на якість відтвореного за цим критерієм зображення.

Оскільки в [8,9] побудова базисних функцій не була пов'язана із особливостями зорової системи людини, то існує можливість отримання непридатного з точки зору прикладного застосування результату.

В рамках перевірки, що проводилась, аналізувалася якість відтвореного зображення як з точки зору спостерігача, так і в термінах співвідношення С/Ш, при зміні степені регулярності фільтрів.

Дослідження проводились із однією з пар розроблених в [9] пар біортогональних фільтрів, значення імпульсної характеристики фільтрів для яких приведені в таблиці 1 (відповідно, 8 та 9) та вигляд $\phi, \psi, \tilde{\varphi}, \tilde{\psi}$ - функцій - на рис.1,а.

В процесі досліджень був використаний набір з трьох тестових 256 градаційних зображень розміром 256×256 елементів, які представлені на рис. 2.

5. ОТРИМАНІ РЕЗУЛЬТАТИ

В ході експериментів контролювалася як візульна якість відновленого зображення контролю, так і фіксувалося співвідношення С/Ш з метою порівняння пари фільтрів, що аналізувалася, із тими, що, згідно [6], є найбільш ефективними для застосування в задачах компресії зображень.

Було встановлено, що використання розроблених за запропонованою в [8,9] методикою пар фільтрів в задачах компресії може привести при великих коефіцієнтах стиску до появи артефактів, до яких є чутливим людське око (рис. 3, а, в середині еліпса), тоді як збільшення регулярності дозволяє практично позбутися цього небажаного ефекту (рис. 3, б) (Для порівняння на рис. 3, в приведена та сама ділянка на зображенні при застосуванні пари фільтрів 9/7).

Отримані числові результати наведені в таблиці 1. Першими представлені відзначені в [6] коефіцієнти 6 пар біортогональних фільтрів, степінь їх регулярності [2,3] та оцінка якості відтворених зображень з тестового набору (в dB) на основі результатів запропонованого трирівневого вейвлетного перетво-рення. Нижче подані аналогічні характеристики для пари фільтрів із [9] та підвищеним степенем регулярності.

Числові результати свідчать, що підвищення степені регулярності дозволило покращити не тіьки візуальну якість відторених зображень, але й на 0,4-0,6 дБ збільшити співвідношення С/Ш для відтворених зображень з використаного тестового набору.

Порівняльний аналіз (таблиця 1) засвідчує той факт, що пара фільтрів із збільшеною степінню регулярності забезпечує практично однакову візуальну якість відновленого зображення та більше співвідношення С/Ш.

6. ВИСНОВКИ

На основі проведених досліджень можна зробити наступні висновки:

1. Встановлено необхідність проведення корекції степені регулярності фільтрів, розроблених в [9], з метою покращення візуальної якості та числових характеристик відновленого зображення.

2. В результаті проведеного порівняльного аналізу встановлено, що отримана пара фільтрів із підвищеною степенню регулярності не поступається своїм аналогам за візуальними показниками та перевершує їх за числовими оцінками.

ЛІТЕРАТУРА

1. G.Davis and Aria Nosratinia, "Wavelet-Based Image Coding: An Overview", Applied and Computational Control, Signals, and Circuits, Vol. 1, No. 1, Spring 1998.(Доступна в електронному варіанті: http://www.utdallas.edu/~aria/papers/dn383.ps.gz.
2. O.Rioul, "Simple regularity criteria for subdivision schemes", SIAM J. Math. Analysis, vol. 23, pp. 1544-1576, Nov.1992.
3. O.Rioul,"Regular wavelets: a discrete time approach", IEEE Trans. on Signal Proc., vol. 41, no 12, pp. 3572-35-79, Dec. 1993.
4. O. Streomme, "On the applicability of wavelet transforms to image and video compression", Ph.D. dissertation, Dep. of Computer Sc., Univercity of Strathclyde, Glasgow, Scotland, Febr. 1999.
5. M.Vetterli and C.Herley, "Wavelets and filter banks: Theory and design", IEEE Trans. on Signal Proc., vol. 40, no 9, pp. 2207-2232, Sept. 1992.
6. J.Villasenor, B.Belzer and J.Liao, "Wavelet filter evaluation for image compression", IEEE Trans. Image Proc., vol.2, pp. 1053-1060, Aug. 1995.
7. X.Wu, "An algorithmic study on lossless image compression". Доступна в електронному варіанті: ftp://ftp.csd.uwo.ca/pub/from_wu/papers/.
8. Коваль О.І. Вейвлетні базиси на основі інтерполяторів, побудованих з використанням теорії полюсів / Матеріали XV відкритої науково-технічної конференції молодих науковців і спеціалістів ФМІ ім. Г.В. Карпенка НАН України "КМН - 2000", Львів, 2000. - С. 82-84.
9. Коваль О.І., Русин Б.П. Побудова біортогональних вейвлетних базисів на основі теорії полюсів / Матеріали шостої Міжнародної конференції "ТЕОРИЯ И ТЕХНИКА ПЕРЕДАЧИ, ПРИЁМА И ОБРАБОТКИ ИНФОРМАЦИИ" (НОВЫЕ ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ), ХГТУРЭ "ТУАПСЕ - 2000", 17-19 вересня 2000 р.
10. Шенен П., Коснар М., Гардан И. и др. Математика и САПР: в 2-х кн. Кн. 1. Пер с франц..- Мир, 1988.-204 с., ил.

Підвищення візуальної якості затемнених деталей на зображенні

Роман Воробель, Ігор Журавель, Наталія Опир

Фізико-механічний інститут НАН України, 79601, Львів, Наукова 5,
vorobel@ah.ipm.lviv.ua, zuravel@ipm.lviv.ua, nopyr@ipm.lviv.ua, тел. (0322) 65 4803

Roman Vorobel, Igor Zhuravel, Natalia Opyr. Increase the visual quality of darken components on image. The article is devoted to methods, connected with enhancement of visual quality of darken components on image. The method of histogram equalization by using a model of image perception and contrast enhancement method based on statistical estimators of local regions are offered. The results of test of proposed image processing are shown.

1. Вступ

Одним з найбільш зручних способів представлення інформації є зображення. Їх часто використовують при неруйнівному контролі, медичній діагностиці та дистанційному зондуванні Землі. Проте часто їх спільним недоліком є низький контраст. Це пояснюється недосконалістю систем отримання та передачі інформації, наявністю завад та ін. Тому підвищення достовірності прийняття рішень на підставі цих зображень вимагає їх покращання.

Системи покращання зображень спрямовані на те, щоб поліпшити візуальне сприйняття зображення або перетворити його в форму більш придатну для візуального чи машинного аналізу. При комп'ютерній обробці покращання зображення як правило має за мету отримання максимальної інформації про об'єкт зацікавлення. Основним споживачем цієї інформації при візуальному аналізі є людина. А система сприйняття зображень людини погано реагує на зміну яскравості в області їх малих значень (затемнених ділянках).

Узагальнюючи методи поліпшення якості зображень зазначимо, що основними підходами до поліпшення зображень є використання непрямих методів підвищення контрастності зображень шляхом табличних, гістограмних чи рангових перетворень та методів нерізкого маскування, а також застосування методів безпосереднього визначення локальних контрастів [1]. Серед цієї множини підходів виділимо два, а саме методи гістограмних перетворень та методи перетворення локального контрасту. Вони пов'язані з можливістю врахування як сприйняття зображень людиною, так і текстури зображень. Опишемо способи підвищення їх ефективності та проаналізуємо отримані результати.

2. Модифікація методу гістограмних перетворень

Підхід до модифікації гістограмних перетворень запропоновано у роботі [2]. Він дозволяє виділяти окремі дрібні деталі та структуру затемнених деталей на зображенні.

Запропоновані гістограмні перетворення [2] використовують модель сприйняття зображення людиною [3], в якій залежність порогу чутливості від яскравості описується аналітичним виразом

$$\Delta L = \begin{cases} \beta(\frac{\Delta L}{L})_{\max} \sqrt{L}\sqrt{L_{x_1}}, & \text{для } L_{\min} > L \geq L_{x_1}, \\ \beta(\frac{\Delta L}{L})_{\max} L, & \text{для } L_{x_2} > L \geq L_{x_1}, \\ \beta(\frac{\Delta L}{L})_{\max} L^2 \frac{1}{L_{x_2}}, & \text{для } L \geq L_{x_2}, \end{cases} \quad (1)$$

де $L_{\min}$ - мінімальна яскравість; інтервал $[L_{\min}, L_{x_1}]$ відповідає області де Вріз-Роз, низької чутливості, близької до абсолютного візуального порогу сприйняття; проміжок $[L_{x_1}, L_{x_2}]$ відповідає області Вебера, а $L \geq L_{x_2}$ – області насичення. Параметр β враховує величину постійного порогу в області Вебера і розміри об'єкта.

Аналітичні вирази для знаходження вихідних яскравостей залежать від інтегралів на відповідних проміжках. Для спрощення запису результуючих виразів введемо позначення:

$$I_1 = \int_{L_{\min}}^{L_{x1}} P_L dL, \quad I_2 = \int_{L_{x1}}^{L_{x2}} P_L dL, \quad I_3 = \int_{L_{x2}}^{L_{\max}} P_L dL,$$

$$I_\Sigma = J_{\max} / \sqrt{2I_2 + I_1 + 4I_3} \quad (2)$$

Для рівномірного закону розподілу вихідних яскравостей і з врахуванням, що він розтягується до інтервалу $[0, J_{\max}]$, для кожного проміжку нами отримані такі аналітичні залежності для вихідних величин яскравостей:

для проміжку $[L_{\min}, L_{x_1}]$ –

$$J = I_\Sigma \left(\int_{L_{\min}}^{L} P_L dL \right)^2 / \left((\sqrt{2I_2 + I_1}) I_1 \right), \quad (3)$$

© UkrObraz'2000

для проміжку $[L_{x_1}, L_{x_2}]$ –

$$J = \left(2\int_{L_{x_1}}^{L} P_L dL + I_1\right) I_\Sigma / \sqrt{(2I_2 + I_1)} \quad , \quad (4)$$

для проміжку $\left[L_{x_2}, L_{\max}\right]$ –

$$J = I_\Sigma \sqrt{2I_2 + I_1 + 4\int_{Lx_2}^{L} P_L dL} \, . \quad (5)$$

Зауважимо, що для побудови числових алгоритмів відповідні інтеграли заміняють сумами:

$$\int_{L_{\min}}^{L} P_L dL = \sum_{W=L_{\min}}^{L} p_W W \, , \quad \text{і т.д.}, \quad (6)$$

де p_W – частота W-ї яскравості.

Щоб повернутися до скоректованих вхідних яскравостей, необхідно використати обернену функцію

$$L^* = F^{-1}(J) \, .$$

Обернену функцію із співвідношень (3-5) не вдається отримати в явному вигляді, тому скористаємося табличним представленням з відповідною інтерполяцією для проміжних точок. Найпростіше використовувати лінійну інтерполяцію. Але з метою зменшення похибок (і знаючи приблизний вигляд функції перетворення) кращий результат дає квадратична інтерполяція.

На рис.1 приведено тестове зображення, що має 256 градацій рівнів сірого, на якому об'єкти зацікавлення є затемнені. Використовуючи

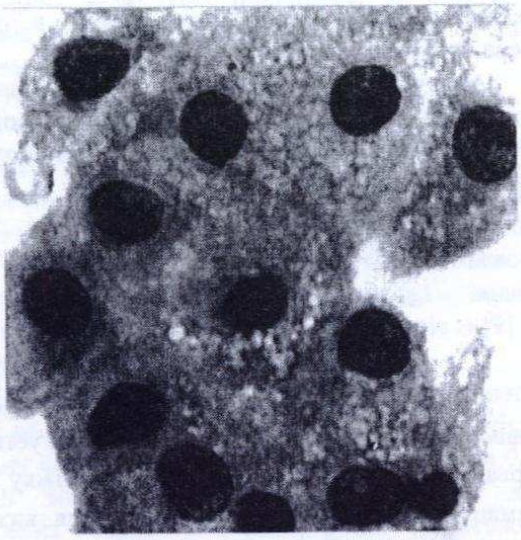


Рис. 1. Вхідне зображення

запропонований метод при L_{x_1}=40 і L_{x_2}=240 отримуємо зображення, що подане на рис. 2.

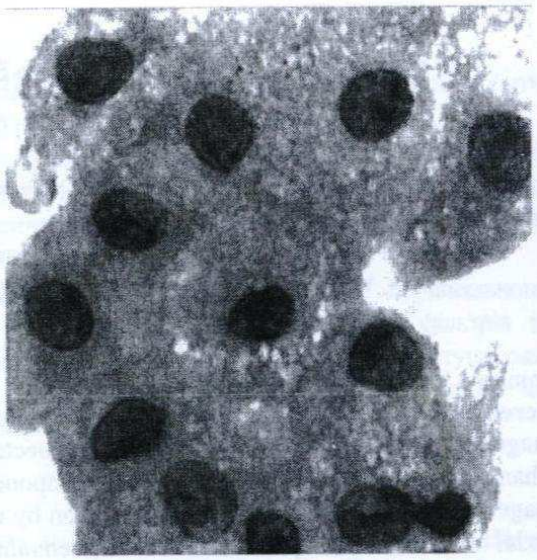
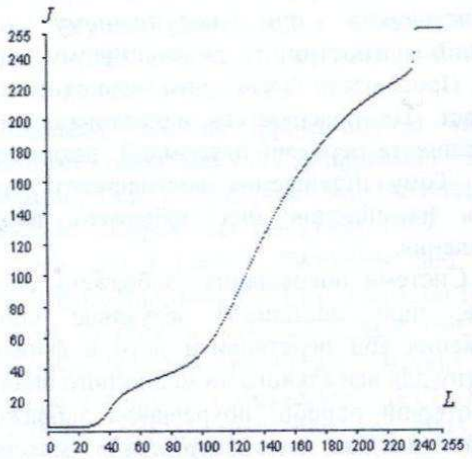
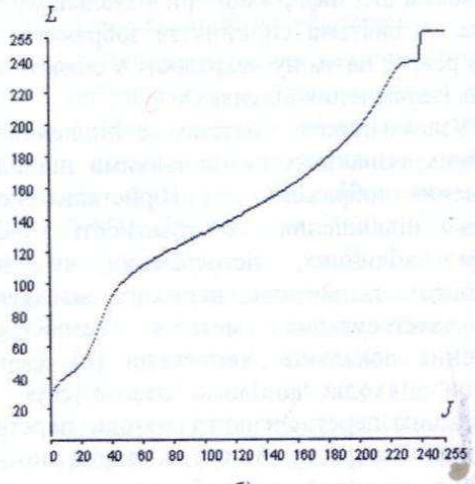


Рис.2. Зображення оброблене запропонованим методом

Для зображення що на рис. 2 відповідні пряма і обернена функції перетворення представлені на рис. 3*а* і 3*б*.



а)



б)

Рис.3. Пряма (*а*) і обернена (*б*) функції перетворення

З аналізу вигляду функції перетворення випливає, що вхідні яскравості які відповідають затемненій області - підсилені. Тому в області малих інтенсивностей об'єкти зацікавлення стали видимими (див. рис 2).

Задаючи параметри L_{x_1} і L_{x_2}, які вибирають в залежності від характеристик монітора і суб'єктивності системи сприйняття оператора, можемо керувати процесом еквалізації гістограм при побудові алгоритмів обробки зображень.

2. Модифікація методу перетворень локального контрасту

В роботі [4] запропонована оригінальна технологія підвищення якості зображень через підсилення локальних контрастів. Її суть полягає у визначенні числового значення локального контрасту для певного елемента зображення, нелінійному його підсиленні та відновленні цього ж елемента зі зміненою яскравістю, яка забезпечує в порівнянні з вхідним зображенням підсилення локального контрасту. Структурна процедура підсилення локального контрасту складається з трьох основних етапів і використовується для кожного елемента $L(i,j)$ з координатами (i,j) вхідного зображення L, $L(i,j) \in L$.

На першому етапі обчислюється локальний контраст елемента. Для цього в [4] запропоновано використовувати вираз

$$C(i,j) = \frac{\left|\bar{L}_1(i,j) - \bar{L}_2(i,j)\right|}{\bar{L}_1(i,j) + \bar{L}_2(i,j)}, \qquad (7)$$

де

$$\bar{L}_1(i,j) = \frac{1}{n^2} \sum_{(i,j)\in R_1} L(i,j), \qquad (8)$$

$$\bar{L}_2(i,j) = \frac{1}{m^2} \sum_{(i,j)\in R_2} L(i,j), \qquad (9)$$

а $m = 3n$, $n > 1$. Як області R_1 і R_2 використовуються ковзні вікна у вигляді квадрата з центром в елементі з координатами (i,j), причому вікно R_1 є в середині вікна R_2.

На другому етапі здійснюється нелінійне підсилення локального контрасту

$$C^*(i,j) = \varphi[C(i,j)], \qquad (10)$$

де $\varphi[C(i,j)]$ - нелінійна монотонна функція, яка має задовольняти умовам

$$C(i,j) \in [0,1],\ \varphi[C(i,j)] \geq C(i,j),\ \varphi[C(i,j)] \in [0,1], \qquad (11)$$

тобто функція $\varphi[C(i,j)]$ є монотонно зростаючою і визначеною на проміжку $[0,1]$. За такі функції приймаються степеневі, експоненційні, логарифмічні та гіперболічні функції.

На третьому етапі відновлюється елемент з координатами (i,j), але з усередненою яскравістю $\bar{L}_1(i,j)$, яка забезпечує підсилення локального контрасту. Для цього проводяться обчислення за формулами

$$\bar{L}_1^*(i,j) = \begin{cases} \bar{L}_2(i,j)\dfrac{1-C(i,j)}{1+C(i,j)} & \text{при } \bar{L}_1(i,j) \leq \bar{L}_2(i,j), \\ \bar{L}_2(i,j)\dfrac{1+C(i,j)}{1-C(i,j)} & \text{при } \bar{L}_1(i,j) > \bar{L}_2(i,j). \end{cases} \qquad (12)$$

Обчислення виразу (12) повторюють для кожного елемента вхідного зображення L. Однак більш ефективним є використання не усередненого значення $\bar{L}_1$ як яскравості елемента, для якого обчислюється локальний контраст (7), а безпосереднього значення яскравості центрального елемента $L(i,j)$ області R_1, що відповідає випадку вибору області R_1 з розміром сторони квадрата $n = 1$.

Надалі вважатимемо, що (при $n = 1$)

$$C(i,j) = \frac{\left|L(i,j) - \bar{L}_2(i,j)\right|}{L(i,j) + \bar{L}_2(i,j)} \qquad (13)$$

та

$$L^*(i,j) = \begin{cases} \bar{L}_2(i,j)\dfrac{1-C(i,j)}{1+C(i,j)} & \text{при } L(i,j) \leq \bar{L}_2(i,j), \\ \bar{L}_2(i,j)\dfrac{1+C(i,j)}{1-C(i,j)} & \text{при } L(i,j) > \bar{L}_2(i,j). \end{cases} \qquad (14)$$

Однак використання описаного методу показує, що його ефективність недостатня для обробки зображень, які містять дрібні деталі. Причиною цього є те, що локальний контраст визначається за формулою (7) чи її частковим випадком (13). Згідно виразу (13) локальний контраст елемента зображення розглядається як міра його відмінності від оточуючого фону. Складовими виразу (13) є безпосередні величини елементів чи їх усереднені значення, що призводить до неповного опису текстури локальної області.

Найбільш повно такі характеристики текстури як однорідність, шорсткість та зернистість описуються статистичними методами.

Одним з найбільш простих методів опису текстури є використання моментів гістограми інтенсивності. Нехай L - випадкова величина, яка визначає дискретну інтенсивність зображення, $H(L(i,j))$ - відповідні значення гістограми. Відомо,

що n-й момент $L(i,j)$ відносно середнього значення визначається формулою

$$\mu_n(L) = \sum_{(i,j)\in R_2} \left(L(i,j) - \overline{L}\right)^2 H\left(L(i,j)\right), \qquad (15)$$

де $\overline{L}$ - середнє значення яскравостей елементів локального околу R_2 (знаходиться аналогічно виразу (9)).

З виразу (15) випливає, що $\mu_0 = 1$, а $\mu_1 = 0$. Другий момент, який називається дисперсією і позначається як $\sigma^2(L)$, важливий для опису текстури. Він є мірою контрасту інтенсивності і застосовується для опису однорідності поверхонь. Так у роботі [5] за міру контрасту текстури запропоновано використовувати величину

$$C(i,j) = 1 - \frac{1}{1 + k\sigma^2(L)}, \qquad (16)$$

де $\sigma^2(L)$ – дисперсія в околі 3×3, k=0.8 – коефіцієнт регуляризації. Величина $C(i,j)$ рівна нулю для околів з постійною інтенсивністю і наближається до одиниці для великих значень $\sigma^2(L)$. Ця властивість виразу (16) повністю відповідає вимогам визначення локального контрасту. Тому, за аналогією з описаним відомим підходом до перетворення локального контрасту, нами запропоновано використовувати в ньому міру контрасту, яка визначається за виразом (16).

Отже, в запропонованому методі на його першому етапі для кожного елемента зображення обчислюємо локальний контраст, використовуючи вираз (16).

На другому етапі здійснюємо таке ж нелінійне перетворення (10) локального контрасту $C(i,j)$.

На третьому етапі відновлюємо зображення шляхом визначення нового значення яскравості $L^*(i,j)$ елемента з координатами (i,j). Для цього використовуємо вираз, який знаходимо з формули (16):

$$L^*(i,j) = \overline{L}(i,j) + \left(\frac{C^*(i,j) * n * m}{1 - C^*(i,j)} - \right.$$

$$\left. - \sum_{\forall (i,j)\in R_2 - R_1} \left(\overline{L}(i,j) - L(i,j)\right)^2 H\left(L(i,j)\right)\right)^{0.5} \qquad (17)$$

Описану процедуру повторюємо для кожного елемента зображення. Використання запропонованого методу проілюстровано на рис. 4, де показано результат обробки вхідного зображення, що на рис. 2.

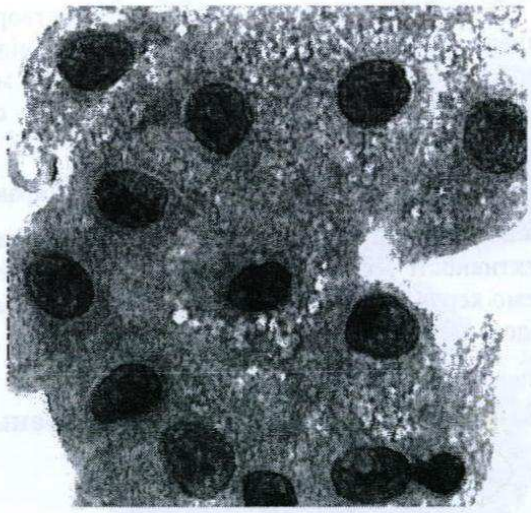


Рис.4. Зображення, що на рис. 2, оброблене методом підсилення локальних контрастів.

3. Висновки

Результати моделювання запропонованих методів модифікації гістограмних перетворень та перетворення локальних контрастів показали, що для виділення затемнених ділянок, що є об'єктом зацікавлення, доцільно спочатку використовувати гістограмні перетворення, які підвищують розрізнюваність деталей об'єктів, а потім проводити перетворення локальних контрастів, що збільшують різницю яскравостей в локальних околах. За допомогою такої технології обробки зростає інформативність слабоконтрастних зображень.

1. Воробель Р.А. Цифрова обробка зображень на основі теорії контрастності. Автореф. дис. … д.т.н. Львів: ДНДІ ІІ, 1999. – 33 с.
2. Опир Н.В. Еквалізація гістограм на основі апроксимаційної моделі сприйняття. Відбір і обробка інформації. 2000. – Вип. 14(90). – С.17-22.
3. Kundu S.K. Pal M.K. A note on gray level-intensity transformation: effect on HVS thresholding// Patern Recognition Letter. 1988. – V.8. – P. 257-269.
4. Dhawan A.P., Buelloni G., Gordon R. Enhancement of mammographic features by optimal adaptive neighbourhvod image processing. IEEE Trans. Med. Imaging. 1986. – V.5. – P. 8-15.
5. Фу К., Гонсалес Р., Ли К. Робототехника. М.: Мир, 1989. - 624 с.

Increasing of low contrast medical multispectral image segmentation sensitivity on base regularized method of the Gram-Schmidt orthogonalization

A. M. Akhmetshin, D. A. Fenoga

Dniepropetrovsk State University
Ukraine, Dniepropetrovsk, 49050, Naukova st., 13
E-mail : akhm@mail.dsu.dp.ua

Annotation

In given work methods, which allow to increase of low contrast medical multispectral image segmentation sensitivity are considered. Algorithm of fuzzy c-means, which allows to rather precisely select areas of image, having uniform structure is described. For increasing of segmentation algorithm sensitivity principal component method and method of pereorthogonalization are used to source images. The practical results were received on base of a multispectral medical image of head brain of the man obtained through system of a nuclear magnetic resonance (NMR).

1. Introduction

The segmentation of tissues of medical images obtained with the help of NMR-systems, is a first step in the analysis and manipulation by interior tissues of a human body. As against other systems of deriving medical images, NMR-systems can generate some images, each of which emphasizes its own parameter of anatomic structure.

As algorithms of segmentation which allow to allocate areas of an image with a homogeneous structure it is possible to choose as precise logic algorithms, and fuzzy logic algorithms.

In methods of precise logic (for instance, precise k-mean) is supposed that each pixel can belong to only one class; this significantly reduces sensitivity of the whole process of segmentation and does not allow to select the finest details of the image. In algorithms of fuzzy logic (for instance, fuzzy c-means) is supposed that pixel can belong to several classes simultaneously that allow to carry out more exact analysis of the initial image.

The complexity of the practical decision of the problem of low contrast areas segmentation can be caused by influence of three factors:

1. The areas representing the greatest interest, frequently are poorly distinctive, or at all not distinctive under direct visual analysis of source images;
2. In accordance with physical nature of a multispectral registration, the most important particularities of areas of medical images can reveal itself only on small part of the set of source images;
3. All snapshots, as the images of the same site of a human body, are force correlated, that can hinder at attempts of visual selection of the most informative sites.

For overcoming these shortages and for increasing of segmentation sensitivity most attractive is the application of various orthogonal transformations, which convert the initial images to new information basis, in which the synthesized images are uncorrelated.

2. Fuzzy c-means method

In the fuzzy c-means method is supposed that each pixel belongs to all classes with the different degree of accessory. Such approach allows to receive more exact results in that cases, when values of samples of several classes are closely located to each other, as well as reduces influence of noise by the result of segmentation.

Let is given a set of a vector-pixels $Y=\{y_k: 1 \le k \le n\}$, where n – a whole number of pixels.

Algorithm of fuzzy segmentation can be described as follows [1]:

Step 1. Arbitrarily choose initial values of centres of clusters z_{qk}, $1 \le q \le c$, $1 \le k \le p$,, where c – number of classes, p – number of images in multispectral image.

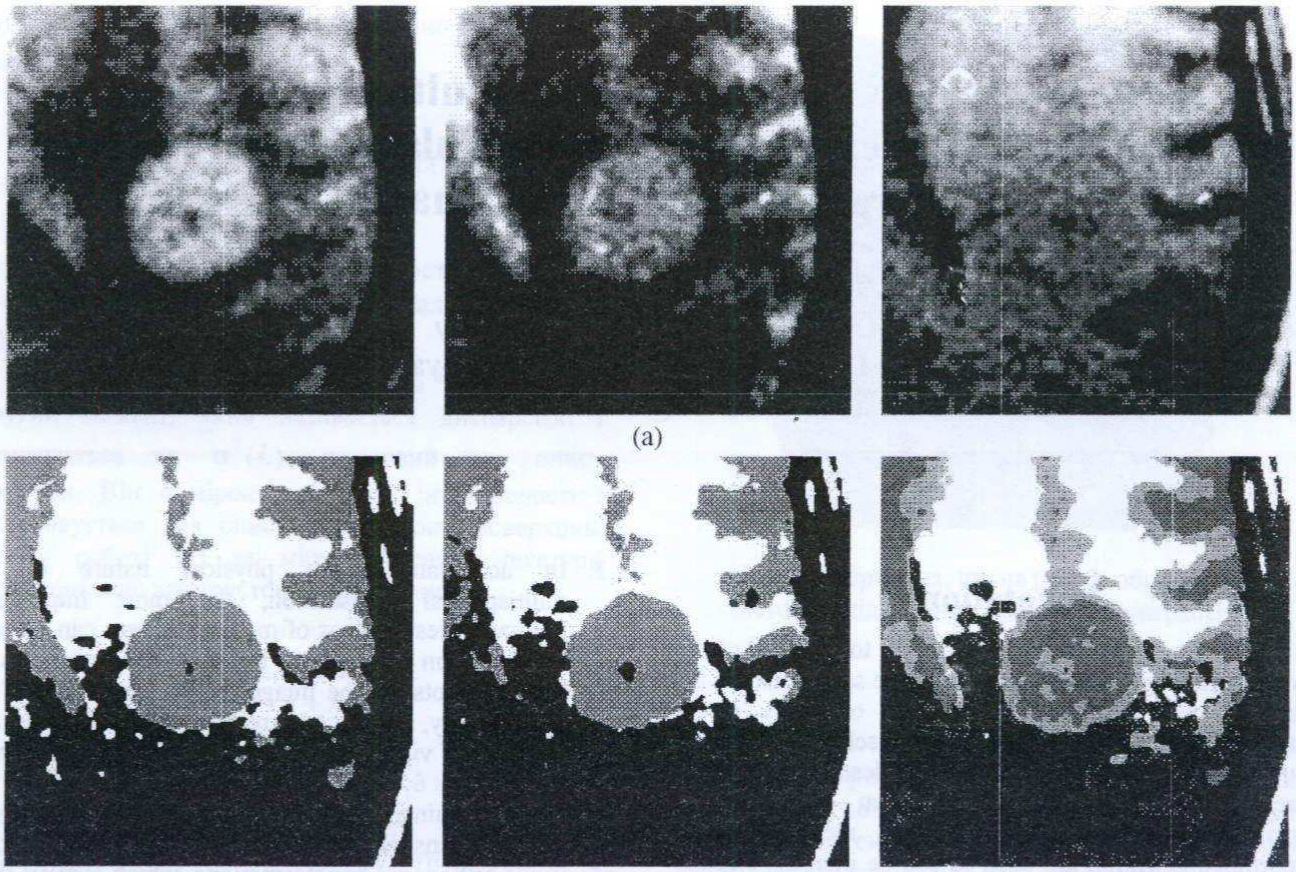
Step 2. Calculate square of Euclidean distance between a vector-pixels y_q and class centres z_k for all classes under formula

$$d_{qk}^2 = \left\| y_q - z_k \right\|^2, \; 1 \le q \le n, \; 1 \le k \le c \qquad (1)$$

Step 3. Calculate matrix of accessory U^t, using formula

$$u_{qk} = \frac{\left[\frac{1}{d_{qk}^2} \right]^{\frac{1}{m-1}}}{\sum_{j=1}^{c} \left[\frac{1}{d_{qj}^2} \right]^{\frac{1}{m-1}}} \qquad (2)$$

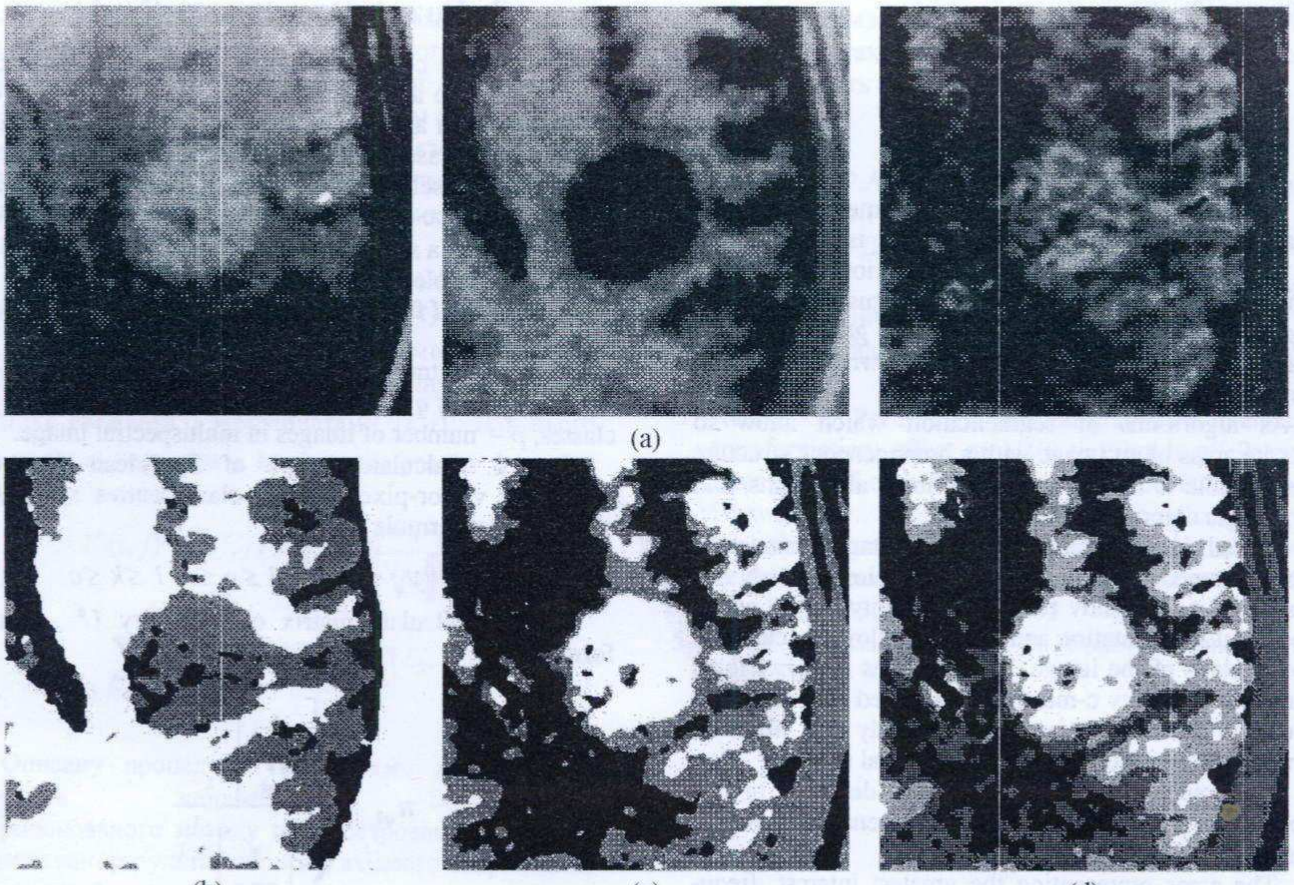
© UkrObraz'2000



(a)

(b) (c) (d)

Fig. 1. a). Source images of part of head brain; b), c), d). Results of using the fuzzy c-means algorithm for segmentation of images in the fig. 1a on 3, 4 and 5 classes accordingly



(a)

(b) (c) (d)

Fig. 2. a). Using the principal component method to source images (fig. 1a); b), c), d). Results of using the fuzzy c-means algorithm for segmentation of images in fig. 2a on 3, 4 and 5 classes accordingly

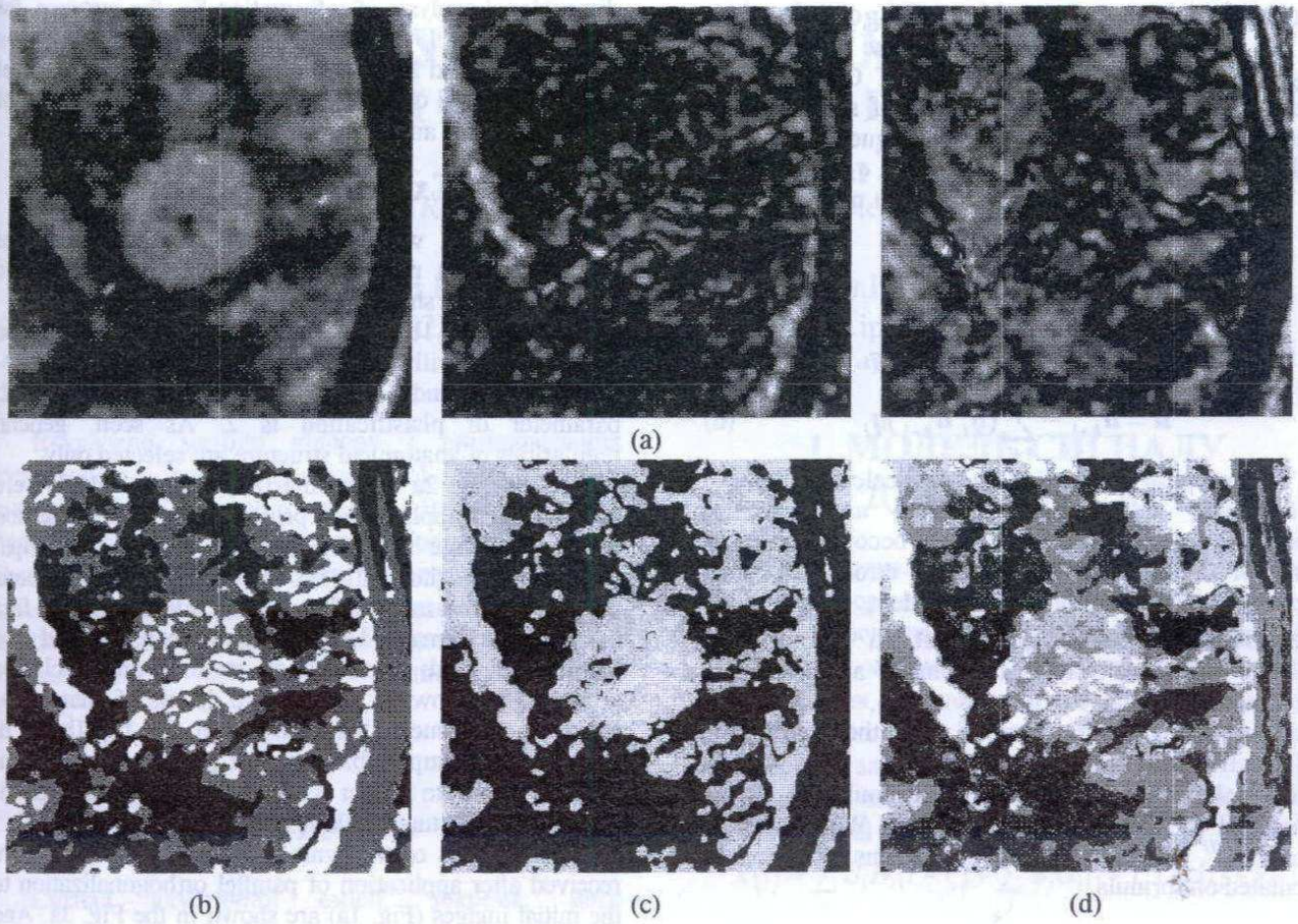


Fig. 3. a). Module of synthesized images, which were derived after using parallel orthogonalization to source images (fig. 1a); b), c), d). Results of using the fuzzy c-means algorithm for segmentation of images in fig. 3a on 3, 4 and 5 classes accordingly

Step 4. Recount values of cluster centres :

$$z_{qk} = \frac{1}{\sum_{j=1}^{n} (u_{jk})^m} \sum_{j=1}^{n} (u_{jk})^m y_{jk} \qquad (3)$$

Step 5. Check up the convergence of segmentation process $\Delta = max||U^{t+1} - U^t||$. If $\Delta > \varepsilon$, then pass to a step 2, otherwise terminate algorithm

Value of parameter of phasification m lies in a range from 1 to ∞ and reduces effect of noise at the calculation of cluster centres.

The direct application of a fuzzy c-means method to force correlated multispectral image allows to select only most common areas of image. Therefore for selection of finer details, i.e. for increasing of segmentation sensitivity, it is necessary beforehand to process image for suppression correlations.

3. Principal components method

One of the methods, which the most often is applied to the suppression of correlations between images, is a principal component method. But, being based on personal experience of using this method in various areas, it is possible to conclude, that at projection of many-dimensional images in this basis, one-two synthesized images $Z_l(x,y)$, which appropriate to the greatest eigenvalues λ_l of correlation matrix $\boldsymbol{R}$ of a cumulative many-dimensional image can be informative only, i.e. the synthesized images are formed on base of expression

$$Z_l(x,y) = \sum_{m=1}^{M} \Phi_l(m) I_m(x,y) \ , \ l=1,\ldots,L \qquad (4)$$

where L – rank of a correlation matrix $\boldsymbol{R}$; M – a number of multispectral image; $\boldsymbol{I_m}$ - initial images; $\boldsymbol{Z_l}$ – synthesized images; $\boldsymbol{\Phi_l}$ – eigenvector (principal component) of correlation matrix $\boldsymbol{R}$.

As the expression (4), as a matter of fact, represents Karunen-Loeve decomposition [2], then, being based on common reasons, it is possible to conclude, that this transformation selects only common regularities peculiar to all multispectral image, that does not correspond in any way to task of segmentation of low contrast sites.

4. Orthogonalization of multispectral image

In the context of problem of segmentation of low contrast areas more perspective is an orthogonalization of the initial images on basis of Gram-Schmidt method, which allows not only to dispose of force correlation between initial multispectral image, as well as considerably increase segmentation sensitivity.

Method of Gram-Schmidt orthogonalization

Let it is necessary to orthogonalize n vectors of height m : $a_1, a_2, \ldots, a_n$. Method of Gram-Schmidt orthogonalization consists in forming of some basis of vectors $q_1, q_2, \ldots, q_n$, so that each subsequent vector q_{k+1} will be orthogonal to previous vectors $q_1, q_2, \ldots, q_k$ [3].

The first step of a method consists in replacement of the first column a_1 on:

$$q_1 = \frac{a_1}{\| a_1 \|} \tag{5}$$

Further, if already derived $q_1, q_2, \ldots, q_k$, , find

$$u = a_{k+1} - \sum_{i=1}^{k} (q_i^T a_{k+1}) q_i \tag{6}$$

Then suppose $q_{k+1}=u/\|u\|$ and pass calculation q_{k+2}

However if the calculations are executed approximately, the described process becomes unstable, moreover the influence of rounding errors that more, than closer columns a_i to linear dependent.

Method of pereorthogonalization

Let's describe a way of deriving of a column q_{k+1} which is orthogonal to columns $q_1, q_2, \ldots, q_k$. This method is named by a method of pereorthogonalization, which is firm to rounding errors [4].

Let column u, received on the formula (6), is not orthogonal to columns $q_1, q_2, \ldots, q_k$. We express it through $u^{(1)}$ and then following columns $u^{(s+1)}$ can be calculated on formula

$$u^{(s+1)} = u^{(s)} - \sum_{i=1}^{k} (q_i^T u^{(s)}) q_i \tag{7}$$

After nonzero column $u^{(s)}$, with due accuracy orthogonal to columns $q_1, q_2, \ldots, q_k$, is derived, suppose $q_{k+1}=u^{(s)}/\|u^{(s)}\|$ and pass calculation q_{k+2} .

Using a method of pereorthogonalization

At choice of technique of orthogonalization procedure there are possible two different approaches:

1. Orthogonalization on columns with following orthogonalization of the received result on rows (in this sense a technique is similar to algorithm of two-dimensional Fourier transformation) – two-dimensional orthogonalization. Thus the received images are real and their number is number of the initial multispectral image, i.e. synthesized image $Z(m,n)$ will be characterized by expression

$$Z(m,n) = O_V \{ O_H \{ I(m,n) \} \quad m=1,\ldots,M; n=1,\ldots,N \tag{8}$$

where O_V and O_H – orthogonal projection operators on columns and on rows accordingly;

2. Independent orthogonalization on rows and on columns – .parallel orthogonalization. In this case synthesized images can be considered as complex and characterized both amplitude-spatial, and phase-spatial characteristics, i.e.

$$\begin{aligned} Z(m,n) &= O_V \{ I(m,n) \} + j O_H \{ I(m,n) \} = \\ &= |Z(m,n)| \exp(j\varphi(m,n)) \end{aligned} \tag{9}$$

where $|Z(m,n)|$ and $\varphi(m,n)$ – amplitude-spatial, and phase-spatial characteristics of synthesized image accordingly.

This allows to pass from multi-dimensional to hyper-dimensional analysis of information So, for instance, for 3 initial images 17 variants of display of the information are possible, and for 4 initial images – 31 variants, that opens additional opportunities for deciding a problem of segmentation of anatomical structure of tissues.

5. Experimental results

Experiments were carried out on an example of processing of a multispectral medical image of head brain, which are shown in the Fig. 1a.

On the Fig. 1b, c, d results of segmentation of the initial images with the help of the fuzzy c-means algorithm on 3, 4 and 5 classes accordingly are presented; parameter of phasification is 2. As seen, general regularities of anatomical structure are selected only.

On the Fig. 2a the synthesized images, which were received after application a principal component method to source image (see the fig. 1a). Using a principal component method allows to eliminate correlations between source images, hereunder, selecting more fine details in each image. Using fuzzy c-means method (see the fig. 2b, c, d; segmentation on 3, 4 and 5 classes accordingly) allows to select more details of anatomical structure of tissues, than images on the fig. 1b, c, d. However it is impossible say that principal component method allow to select very small particularities of anatomical structure of tissues.

The modules of the synthesized images, which were received after application of parallel orthogonalization to the initial images (Fig. 1a) are shown in the Fig. 3a. And already these images allow to select a number of additional particularities, not visible in the Fig. 1a and 2a. Segmentation of the received images on 3, 4 and 5 classes by the fuzzy c-means method is presented in Fig. 3b, c, d accordingly. At segmentation the finer areas of a tissue having homogeneous structure are precisely selected. Sufficiently exactly stand out areas, in which values of centres of clusters are located close to one another.

6. Conclusion

In given work methods of processing medical multispectral image, which allow to increase segmentation sensitivity of low contrast sites of the initial images, are considered.

The results of experimental researches allow to conclude, that the transition to Grama-Shmidt orthogonal basis allows to considerably increase low contrast multispectral medical image segmentation sensitivity.

Literature

[1] J. C. Bezdek. "*A review of probabilistic, fuzzy, and neural models for pattern recognition*". Fuzzy Logic and Neural Network Handbook. C. H. Chen (eds.), Ch. 2. – McGraw-Hill, 1996 – 542 p.

[2] Фукунага К. "*Введение в статистическую теорию распознавания образов*". – М.: Наука, 1979. – 368 стр.

[3] Ланкастер П. "*Теория матриц*". – М.: Наука, 1982. – 272 стр.

[4] Беклемишев Д. В. "*Дополнительные главы линейной алгебры*". – М.: Наука, 1983. – 336 стр.

СИГНАЛ АКУСТИЧНОЇ ЕМІСІЇ МАТЕРІАЛІВ І КОНСТРУКЦІЙ -ВИМІРЮВАЛЬНІ ПЕРЕТВОРЕННЯ

О.П.Бухало, Б.П.Клим, Г.В.Микитин, Є.П.Почапський, Р.Ф.Федорів

Фізико-механічний інститут НАН України,79601,Львів,вул.Наукова,5,тел.(0322) 65-43-51, факс (0322) 64-94-27,E-mail: dep23@ah.ipm.Lviv.ua

Викладено методи відбору і вимірювальних перетворень сигналу акустичної емісії (АЕ) матеріалів і конструкцій. Використовуючи запропоновану модель сигналу розглянуто його ймовірнісні характеристики наведено алгоритми фільтрації сигналу, оцінок інформативних параметрів амплітудних розподілів і параметра заникання сигналу АЕ в часі, а також відповідні похибки оцінок.

ВСТУП

Сигнал акустичної емісії, зокрема, його амплітудні, часові, частотні, ймовірнісні та інші характеристики містять інформацію про дефектність матеріалів і конструкцій.

Зв'язок характеристик сигналу АЕ з параметрами досліджуваних об'єктів знаходять використовуючи математичні моделі утворення і поширення дефектів у твердих тілах, результати досліджень розповсюдження різних типів акустичних хвиль, характеристики перетворювачів акустичних хвиль в електричний сигнал, експериментальні дослідження.

Так, наприклад, амплітуда імпульсу сигналу АЕ пропорційна довжині стрибка тріщини, кількість зареєстрованих імпульсів пропорційна сумарному розкриттю тріщини [1], амплітудний розподіл імпульсів сигналу АЕ характеризує тип руйнування [2,3], положення максимуму і ширина частотного спектру визначають стадії деформації [4], інтенсивність потоку імпульсів і функція її зміни в часі характеризують міцність конструкції [1,5,6], реєстрація моментів появи імпульсів на елементах пьєзоантени дозволяє локалізувати місце знаходження дефекту [7] і т.д.

Для підвищення достовірності досліджень часто використовують простори інформативних параметрів сигналу, наприклад, медіанна частота – середня енергія, відносна ентропія – середня амплітуда та ін.

Розробка алгоритмів відбору і обробки АЕ з метою визначення його інформативних параметрів і оцінка похибок вимірювальних перетворень мають вирішальне значення для задачі неруйнівного контролю методами АЕ.

1. МОДЕЛЬ СИГНАЛУ ДЕТЕКТОРА АЕ

Пружні коливання, що виникають внаслідок актів АЕ, перетворюють відповідним детектором у випадкову послідовність імпульсів деякої форми, яка залежить від механізму і місця утворення дефекту, характеристик детектора, типу акустичної хвилі і т.д. Сигнал детектора містить також імпульсну і неперервну завади. Тому в загальному випадку імпульсний сигнал детектора АЕ може бути представлений у вигляді моделі

$$X(t)=\sum_i U_i H_i(t-t_i)+\sum_j B_j G_j\left(t-t_j^{*}\right)+s(t)+\gamma(t),$$

де U_i - амплітуда i-го інформаційного імпульсу, що з'явився в момент t_i, $H_i(\bullet)$- функція форми імпульсу, яка дорівнює одиниці в максимумі; $B_j, G_j(\bullet)$, t_j^{*} - аналогічні значення для імпульсної завади; $s(t)=s_0 Sin\omega t$ - неперервна синусна завада, де s_0 - її амплітуда, а - ω-частота; $\gamma(t)$-стаціонарна неперервна випадкова завада, яка може бути описана в рамках кореляційної теорії; $G(\bullet)$ *і* $H(\bullet)$ і дорівнюють нулю при від'ємному значенні аргументу.

В даній моделі для вимірювальних перетворень використовують лише інформаційну складову, а завади повинні бути відфільтровані чи скомпенсовані.

Амплітуди інформаційних імпульсів є випадковими, характер слідування імпульсів в часі описується пуассоновою моделлю [7].

Таким чином, інформаційну компоненту сигналу характеризують густиною розподілу амплітуд $p_\alpha(u)$, вектором її параметрів α, моментами $\overline{u}$, $\overline{u^2}$, ентропією L, ймовірністю $P(N)$ появи N імпульсів за час T $P(N)=(N!)^{-1}\left[\int_0^T n(t)dt\right]^N \exp\left(-[\bullet]\right)$, де $n(t)$ - інтенсивність імпульсного потоку, тривалістю імпульсів, часом їх зростання і заникання, частотним спектром $K(\omega)$, який є перетворенням Фур'є функції $H(t)$, спектральною густиною $Q(\omega)=n\overline{u^2}\left|K(\omega)\right|^2$ (для $n=const$), значенням чи параметрами функції інтенсивності $n(t)$.

© UkrObraz'2000

Аналогічним чином може бути характеризована імпульсна завада. Неперервну синусну заваду характеризують значеннями s_0 і ω, а неперервну випадкову – її густиною розподілу $\eta(\gamma)$ і кореляційною функцією $R_\gamma(\tau)$, де τ - часовий зсув.

2. ФІЛЬТРАЦІЯ СИГНАЛУ

Виключення імпульсної завади з сигналу $X(t)$ проводять за формою, амплітудою чи тривалістю імпульсів, блокуванням вимірювальних каналів під час ударів, роботи механізмів [8], за співпадінням імпульсів на виходах основного і допоміжного детекторів, які одночасно реєструють електромагнітні завади [9], виключенням неправдивих імпульсів шляхом математичної обробки різниць часів появи імпульсів АЕ на елементах акустичної антени [7] і т.д.

Неперервні завади виключають методами пасивної [10], адаптивної [11] чи прогнозуючої [12] фільтрації. В останньому випадку, наприклад, компенсацію неперервної завади під час дії інформаційного імпульсу тривалістю θ в інтервалі $(t_i, t_i+\theta)$ виконують за алгоритмом

$$X^*(t) = U_i H_i(t-t_i) + s(t) + \gamma(t) -$$
$$-\sum_{k=0}^{1}\left[s^{(k)}(t_i)+\gamma^{(k)}(t_i)\right](t-t_i)^k,$$

де абсолютне значення завади і її похідної в момент t_i використовують для прогнозу і компенсації завади під час дії імпульсу.

3. ОЦІНКА АМПЛІТУДНИХ РОЗПОДІЛІВ І ЇХ ПАРАМЕТРІВ

Амплітудний розподіл інформаційної імпульсної компоненти визначають за допомогою багатоканального амплітудного аналізатора з l_0 суміжними амплітудними "вікнами" шириною Δu кожна. Гістограма апаратного амплітудного розподілу $q(l)$, а також стандартні оцінки моментів і ентропії L мають вигляд

$$q(l) = N_l\left(\sum_{l=1}^{l_0} N_l\right)^{-1}; \qquad \hat{u} = \Delta u\sum_{l=1}^{l_0} lq(l);$$

$$\hat{u^2} = (\Delta u)^2\sum_{l=1}^{l_0} l^2 q(l); \quad \hat{L} = -\sum_{l=1}^{l_0} q(l)\log\left[q(l)\right];$$

де N_l - кількість відліків в l - му каналі.

Оцінка густини розподілу і моментів за гістограмою дає похибку зміщення [13], а якщо, крім того, амплітудний розподіл імпульсів виходить за межі шкали аналізатора, то гістограма зрізаного апаратного розподілу амплітуд не дає можливості оцінки дійсних значень моментів і ентропії реального розподілу $p(u)$.

Із багатьох експериментальних даних відомо, що на різних етапах руйнування різноманітних матеріалів густина розподілу амплітуд $p_\alpha(u)$ має конкретний вигляд (Гауса, Релея, експоненційний, показниковий) [14, 15, 16], де невідомими (інформативними) є значення вектора його параметрів α. Тоді у кожному конкретному випадку можна оцінити дійсні значення параметрів густини розподілу навіть за зрізаною гістограмою і тим самим отримати інформацію про реальні амплітудні характеристики сигналу АЕ.

Для цього можна використати метод максимальної правдоподібності, приймаючи до уваги, що відліки N_l в каналах аналізатора є незалежними. Тоді значення сумісної ймовірності $P\left(N_1, N_2, \ldots N_{l_0}\middle|N_\Sigma\right)$ розподілу кількості відліків N_l по каналах при умові, що їх сумарна кількість дорівнює N_Σ, має вигляд

$$P\left(N_1, N_2, \ldots N_{l_0}\middle|N_\Sigma\right) =$$
$$= \left[P\left(N_\Sigma\right)\right]^{-1}\prod_{l=1}^{l_0} P\left(N_l\right) = N_\Sigma!\prod_{l=1}^{l_0}\left(\frac{\overline{N}_l}{\overline{N}_\Sigma}\right)^{N_l}\frac{1}{N_l!},$$

$$\overline{N}_l = \overline{N}_\Sigma\int_{(l-1)\Delta u}^{l\Delta u} p_\alpha(u)du\left(\int_0^{l_0\Delta u} p_\alpha(u)du\right)^{-1}, \quad \overline{N}_\Sigma = \int_0^T n_\Sigma(t)dt,$$

T - час вимірювань, $n_\Sigma(t)$ - інтегральна інтенсивність потоку імпульсів.

Прирівнюючи нулю похідні функції правдоподібності $\ln P(\bullet)$ по кожному з параметрів α_r вектора α (r=1, 2, ... r_0, де r_0 - кількість невідомих параметрів), отримуємо систему рівнянь для оцінок значень α_r, яка в загальному випадку має вигляд

$$\sum_{l=1}^{l_0} N_l\frac{F'_{\alpha_r}\left[l\Delta u\right] - F'_{\alpha_r}\left[(l-1)\Delta u\right]}{F_\alpha\left[l\Delta u\right] - F_\alpha\left[(l-1)\Delta u\right]} - N_\Sigma\frac{F'_{\alpha_r}\left(l_0\Delta u\right)}{F_\alpha\left(l_0\Delta u\right)} = 0,$$

де $N_\Sigma = \sum_{l=1}^{l_0} N_l$; $F_\alpha(z) = \int_0^z p_\alpha(u)du$, а похідна функції $F_\alpha(\bullet)$ береться по невідомому параметру α_r.

Для практично важливого випадку вузьких амплітудних "вікон" ($\sqrt{D_u} > 5\Delta u$, де D_u - дисперсія амплітудного розподілу) останні рівняння перетворюються до вигляду

$$\sum_{l=1}^{l_0} N_l\frac{p'_{\alpha_r}\left[(l-0{,}5)\Delta u\right]}{p_\alpha\left[(l-0{,}5)\Delta u\right]} - N_\Sigma\frac{F'_{\alpha_r}\left(l_0\Delta u\right)}{F_\alpha\left(l_0\Delta u\right)} = 0,$$

r = 1, 2, ... r_0, а похідна береться по шуканому параметру α_r.

Останні рівняння для всіх практично важливих випадків мають простий вигляд і розв'язуються методом ітерацій, причому на першому кроці ітерацій другий член в цих рівняннях необхідно прийняти рівним нулю.

Так, наприклад, для показникового розподілу Парето $p_m(u) = (m-1)u_0^{-1}\left(uu_0^{-1}\right)^{-m}$, де u_0 - поріг дискримінації амплітуд, $u>u_0$, $m>1$, отримуємо рекурентний алгоритм оцінки інформаційного параметра m:

$$\hat{m} = \left[1 - \left(1 + \frac{l_0 \Delta u}{u_0}\right)^{-\hat{m}_{k-1}+1} \ln\left(1 + \frac{l_0 \Delta u}{u_0}\right)^{\hat{m}_{k-1}-1}\right] N_\Sigma \times$$

$$\times \left(\sum_{l=1}^{l_0} N_l \ln\left[1 + \frac{(l-0{,}5)\Delta u}{u_0}\right]\right)^{-1} + 1,\ k = 1{,}2{,}3\ldots\ ;\ \hat{m}_0 = 1\ .$$

Для експоненційного розподілу амплітуд $p_A(u)=A^{-1}\exp(-A^{-1}u)$, отримуємо, відповідно

$$\hat{A}_k = \left\{1 - \frac{l_0 \Delta u}{\hat{A}_{k-1}}\left[\exp\left(\frac{l_0 \Delta u}{\hat{A}_{k-1}}\right) - 1\right]^{-1}\right\}^{-1} \frac{\Delta u}{N_\Sigma} \sum_{l=1}^{l_0} (l-0{,}5) N_l,$$

$$k = 1{,}2{,}3,\ \ldots\ ;\ \hat{A}_0 = 0\ .$$

Таким самим чином можуть бути знайдені алгоритми оцінки невідомих параметрів інших законів розподілів (Релея, нормального та ін.)

Процес ітерацій закінчують, коли відносна зміна оцінки параметра α_r стає меншою за відносну статистичну помилку $\delta\alpha_r$. Остання має вигляд

$$\delta\hat{m} = (m-1)m^{-1}N_\Sigma^{-0{,}5}; \qquad \delta\hat{A} = N_\Sigma^{-0{,}5}.$$

Підставляючи отримані значення оцінок параметрів α_r у формули відповідних густин розподілів можна визначити їх реальний вигляд, значення моментів і диференціальної ентропії. Враховуючи також похибку $\Delta\hat{\alpha}_r$ оцінки параметра α_r ($\Delta\hat{\alpha}_r = \hat{\alpha}_r\left[\left(\delta\hat{\alpha}_r\right)^2 + \delta_{\hat{\alpha}_r}^2\right]^{0{,}5}$, де $\delta_{\hat{\alpha}_r}$-відносна апаратна систематична похибка оцінки параметра α_r) легко оцінити обумовлену нею похибку визначення форми густини розподілу

$$\Delta p = \int_0^\infty \left| p_{\hat{\alpha}_r}(u) - p_{\hat{\alpha}_r + \Delta\hat{\alpha}_r}(u)\right| du \approx \Delta\hat{\alpha}_r \int_0^\infty \left|\frac{dp_{\hat{\alpha}_r}(u)}{d\hat{\alpha}_r}\right| du,$$

похибку оцінки моментів $\Delta\overline{u^k(\hat{\alpha}_r)} \approx \left|\frac{d\overline{u^k(\hat{\alpha}_r)}}{d\hat{\alpha}_r}\right|\Delta\hat{\alpha}_r$, а також похибку оцінки ентропії (еквівалентну втраті інформації за рахунок невизначеності параметра α_r)

$$\Delta L = \left|L(\hat{\alpha}_r) - L(\hat{\alpha}_r + \Delta\hat{\alpha}_r)\right| \approx \left|\frac{dL}{d\hat{\alpha}_r}\right|\Delta\hat{\alpha}_r,$$

де $L = -\int_0^\infty p_{\hat{\alpha}}(u)\log\left[p_{\hat{\alpha}}(u)\right]du$. Так, наприклад, для розподілу Парето $\Delta p = \frac{m}{m-1}\sqrt{(\delta\hat{m})^2 + \delta_{\hat{m}}^2}$,

$$\Delta\overline{u^k(\hat{m})} = \frac{k\hat{m}u_0^k}{(\hat{m}-k-1)^2}\sqrt{(\delta\hat{m})^2 + \delta_{\hat{m}}^2}, \qquad \hat{m} > k = 1{,}2\ ;$$

$$\Delta L = \left(\frac{m}{m-1}\right)^2 \log e\sqrt{(\delta\hat{m})^2 + \delta_{\hat{m}}^2},$$

а для експоненціального розподілу амплітуд

$$\Delta p = \frac{2}{e}\sqrt{\left(\delta\hat{A}\right)^2 + \delta_{\hat{A}}^2}, \qquad \Delta\overline{u^k(\hat{A})} = \left(k\hat{A}\right)^k\sqrt{\left(\delta\hat{A}\right)^2 + \delta_{\hat{A}}^2},$$

$$\Delta L = \log e\sqrt{\left(\delta\hat{A}\right)^2 + \delta_{\hat{A}}^2}.$$

4. ОЦІНКИ ІНТЕНСИВНОСТІ ІМПУЛЬСНОГО ПОТОКУ І ПАРАМЕТРА ЇЇ ЗАНИКАННЯ В ЧАСІ

Інтенсивність потоку імпульсів сигналу АЕ оцінюють, як правило, на ділянках локальної стаціонарності де $n(t)=n=const$ шляхом накопичення кількості відліків N за час T. Тоді $\hat{n} = NT^{-1}$, $\delta\hat{n} = N^{-0{,}5}$. Максимальне значення інтенсивності має місце здебільшого на ділянці пластичної деформації [1].

При ступінчастому навантаженні композитних матеріалів [5] чи при попередній витримці металевих виробів в атмосфері водню [6] інтенсивність $n(t)$ потоку імпульсів сигналу АЕ спадає за експоненційним законом $N(t) = n_0\exp\left(tT_0^{-1}\right)$, де n_0 – початкове значення інтенсивності, а інформативним параметром є значення T_0 - постійної часу її заникання. Тоді реєструючи кількість відліків M_l в кожному з l_0 послідовних часових "вікон" тривалістю ΔT отримуємо розподіл чисел M_l по часових вікнах (l=1,2, ... l_0).

Ймовірність $P\left(M_1, M_2, \ldots M_{l_0}\middle|M_\Sigma\right)$ розподілу відліків M_l по каналах ("вікнах") при загальній кількості відліків M_Σ має вигляд

$$P\left(M_1, M_2, \ldots M_{l_0}\middle|M_\Sigma\right) = M_\Sigma!\prod_{l=1}^{l_0}\left(\frac{\overline{M_l}}{\overline{M_\Sigma}}\right)^{M_l}\left(M_l!\right)^{-1},$$

де $\overline{M_l} = \int_{(l-1)\Delta T}^{l\Delta T} n(t)dt$; $\qquad \overline{M_\Sigma} = \int_0^{l_0\Delta T} n(t)dt$.

Диференціюючи функцію правдоподібності $\ln P(\bullet)$ по параметру T_0 і прирівнюючи результат нулю отримуємо точне трансцендентне рівняння для оцінки значення T_0

$$\left[1 - \exp\left(-\frac{\Delta T}{\hat{T}_0}\right)\right]^{-1} - l_0\left[\exp\left(l_0\frac{\Delta T}{\hat{T}_0}\right) - 1\right]^{-1} = M_\Sigma^{-1}\sum_{l=1}^{l_0} lM_l,$$

яке можна розв'язати методом табулювання чи ітерацій.

Імітаційне моделювання на ЕОМ роботи описаних вище алгоритмів оцінок параметрів зрізаних густин розподілів амплітуд і параметра заникання сигналу АЕ дало значення систематичних похибок не більше ніж 3-5%.

5. ВИМІРЮВАЛЬНА СИСТЕМА

Згадані вище та інші алгоритми відбору і обробки сигналу АЕ реалізовані в розробленій інформаційно-вимірювальній системі для експериментальних досліджень характеристик конструкційних матеріалів методом АЕ.

Система має різноманітні засоби програмного і ручного керування, самоконтролю, фільтрації сигналу і дозволяє реєструвати сигнал АЕ в реальному масштабі часу. Обробка сигналу проводиться за різними вимірювальними процедурами, включаючи оцінку амплітудних і часових розподілів сигналу та їх параметрів, оцінку частотного спектру сигналу, його автокореляційної функції, визначення координат джерела АЕ, візуалізацію первинної інформації і результатів обробки і т.д.

Для контролю реальних значень систематичних похибок оцінок параметрів α_r і T_0 в інформаційно-вимірювальній системі було використаний метод взірцевої стохастичної міри і відповідні технічні засоби, що описані в [17, 18]. Експериментальна оцінка систематичних похибок визначення величин α_r і T_0 дала значення в межах 3-5%.

ВИСНОВКИ

Дослідження характеристик матеріалів і конструкцій методом АЕ грунтується на вивченні ефектів утворення і поширення дефектів у твердих тілах, виборі моделі сигналу, виявленні його інформативних параметрів, створення методів і засобів відбору і обробки інформаційного сигналу, обчисленні похибок вимірювальних перетворень, створенні відповідних вимірювальних систем, а також методів і засобів їх метрологічного забезпечення.

ЛІТЕРАТУРА

1. Бунина Н.А. *Исследование пластической деформации металлов методом акустической эмиссии.* -Л.: Изд. ЛГУ, 1990.-156 с.
2. Юдин А.А., Иванов В.И. *К теории акустической эмиссии при пластической деформации металлов* // Акустическая эмиссия материалов и конструкций. Ч.1 (1-я Всесоюзная конференция). – Ростов-на-Дону: Изд. Ростовского университета.-1989. - С.138-142.
3. Буйло С.И. *Об интерпретации максимумов и достоверности оценки вида амплитудного распределения АЭ* // Техн. диагностика и неразруш. контроль.-1995.-№1. - С.31-38.
4. *Спектральный анализ сигналов акустической эмиссии растущей трещины*/ А.Е.Андрейкив, Н.В.Лысак, В.Р.Скальский и др. // Техн. диагностика и неразруш. контроль.-1993.-№1. - С.75-84.
5. Надолинный Б.А., Тихий В.Г., Санин Ф.П. *Акустико-эмиссионный метод прогнозирования давления разрушения* // Акустическая эмиссия материалов и конструкций. Ч.2. (1-я Всесоюзная конференция). - Ростов-на-Дону: Изд. Ростовского университета.-1989. - С.101-105.
6. Лысак Н.Б., Скальский В.Р., Сергиенко О.Н. *Определение водородной поврежденности сталей котельных труб импульсным прозвучиванием* // Техн. диагностика и неразруш. контроль.-1993.-№3. - С.58-65.
7. *Акустическая эмиссия и ее применения для неразрушающего контроля в ядерной энергетике* / В.И.Артюхов, К.Б.Вакар, В.И.Макаров и др.-М.: Атомиздат, 1980.-216 с.
8. Вакар К.Б. *Приборы и информационные системы регистрации и обработки сигналов акустической эмиссии материалов и конструкций* // Акустическая эмиссия материалов и конструкций (1-я Всесоюзная конференция).- Ростов-на-Дону: Изд. Ростовского университета.-1989. - С.108-114.
9. Патент 2419, Украина, МКИ G01N29/14. *Способ контроля роста трещин в образцах материалов* / Андрейкив А.Е., Скальский В.Р., Лысак М.В. – Опубл. 26.12.94. Бюл.№5-1.
10. Скальский В.Р. *Прибор для регистрации сигналов акустической эмиссии СВР-4* // Техн. диагностика и неразруш. контроль.-1995.-№1.С.71-79.
11. *Акустико-эмиссионная система диагностики состояния ответственных металоизделий* / В.В.Шип, Г.А.Бигус, Е.Г.Дорохова и др. // Техн. диагностика и неразруш. контроль.-1997.-№3. - С.56-59.
12. Бухало О.П. *Подавление непрерывных помех в каналах передачи стохастических импульсных сигналов* // Метрология.-1990.-№6. - С.11-19.
13. Бендат Дж., Пирсол А. *Прикладной анализ случайных данных*.-М.:Мир, 1989.-540 с.
14. Буйло С.И. *Использование статистических моделей для количественной оценки амплитудных распределений сигналов акустической эмиссии и повышения достоверности АЭ метода* // Дефектоскопия.-1996.-№5. - С.26-34.
15. Манохин А.И., Маслов Н.И., Белов А.В. *Энергетический анализ природы сигналов акустической эмиссии* // Акустическая эмиссия материалов и конструкций. Ч.1. (1-я Всесоюзная конференция). – Ростов-на-Дону: Изд. Ростовского университета.-1989. - С.154-161.
16. *Применение метода акустической эмиссии для анализа процессов деформации и разрушения композиционных материалов* / Е.А.Кулешова, В.А.Куликов, О.В.Гусев и др. // Акустическая эмиссия материалов и конструкций. Там же. - С.169-180.
17. *Metrological support of NDT by acoustic emission method* / O.Bukhalo, B.Klym, G.Mykytyn at al // Computer Methods and Inverse Problems in Nondestructive Testing and Diagnostics.-Minsk, October 20-23,-1998. - P.305-310.
18. Бухало О.П. *Образцовый импульсный стохастический сигнал – применение, воспроизведение, аттестация* // Измерительная техника.-1997.-№6. - С.24-33.

МАТЕМАТИЧНІ МОДЕЛІ АНАЛІЗУ КОНТУРІВ ПЛОСКИХ СИМЕТРИЧНИХ ЗОБРАЖЕНЬ

Антоніна Буда, Тетяна Мартинюк, Генадій Лисенко, Ольга Король, Сергій Буда

Вінницький державний технічний університет, м.Вінниця, вул. Хмельницьке шосе, 95, тел.(0432) 440-663, факс (0432) 465-772, E-mail: laser@lili.vstu.vinnica.ua

Досліджуються полярні моменти, створюється система ознак, зручна для аналізу форм симетричних двомірних об'єктів. Розглядаються зображення об'єктів, контури яких вже визначені. За точку прив'язки, відносно якої ведуть розрахунки, приймають центр тяжіння зображення. Відслідковується сукупність числових значень моментів, на підставі яких можна проводити класифікацію.

ВСТУП

Сучасні класичні методи аналізу та розпізнавання зображень орієнтовані до наближення людського сприйняття з адекватним представленням інформації, що закладена в точному описі зображення. Пряма та криві лінії можно розглядати як найпростіші елементи великих структур, наприклад таких, як прямокутники, трикутники, кола та області довільної форми. Ці структури можна описати або проаналізувати за допомогою характеристик їх форми: метричних, топологічних, та аналітичних [1].

Аналітичні характеристики форми — це математичні моделі координатного опису проекції. До них відносять перетворення зображення по методу Фур'є, геометричні та полярні моменти в декартовій системі координат. Відомі напрямки дослідження внутрішньої форми об'єкту із застосуванням теорії моментів базуються на створенні математичних моделей:

1. Нормалізації зображення та виконання перетворень зсуву, орієнтації та маштабування зображення [2].

2. Інваріантних моментів, які є незмінними до виконання означених перетворень [3].

3. Зрівноваження моментів більш високих порядків та отримання стислої характеристики «центр-образ» [4].

Виконано машинне моделювання математичного опису геометричних моментів в полярній сістемі координат з метою дослідження зовнішньої форми плоских бінарних зображень та побудови ознак симетричних зображень.

МАТЕМАТИЧНІ МОДЕЛІ КОНТУРНИХ ЗОБРАЖЕНЬ

Обчислення полярних моментів в загальному вигляді може бути представлено як

$$C_{nm} = \iint r^{n+1} \cos m\theta \bullet f(r\cos\theta + x_0, r\sin\theta + y_0)\, dr d\theta, \quad (1)$$

$$S_{nm} = \iint r^{n+1} \sin m\theta \bullet f(r\cos\theta + x_0, r\sin\theta + y_0)\, dr d\theta, \quad (2)$$

де r та θ — відповідно радіус та кут в полярних координатах відносно центра тяжіння (x_0,y_0) зображення об'єктів; m, n — показники степені моментів (m, $n = 0,1,..$).

Ці вирази (1) та (2) передбачають дослідження замкненого контуру об'єкта складної форми. Якщо ж врахувати дискретизацію зображення та те, що функція яскравості світла $F_{(x,y)}$ області зображення дорівнює одиниці, а зовні нулю, то математична модель полярних моментів (5) значно спрощується:

$$C_{mb} = \sum_k r_k^{n+1} \cos m\theta_k , \quad (3)$$

$$C_{mb} = \sum_k r_k^{n+1} \cos m\theta_k . \quad (4)$$

Де число k представляє точки контура.

Центр тяжіння в цьому випадку визначається як:

$$x_0 = \Sigma \frac{i}{N} , \quad (5)$$

$$y_0 = \Sigma \frac{j}{N} , \quad (6)$$

де N — число точок контура, i — абсциса цих точок, j — координата.

Оскільки головна інформація щодо конфігурації фігур міститься в її контурній лінії, то беруться до уваги зміни радіусів-векторів r від центра тяжіння зображення до найбільш та найменьш віддалених точок контура (рис.1).

© UkrObraz'2000

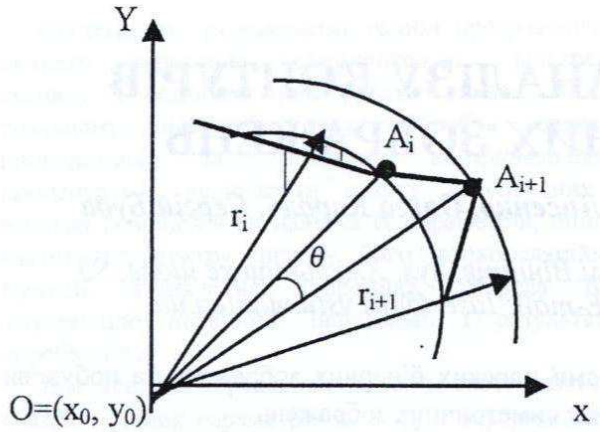


Рис. 1

Запропонований математичний опис зображення був застосований для контурів зображення, показаних на рис. 2.

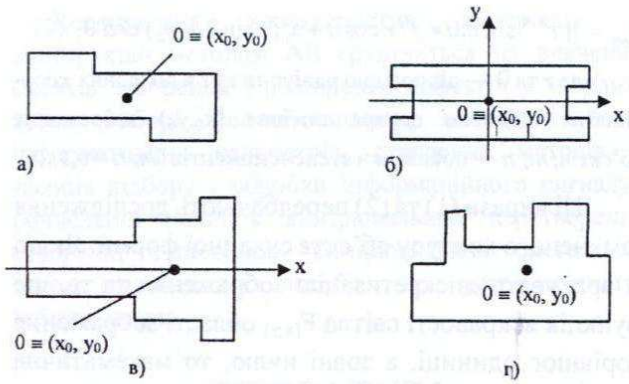


Рис.2

З метою класифікації ставилась у відповідність деяка сукупність чисел, що відповідає певним числовим значенням або нулю.

При обчисленні до уваги були прийняті змішані показники степені n та m, сума яких не перевищувала шести.

ОЗНАКИ КОНТУРІВ ЗОБРАЖЕНЬ

Числові значення двох різновидів полярних моментів C_{nm} та S_{nm} для показаних зображень можна класифікувати так:

1. Зображення, яке має лише центр (рис.2,а). Сукупність моментів $\{C_{11}, C_{12}, C_{22}, S_{11}, S_{12}, S_{22}\}$ характеризується певними числовими значеннями, інші дорівнюють нулю.

2. Зображення, яке має центр симетрії та дві ортогональні осі X та Y (рис.2,б).

Моменти першого різновиду C_{nm} (C_{12}, C_{22}) характеризується певними числовими значеннями, моменти другого різновиду S_{nm} (S_{12}, S_{22}) дорівнюють нулю.

3. Зображення, яке має одну вісь симетрії та може бути орієнтовано в координатній системі XOY (рис.2,в).

3.1. Паралельно до осі X — вся сукупність першого різновиду моментів $\{C_{nm}\}=0$, а другі різновиди моментів S_{13}, S_{33} мають числові значення;

3.2. Паралельно до осі Y — вся сукупність $\{C_{nm}\}$ характеризується числовими значеннями, відмінними від нуля, а вся сукупність $\{S_{nm}\}=0$.

Зображення без ознак симетрії (рис. 2,г). Сукупності двох різновидів полярних моментів $\{C_{nm}, S_{nm}\}$ характеризуються числовими значеннями, відмінними від нуля.

ВИСНОВКИ

Запропонована сукупність ознак контуру дозволяє:

1) вводити класифікацію зображень об'єктів з розширенням її до підкласів зображень шляхом дослідження змін $\{C_{nm}\}$ та $\{S_{nm}\}$, що мають числові значення на кожному кроці кута зображення;

2) визначати орієнтацію осі симетричного об'єкту зображення;

3) звернутись до пошуків окремих складових частин зображення (крок дискритизації, загальне число кроків дискредитації, дискретні значення їх радіус-векторів) з метою отримання додаткових ознак.

ЛІТЕРАТУРА

1. **Претт У.** Цифровая обработка изображений: Пер. с англ.— Т. 2.— М.: Мир, 1982.— 786 с.

2. **Анисимов Б. В., Курганов В. Д., Злобин В. К.** Распознавание и цифровая обработка изображений.— М.: Высш. шк., 1987.— 295 с.

3. **S. Dudani et al.** Aircraft idetification by moment invariant //IEEE Trans, Comput. C-26. 1977. P. 39–45.

4. **Буда А. Г.** Розробка моделей та дослідження прикладних методів обробки геометричних зображень на підставі моментних характеристик: Автореф. дис. … канд. техн. наук.— Вінниця, 1993.— 25 с.

5. **Куафе Ф.** Взаимодействие робота с внешней средой.: Пер. с франц.— М.: Мир, 1985.— 285 с.

THE PARTICULARITIES OF RADIO HEAT IMAGES OF CLOUDY FIELDS AS A PRIORI INFORMATION FOR REMOTE SENSING ALGORITHMS

N.V. Ruzhentsev, Yu.M. Gerasimov, A.V. Antonov

Institute of Radio Astronomy, National Academy of Sciences
4 Krasnoznamennaya Str., 61002 Kharkov, Ukraine
Fax: 38 0572 476506 E-mail: ruzh@rian.kharkov.ua

The particularities of earth cover and cloudiness contrasted unhomoginaities at millimetre wave range are shown.

Introduction

Processing of radio thermal and radar images of earth surface and cloudy fields with the purpose of their quality improvementing is one of major procedures at restoring of environmental parameters. As a result of such image processing is decreasing of negative influence of multiplicative (external origin) and additive (apparatus origin) disturbances. Such procedure carrying out provides an improvement of perception of forms and boundaries of earth covers and atmosphere contrasted formations as well as their radio and physical properties retrieving.

However it is necessary to note that in all these cases in different wavelength ranges there are very frequently the situations when the efficiency of evocative methods of processing is reduced (and even in a number of situations its may be reduced to the zero). It is connected with a negative influence of atmospheric discontinuities especially essential in millimetre and sub-millimetre wavelength ranges then earth cover parameters are sounded. On the other hand then atmosphere parameters are sounded we should collide with earth cover destructive influence. We think the discrimination of radiometric contribution between ground surface and atmosphere is one of the most important issues in passive microwave remote sensing from air and space. Especially in millimetre wavelength there is not so much data existed at present. In these situations while as a rule not surmounted to the present time with the help of the conventional approaches (for a case of remote sensing over surface of the Land and sea ice fields) it is necessary to search for the new approaches and ways of mentioned problem overcoming.

Results

This report is dedicated to searching and investigating of above mentioned differences.

The helicopter trace measurements of space changes of earth cover radio brightness temperatures at 3 millimetre-waves /1 –2/, have shown a continuous character of sizes distribution of contrasted formations, and also what absence of preferable direction in azimuth orientation for these formations /1/. In same time, as a result of one-year cycle of observations of atmospheric fronts passing above our terrestrial observatory (49N 37E) we marked some periodic character of space discontinuities of atmospheric attenuation at millimetre wave range /1, 3/.

Such differences (in case their determination) could be useful for secondary processing of radio images by the methods of texture analysis for example. Here we speak about usage of certain guessed statistical differences between discontinuities of earth surface and atmosphere in their form, group orientation and sizes. You see if it would possible to confirm a systematic character of such differences in the form, sizes and orientations, these statistical data could be applied as some a priory information useful for creation of new method of processing. It is

© UkrObraz'2000

understandable, that for research of such legitimacies the realisation of the analysis of radio heat images for spatially extended areas is necessary. As a source of such data we selected by the radio images over tropics obtained by TRMM satellite (NASA-NASDA). The radiometric instrumentation of the given satellite for today differs by the greatest space resolution of antenna at 3мм and 8мм ranges, and also has the infrared channel. At processing of data presented by Goddard Distributed Active Archive Center we used the radiometer images of areas with a size of 800x2000 km obtained at infra-red, 3мм and 8мм wave ranges. Thus we used the specially designed by us for this purpose a soft-ware - algorithmic support, which allows:

- to conduct a conversion of the satellite maps from a radio brightness sort to a binary pattern (on which selection of contrasting formations of cloudy fields is yielded by a method sliding average – Fig. 1a);
- to calculate sizes of effective lengths and widths for each of the allocated contrasting formations;
- to carry out a calculation and statistical analysis of lengths, widths, ellipticity and orientation for formations contrasted on the radio images.

The Figures 1 and 2 shows the results of such processing for two typical tropical frontal zones at Indian Ocean (14S 60E, the Madagascar's area) and at Pacific Ocean (20,75N 156,6W, the Hawaii's area).

In both figures it is easy to note the availability of some prevailing direction in orientation of cloudy formations (here zero is a direction the North - South). This prevate direction is well allocated not only on a space scale of 700x800 km (fig. 1b) but also on the greater scale 2000x800км (fig. 1c). In this case allocation of directions of orientation practically coincides for IR and for two points of millimetre wave range. Cases however are frequent, when the indicated prevailing directions differ for a little between IR and millimetre ranges, or in IR range they are expressed not so brightly, as in millimetre waves.

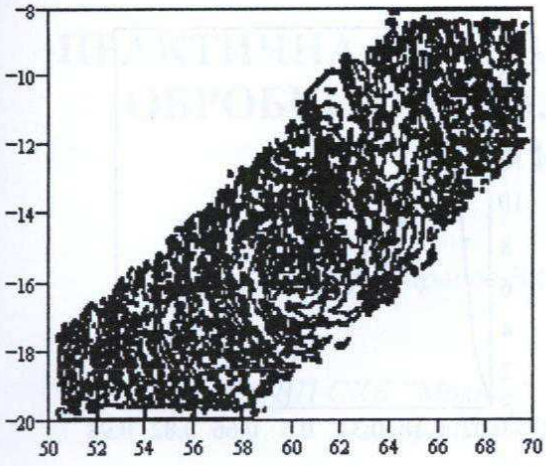
We can also mark availability of certain reference values of relations of lengths and widths (ellipticity) for the allocated formations in cloudiness fields, viewed by us and also availability some reference for each wavelength range of sizes of length of cloudy discontinuities. All this basically is tags (an priori information), which usage can be useful at sharing interaction influence of earth surface and atmosphere radiation at processing of radio heat images.

It is necessary also to note a frequent availability of some harmonically (multiplicity to two or three) in distribution of lengths of clouds formations /Fig. 1e). The same features were marked earlier by us at a data analysis of terrestrial observations of passing of atmospheric fronts in middle latitudes /1/. It is easy to see the maximums at distributions of "ellipticity" values for both different cyclones (Fig. 1d and 2b). At that these characteristic maximums have values near 0,2 for both different atmosphere events observed for different areas of world.

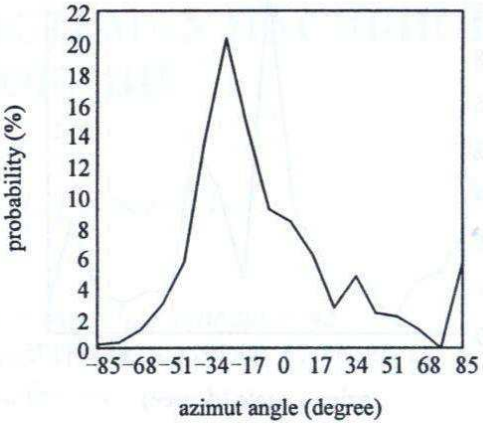
Conclusions

Thus as a result of this work, the availability chain expressed features in a morphological structure of cloudy fields is confirmed. These features can be used as the a priori information at build-up of new algorithms of retrieving of cloudy coverage parameters observed on the phone of earth surface, and also at retrieving of earth surface parameters apparent through a cloudy cover in a millimetre waves.

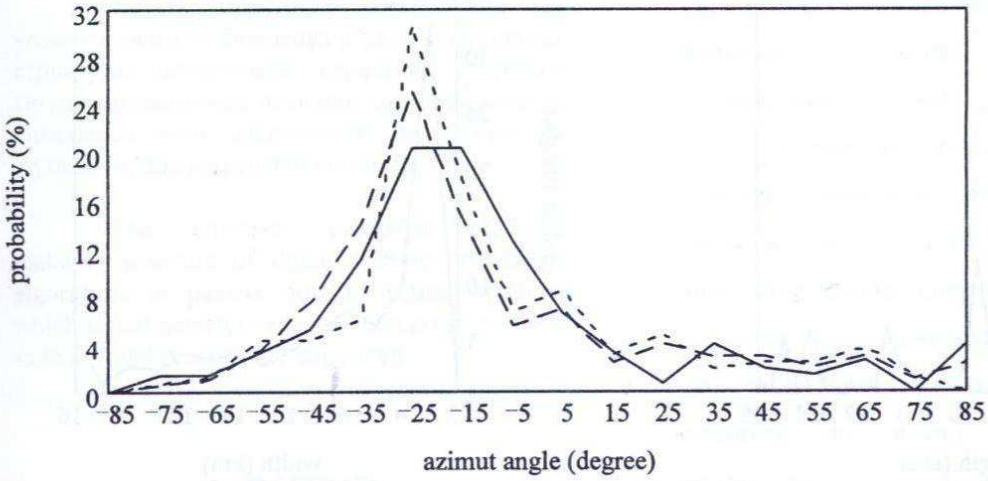
At realisation of further operations in this direction we plan to receive more complete statistical grounding of the quantitative evaluations for above mentioned features in different seasonal and synoptic situations. After that it will be possible to initiate directly with development of algorithms of secondary processing of air-space radio heat images.



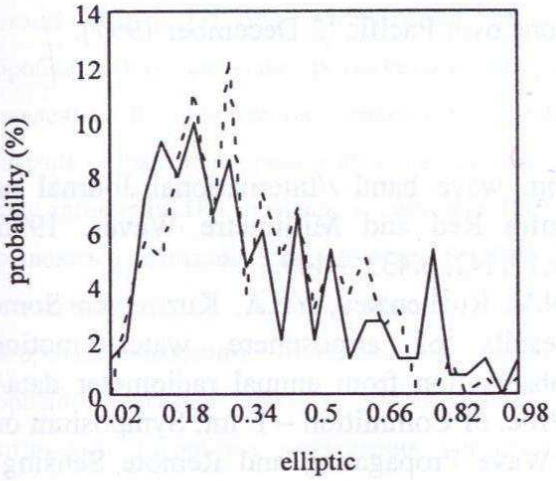
a)



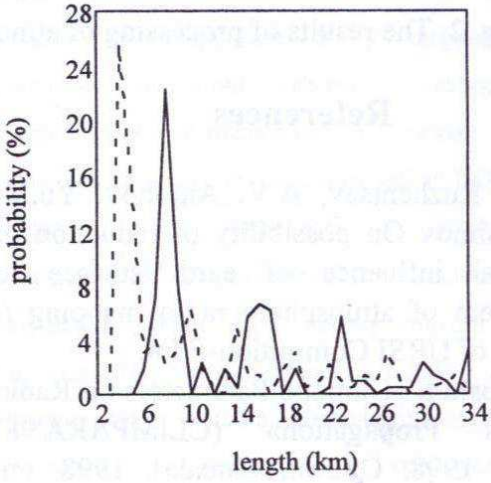
b)



c)



d)



e)

Fig. 1. The results of processing of the Hurricanece over Indian Ocean (2 October, 1998)

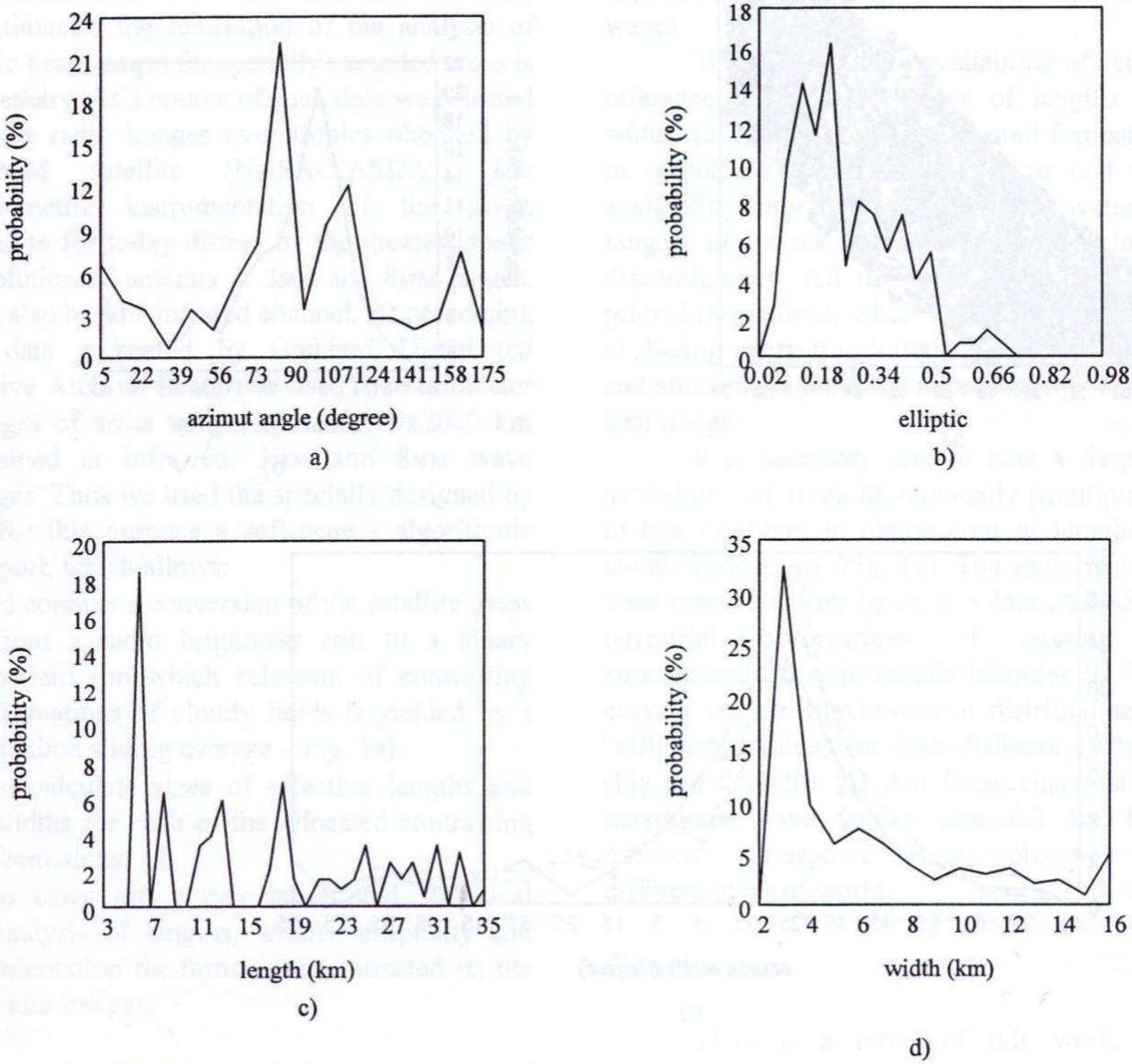


Fig. 2. The results of processing of atmosphere cyclone over Pacific (2 December 1997).

References

1. N.V. Ruzhentsev, A.V. Antonov, Yu.M. Gerasimov On possibility of reduction of mutual influence of earth surface to problem of atmosphere radio mapping // Proc. of URSI Commition-F Int. Symposium «Climatic Parameters on Radio Waves Propagation» (CLIMPARA'98) (April 1998, Ottawa, Canada), 1998, pp. 216 – 219.
2. N.V. Ruzhentsev, V.P. Churilov Terrain radiation: - measurement investigation at 3 mm wave band //International Journal on Infra Red and Millimetre Waves. 1996, v.17, N2, p.433 – 449.
3. N.V. Ruzhentsev, Yu.A. Kuzmenko Some results of atmosphere water motion observation from annual radiometer data// Proc. of Commition – F Int. Symposium on "Wave Propagation and Remote Sensing" (22 September, 1998, Aveiro, Portugal) 1998,p.247-250

ПРАКТИЧНА РЕАЛІЗАЦІЯ АЛГОРИТМІВ ПЕРВИННОЇ ОБРОБКИ СИГНАЛІВ У СИСТЕМАХ ПАСИВНОЇ ЗВУКОВОЇ ЛОКАЦІЇ

Тарасов В.О., Кропачев Д.О., Коржов В.І.

КП СКБ "Молнія", 65005, м. Одеса, вул. Прохорівська, 45
Телефон: (048) 732-31-64, факс: (048) 732-49-06
E – mail: kpmo@farlep.net

Подана практично реалізована на універсальному обчислювачі квазі-оптимальна структура алгоритмів первинної обробки сигналів у системах пасивної звукової локації, виявлювач якої адаптується не тільки до шумового, але і до сигнального процесу.

The universal computer practical realizing structure of signal primary processing algorithms in passive sound ranging systems, which signal detectors adapt as to a noise process, as to a signal process, are suggested.

1. ВСТУП

Традиційно в системах пасивної звукової локації (ПЗЛ) [1] під алгоритмами первинної обробки (ПО) сигналів розуміються алгоритми виявлення й оцінювання параметрів корисних сигналів. Практично реалізовані на сьогоднішній день алгоритми ПО сигналів в системах ПЗЛ, що працюють у реальному масштабі часу (РМЧ), є квазі-оптимальними [2], тобто класичні оптимальні алгоритми виявлення й оцінювання параметрів корисних сигналів (або в узагальненому виді оптимальні алгоритми розрізнення сигналів) [3], спрощуються розроблювачем із метою забезпечення їхній практичної реалізуємості в умовах обмеженості апаратних і обчислювальних ресурсів. При цьому розроблювачі, як правило, жертвують оптимальністю алгоритмів виявлення. Це приводить до зниження імовірності правильного виявлення і розв'язуючої спроможності. Алгоритми ж оцінювання параметрів корисних сигналів прагнуть реалізувати оптимальним способом.

Головною задачею розроблювача, що створює практично реалізуємі алгоритми ПО сигналів у системах ПЗЛ, залишається еволюційне (у міру збільшення апаратних та обчислювальних ресурсів) наближення квазі-оптимальних алгоритмів до оптимальних. На цьому шляху еволюційного розвитку в зв'язку зі швидким ростом продуктивності універсальних обчислювачів представляється доцільним мінімізувати апаратне опрацювання, обмежив його задачі прийомом сигналів, перетворенням їх у цифровий вид, передачею по лінії зв'язку і введенням в універсальний обчислювач. У зв'язку з цим з'являється можливість основний обсяг ПО сигналів у системі ПЗЛ перекласти на універсальний обчислювач і реалізувати ПО в основному на програмному рівні. При такому підході поява в розпорядженні розроблювача більш потужних обчислювальних засобів, дозволяє реалізувати нові структури алгоритмів ПО, які усе більш наближаються до оптимальних, шляхом зміни, в основному, тільки програмно-математичного забезпечення (ПМЗ) ПО сигналів у системах ПЗЛ.

У даній роботі подана практично реалізована на універсальному обчислювачі квазі-

© UkrObraz'2000

оптимальна структура алгоритмів ПО сигналів в системах ПЗЛ, виявлювач якої адаптується не тільки до шумового, але і до сигнального процесу.

2. ПРАКТИЧНА РЕАЛІЗАЦІЯ АЛГОРИТМІВ ПЕРВИННОЇ ОБРОБКИ СИГНАЛІВ

У роботі [2] показано, що однією з найбільш перспективних для практичної реалізації структур є структура алгоритмів ПО на основі комбінації процедур адаптивного виявлення сигналів, попередньої фільтрації і кросс-коррелятора. У роботі [2] також описана практична реалізація квазі-оптимальної ПО сигналів в системах ПЗЛ, котра заснована на комбінації процедур виявлення сигналів адаптивним (до рівня шуму) виявлювачем і оцінювання параметрів сигналів кросс-коррелятором і реалізована у виді комплексу ПМЗ, що працює в реальному масштабі часу на універсальному обчислювачі IBM PC 486. Макет, що реалізує цю структуру алгоритмів квазі-оптимальної ПО сигналів в системі ПЗЛ був випробуваний у польових умовах. Результати іспитів показали, що по основним тактико-технічним параметрам макет задовольняє запропонованим сьогодні до нього вимогам. Однак відповідно до результатів роботи [2] наступним закономірним кроком по шляху розвитку поданої структури алгоритмів квазі-оптимальної ПО сигналів в системі ПЗЛ є доробка виявлювача поданої структури ПО сигналів в системі ПЗЛ , тобто адаптація виявлювача не тільки до шумового, але і до сигнального процесу і, таким чином, еволюційний перехід до повноцінного адаптивного виявлювача.

Відповідно до концепції еволюційного (у міру збільшення апаратних і обчислювальних ресурсів) наближення квазі-оптимальних алгоритмів до оптимальних на сьогоднішній день можна запропонувати наступну практичну реалізацію квазі-оптимальної структури алгоритмів ПО сигналів в системах ПЗЛ, виявлювач якої адаптується не тільки до шумового, але і до сигнального процесу і яка може бути реалізована у виді комплексу ПМЗ, що працює в РМЧ на універсальному обчислювачі (рис.1).

У поданій на рис.1 структурі квазі-оптимальних алгоритмів ПО сигналів в системі ПЗЛ можна виділити наступні основні етапи обробки:

- прийом акустичних сигналів звукоприймальними пристріями (ЗП) на лінійній акустичній базі (АБ), перетворення сигналів ЗП у цифровий вид у частотному діапазоні 5 - 120 Гц, передача сигналів ЗП по лініях зв'язку і введення їх в буфер обчислювача (блоки 1 і 2);
- фільтрація фільтром із смугою пропускання 5-40Гц (блок 3), оцінка рівнів шумів і адаптація до них рівнів виявляючих порогів (блок 4), виявлення, оцінка моменту приходу і виділення корисних сигналів, оцінка відношення сигнал/шум (блок5);
- відсівання сигналів балістичної хвилі, що використовує сигнальні фрагменти в діапазоні 5 - 120 Гц (блоки 6 і 7), і відсівання інших "завадових" сигналів (блоки 8 і 9);
- класифікація вибірки зареєстрованих корисних сигналів по переважному значенню відношення сигнал/шум (блок 10);
- накопичення заданого числа виділених корисних сигналів і тимчасове припинення уведення вхідної інформації (блок 11);
- формування можливих пар сигналів, що відповідають тому самому джерелу (блок 12), визначення класу виділених корисних сигналів по їх тимчасовій і частотній структурі (блок 13), визначення класу пари (блок 14), виділення найбільш ймовірної пари сигналів із сформованих пар і визначення необхідності взаємо-корреляційної обробки сигналів пари (блок 15);
- оцінювання параметрів сигналів і взаємо-корреляційне оцінювання затримки між сигналами

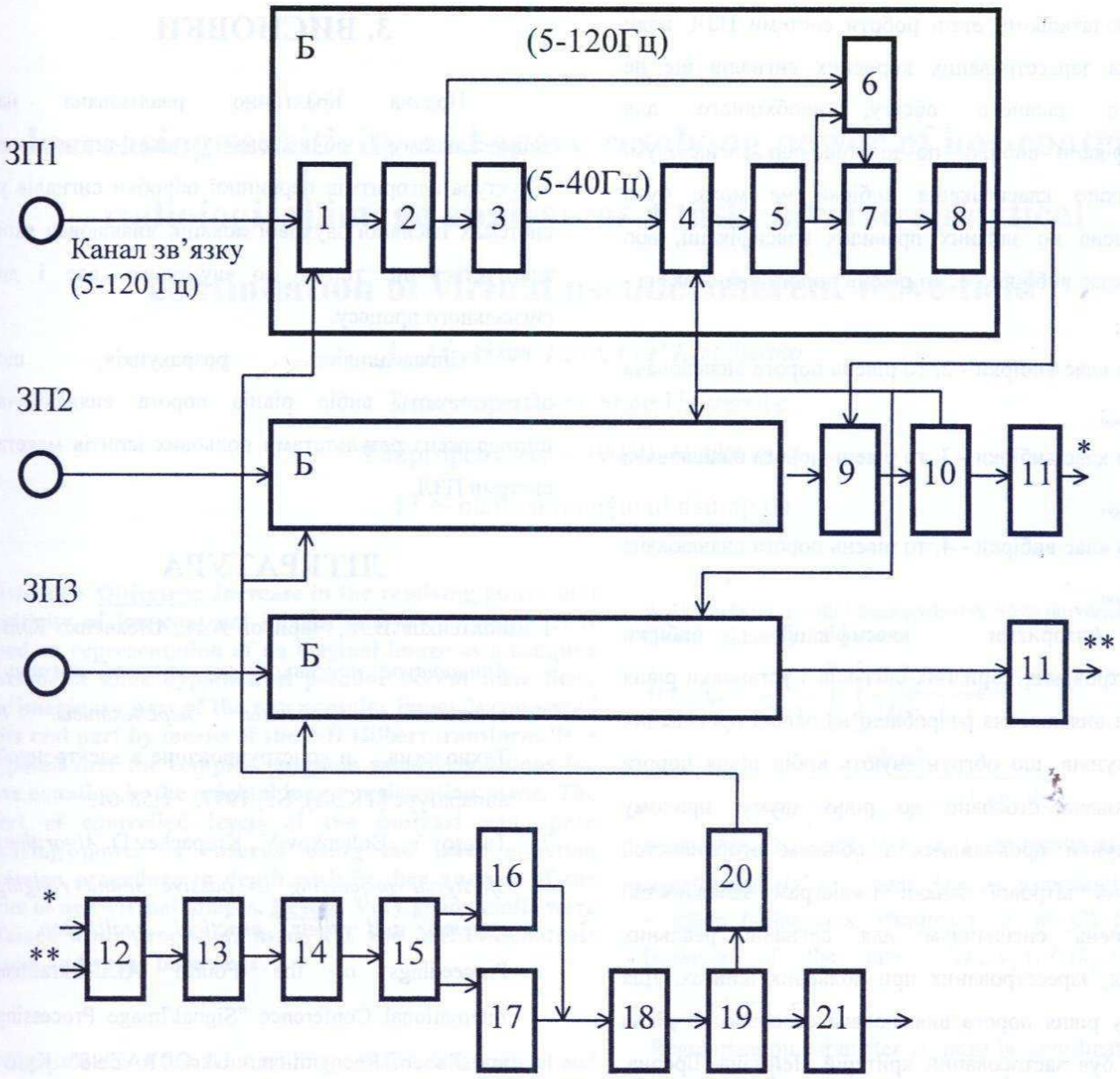


Рис.1. Структура ПО сигнали в системі ПЗЛ, заснована на комбінації процедур виявлення сигналів адаптивним виявлювачем і оцінювання параметрів сигналів кросс-коррелятором і реалізована у виді комплексу ПМЗ, що працює в РМЧ на універсальному обчислювачі.

(блоки 16 і 17), формування вихідних кодограм (блок 18) і передача їх через вихідний буфер (блок 19) на універсальний обчислювач вторинної обробки (блок 21);

- очищення робочих масивів пам'яті й ініціалізація подальшого введення вхідної інформації (блок 20).

У блоці 10 класифікація вибірки зареєстрованих корисних сигналів за значенням відношення сигнал/шум здійснюється в такий спосіб:

- якщо вибірка зареєстрованих корисних сигналів заданого об'єму є рівно ймовірною по відношенню сигнал/шум, то клас вибірки - 1;
- якщо у вибірці зареєстрованих корисних сигналів заданого об'єму переважають сигнали з відношенням сигнал/шум > 15, то клас вибірки - 2;
- якщо у вибірці зареєстрованих корисних сигналів заданого об'єму переважають сигнали з відношенням $8 \leq$ сигнал/шум ≤ 15, то клас вибірки - 3;
- якщо у вибірці зареєстрованих корисних сигналів заданого об'єму переважають сигнали з відношенням сигнал/шум < 8, то клас вибірки - 4.

У блоці 4 рівень порога виявлювача задається в такий спосіб:

- на початковому етапі роботи системи ПЗЛ, коли вибірка зареєстрованих корисних сигналів ще не досягла заданого обсягу, необхідного для класифікації вибірки по відношенню сигнал/шум, або якщо класифікація вибірки не може бути проведена по заданих правилах класифікації, або якщо клас вибірки - 1, то рівень порога виявлювача - $7\sigma_{шуму}$;
- якщо клас вибірки - 2, то рівень порога виявлювача - $8\sigma_{шуму}$;
- якщо клас вибірки - 3, то рівень порога виявлювача - $5\sigma_{шуму}$;
- якщо клас вибірки - 4, то рівень порога виявлювача - $4\sigma_{шуми}$.

Алгоритми класифікації вибірки зареєстрованих корисних сигналів і установки рівня порога виявлювача розроблені на основі проведених розрахунків, що обгрунтовують вибір рівня порога виявлювача стосовно до рівня шуму, причому розрахунки проводилися з обліком особливостей реальної вітрової завади і діаграм залежностей відношень сигнал/шум для сигналів реальних джерел, зареєстрованих при польових іспитах. Для вибору рівня порога виявлювача стосовно до рівня шуму був застосований критерій Неймана-Пірсона, який наказує добиватися забезпечення порогом виявлювача мінімальної імовірності пропуску сигналів при обмеженні зверху на імовірність помилкової тривоги.

Справедливість розрахунків, що обгрунтовують вибір рівнів порога виявлювача підтверджена результатами польових іспитів макета системи ПЗЛ, у якому була реалізована структура ПО сигналів в системі ПЗЛ, заснована на комбінації процедур виявлення сигналів адаптивним до рівня шуму виявлювачем і оцінювання параметрів сигналів кросс-корелятором [2], зокрема результатами польових іспитів макета при рівнях порога виявлювача - $7\sigma_{шуму}$ і $4\sigma_{шуму}$.

3. ВИСНОВКИ

Подана практично реалізована на універсальному обчислювачі квазі-оптимальна структура алгоритмів первинної обробки сигналів у системах пасивної звукової локації, виявлювач якої адаптується не тільки до шумового, але і до сигнального процесу.

Справедливість розрахунків, що обгрунтовують вибір рівнів порога виявлювача підтверджена результатами польових іспитів макета системи ПЗЛ.

ЛІТЕРАТУРА

1. Болтенков В.А., Чариков А.Н., Шевченко Ю.В. *Аппаратура пассивной звуковой локации: состояние, проблемы, перспективы.* // Технология и конструирование в электронной аппаратуре (ТКЭА), N3, 1997, с.58-61.
2. Tarasov V., Boltenkov V., Kropachev D. *Algorithms of signal processing in passive sound ranging systems and their practical realization.* // Proceedings of the Fourth ALL-Ukrainian International Conference "Signal/Image Processing and Pattern Recognition - UkrOBRAZ'98". Kyjiv, 1998, P. 141-142.
3. Сосулин Ю.Г. *Теория обнаружения и оценивания стохастических сигналов.* - М.: Сов. радио, 1978. - 320с.

Increasing sensitivity and space resolving power of low contrast radiological image analysis as a task inverse analytical continuation of virtual pseudocoherent wave field

A. M. Akhmetshin, O.P.Kirichenko

Dniepropetrovsk State University,

Dnipropetrovsk – 49050, Naykova st.,

13 e–mail: akhm@mail.dsu.dp.ua

Abstract **- Objective: Increase in the resolving power and sensitivity of low-contract images analysis. Method: It is based on representation of an original image as a complex envelope of some hypothetical pseudocoherent wave field. The imaginary part of the new complex image is connected to its real part by means of the 2-D Hilbert transform. It is supposed that the complex envelope satisfy conditions for wave equation in the original image registration plane. The effect of controlled levels of the contrast and space resolving power is ensured using the wave equation inversion procedure in depth with further analysis of the series of new virtual images. Result: Very good results were obtained while processing many CT and MR low-contrast images and space born ones.**

The problem of low contrast image (LCI) enhancement and segmentation is one of the most important for medical diagnostics and remote sensing.

Very often the task of a LCI enhancement can be represented as an one of inverse filtering within the framework of the equation of convolution [1].

$$i(x,y)_{obs} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x-u,y-v)i(u,v)_{id}\,dudv, \qquad (1)$$

where observed image $i(x,y)_{obs}$ being considered as a result of the influence of some distortion operator $H(x,y)$ onto an unknown ideal image $i(x,y)_{id}$.

In this case a problem of the LCI enhancement can be formulated as that of inverse filtering.

The inaccuracy of distortion of the operator H can be reduced to the influence of a structural and measurement noise $n(x,y)$. Then the inverse filtering problem is reduced to a search of a corresponding regularized solution $\hat{i}(x,y)_{id}$ [2].

A.Akhmetshin is with Dnepropetrovsk State university, Ukraine. E-mail: akhm@rff.dsu.dp.ua

$$\hat{i}(x,y)_{id} = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{I(k_x,k_y)}{H(k_x,k_y)}e^{-j(k_x x+k_y y)} \times \frac{dk_x dk_y}{1+\alpha(k^2{}_x+k^2{}_y)e^{-\beta(k^2{}_x+k^2{}_y)^{1/2}}} \qquad (2)$$

where $I(k_x,k_y)$ and $H(k_x,k_y)$ are results of the Fourier transform of $i(x,y)_{obs}$ and $H(x,y)$ accordingly; k_x and k_y – space frequencies. Parameter β in (2) depends on behavior of the ratio $I(k_x,k_y)/H(k_x,k_y)$ when $(k_x^2+k_y^2)^{1/2}$.

Regularization parameter β must be coordinated with the accuracy of the image $i(x,y)_{obs}$ registration.

There are many publications (the comprehensive bibliography is given in [3]) devoted to consideration of various approaches intended for solving the equation (1), but, however, under the condition that analytical description of the distortion operator H in some form is known. However, in application to tomographic LCI (especially for MRI), the principal problem is the fact that the form of $i(x,y)_{obs}$ depends on too many unknown physical and technical factors [4]. The latter means that it is practically impossible to consider the problem of increasing the quality of the tomographic LCI $i(x,y)_{obs}$ (as example) on base of equation (2), because the analytical description of the distortion operator $H(x,y)$ is unknown. This very fact conditions availability of a great deal of the purely mathematical algorithms of the secondary processing of LCI [5].

In the paper, the information possibilities of the new LCI processing method are considered. The presented

© UkrObraz'2000

approach is based on the physical inversion method which does not require any *a priory* analytical description of $H(x,y)$ and providers effective simultaneous segmentation of LCI via using of new phase-space characteristics. The method is based on some assumptions:

Assumption 1. The original LCI $i(x,y)_{obs}$ is the real part of a complex envelope $s(x,y)$ of an unobserved coherent wave field $g = g(x,y,z,t)$ which spreads up along the axis z from $z = -\infty$ to $z = 0$:

$$g(x,y,z,t) = s(x,y)\exp[j(k_z z - \omega t)] \quad (3)$$

$$k_z = 2\pi/\lambda; \qquad \omega = 2\pi f;$$

$$i(x,y) = \mathrm{Re}\{s(x,y)\},$$

where f is the frequency and λ -- the wavelength of the hypothetical wave field. In this sense $g = g(x,y,z,t)$ can be considered as a pseudocoherent (virtual) wave field (PWF). It is supposed that complex envelope $s(x,y)$ of the PWF be "frozen" in the plane $z = 0$ of the observation of the original image $i(x,y)_{obs}$ at the moment $t = 0$, i. e.

$$i(x,y)_{obs} = i(x,y) = \mathrm{Re}\{s(x,y,z=0,t=0)\}, \quad (4)$$

Assumption 2. The imaginary part of the complex envelope of the PWF $q(x,y) = \mathrm{Im}\{s(x,y)\}$ is connected to its real part $i(x,y) = \mathrm{Re}\{s(x,y)\}$ by means of the 2-D Hilbert transform

$$q(x,y) = \frac{1}{\pi^2} P \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{i(u,v)}{(u-x)(v-y)} du dv, \quad (5)$$

where P stands for the Cauchy principal value. This allows to express the complex envelope of the PWF in the form

$$s(x,y) = i(x,y) + jq(x,y) = a(x,y)\exp(j\varphi(x,y)) \quad (6)$$

$$a(x,y) = \sqrt{i^2(x,y) + q^2(x,y)},$$

$$\varphi(x,y) = arctg[q(x,y)/i(x,y)], \quad (7)$$

where $a(x,y)$ and $\varphi(x,y)$ denote the magnitude-space and phase-space characteristics, respectively.

The Fourier transform of the $i(x,y)_{obs} = i(x,y)$

$$I(k_x, k_y) = F\{i(x,y)\} =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} i(x,y)\exp[-j(k_x x + k_y y)]dxdy \quad (8)$$

is connected to the complex envelope $s(x,y) = s(x,y,z=0,t=0)$ of the PWF by the expression

$$s(x,y) = 2\int_{0}^{\infty}\int_{0}^{\infty} I(k_x,k_y)\exp[j(k_x x + k_y y)]dk_x dk_y \quad (9)$$

In equation (8), k_x and k_y denote space frequencies, F – the operator of the Fourier transform.

The complex envelope characteristics $a(x,y)$, $\varphi(x,y)$, $\mathrm{Re}\{s(x,y)\}$ and $\mathrm{Im}\{s(x,y)\}$ are considered as the sought-for information in the framework of the proposed method.

Remark. On this stage is possible to realize an algorithm for providing effect of a virtual lens using with varying focus depth. As it is known, the complex envelope of a coherent optical wave on the output of the lens with transfer function $T(x,y)$ has form [6]

$$S_{out}(x,y) = T(x,y)S(x,y), \quad (10)$$

where

$$T(x,y) = C\exp(-jk\rho^2/2f_0);$$

$$C = \exp(jk\eta d) \quad (11)$$

In (11) $k = 2\pi/\lambda$; d – maximum thickness of the virtual lens; η - virtual coefficient of refraction; f_0 – focus depth of the lens; $\rho = (x^2 + y^2)^{1/2}$. From equation (11) is seen that virtual lens $T(x,y)$ providers only phase correction of the input complex envelope $S(x,y)$.

From principle point of view, a digital variation of the focus depth f_0 (as controlled parameter) allows enhance the quality of a virtual analysis of the LCI, but there is one practical difficulty. The $S(x,y)$ is given in the rectangular coordinate system, but $T(x,y)$ is computed in the polar co-ordinate system. The influence of a pixel structure of the $S(x,y)$ will be displayed in appearance of artifacts (there is some analogy with the slice theorem in tomography). Experiments had shown that ones influence on the phase-space characteristics of the $S(x,y)_{out}$ was very significant. For this reason the other method of the phase-correction was used.

Assumption 3. PWF $g(x,y,z,t)$ satisfies the wave equation

$$\frac{\partial^2 g}{\partial t^2} = c^2\left(\frac{\partial^2 g}{\partial x^2} + \frac{\partial^2 g}{\partial y^2} + \frac{\partial^2 g}{\partial z^2}\right) \quad (12)$$

with boundary conditions for PWF complex envelope $\mathrm{Re}\{s(x,y,z=0,t=0)\} = i(x,y)_{obs}$. In (12) c stands for the phase velocity of the PWF and in general case $c = c(x,y)$.

Method:. The solution to the problem of increasing the LCI contrast and spatial resolving power is inversion of the wave equation (12) in depth from plane of the original image registration $z = 0$ to $z = z_i$, $i = 1,2,...,N$, i.e. the problem is to restore a complex wave field $g(x,y,z)$ in the time moment $t = 0$ from the wave field $g(x,y,z=0,t=0)$ which exists in the plane of original observation.

There are various possible approaches to the problem of the wave equation (12) inversion. The used method's outline description in application to the idealized situation $c = const$ is as follows. Using the properties of the Fourier transform, from the equation (12) we obtain

$$F\left\{\frac{\partial^2 g}{\partial t^2}\right\} = (j\omega)^2 G(k_x,k_y,k_z,\omega),$$

$$F\left\{\frac{\partial^2 g}{\partial y^2}\right\} = (jk_y)^2 G(k_x,k_y,k_z,\omega), \qquad (13)$$

$$F\left\{\frac{\partial^2 g}{\partial x^2}\right\} = (jk_x)^2 G(k_x,k_y,k_z,\omega),$$

$$F\left\{\frac{\partial^2 g}{\partial z^2}\right\} = (jk_z)^2 G(k_x,k_y,k_z,\omega).$$

Taking the Fourier transform of both parts of the equation (12) along co-ordinates x, y and t, the following expression can be obtained:

$$\frac{d^2 G(k_x,k_y,z,\omega)}{dz^2} =$$

$$= \left(\frac{\omega^2}{c^2} - k_x^{\,2} - k_y^{\,2}\right) G(k_x,k_y,z,\omega) \qquad (14)$$

Equation (14) is simple differential equation of second order, which has two independent solutions corresponding to waves of two types. For our problem only the solution for the downward extrapolation from $z = 0$ to $z = z_i$ is of interest. The solution of the extrapolated PWF has the form

$$G(k_x,k_y,z+\Delta z,\omega) =$$

$$= \exp\left(j\Delta z\sqrt{\frac{\omega^2}{c^2} - k_x^{\,2} - k_y^{\,2}}\right) G(k_x,k_y,z,\omega) \qquad (15)$$

or

$$G(z+\Delta z) = KG(z),$$

where

$$K = \exp\left(j\Delta z\sqrt{\frac{\omega^2}{c^2} - k_x^{\,2} - k_y^{\,2}}\right), \quad |K| = 1. \qquad (16)$$

1) From equation (16), it is seen that PWF extrapolation is based on the PWF front phase correction.

From dependence $G(k_x,k_y,z,\omega)$, one can obtain needed solution $g(x,y,z,t=0)$ applying the 2-D Fourier inverse transform in the plane of space frequencies k_x and k_y with subsequent summation in time frequencies ω [7]:

$$g(x,y,z,t) = \sum_{k_x}\sum_{k_y}\sum_{\omega} G(k_x,k_y,z,\omega)$$

$$\times\exp[j(k_x x + k_y y)] \qquad (17)$$

From equations (15) and (16) it is seen that specific solution depends on the choice of the PWF virtual wavelength, depth of the PWF extrapolation and spatial dependence of the PWF phase velocity $c = c(x,y)$.

Result.

Example of the X-ray CT image of a coxa joint is shown in fig.1.

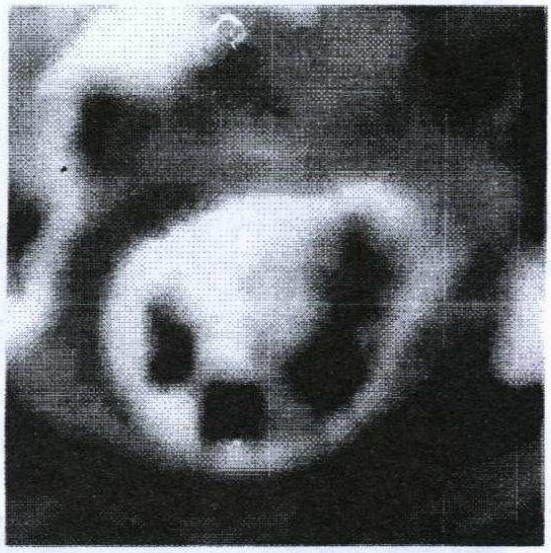


Fig 1. Example of low contrast biomedical image X-ray CT

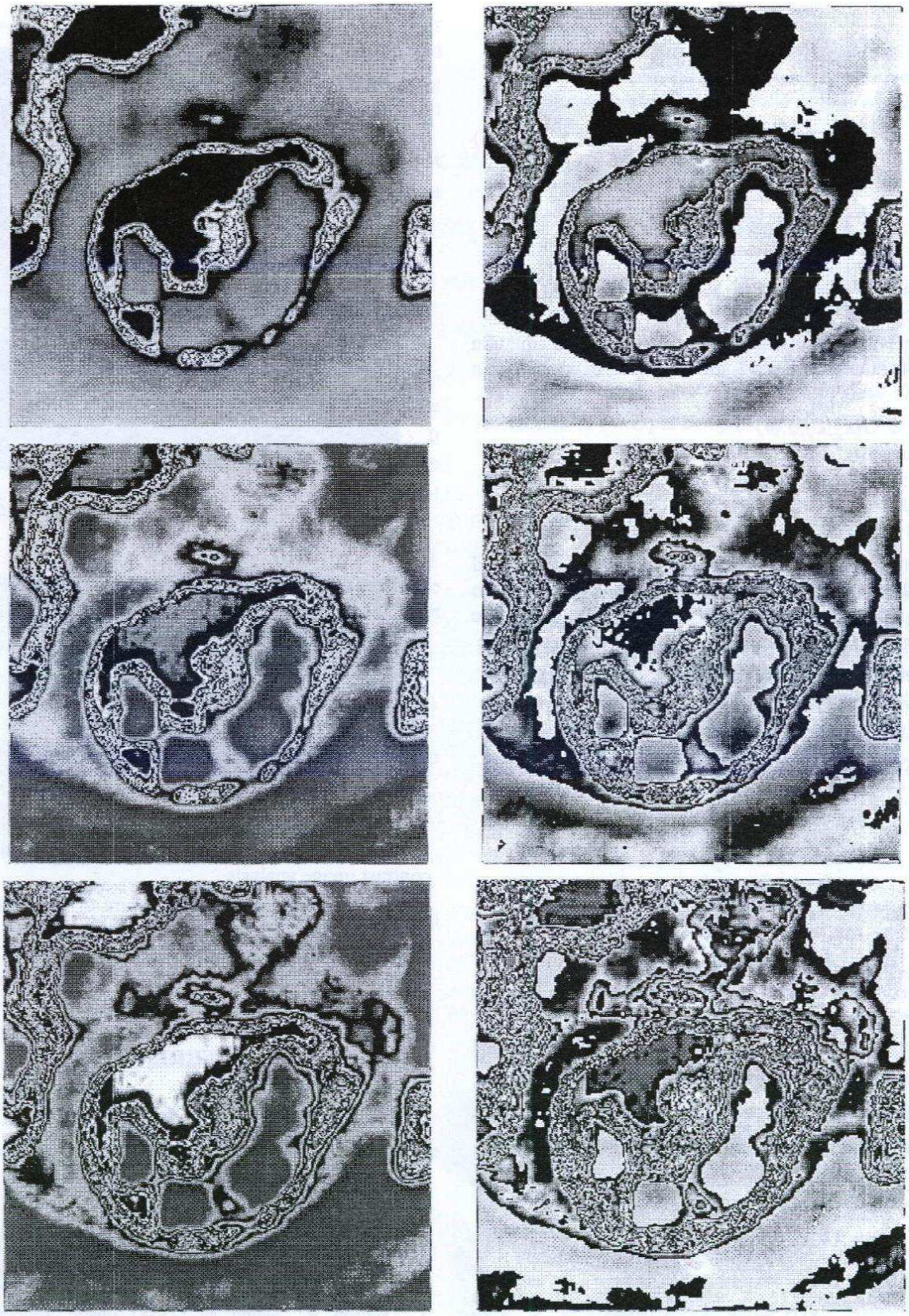
Very important practical feature of the PWF inversion method is the possibility of aiming at the most interesting LCI domains for their detailed analysis.

CONCLUSION

1. The experimental results confirmed the practical expedience of using the PWF inversion method for low-contrast image processing.
2. The PWF inversion method ensures high spatial resolving power and high-sensitive segmentation for low-contrast image analysis.
3. Very important practical feature of the method lies in the possibility of the aiming at the most interesting domains of the low-contrast images by means of proper choice of the PWF extrapolation depth. From physical point of view, such an effect is equivalent to introduction of the variable focus depth for the new virtual images.

REFERENCES

1. Bates R.H.T., McDonnell M.J. (1986) Image restoration and reconstruction. Oxford: Clarendon Press.
2. Tikhonov AN, Arsenin VJ (1977) Solutions to ill-posed problems. New York: Halsted.Ltd.
3. Vasilenko GI, Tavatorin AM (1986) Image reconstruction. Moscow: Radio and Sviaz.
4. Leach MO (1991) MRI obtention with space localization. In Webb S. (ed.): The Physics of Medical Imaging. Moscow: Mir, v.2, p.105-231.
5. Wilson R., Spann M (1988) Image segmentation and uncertainty. New York: Research Studies Press.
6. Yu F.T.S. (1973) Introduction to difraction information processing and holography. Cambridge: MIT Press.
7. Claerbout J (1976) Fundamentals of geophysical data processing. New York: McGrow-Hill Pub.



1 2

Fig.2 The complex envelope S(x,y) for various values of z_i of the extrapolated PWF for the CT image: 1 – a(x,y); 2 – φ(x,y).

THE COMBINED SYSTEMS OF SUPERVISION

D.V. Dovnar, K.G. Predko, O.N. Lysikova
Institute of Applied Optics, National Academy of Sciences of Belarus, 11 Belynitskogo-Biruli, Mogilev 212793, Belarus
Tel.: 375-0222-267124; E-mail: ipo@physics.belpak.mogilev.by

Abstract. The algorithm allowing optimum is offered to restore object under its discrete images generated by several linear systems. The cases are given, when its application allows to receive considerably accurate(up to tenfolds) results of restoration in comparison with results of restoration of the image generated by one, even best, system.

1. INTRODUCTION

The numerical processing of the images with the purpose of reception of the substantial information about the optical characteristics of observable object is usually connected to the solutionof the Fredholm integrated equation of the first kind. In practice, the object can be observed by several not ideal systems. However analyze is going only with results of supervision by the best system and, hence, by solving the equation, appropriate to it. But the systems of supervision can be nonideal on any other business and even the "bad" systems can contain such information on object, which lacks in the best ones. Therefore, we believe, that would be useful to develop the algorithm of supervision of the same object by several systems and to restore the information on object optimum by results of supervision by these systems. Let's name such systems of supervision of objects as the combined systems.

2. Optimum algorithm of restoration of objects under the images generated by combined systems.

The process of formation of the data by the combined system is simulated by the following equation

$$\int_{-\vec{S}}^{\vec{S}} z(\vec{\xi})K(\vec{x},\vec{\xi},p)d\vec{\xi} = f(\vec{x},p) + \qquad (1)$$

$$\gamma(\vec{x},p) = F(\vec{x},p), \quad |\vec{x}| \le \vec{R},$$

Where $z(\vec{\xi})$ are the required characteristics of object,

$$\vec{\xi} = (\xi_1,\xi_2,\ldots,\xi_M),\ d\vec{\xi} = d\xi_1 d\xi_2 \ldots d\xi_M$$

$$\vec{S} = (S_1,S_2,\ldots,S_M),\ \vec{x} = (x_1,x_2,\ldots,x_N),$$

$$d\vec{x} = dx_1 dx_2 \ldots dx_N,\ \ \vec{R} = (R_1,R_2,\ldots R_N).$$

The inequalities of a type $|\vec{x}| \le \vec{R}$ designate $|x_i| \le R_i$ for everyone i. Integer variable p designates number of separate system. The right part of expression (1) is known approximately as $f(\vec{x},p)$ -is exact value of the right part, $\gamma(\vec{x},p)$ is error of its given value (noise). Square-integrated kernel $K(\vec{x},\vec{\xi},p)$ of the equation (1) we shall define (determine) by the following expression:

$$K(\vec{x},\vec{\xi},p) = \begin{cases} A_1 K(\vec{x},\vec{\xi},1) \\ A_2 K(\vec{x},\vec{\xi},2) \\ \ldots\ldots\ldots\ldots \\ A_P K(\vec{x},\vec{\xi},P) \end{cases} \qquad (2)$$

Here - A_p, $p = 1,2,\ldots,P$ normalised (in case of the different registrators they have- different dimensions) factors chosen so that dispersion of noise of all systems of supervision was identical. As algorithm of data processing registered by the combined systems is chosen stabilized [1] algorithm of restoration of objects, as it at any discretisation of the initial data is optimum with a rating of accuracy in mean-squared metrics and is applicable even if the equation (1) has no single-valued solution . The kind of it practically does not vary for the combined systems with any nucleuses such as (2). For the explanatory told we shall give it completely.
The required object is represented as an expansion () on any system of basic functions

$$z(\vec{\xi}) = \sum_{k=1}^{\infty} c_k \psi_k(\vec{\xi}), \quad \left|\vec{\xi}\right| \le \vec{S}. \qquad (3)$$

Then, the "approached" decision of the equation (1) is represented as

$$z(\vec{\xi},\vec{\beta}) = \sum_{k=1}^{\infty} c_k(\vec{\beta})\psi_k(\vec{\xi}) = \qquad (4)$$

$$= \sum_{m=1}^{\infty} \frac{\sum_{l=1}^{m} d_{lm}(\vec{\beta})\psi_l(\vec{\xi}) \sum_{k=1}^{m} d_{km}(\vec{\beta})(\varphi_k,F)}{\beta_m + \sum_{k=1}^{m} d_{km}(\vec{\beta})(\varphi_k,\varphi_m)}.$$

Here $\varphi_k(\vec{x},p) = \int_{-S}^{S} \psi_k(\vec{\xi})K(\vec{x},\vec{\xi},p)d\vec{\xi}$

© UkrObraz'2000

are the images of basic functions, and

$$(\varphi_k,\varphi_l)=\sum_{x,p}\varphi_k(\vec{x},p)\varphi_l(\vec{x},p)$$

are their scalar products. (The summation will be carried out(spent) on all values discretised for input in the computer of arguments of the image). The "approached" values of factors of decomposition (3) are calculated under the formulae

$$c_l(\vec{\beta})=$$

$$=\sum_{m=l}^{\infty}d_{lm}(\vec{\beta})\frac{\sum_{k=1}^{m}d_{km}(\vec{\beta})(\varphi_k,F)}{\beta_m+\sum_{k=1}^{m}d_{km}(\vec{\beta})(\varphi_k,\varphi_m)}.$$

The factors $d_{ln}(\vec{\beta})$ are calculated on by the recurrent formula

$$d_{ln}(\vec{\beta})=$$

$$=-\sum_{m=l}^{n-1}d_{km}(\vec{\beta})\frac{\sum_{k=1}^{m}d_{km}(\vec{\beta})(\varphi_k,\varphi_n)}{\beta_m+\sum_{k=1}^{m}d_{km}(\vec{\beta})(\varphi_k,\varphi_m)},$$

$$l=1,2,\ldots,n-1;\ n=1,2,\ldots.$$

Thus $d_{kk}(\vec{\beta})=1,\ k=1,2,\ldots,$ and if $l>n$, then $d_{ln}(\vec{\beta})=0$.

The "approached" solution (4) depends on stabilizing vector parameter $\vec{\beta}=(\beta_1,\beta_2,\ldots)$.

. Its optimum value is calculated from a condition of a minimum mean-squared error

$$\rho^2=\left\langle\int_{-\vec{S}}^{\vec{S}}\left[z(\vec{\xi},\vec{\beta})-z(\vec{\xi})\right]^2d\vec{\xi}\right\rangle$$

At statistical conditions:

à) $\langle c_kc_i\rangle=\langle c_k^2\rangle\delta_{ik}$, á) $\langle c_k\gamma_i\rangle=0$,

â) $\langle\gamma_i\gamma_k\rangle=\gamma_*^2\delta_{ik}$,

(uncorrelated object, object and noise are uncorrelated, noise is uncorrelated). Here γ_*^2 -is dispersion of an error with zero average at definition of values $f(\vec{x},\vec{\xi},p)$ for input in the computer, the angular brackets mean averaging on set of realizations.

Optimum values of stabilizing parameter we shall mark by an asterisk. Their values: $\beta_m^*=\gamma_*^2/\langle c_m^2\rangle$ .. At such values of stabilizing parameter minimal mean-squared values of an error are as Kind

$$\langle\Delta c_l^2(\vec{\beta}^*,\gamma_*)\rangle=\left\langle\left[c_l(\vec{\beta}^*,\gamma_*)-c_l\right]^2\right\rangle=$$

$$=\sum_{m=1}^{\infty}\frac{d_{lm}^2(\vec{\beta}^*)\langle c_m^2\rangle\gamma_*^2}{\gamma_*^2+\langle c_m^2\rangle\sum_{k=1}^{m}d_{km}(\vec{\beta}^*)(\varphi_k,\varphi_m)},\quad(5)$$

$$\rho^2(\vec{\beta}^*,\gamma_*^2)=$$

$$=\sum_{m=1}^{\infty}\frac{\langle c_m^2\rangle\gamma_*^2\sum_{l=1}^{m}d_{lm}^2(\vec{\beta}^*)}{\gamma_*^2+\langle c_m^2\rangle\sum_{k=1}^{m}d_{km}(\vec{\beta}^*)(\varphi_k,\varphi_m)}.\quad(6)$$

Last two formulas allow to estimate increase of accuracy of definition of object at use of the combined system in comparison with results of restoration of the image generated by any one individual system. So, for example, if as results of measurement by the combined system to consider(examine) two independent images of object, blurred at the expense of rectilinear uniform movement during an exposition on mutually perpendicular to directions, the account under the formulas (5), (6) shows increase of accuracy in tenfolds (at small γ_*^2). This fact is caused by that in the image, blurred on a direction x of value of a spatial spectrum of object on frequencies $(0,\omega_y)$ remain undistorted. In the second image, on the contrary, keep the values of a spectrum of object on frequencies $(\omega_x,0)$.. Thus, at use of optimum algorithm of restoration, the mistakes are sharply reduced at joint numerical processing of two images. Other example gives a roentgen tomography, which too can be considered(examined) as the combined system, on one projection to receive 3-dimentional distribution because it is impossiblewith a help of one projection results. .

3. CONCLUSION

The results of work show, what not always it is necessary to ignore "bad" systems of supervision. Their sharing with usual ones can essentially improve accuracy of restoration of an object, so and raise probability of its recognition.

REFERENCES

1.D.V.Dovnar, K.G. Predko «Approximate reconstraction of object through use of equations lacking single-valued solution».*Optoelectron.Instrum.. Data Process.*1989. No.6. Pp. 1-9.

РЕСТАВРАЦІЯ ОБРАЗІВ У ЧАСТОТНІЙ ОБЛАСТІ МЕТОДОМ ЛОКАЛЬНО-АДАПТИВНОЇ УМОВНОЇ ДЕКОНВОЛЮЦІЇ

КОРОЛЬОВ В.Ю.
Національний Технічний Університет України /КПІ/, радіотехнічний факультет, кафедра теоретичних основ радіотехніки, м. Київ 252056, вул. Політехнічна 12, т. 441-12-71

Розглядається новий алгоритм адаптивної лінійної фільтрації у частотній області з локальною регуляризацією. За критерій правильності реставрації прийнято мінімальне середньоквадратичне відхилення оцінки від початкового образа та неперевищення енергії шума у кожній відновленій трансформанті.

ВСТУП

Реставрація образів — це науковий напрям по розробці методів та засобів компенсації спотворень, що вносяться у зображення в процесі його формування різними системами та джерелами. Значна кількість систем реєстрації зображень мають лінійні інваріантні до зсуву характеристики. Математична модель такої системи — це двомірна згортка [1,2]:

$$g(x,y)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)h(x-\xi,y-\eta)d\xi d\eta+v(x,y), \quad (1)$$

де g(x,y) – спотворене зображення; h(x,y) – імпульсна характеристика системи; f(x,y) – початковий "ідеальний" образ; $v(x,y)$ – невідома реалізація шума.

1.РЕСТАВРАЦІЯ В ОБЛАСТІ КООРДИНАТ

Після процедури дискретизації модель (1) набуває вигляду:

$$G=HF+V, \quad (2)$$

де G – вектор-стовпець дискретних відліків образа g(x,y), послідовно зчитаний по стовцях розміром M^2x1 (для спрощення позначень приймемо, що образи квадратні MxM); F – вектор-стовпець вхідного сигнала f(x,y) розміром M^2x1; H – матриця конволюції розміром M^2xM^2; V – вектор-стовпець невідомої реалізації шума.

В результаті розв'язку задачі отримують оцінку $f_0(x,y)$, яка мало відрізняється від початкового образа по деякому апріорно заданому критерію точності на основі g(x,y), h(x,y) та статистичних характеристик шума $v(x,y)$. Якщо таким критерієм є мінімум середньоквадратичного відхилення оцінки початкового зображення від його оцінки F_0 та обмеження на величину дисперсії шума, то розв'язок задачі (2) набуває вигляду [1,2]: $F_0=\left(H^TH+\alpha C^TC\right)^{-1}H^TG$, де C – матриця регуляризуючого оператора [3] оцінки F_0 розміром M^2xM^2, а α – параметр регуляризації. Розмір матриці $\left(H^TH+\alpha C^TC\right)^{-1}$ (зі змінним емпірично оптимізуємим параметром α), яку необхідно обертати на кожній ітерації складає M^2xM^2. Для телевізійного образа форматом 1024x1024 пікселей $M^2>10^6$. Складність обертання такої матриці можна оцінити по кількості операцій множення-ділення, кількість яких $(M^2)^3>10^{18}$. Зрозуміло, що отримати розв'язок на серійних ПЕОМ середньої швидкодії у прийнятних часових інтервалах неможливо.

2.РОЗВ'ЯЗОК В ОБЛАСТІ ПРОСТОРОВИХ ЧАСТОТ

Суттєво скоротити кількість обчислень дозволяє перехід в область трансформант Фур'є з використанням теореми про згортку оригіналів для (2). В цьому випадку модель спотворення для кожної спектральної складової операндів в (2) буде

$$\dot{G}(m,n)=\dot{H}(m,n)\dot{F}(m,n)+\dot{v}(m,n) \quad (3)$$

де $\dot{G}(m,n)$ та $\dot{F}(m,n)$ – дискретне перетворення Фур'є (ДПФ) від спотвореного g(x,y) та початкового f(x,y) образів; $\dot{H}(m,n)$ – ДПФ від імпульсної характеристики реєструючої системи h(x,y); $\dot{v}(m,n)$ – ДПФ від невідомої реалізації шума. Знайдемо оцінку $\dot{F}_0(m,n)$ початкового образа $\dot{F}(m,n)$. За критерій вірності реставрації приймемо мінімальне середньоквадратичне відхилення оцінки від початкового образа: $\sum_m\sum_n\left|\dot{F}(m,n)-\dot{F}_0(m,n)\right|^2$ та не перевищення енергії є початкового шума $\dot{v}(m,n)$ в оцінці $\dot{F}_0(m,n)$. Такий підхід приводить до системи рівнянь виду [1,2]:

$$\begin{cases}\sum_m\sum_n\left|\dot{G}(m,n)-\dot{H}(m,n)\dot{F}(m,n)\right|^2=\varepsilon & (4a)\\ \sum_m\sum_n\left|\dot{C}(m,n)\dot{F}(m,n)\right|^2\to\min & (4б)\end{cases},$$

де (4а) — умова гладкості розв'язка, $\dot{C}(m,n)$ – регуляризуючий оператор, що реалізує цю умову; (4б) — умова обмеженності дисперсії шума у розв'язку. Методи визначення енергії шума наведені в [4].

Таким чином, розглянута проблема звелась до задачі оптимізації — класичної задачі на умовний екст-

ремум, яку можна розв'язувати за методом множників Лагранжа. Отже, умова (4а) — це цільова функція, а умова (4б) — допустима множина розв'язка. Для того, щоб розв'язати систему рівнянь (4) побудуємо функцію Лагранжа

$$L(\dot{F}_0(m,n)) = \sum_m \sum_n |\dot{C}(m,n)\dot{F}_0(m,n)|^2 + \cdots$$

$$\cdots + \lambda\left[\sum_m \sum_n |\dot{G}(m,n) - \dot{H}(m,n)\dot{F}_0(m,n)|^2 - \varepsilon\right],$$

де λ – множник Лагранжа. Оскільки, коефіцієнти $\dot{G}(m,n)$, $\dot{F}_0(m,n)$ та $\dot{H}(m,n)$ – комплексні числа, то мінімум функції Лагранжа будемо шукати, прирівнюючи нулю частинні похідні від неї за дійсною та уявною частинами $\dot{F}_0(m,n)$ [8]. В результаті отримаємо:

$$\frac{\partial L[\dot{F}_0(m,n)]}{\partial \mathrm{Re}[\dot{F}_0(m,n)]} = \sum_m \sum_n 2\,\mathrm{Re}[\dot{F}_0(m,n)]|\dot{C}(m,n)|^2 + \cdots$$

$$\cdots + \lambda \sum_m \sum_n \left[-\dot{G}(m,n)\dot{H}^*(m,n) - \dot{H}(m,n)\dot{G}^*(m,n) - \cdots\right.$$

$$\left.\cdots - 2\,\mathrm{Re}[\dot{F}_0(m,n)]|\dot{H}(m,n)|^2\right] = 0$$

$$\frac{\partial L[\dot{F}_0(m,n)]}{\partial\, \mathrm{Im}[\dot{F}_0(m,n)]} = \sum_m \sum_n 2\,\mathrm{Im}[\dot{F}_0(m,n)]|\dot{C}(m,n)|^2 + \cdots$$

$$\cdots + \lambda \sum_m \sum_n \left[j\dot{G}(m,n)\dot{H}^*(m,n) - j\dot{H}(m,n)\dot{G}^*(m,n) - \cdots\right.$$

$$\left.\cdots - 2\,\mathrm{Im}[\dot{F}_0(m,n)]|\dot{H}(m,n)|^2\right] = 0$$

Розділимо останню частинну похідну на –j, складемо її з першою. Оскільки система функцій Фур'є є власним базисом для лінійної системи, то система рівнянь буде виродженою і її розв'язок знаходиться з алгебраїчного рівняння:

$$2\left[\{\mathrm{Re}[\dot{F}_0(m,n)] + j\,\mathrm{Im}[\dot{F}_0(m,n)]\}|\dot{C}(m,n)|^2\right] + \cdots$$

$$\cdots + 2\lambda\left[-\dot{G}(m,n)\dot{H}^*(m,n) - \cdots\right.$$

$$\left.\cdots - \{\mathrm{Re}[\dot{F}_0(m,n)] + j\,\mathrm{Im}[\dot{F}_0(m,n)]\}|\dot{H}(m,n)|^2\right] = 0$$

Позначимо: $\alpha = -\frac{1}{\lambda}$. Після зведення подібних отримаємо: $\dot{F}_0(m,n) = \dfrac{\dot{G}(m,n)\dot{H}^*(m,n)}{|\dot{H}(m,n)|^2 + \alpha|\dot{C}(m,n)|^2}$ (5).

Виділимо з (5) вираз для реставруючого лінійного фільтра:

$$\dot{K}(m,n) = \frac{\dot{H}^*(m,n)}{|\dot{H}(m,n)|^2 + \alpha|\dot{C}(m,n)|^2}. \qquad (6)$$

Якщо відсутня апріорна інформація про поведінку енергетичних спектрів початкового зображення та шума, то на практиці використовують тихонівські регуляризуючі множники зі стабілізаторами задачі вида $|\dot{C}(m,n)|^2 = m^{2p} + n^{2p}$, p – натуральне число. Недоліком такого підхода, як буде показано далі, є надмірне зглажування дрібних деталей, тобто дефокусування образа без процедури підбора параметра α [3,4].

2.1.ПОБУДОВА РЕГУЛЯРИЗУЮЧОГО ОПЕРАТОРА В ЧАСТОТНІЙ ОБЛАСТІ ЗА МЕТОДОМ НАЙМЕНШИХ КВАДРАТІВ

Розглянемо інший підхід до побудови регуляризуючого оператора $\dot{C}(m,n)$ в (4) — зглажування за методом найменших квадратів (МНК). Можна показати [4,5], що всереднення емпіричних данних за МНК за допомогою прямих ліній у двомірному випадку в частотній області приводить до розділимого КІХ фільтра:

$$\dot{C}(m,n) = \frac{\sin\left(\frac{Nax}{Ntx}\pi m\right)}{Nax \cdot \sin\left(\frac{\pi m}{Ntx}\right)} \times \frac{\sin\left(\frac{Nay}{Nty}\pi n\right)}{Nay \cdot \sin\left(\frac{\pi n}{Nty}\right)}.$$

Де Nax, Nay — кількість точок, в околицях яких проводиться зглажування; Ntx, Nty — загальна кількість точок; індекси x, y означають напрям відповідної просторової частоти m та n. Чисельні експерименти показали, що при Na = Nt – 1 (у "зглажуванні" значення в одній точці використовується інформація з усіх інших) використання громізких формул усереднення за допомогою парабол другої та четвертої степені не приводить до покращення результата. Крім того, використання пропонуємого оператора замість тихонівського дозволяє отримувати негіршу якість реставрації образа (α=1) без процедури підбора параметрів регуляризації при еволюційній фільтрації [3].

2.2.ПРОЦЕДУРА ВИБОРА ПАРАМЕТРА РЕГУЛЯРИЗАЦІЇ

Розглянемо процедуру вибора параметра регуляризації α. Теоретично знайти параметр регуляризації α можна підстановкою розв'язка задачі (5) в систему рівнянь (4). Знайдене значення α буде оптимальним у середньоквадратичному сенсі. Так при малих значеннях α фільтр K(m,n) (6) вироджується в інверсний фільтр, тобто фільтрація виконується без урахування шумової компоненти, що приводить до значного зашумлення образа замість відновлення. Великі значення α, навпаки, призводять до повного подавлення складової $|\dot{H}(m,n)|^2$ у знаменнику (6), тобто до сильного дефокусування зображення. Крім того, відомо, що середньоквадратичний критерій оптимальності розв'язку системи (4) не співпадає з психовізуальним. Тому на практиці параметр глобальної регуляризації α емпірично добирається оператором на основі суб'єктивних критеріїв якості відновлення образа. Покращення якості реставрації спотвореного образа, зрозуміло, можливе вибором частотнозалежного параметра регуляризації α. Ідея визначення α окремого для кожної трансформанти Фур'є, тобто переходу від процедури глобальної регуляризації до локальної на основі додаткової інформації подана в монографії [7]. Для того, щоб побудувати алгоритм реставрації, оптимальний для конкретного образа [8],

в данній роботі пропонується використати інформацію про енергетичний спектр шума W(m,n). Тоді перше рівняння системи (4) можна переписати у вигляді:

$$\sum_m \sum_n |G(m,n) - H(m,n)F_0(m,n)|^2 = \sum_m \sum_n W(m,n) \quad (8).$$

Позначимо: $N(m,n) = \sqrt{W(m,n)}$. Тоді з (6) та (8) легко отримати: $\alpha(m,n) = \frac{|C(m,n)|^2 [|G(m,n)| - N(m,n)]}{|G(m,n)|^2 N(m,n)}$.

Результати обчислень (початкове зображення, спотворенний образ, відновленний еволюційним фільтром [3], реставрований за пропонуємим методом при $\alpha = 1$, $\alpha = \alpha(m,n)$, інверсна фільтрація) приведені на рис. 1 – 6 відповідно. Прийнято, що імпульсна характеристика системи $h(x,y) = e^{-(x^2+y^2)10^{-2}}$, а відношення сигнал/завада = 10^3. Видно, що описаний алгоритм зваженої фільтрації забезпечує кращу якість фільтрації високочастотних компонент образа ніж відомі підходи. Для порівняльної ілюстрації можливостей алгоритма на рис. 7 – 12 наведені результати обробки одномірного сигнала: початковий сигнал, спотворений, результат відновлення еволюційним фільтром [3], реставрований за методом умовної деконволюції α=const, відновленний за описаним алгоритмом при α=α(ω), відновлений за методом інверсної фільтрації. При імпульсній характеристиці системи $h(t) = e^{-t^2 \cdot 10^{-2}}$ та відношенні $C/З=10^2$ коефіцієнти λ(ω) наведені на рис. 13, на рис. 14 – зворотні величини $\alpha(\omega)=\lambda^{-1}(\omega)$. Обернене перетворення Фур'є від α(ω) наведене на рис. 15.



Рис. 1. Початковий образ.



Рис. 2. Спотворений образ.



Рис. 3. Відновлення еволюційним фільтром.



Рис. 4. Відновлення за методом умовної деконволюції.



Рис. 5. Реставрація за методом локально-адаптивної умовної деконволюції



Рис. 6. Результат інверсної фільтрації.

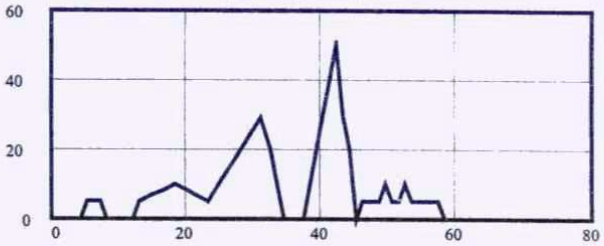


Рис. 7. Початковий сигнал.

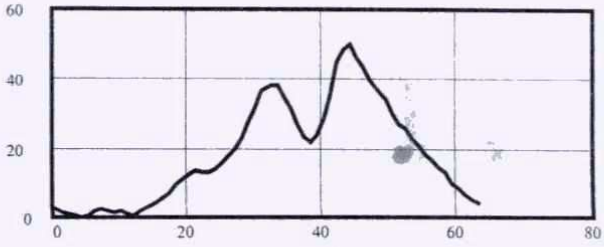


Рис. 8. Спотворений сигнал.

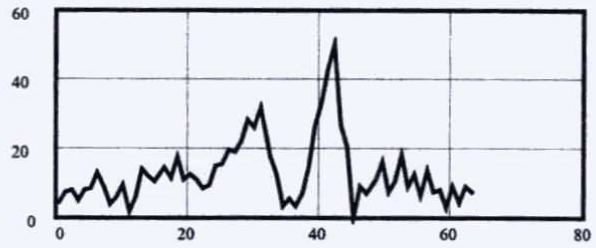


Рис. 9. Відновлення еволюційним фільтром.

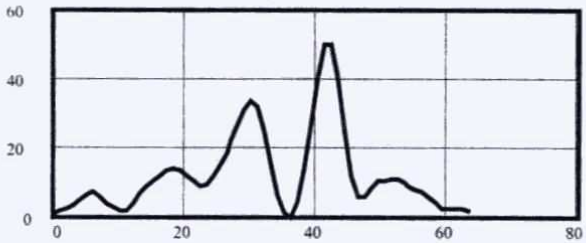


Рис. 10. Відновлення за методом умовної деконволюції.

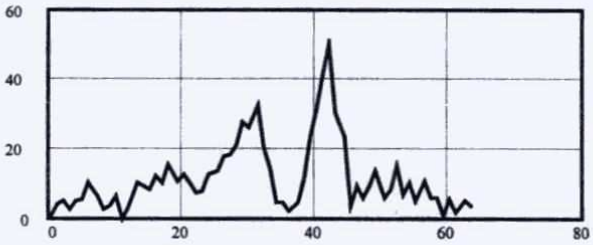


Рис. 11. Реставрація за методом локально-адаптивної умовної деконволюції.

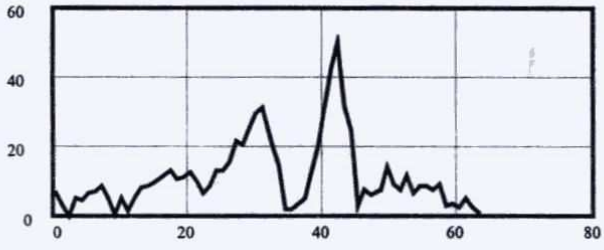


Рис. 12. Результат інверсної фільтрації.

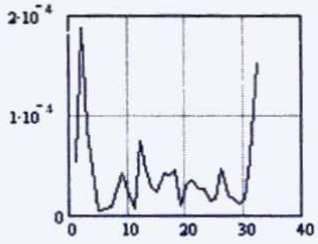


Рис. 13. Графік коефіцієнтів регуляризації λ(ω).

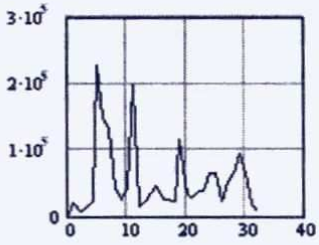


Рис. 14. Графік коефіцієнтів регуляризації α(ω).

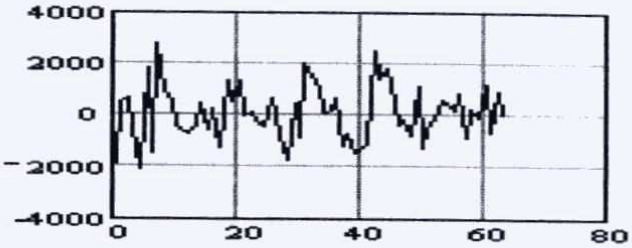


Рис. 15. Графік оберненого перетворення Фур'є для коефіцієнтів регуляризації α(ω).

ВИСНОВКИ

Описан новий алгоритм локально-адаптивної реставрації образів у частотній області. Запропановано нові способи побудови регуляризуючого оператора та визначення параметра локальної регуляризації. В результаті чого покращилась якість відновлення дрібних деталей спотвореного образа порівняно з широко відомими підходами. Описаний алгоритм відновлення зображень був реалізован на ПЕОМ PI-233 MMX. Час на реставрацію образа формата 181x217 завдяки використанню алгоритма швидкого перетворення перетворення Фур'є та теореми про згортку в частотній області склав менше однієї хвилини.

ЛІТЕРАТУРА

1. Рибін О.І.. Корольов В.Ю. Реставрація образів методом умовної деконволюції в області просторових частот //Вісник Технологічного університету Поділля.— 2000.— №1.— С.145-147.
2. Рыбин А.И., Королёв В.Ю. Алгоритм условной деконволюции изображений в частотной области //Радиоэлектроника. — 2000.— № 4.— С. 51-55. (Изв. высш. учеб. заведений).
3. Василенко Г.И., Тараторин А.М. Восстановление изображений. — М.: Радио и связь, 1986. — 304 с., ил.
4. Ярославский Л.П. Введение в цифровую обработку изображений. — М.: Сов. радио, 1979. — 312 с., ил.
5. Хемминг Р.В. Цифровые фильтры: Пер. с англ./ Под ред. А.М. Трахтмана. — М.: Сов. радио, 1980. — 224 с., ил.
6. Ярославский Л.П. Цифровая обработка сигналов в оптике и голографии: Введение в цифровую оптику. — М.: Радио и связь, 1987. — 296 с., ил.
7. Тихонов А.Н., Арсенин В.Я., Тимонов А.А. Математические задачи компьютерной томографии. М.: Наука. Гл. ред. физ-мат. лит., 1987.— Проблемы науки и технического прогресса.— 160 с., ил.
8. Виткус Р.Ю., Ярославский Л.П. Адаптивные линейные фильтры для обработки изображений // Адаптивные методы обработки изображений. /Под. ред. В.И. Сифорова, Л.П. Ярославского. — М.: Наука, 1988. — 248 с., ил. С. 6-35.

РОЗПІЗНАВАННЯ ОБРАЗІВ З САМОНАВЧАННЯМ ЗА ДОПОМОГОЮ ДЕТЕРМІНОВАНОЇ ФРАКТАЛЬНОЇ МОДЕЛІ

Г.М. Востров, М.В. Полякова, В.В. Любченко
ОДПУ, 65044, Україна, Одеса, пр. Шевченка, 1, (0482) 288-478, (0482) 685-152,
vvl@lmpo.ospu.odessa.ua

У представленій доповіді розглядається такий аспект структурного аналізу як спосіб побудови компонент алгоритмів автоматичної класифікації. Ми пропонуємо в якості компонент алгоритмів автоматичної класифікації використати отримані при рішенні задачі фрактального стиснення сигналу коефіцієнти систем ітерованих функцій.

1. ВСТУП

Останнім часом нарівні з Фур'є-аналізом широке поширення отримали методи, засновані на масштабних властивостях відліків часового ряду або його приростів, зокрема, циклічний аналіз, фрактальні методи і вейвлет-аналіз. Це пов'язано з тим, що спектральний аналіз часто використовується для знаходження домінуючих частот часового ряду, що розглядаються як гармонічні сигнали в сумі з випадковим шумом, який, як передбачається, має безперервний спектр. Однак, в реальних даних часто скрутно відділити гладкий сигнал від шуму, отриманим гармонікам не завжди можна дати змістовну інтерпретацію, а шум в реальних даних часто можна охарактеризувати масштабною інваріантністю над деяким інтервалом масштабів.

Однією з часто виникаючих прикладних задач є задача структурного аналізу випадкового процесу, під якою тут мається на увазі задача автоматичної класифікації станів випадкового процесу. У основу структурного аналізу випадкових процесів покладемо припущення, що якщо випадкова величина x на i-ом і j-ом інтервалах породжується одним і тим же комплексом умов, то відповідні ним два вектори стану відносно деякої метрики в початковому просторі будуть більш близькими, ніж в тому випадку, коли вони породжувалися двома різними комплексами умов. Звідси слідує, що задача структурного аналізу може бути зведена до задачі автоматичної класифікації станів випадкового процесу. У цьому випадку якість результатів, що отримуються, істотно залежить від вибору метрики, формалізації поняття класу і способу побудови компонент алгоритмів автоматичної класифікації.

У даній роботі розвивається останній аспект рішення задачі автоматичної класифікації. Ми пропонуємо використати отримані при рішенні задачі стиснення коефіцієнти систем ітерованих функцій (СІФ) в якості компонент алгоритмів автоматичної класифікації.

При фрактальному стисненні більшість СІФ використовують афіні перетворення. Однак, використання неафінних стискаючих відображень, наприклад, лінійного вейвлет-перетворення [2], дозволило досягнути кращого в порівнянні з афінними якості без істотного погіршення міри стиснення, та, отже, кращої концентрації енергії сигналу. Тому в якості компонент алгоритмів автоматичної класифікації пропонується використати коефіцієнти СІФ на основі вейвлет-розкладання, лінійного або нелінійного.

2. ПОБУДОВА СІФ НА ОСНОВІ ВЕЙВЛЕТ-МОДЕЛЕЙ

Згідно з теорією фрактального стиснення [1] часовий ряд представляють як фіксовану точку кінцевої множини стискаючих відображень СІФ в просторі непустих компактних підмножин евклідового простору. Часовий ряд розділяють на непересічні блоки, дочірні, і більші за розміром, можливо пересічні, батьківські. Відображення будуються між батьківським блоком після усереднення або децимації його до розмірів дочірнього і дочірніми. Об'єднання цих відображень називають образом. Можливість фрактальної апроксимації влаштовується в теоремі про образ [1], яка виражає реальну помилку декодування в термінах помилки образу (кодування).

У загальному випадку відображення часового ряду у вигляді даних являє собою множину точок вигляду $\{(x_i, F_i)\in R^2, i=0, 1, \ldots, N_x\}$, де $x_0<x_1<\ldots<x_{N_x}$, N_x $\{(x_i, F_i)\in R^2, i=0, 1, \ldots, N_x\}$, де $x_0<x_1<\ldots<x_{N_x}$, N_x — кількість відліків. Інтерполяційна функція, відповідна даній множині точок, безперервна функція f: $[x_0, x_{N_x}]\to R$ така, що $f(x_i)=F_i$ для $i=0, 1, \ldots, N_x$. Маючи множину $\{(x_i,F_i)\in R^2, i=0, 1, \ldots, N_x\}$, можна побудувати СІФ на R^2, атрактор якої — графік безперервної функції f, яка інтерполює відліки вхідних даних.

Нехай f визначена на відрізку $I=[0,1]$. f: $I\to R$, вважаємо, $f\in L^2(I)$. Метричний простір $L^2(I)$ з метрикою $d_{rms}(f,g)=(\int_I(f(x)-g(x))^2dx)^{1/2}=\|f-g\|_2$, де f, g —

елементи простору $L^2(I)$, повне, отже, задовольняє умові принципу стискаючих відображень. Будемо вважати областю значень f інтервал I, тому передусім f перетворюємо на [0,1] з використанням афіного перетворення.

Звичайно фрактальний оператор являє собою композицію геометричного перетворення між координатами x і власне фрактальної інтерполяційної функції між значеннями f. Согласно теоремі про образ, фрактальний оператор повинен задовольняти умові стисливості, а композиція двох відображень є стискаючим відображенням, якщо кожне вхідне в неї відображення є стискаючим. Тому геометричні перетворення і фрактальна інтерполяційна функція будуються таким чином, щоб вони були стискаючими. Область визначення f розіб'ємо на M непересічних дочірніх сегментів. Визначимо геометричні перетворення між батьківським і m-им дочірнім сегментом $w_m(x)=\alpha_m x+\beta_m$ з параметрами α_m, β_m, m=1, …, M, які децимують батьківський сегмент до розміру дочірнього. Щоб $w_m(x)$ були стискаючими, α_m вибираються з (0,1), m=1, …, M.

Основні математичні моделі, що використовуються в даній роботі для представлення часового ряду — лінійна і нелінійна вейвлет-моделі. Представлення часового ряду у вигляді

$$f(x)=\sum_k < f,\varphi_{j_0,k} > \varphi_{j_0,k}(x)+\sum_{j,k} < f,\psi_{j,k} > \psi_{j,k}(x), \quad (1)$$

будемо називати лінійною вейвлет-моделлю часового ряду. Відмінність нелінійної вейвлет-апроксимації від лінійної полягає в тому, що нелінійна вейвлет-апроксимація являє собою суму тільки тих базисних функцій з розкладання (1), коефіцієнти при яких вносять найбільший внесок в $\|f\|_{L^2}$, тобто мають найбільший модуль [3].

У загальному випадку вейвлет-розкладання містить нескінченне число рівнів (масштабів), яке при практичному застосуванні повинне бути обмежене. Нехай j_0 та j_{max} — мінімальний і максимальний рівні вейвлет-апроксимації. Зазначимо, що в основу запропонованої фрактальної інтерполяційної функції належить розкладання по ортонормованому вейвлет-базису $\{e_n\}$ на R. Для того, щоб визначити, які функції базису будуть мати ненульові коефіцієнти в розкладанні, необхідно знати область визначення f. При виборі у якості області визначення сегмента I ненульові коефіцієнти будуть мати тільки елементи базису, що належать множині $S=\{e_n: \mathrm{supp}(e_n)\cap(0,1)\neq 0\}$. Тут під supp розуміється носій, у разі вибору вейвлет-базису з компактним носієм, та δ-носій, у разі вибору вейвлет-базису з швидким розпадом. В дослідженні використовувалися вейвлети Добеші з мінімальною фазою 2-го і 3-го порядків [4]. В цьому випадку $S=\{\varphi_{j_0,k}: -2n+2\le k\le 2^{j_0}-1\}\cup\{\psi_{j,k}: j_0\le j\le j_{max}, -2n+2\le k\le 2^j-1\}$, где n — порядок вейвлету, функції $\varphi_{j_0,k}$ та $\psi_{j,k}$ отримані з стандартної скейлінгової функції φ і вейвлет-функції ψ з використанням співвідношення $f_{j,k}=2^{j/2}f(2^j\,.\,-k)$ [4]. Як фрактальну інтерполяційну функцію T_m, m=1, …, M, для $g\in L^2(I)$ можна використати одну з апроксимацій

$$T_m(g(x))=\sum_k \alpha_{j_0,k}\varphi_{j_0,k}(g(x)), \quad (2)$$

$$T_m(g(x))=\sum_{k\in E_N}\alpha_{j_0,k}\varphi_{j_0,k}(g(x))+\sum_{k\in E_N}\beta_{j,k}\psi_{j,k}(g(x)), \quad (3)$$

де (2) — СІФ на основі лінійної вейвлет-апроксимації скейлінговими функціями [2];

(3) — СІФ на основі нелінійної вейвлет-апроксимації [5];

E_N — множина індексів N найбільших по модулю коефіцієнтів в розкладанні по вейвлет-базису.

Нехай

$\tilde{g}_m(x)=g(w_m^{-1}(x))\,\chi_{A_m}(x)$ и $g_m(x)=g(x)\,\chi_{A_m}(x)$,

де A_m — m-ий дочірній сегмент, m=1, …, M;

$\tilde{g}_m$ — g після геометричного перетворення (децимації) з носієм A_m;

g_m — обмеження g до A_m;

$\chi_A(x)$ — характеристична функція множини A.

Тоді оператор $T(g(x))=\sum_{m=1}^{M}T_m(\tilde{g}(x))$ — апроксимація функції g, що дає множину відповідних коефіцієнтів стискаючого оператора T. Коеффіцієнти фрактальної інтерполяційної функції будемо шукати шляхом мінімізації $d(g,T(g))$, де d — метрика (наприклад, d_{rms}), при умові, що $T(g)$ — оператор стиснення (щоб задовольнити теоремі про образ [1]). Цю задачу можна звести до задачі знаходження коефіцієнтів функції $T_m(g)$ шляхом мінімізації $d(g,T_m(g))$ по відношенню до відповідних параметрів при умові стисливості оператора T_m.

Отримані коефіцієнти фрактальної інтерполяційної функції пропонується використати в якості компонент алгоритмів автоматичної класифікації при рішенні прикладних задач.

3. АВТОМАТИЧНЕ РОЗДІЛЕННЯ СИГНАЛУ ПРО СТАН МОРСЬКОЇ ПОВЕРХНІ

3.1. Опис об'єкта дослідження і початкових даних

У цей час основним способом отримання інформації про стан морської середи є використання методів радіолокації. У проведеному дослідженні використовувалися дані, сформовані в двовимірний (дальність, час (азимут)) масив комплексних огинаючих (квадратурних сигналів) адитивної суміші відображень від об'єктів, що досліджуються, які залежать від метеорологічних умов і параметрів радіолокаційних сигналів.

Метою розділення такого радіолокаційного сигналу про стан морської поверхні є виявлення формацій, характерних для об'єктів, що розрізнюються. Характерною особливістю радіолокаційний портрета є та обставина, що доступна для спостереження інформація реєструється одночасно по декількох каналах і може бути представлена лише у вигляді багатомірний експериментальної кривої. Структурні особливості радіолокаційний портрета як багатомірної кривої відображають положення кордонів між окремими об'єктами морської поверхні. У процесі обробки сигналу ми будемо спиратися на два основних припущення.

Перше полягає в тому, що залежність між потоками випадкових подій на окремих компонентах багатомірної кривої найбільш яскраво виражена на сусідніх компонентах і слабшає по мірі збільшення абсолютної величини різниці їх індексів. Друге припущення торкається характеру реакції кожної компоненти на чергову подію. Ми обмежимося тут найпростішою моделлю реакції на подію, розглядаючи кожну криву в інтервалі між двома подіями як однорідні випадкові процеси, зв'язуючи характер коливань з подією, що відкриває інтервал. Якщо умовно інтерпретувати значення сигналу як яскравість відповідних пікселів площини, то прийняті припущення дозволяють розглядати експериментальну криву як плоске зображення огинаючих відображених сигналів з кольоровим кодуванням амплітуди або фази.

Вейвлет-аналіз на площині досить складений як з математичної точки зору, так і з точки зору обчислювальної реалізації. Тому можна використати підхід, заснований на розгляді кожної радіолокаційної траси. Однак необхідно враховувати і наявність поперечної залежності відліків від комплексів умов, що визначають події. Тому виникає необхідність у використанні спеціальних механізмів, що дозволяють розглядати всі радіолокаційний траси одночасно. Один з можливих варіантів рішення — використання методів перетворення двовимірних полів даних в одновимірні. Якщо прочитати радіолокаційний портрет в ОЗП ЕОМ інформаційно-вимірювальної системи з файла в порядку зростання адресу, то буде отримана телевізійна розгортка по дальності. Тому при реалізації алгоритму розділення сигналу будемо використати розгортку цього типу.

3.2. Перевірка адекватності моделі реальним даним

Алгоритми фрактального стиснення засновуються при використанні довгострокових кореляцій між відліками сигналу. У контексті фрактального стиснення ці кореляції являють собою часткові самоподібності (можливо, узагальнені), які виникають через те, що фрагменти початкового сигналу можуть бути знайдені з використанням функціональної залежності між відліками цього сигналу. Відомо, що у разі хаотичної динаміки, коли можлива фрактальна апроксимація процесу, кореляції всередині процесу приводять до того, що експериментальна вибірка формує деяку множину розмірності меншої, ніж розмірність всього початкового простору. На цих фактах побудоване визначення фрактальної і інформаційної розмірності, що описує властивість масштабної інваріантності моменту першого порядку для приростів випадкового процесу

$$E[|X(t,\Delta t)|^q]=c(q)\,\Delta t^{\tau(q)+1},$$

де $X(t)$ — початковий випадковий процес на інтервалі $[0,T]$, $X(t,\Delta t)=X(t+\Delta t)-X(t)$, q вибирається з околиці одиниці.

Якщо дані задовольняють наступним вимогам: прирости стаціонарні і задовольняють властивості масштабної інваріантності принаймні для моменту першого порядку, то вони можуть бути описані за допомогою стохастичної фрактальної моделі, перевірка адекватності якої описується в [6]. Таким чином, для того, щоб можна було виконати фрактальну апроксимацію процесу, він повинен задовольняти вимогам стохастичної фрактальної моделі принаймні першого порядку. Отже, виходячи з цих емпіричних міркувань, перевірка адекватності детермінованих фрактальних моделей алгоритмів фрактального стиснення може бути виконана за допомогою процедури перевірки адекватності стохастичної фрактальної моделі для моменту першого порядку. Приведемо її короткий опис. Передбачаючи, що початковий випадковий процес $X(t)$ заданий на інтервалі $[0,T]$, розділимо його на N інтервалів довжини Δt і обчислимо

$$S_q(T,\Delta t)=\sum_{i=0}^{N-1}|X(i\Delta t,\Delta t)\Delta t|^q . \qquad (4)$$

Щоб перевірити (4), будуємо графіки залежності $\log S_q(T,\Delta t)$ від $\log\Delta t$ для різних q з околиці одиниці. Наприклад, ми вибирали q з інтервалу [0.5, 1.5] з кроком 0.2. Лінійність цих графіків для різних q і є тестом адекватності моделі. В нашому дослідженні лінійність оцінювалася за допомогою коефіцієнта детермінації. Для даних, що використовуються, він виявився в межах 0.87 — 0.99, що свідчить про адекватність моделі початковим даним.

3.3. Побудова компонент алгоритмів класифікації

Для рішення задачі автоматичного розділення сигналу про стан морської поверхні необхідно передусім виконати сегментацію сигналу на однорідні сегменти. Особливості в рішенні визначаються тим, що, по-перше, початкові дані представлені у вигляді векторної експериментальної кривої, і, по-друге, результат представляє матрицю (а не вектор) сегментації. Це приводить до необхідності введення підготовчого і заключного етапів. На підготовчому етапі виконується перетворення двовимірного радіолокаційного

портрету в одновимірну експериментальну криву. На заключному етапі вектор сегментації перетворюється в двовимірну матрицю. Відомо, що експериментальна крива являє собою нестаціонарний процес, що повільно міняється. Оскільки часто зустрічаємою особливістю сигналу є короткочасні високочастотні сплески (списи), найбільш ефективним на наш погляд для сегментації даних буде використання модифікації вейвлет-методу виявлення змін властивостей, описаної нами в [7]. Основна ідея методу полягає в застосуванні порога в процентах від початкової дисперсії сигналу до коефіцієнтів дискретного вейвлет-перетворення.

Перевага віддається вейвлет-методам виявлення змін виходячи з наступних міркувань. При використанні в основі фрактального оператора вейвлет-апроксимації для задоволення умовам стисливості [2,5] бажано, щоб значущих вейвлет-коефіцієнтів в розкладанні було як можна менше і вони були невеликими по абсолютній величині. Отже, відліки одного інтервалу сегментації повинні бути однорідними, оскільки якщо інтервал включає в себе особливість, з'явиться велике число значущих коефіцієнтів, які будуть нести інформацію про цю особливість. Таким чином, ефективне кодування при використанні вейвлет-моделей досягається у разі однорідності даних.

Вейвлет-методи визначення моментів зміни властивостей допускають швидкі обчислювальні алгоритми і є адаптивними. Вони володіють також деякою універсальністю, дозволяючи виявляти зміни декількох характеристик одночасно при досить загальних припущеннях відносно властивостей випадкового процесу, а також добре реагують на несподівані сінгулярності. Саме такими властивостями повинен володіти алгоритм сегментації часового ряду при рішенні задачі фрактального аналізу.

Внаслідок застосування вейвлет-методів визначення моментів зміни властивостей віипадкового процесу отримуємо сегментацію сигналу на однорідні дільниці різної довжини. Якщо використати в якості компонент алгоритмів автоматичної класифікації безпосередньо ці сегменти, це приведе до необхідності виконувати класифікацію в просторі змінної метрики. Щоб цього уникнути, в якості компонент алгоритмів класифікації можна використати коефіцієнти фрактальних інтерполяційних функцій на основі лінійної або нелінійної вейвлет-апроксимації. Побудова фрактальної інтерполяційної функції детально розглядається в [2,5] і стисло характеризувалася у другому розділі. Кількість коефіцієнтів фрактальних інтерполяційних функцій для сегментів різної довжини буде однаковою, якщо використати для кожного сегмента одні і ті ж рівні вейвлет-апроксимації. У разі нелінійної вейвлет-апроксимації коефіцієнти функцій, що не входять в розкладання, покладаються рівними нулю. Отримані вектора коефіцієнтів, що характеризують кожний сегмент часового ряду, класифікуємо відомими методами. Ми використали, наприклад, спочатку алгоритм "Спектр" для попередньої класифікації, а потім відомий алгоритм *k*-середніх [8].

4. ВИСНОВКИ

Застосування запропонованого підходу до радіолокаційних даних про стан морської поверхні підтвердило його працездатність на реальних даних, однак, нами отримані поки тільки попередні результати, запропоновані алгоритми вимагають подальшого дослідження. Основними подальшими напрямами досліджень можуть стати методи перевірки адекватності реальних даних фрактальної моделі, а також інтерпретація результатів класифікації.

ЛІТЕРАТУРА

1. Barnsley M., Fractals Everywhere, Academic Press, Boston, 1988.

2. Saeed Asgari, Nguyen T.Q., Sethares W.A. Wavelet-based approach to fractal image compression. //Proceeding of NORSIG. 1996. P. 323 326.

3. Cohen A., D'Ales J.-P., "Nonlinear approximation of random functions", SIAM J. Appl. Math., vol. 57, pp.518-540, April 1997.

4. Daubechies I., Ten Lectures on Wavelets, Philadelphia, PA: SIAM Press, 1992.

5. Востров Г.Н., Полякова М.В., Любченко В.В. Фрактальное сжатие одномерных временных рядов с использованием нелинейной вейвлет-аппроксимации // Тр. Одес. политехн. ун-та. — Одеса, 1999. — Вип. 3. — С. 194 —198.

6. Davis A., Marshak A., Wiscombe W. Wavelet-based multifractal analysis of non-stationary and/or intermittent geophysical signals //Wavelets in Geophysics, Efi Foufoula, Praveen Kumar (eds.): Academic Press, August, 1994. — P. 249 — 298.

7. Полякова М.В., Любченко В.В. Структурный анализ временных рядов со скачками среднего значения //Оптимизация управления, информационные системы и компьютерные технологии: Труды Украинской академии экономической кибернетики (Южный научный центр). — Київ-Одеса: ИСЦ, 1999. — Вип.1. — Ч.1. — С. 174 — 179.

8. Айвазян С.А., Енюков И.С., Мешалкин Л.Д. Прикладная статистика: Классификация и снижение размерности. - М.: Финанси та статистіка, 1985. – 487с.

ЗАСТОСУВАННЯ ТОМОГРАФІЧНОГО ПІДХОДУ В ЗАДАЧАХ ВІДНОВЛЕННЯ РАДІОЗОБРАЖЕНЬ КОСМІЧНИХ РАДІОДЖЕРЕЛ

В.В. Кошовий, А.Б. Лозинський, Б.А. Лозинський

Фізико-механічний інститут ім. Г.В. Карпенка НАН України, Львів 79053, вул. Наукова 5, тел. (0322) 65-48-74, факс (0322) 64-94-27, E-mail: Lozynsky@ah.ipm.lviv.ua

В роботі показується, що задача відновлення радіозображень може бути зведена до класичної томографічної задачі відновлення функції двох змінних по її радонівському образу. На відміну від вимоги квазімонохроматичності при відновленні радіозображень на основі вимірювання поперечної кореляційної функції, запропонований підхід дозволяє використовувати широкі смуги приймальних трактів.

Слід очікувати, що методи відновлення зображень на основі поздовжньої кореляційної функції будуть більш стійкі до перекручень та впливу завад, якими так багаті низькочастотні радіодіапазони.

1. ВСТУП

Сучасна томографія для отримання інформації використовує випромінювання найрізноманітнішої природи. Це і ультразвук, і радіо- та оптичні сигнали, рентгенівські та гама-промені, різного роду корпускулярне випромінювання та ін. Для кожного виду випромінювання характерні свої специфічні особливості, які проявляються в постановці томографічного експерименту та в його апаратній реалізації.

Дана робота присвячена огляду потенційних можливостей застосування томографічного підходу в радіоастрономічних задачах відновлення зображень космічних радіоджерел.

Традиційно в радіоастрономії при вирішенні задачі відновлення зображень космічних радіоджерел використовується теорема Ван Ціттерта-Церніке, яка пов'язує Фур'є-перетворенням кутовий розподіл радіояскравості джерела з поперечною кореляційною функцією [1]. Застосування цього методу в радіоінтерферометрії з наддовгою базою (РНДБ), особливо в декаметровому (ДМ) діапазоні, наштовхується на великі труднощі, пов'язані зі значними фазовими спотвореннями при проходженні радіохвиль через іоносферу Землі – так званою "фазовою проблемою".

Разом з тим, існує реальна можливість відновлення радіозображення вимірюючи часову взаємну кореляційну функцію для двох рознесених у просторі точок (поздовжня кореляційна функця) і система РНДБ добре вписується в рамки такого підходу [2].

Зауважимо, що радіоастрономічну науку цікавлять в першу чергу космічні об'єкти, відстань до яких дуже велика. Тому під відновленням радіозображення розуміється визначення кутового розподілу радіояскравості випромінювання, яке надходить з різних ділянок неба. Це зручно інтерпретувати як розміщення зображень космічних радіоджерел на поверхні сфери з одиничним радіусом – небесної сфери.

Далі будемо дотримуватись наступних домовленостей, які визначаються специфікою поставленої задачі, спрощують виклад, але зовсім не виключають можливості узагальнення:

– з уваги на віддалі, з якими маємо справу в радіоастрономії, вважаємо, що хвиля, яка падає на радіоінтерферометричну систему є плоскою;

– більшість цікавих для радіоастрономії джерел є некогерентними (випромінювання будь-якої пари світних елементів є некорельованим) і випромінюють в широкому спектрі частот. Реєструємо ж ми випромінювання в настільки вузькій (але відмінній від нуля) частотній смузі, що можна вважати спектр прийнятого сигналу плоским, однаковим як для всіх елементів джерела так і для джерела в цілому, а дисперсією в середовищі поширення можна знехтувати.

2. ПРОЕКЦІЙНІ ДАНІ

Будемо розглядати проекцію однієї з небесних півсфер на площину екватора – в радіоастрономії її прийнято називати площиною $u'v'$ [1]. Розглянемо радіоінтерферометр, утворений двома антенами, рознесеними у напрямку схід-захід. Зрозуміло, що при при обертанні Землі вектор бази

© UkrObraz'2000

радіоінтерферометра також обертається відносно нерухомих зір, причому не виходячи з площини, в якій описує коло. Ця площина паралельна площині екватора, за згаданого наближення плоскої хвилі їх можна сумістити. Геометрична затримка τ радіоінтерферометра при поширенні хвилі до двох антен однакова для всіх тих напрямків, які складають кут φ до вектора бази, і рівна $\frac{D}{c}\cos(\varphi)$ (рис. 1).

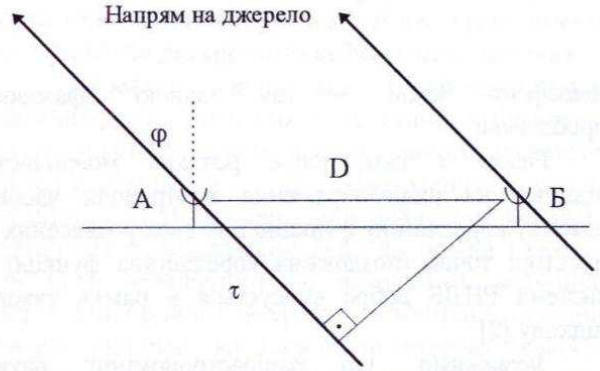


Рис. 1. Геометрична схема радіоінтерферометра.

Разом з тим, сукупність напрямків, які складають один і той же кут до певної осі, утворює конус (рис. 2). В нашому випадку вісь конуса співпадає з вектором бази, а лінія перетину конуса з небесною сферою є колом, центр якого лежить на продовженні вектора бази. Проекція цього кола на *u'v'*-площину є прямою, перпендикулярною вектору бази. Кожному куту розкриву конуса відповідає своя проекційна пряма.

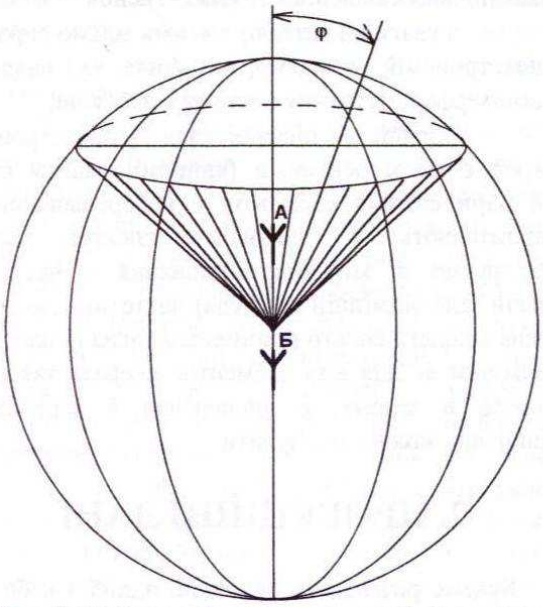


Рис. 2. Небесна сфера і конус, який відповідає певному значенню затримки τ.

Розглянемо тепер поздовжню кореляційну функцію – відгук інтерферометра АБ. Допустимо, що випромінювання елементів джерела, а, тим більше, різних джерел, є δ-корельованим [1, 3]. Тоді

$$R_{АБ}(t) = \int X_A(t) X_Б(t-t)dt ,$$

тобто значення $R_{АБ}(\tau)$ для конкретного значення τ можна отримати, як інтегральну інтенсивність випромінювання небесної сфери вздовж лінії, геометрична затримка для елементів якої становить τ. Оскільки геометрична затримка однозначно визначається кутом, утвореним вектором бази і напрямом на елемент випромінювання, ця лінія є колом – результатом перетину конуса з поверхнею кулі, коли вісь конуса проходить через центр кулі (рис. 2). Перейшовши до проекції на площину *u'v'* та врахувавши поворот вектора бази інтерферометра при обертанні Землі, відгук інтерферометра можна представити у вигляді

$$R_{АБ}(\tau, T) = \int f(\tau\cos T - y\sin T, \tau\sin T - y\cos T)dy ,$$

що являє собою проекцію функції розподілу радіояскравості $f(u', v')$ на пряму, яка відповідає годинному куту T. Таке проектування співпадає по своїй структурі з перетворенням Радона [3], а тому можна для відновлення функції $f(\alpha, \delta)$ на основі відомої $R(\tau, T)$ скористатися томографічними методами.

Підкреслимо, що ми ніяк не обмежували смугу частот приймального тракту інтерферометра. Якщо ж прийняти, що реєстрований спектр випромінювання по площі зображення джерела у всіх його точках однаковий, то поздовжня кореляційна функція є згорткою нормованої автокореляційної функції випромінювання елемента джерела з проекцією кутового розподілу яскравості на продовження вектора бази. Тому при переході від Радонівського до Фур'є-образу зображення отримується широкий трек, ширина якого пропорційна відношенню ширини смуги частот приймальної системи до центральної частоти прийому. При квазімонохроматичному наближенні такий трек вироджується в лінію – відповідає теоремі Ван Ціттерта-Церніке, а при нескінчено широкій смузі частот для відновлення зображення було б досить вимірювань на одній базі.

Опираючись на теорему про центральний переріз [4] можна представити взаємозв'язок поперечної та поздовжньої кореляційних функцій, місце томографічних методів у відновленні радіозображень. Як видно, якщо вимірювання поперечної кореляційної функції переводять зображення в Фур'є-простір, то поздовжньої – в простір Радона. Використання можливості відновлення зображення, крім очевидного – через Фур'є простір, зворотнім проектуванням, тобто з використанням сумарного зображення, представляє інтерес у плані пошуку можливостей подолання фазової проблеми. В Радонівському просторі вплив середовища поширення проявляється, в першу

чергу, у спотворенні параметра геометричної затримки τ, яке, при спробі перейти у простір Фур'є, відразу ж згубно позначається на фазі по причині її циклічності. Проаналізуємо можливості відновлення зображень іншими методами в умовах завад.

Враховуючи значне перекручення фази при малих амплітудних спотвореннях [1] можна запропонувати наступну модель реєстрованих даних

$$R_B(\tau,T) = R_{AB}\left(\tau + n(T),T\right),$$

де $n(T)$ - випадкова функція.

Тоді сумарне зображення

$$S(u',v') = \frac{1}{\pi}\int_0^{\pi} R_{AB}\left(u'\cos T + v'\sin T + n(T),T\right)dT .$$

Очевидно, шо з таких проекційних даних точне відновлення функції $f(x,y)$ здійснити неможливо. Тоді природньо задачу реконструкції переформулювати наступним чином: вимагається по наближених проекційнх даних знайти наближену функцію $f_B(x,y)$, яка в певному розумінні добре описувала б шукану функцію $f(x,y)$.

Для цього зауважимо, що випадкові зміщення проекційних даних приводять до розсіювання точки в сумарному зображенні. При представленні вихідного зображення у вигляді суми світних точок, образ кожної буде аналогічно розсіяний в сумарному зображенні, тобто сумарне зображення буде згорнуте з функцією розсіювання точки, яка залежить від кореляційних характеристик спотворюючого середовища, в нашому випадку – іоносфери, і може бути врахована.

Таким чином, випадкові зміщення проекційних даних вздовж координати τ еквівалентні просторовій фільтрації сумарного зображення фільтром з випадковим імпульсним відгуком $H(u',v')$. Цей факт підкреслює інтуїтивно відчутну можливість проведення самокалібрування, коли на шуканому зображенні є зосереджена яскрава деталь, як квазар у випадку, наприклад, Крабовидної туманності.

Накопичення сумарних зображень приводить до підвищення ваги регулярної і зниження ваги випадкової компоненти імпульсного відгуку $H(u',v')$, що дозволяє відновлювати наближення шуканого зображення шляхом фільтрації просторовим фільтром з імпульсним відгуком $\frac{1}{H(u',v')}$.

Отже, використання для задач відновлення радіозображень поздовжньої кореляційної функції дозволяє отримати проекційні дані – Радонівський образ проекції шуканого зображення на $u'v'$-площину. Застосування томографічних методів відновлення зображень дозволяє використати накопичення вхідних даних і понизити чутливість вимірювань до спотворення в середовищі поширення хвиль.

Принциповим обмеженням можливих областей застосування томографічного підходу є допущення про однаковий спектр випромінювання елементів джерела. У випадку вирішення радіоастрономічних задач це означає відновлення радіозображень об'єктів з достатньо рівномірним розподілом спектрального індекса. В іншому разі згадане допущення буде вимагати вибору такої ширини смуги пропускання приймальних трактів радіоінтерферометричної системи, щоб нерівномірністю спектру випромінювання по площі радіоджерела можна було б знехтувати. Враховуючи, що зі зміною частоти в рамках одного діапазону зображення радіоджерел змінюється повільно, таке обмеження не буде суттєво відбиватись на можливостях вибору об'єкта спостережень.

В загальному випадку під задачею відновлення радіозображення розуміється побудова зображення по виміряних значеннях двовимірного Фур'є-спектру просторових частот для різних баз. У випадку орієнтації бази радіоінтерферометра Схід-Захід використання обертання Землі дозволяє отримати значення просторового спектру вздовж дуги еліпса на площині баз. Вибираючи значення просторового спектру для різних довжин баз для певної взаємної орієнтації інтерферометра і джерела отримаємо центральний переріз Фур'є-образу шуканого зображення, який, згідно теореми про центральний переріз, пов'язаний одновимірним перетворенням Фур'є з відповідною паралельною проекцією зображення в радонівському просторі. Кожен такий переріз дозволяє знайти одну проекцію, а знаходження однієї проекції представляє собою одновимірний випадок відновлення, який і розглянемо докладніше.

3. ОДНОВИМІРНИЙ ВИПАДОК

Розглянемо задачу визначення кутового розподілу яскравості випромінювання джерела $f_k(\varphi)$ в одновимірному випадку. Представимо $f_k(\varphi)$ у вигляді розбиття на дрібні елементи, випромінювання від яких є некорельованим. Кожна з антен приймає сумарний сигнал від усіх елементів, причому для кожного елемента зі своєю різницею ходу до антен. Тоді взаємна поздовжня кореляційна функція являтиме собою суму автокореляційних функцій випромінювання елементів розбиття $f_k(\varphi)$.

Перейдемо від дискретного до неперервного представлення $f_k(\varphi)$. Сформулюємо та доведемо важливу теорему.

Теорема.

В одновимірному випадку поздовжня кореляційна функція є згорткою нормованої автокореляційної функції випромінювання

елемента джерела з функцією кутового розподілу яскравості.

Доведення.

Виходячи з геометрії задачі, аналогічно з вищеприведеними міркуваннями представимо функцію кутового розподілу $f_k(\varphi)$ у зручнішому для нас вигляді $f(\tau)$. Потужність випромінювання елемента джерела e_i з координатою τ_{ei} і розміром $\Delta\tau_{ei}$ можна записати так: $f(\tau_{ei})\Delta\tau_{ei}$. Позначимо через $r(\tau)$ нормовану автокореляційну функцію (однакову для всіх елементів). Тоді, можна записати:

$$R_{АБ}(\tau) = \lim_{\Delta\tau_{ei}\to 0}\sum_i R_{ii}(\tau-\tau_{ei}) =$$

$$= \lim_{\Delta\tau_{ei}\to 0}\sum_i r(\tau-\tau_{ei})f(\tau_{ei})\Delta\tau_{ei},$$

або при переході до інтегралу:

$$R_{АБ}(\tau) = \int_{джерело} r(\tau-\tau_e)f(\tau_e)d\tau_e. \qquad (1)$$

Оскільки функція $f(\tau)$ визначена тільки в принципово обмеженій $\pm\tau_{max}=D/c$ (з геометричних міркувань) області, покладемо, що $f(\tau)$ рівна нулю за її межами, що дозволить розширити межі інтегрування до $\pm\infty$. Легко зауважити, що в цьому випадку інтеграл в (1) являє собою згортку

$$R_{АБ}(\tau) = r(\tau)**f(\tau), \qquad (2)$$

що і треба було довести.

Наслідок.

Використовуючи сформульовану теорему про згортку ***з поздовжньої кореляційної функції можна визначити кутовий розподіл яскравості*** $f_k(\varphi)$ (через $f(\tau)$).

Позначимо $R_{АБ}(\tau)\Leftrightarrow G(\omega)$, $r(\tau)\Leftrightarrow g(\omega)$ та $f(\tau)\Leftrightarrow F(\omega)$, де знак $\Leftrightarrow$ означає Фур'є-перетворення, причому зауважимо, що $R_{АБ}(\tau)$ – вимірювана функція, $r(\tau)$ і $g(\tau)$ – наперед відомі. Тоді на підставі (2) можна записати:

$$G(\omega) = g(\omega)F(\omega),$$

$$F(\omega) = \frac{G(\omega)}{g(\omega)}. \qquad (3)$$

Залишається тільки отримати $f(\tau)$ з її Фур'є-образу $F(\omega)$ і задача розв'язана. Слід зауважити, що дана задача відноситься до класу некоректних обернених задач і при практичній реалізації розв'язку (3) необхідно застосовувати методи регуляризації.

Розглянемо деякі випадки різного вигляду $r(\tau)$.

1. Випадок нескінчено широкого рівномірного спектру випромінювання. В цьому разі кореляційна функція являє собою дельта-імпульс $r(\tau)=\delta(\tau)$ і $R_{АБ}(\tau)=\delta(\tau)**f(\tau)=f(\tau)$, тобто тривіальний випадок, вимірювана поздовжня кореляційна функція власне і є шуканою $f(\tau)$.

2. Випадок квазімонохроматичного наближення. Тоді $r(\tau)=\cos(\omega_0\tau)$, а $g(\omega)=\delta(\omega_0)$, де ω_0 визначається відношенням бази інтерферометра до довжини хвилі. В результаті застосування (3) можливо визначити тільки одну гармоніку $F(\omega_0)$ для однієї бази на одній частоті - по своїй суті це теорема ван Ціттерта-Церніке. Для відновлення зображення потрібні вимірювання на наборі баз.

3. Випадок, коли $r(\tau)$ - гаусівська крива, тоді $g(\omega)$ теж буде мати форму гаусівської кривої і, чим ширший спектр, то тим дальші від центру значення $G(\omega)$ можна ділити на $g(\omega)$ без втрати точності – ілюстрація (еквівалент) обмеження роздільчої здатності – вищі просторові частоти несуть інформацію про дрібніші деталі зображення.

4. ВИСНОВКИ

Тоді як задача відновлення зображення на основі поперечної кореляційної функції вимагає переходу до квазімонохроматичного наближення, тобто принципово не є точною, в термінах поздовжньої кореляційної функції можна точно оперувати інформацією в частотному континуумі.

На відміну від вимоги квазімонохроматичності при відновленні радіозображень на основі вимірювання поперечної кореляційної функції, запропонований підхід дозволяє використовувати широкі смуги приймальних трактів. Чим ширший спектр реєстрованого сигналу, тим вищу можна отримати роздільчу здатність і, що особливо важливо в радіоастрономії, тим вищу чутливість. Крім того цей метод вигідний можливістю компенсувати недостатню кількість баз інтерферометрів розширенням частотної смуги прийому. В принципі для відновлення радіозображення достатньо однієї довжини бази при використанні апертурного синтезу за рахунок обертання Землі.

Тому слід очікувати, що методи відновлення зображень на основі поздовжньої кореляційної функції будуть більш стійкі до перекручень та впливу завад, якими так багаті низькочастотні радіодіапазони.

ЛІТЕРАТУРА.

1. *Томпсон Р., Моран Дж., Свенсон Дж.* Интерферометрия и синтез в радиоастрономии: Пер. с англ. - М.: Мир, 1989. - 568 с.
2. *Кошовий В.В., Лозинський А.Б.* Відновлення радіозображень космічних радіоджерел на основі поздовжньої кореляційної функції // Відбір і обробка інформації. - 1998. - №12. - С. 37-41.
3. *С. А. Ахманов, Ю. Е. Дьяков, А. С. Чиркин.* Введение в статистическую радиофизику и оптику. - М.: Наука, 1981. - 640 с.
4. *Троицкий И.Н.* Статистическая теория томографии. - М.: Радио и связь, 1989. - 240 с.

Узагальнений підхід швидкого трансформування класу Фур'є на основі згорток

І.О. Процько, В.А. Радомський

Інститут конструювання моделювання швейних виробів, м. Хмельницький, вул. Львівське шосе, 43/2, тел. (0382)557799, e-mail:khikmshv@ifocom.khmelnitskiy.ua

Узагальнений підхід для обчислення дискретного трансформування Фур'є (ДТФ) послідовностей довільного обсягу на основі циклічних згорток одержано в даній роботі. Синтез швидкого трансформування Фур'є (ШТФ) базується на декомпозиції експоненціального базису (ЕБ) ДТФ на матрицю показників і матриці знаків синуса і косинуса. Доведено еквівалентність блочно-матричних структур матірці показників і відповідних матриць знаків. В результаті, ДТФ послідовностей довільного обсягу зводиться до обчислень циклічних згорток для синусної і косинусної частини ЕБ. Це спрощує індексацію послідовностей і синтез швидких алгоритмів дискретного трансформування Фур'є або Хартлі послідовностей довільного обсягу.

ВСТУП

Існують задачі аналізу інформаційних даних, що значно ефективніше розв'язуються в області спектрального представлення. Для цього використовують ШТФ вхідного масиву даних. ШТФ є потужним інструментом, що базується на ефективних алгоритмах реалізованих спеціальними програмними або апаратними засобами.

Теорія ШТФ почала свій розвиток з появою алгоритмів типу Кулі-Тюкі [1]. Особливого поширення алгоритми даного класу одержали для значень обсягів даних рівних цілій степені двійки. Але фізична природа багатьох явищ характеризується обсягом інформації, що може мати довільну кількість даних. Тому актуальним є довільний обсяг вхідних даних ефективного обчислення ДТФ масиву даних з кількістю відліків N (1), що належать множині натуральних чисел.

Виконання ДТФ послідовностей довільного обсягу, що проводиться засобами ШТФ з обсягами інформаційних даних рівних цілій степені двійки вимагає додаткових операцій. Вони полягають в зміні частоту дискретизації вхідного інформаційного сигналу в процесі аналізу. Збільшення числа відліків за рахунок збільшення частоти дискретизації при незмінному інтервалі аналізу сигналу вимагає змінення тривалості такту роботи аналого-цифрового перетворювача. При цьому практично складно отримати значення кількості вибірок рівного цілій степені два. В більшості випадків одержані вхідні дані доповнюють нулями до кількості, що дорівнює цілій степені двійки. Розгляд питання використання підходу доповнення нулями послідовності ДТФ до необхідної (рівної цілій степені два) кількості відліків функції, дозволяє отримати інтерпольоване перетворення більшого числа компонент функції ДТФ та підвищити точність оцінки частотних піків на спектральній характеристиці [2]. Але при збільшені обсягу, зростає величина обчислювальних затрат і, відповідно, час виконання ШТФ.

Для довільного обсягу N інформаційних даних, одержано різноманітні форми узагальненого ШТФ [3,4]. Існуючі підходи мають як свої переваги, так і особливості, що полягають в специфічній складності синтезу швидкого алгоритму для певних значень обсягів. Тому є актуальним дослідження і розвиток загальних підходів ефективного обчислення ДТФ послідовностей довільного обсягу, що дасть можливість створювати гнучкі та універсальні інструменти спектрального представлення інформації.

1. ПРИВЕДЕННЯ ОБЧИСЛЕННЯ ДТФ ДО ЦИКЛІЧНИХ ЗГОРТОК

Одним з підходів в теорії синтезу алгоритмів ШТФ є приведення обчислення до циклічних згорток. Тобто, алгоритми швидких згорток визначають ефективне обчислення ДТФ. Цей підхід ефективного обчислення ДТФ започатковано в роботі Рейдера [5]. ДТФ зводилось до обчислення циклічної згортки для послідовності обсягом, що дорівнює простому числу. В основу приведення ДТФ:

$$X(k)=\sum_{n=0}^{N-1} x(n)W^{nk}, \qquad (1)$$

де n,k=0,...N-1, W=exp(-j2π/N), до циклічної згортки покладено переіндексацію порядку вхідної послідовності x(n). Для цього застосовано первісний корінь g з відповідним показником степені, що зводить добуток індексів до додавання показників степені числа g. Тобто, обчислення ДТФ набуває виду:

$$X(g^l)=\sum_{m=0}^{N-2} x(g^{-m})W^{g^{(l-m)}} + x(0) \qquad (2)$$

© UkrObraz'2000

що відповідає циклічній згортці без x(0) [6].
При цьому враховується, що $W^{nk} = W^{(nk)modN} = W^{g\,l\,g\,m} = W^{g\,(l+m)} = W^{N}=1$ Числу g відповідає первісний корінь з властивостями $g^{N-1}=1$, $g^{k}\neq 1$, для $0<k<N-1$.

Дальше підхід розвинуто та узагальнено в роботах Винограда [7] для послідовностей обсягом, що дорівнює простому числу та степені простого числа, та одержано значення мінімальної мультиплікативної обчислювальної складності. В цих алгоритмах використовуються для певного обсягу конкретні обчислення з використанням китайської теореми про залишки, властивостей прямого добутку матриць та алгоритмів циклічної згортки. В роботі [8] розглянуто показники елементів матриці ЕБ ДТФ, що дорівнюють (nk)mod N, при n,k=1,2,...,N-1, і виконано відносно них переставлення (переіндексацію), з метою одержання циклічної зліва матриці показників степені ЕБ ДТФ. Це дало можливість більш зручніше проаналізувати структуру матриці показників ЕБ ДТФ для довільних значень обсягів.

2. ДЕКОМПОЗИЦІЯ ЕКСПОНЕНЦІАЛЬНОГО БАЗИСУ ДТФ

Проведемо декомпозицію та детальний аналіз елементів матриці ЕБ ДТФ (1). Функція W^{rk}_{N}, де r,k =0(1)N-1, періодична відносно N, тому, можна записати квадратну матрицю показників експоненціального базису порядку (N×N) у вигляді:

$$V=[v(r,k) \bmod N], \qquad (3)$$

де v(r,k) -значення елементів матриці по r=1(1)N-1 рядку та k=1(1)N-1 стовпцю. Враховуючи симетричність функцій ЕБ ДТФ W_N^l, де l=1(1)N-1, її можна повністю визначити для l=0(1)N/4, доповнюючи матрицями знаків Zs синуса і Zc косинуса [9], при цілих N, кратних 4-ом, і для парних N , що є дільниками 360 :

$$V=[v(r,k) \bmod (N/4)], \quad r,k=1(1)(N/4), \qquad (4)$$

для решти N, (непарних, парних не кратних 4 і парних не дільників 360) визначаємо при l=0(1){N/2} :

$$V = [v(r,k) \bmod (N/2)], \quad r,k=1(1)\{N/2\}, \qquad (5)$$

при N-непарному, береться ціла частина від ділення {N/2}. Порядок матриць (4,5) при цьому зменшується.

Матриці знаків синуса Zs і косинуса Zc формуються на основі елементів матриці (3) і визначаються за нерівностями:

$$Zs[r,k]=\begin{cases} +1, & \text{якщо } v(r,k)< N/2 \\ -1, & \text{якщо } v(r,k)> N/2 \ ; \end{cases} \qquad (6)$$

$$Zc[r,k]=\begin{cases} +1, & \text{якщо } 3N/4<v(r,k)<N/4 \\ -1, & \text{якщо } N/4<v(r,k)<3N/4 \ , \end{cases} \qquad (7)$$

де r,k=1(1)N/4 для (4) і r,k=1(1){N/2} для (5).

На рис. 1 показано розміщення симетрії значень матриць знаків синуса, косинуса, де для Zs - -антисиметричне відображення значень синуса; для Zc + - симетричне відображення значень косинуса; для Zs ▬ -антисиметричне відображення значень парних лінійок і парних стовпців матриці значень знаків синуса; для Zc ▬ - антисиметричне відображення значень непарних лінійок і непарних стовпців матриці значень знаків косинуса.

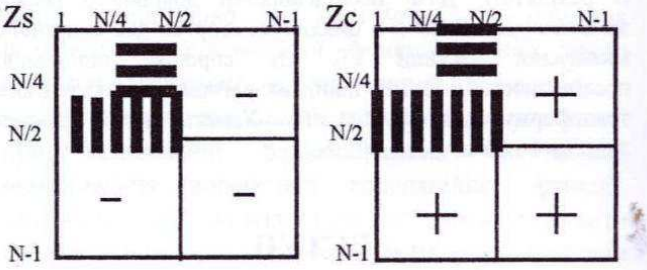


Рис.1. Симетричність розміщення елементів у матрицях знаків Zs синуса та Zc косинуса.

Аналіз одержаних матриць V (4,5) показників ЕБ ДТФ, що містять цілі числа, у випадку для N простого, показує, що матриця є нормалізований латинський квадрат. В інших випадках, довільних значень обсягів N, з матриці V (4,5) може бути виділений набір латинських квадратів, де кожен одержаний квадрат не містить елементів інших квадратів.

Визначення матриць показників ЕБ ДТФ у випадку довільного обсягу вимагає затрат часу при обчисленні за формулами (4,5) для кожного конкретного значення обсягу. Тому ефективніше формувати матриці показників за допомогою переставлення елементів кожного наступного рядка матриці на основі попереднього. Для цього знайдено форму підстановки, що визначається на основі елементів першого та другого рядка квадратної матриці показників. Підстановку можна описати за допомогою циклічного розкладу, який характеризується числом циклів циклічного розкладу $P(z)=(z_1)(z_2)...(z_i)$, що є конкретним для кожного довільного значення обсягу. Наприклад, для обсягу N=17, P(z)=(1 2 4 8)(3 6 5 7); N=30, P(z)=(1 2 4 7)(3 6) (5); N=44, P(z)= (1 3 9 5 7)(2 4 8 6 10); N=64, P(z)=(1 3 9 5 15 13 7 11)(2 6 14 10)(4 12)(8).

Отже, сформовані Vp матриці показників за циклічним розкладом підстановки P(z) визначають структури матриць показників степені ЕБ ДТФ послідовностей довільного обсягу і відповідні еквівалентно-структурні матриці знаків Z'c, Z's, які в сукупності становлять цілісність ЕБ ДТФ. Під еквівалентно-структурними матрицями розуміються матриці одного порядку з однаковими властивостями, що складаються з різних елементів, і містять відповідні підматриці з однаковими порядками та властивостями.

Розглянемо структуру, сформованих за циклічним розкладом підстановки, матриць показників Vp степені ЕБ ДТФ послідовностей довільного обсягу та відповідних матриць знаків Z's синуса і Z'c косинуса, що містять елементи ±1. Кожен одержаний квадрат, що побудований на основі підстановки z_i, з циклічного розкладу P(z), містить елементи рівні між собою розміщені паралельно бічній діагоналі v[i,j]=v[k,l], при i+j=k+l, де i,j,k,l ∈{1,2,..., n}, v[i,j] ∈Vp або рівні всі елементи симетрично розміщені відносно головної діагоналі.. Тобто, такі квадратні матриці називаються Ганкелевими (Hankel), що повністю визначаються своїм першим рядком та останнім стовпчиком. Окрім того, у даної Ганкелевої матриці Vp кожен наступний рядок одержаний з попереднього циклічним зсувом вліво. Тобто, Ганкелева матриця повністю визначається своїм першим рядком і такі матриці називають ліво-циркулянтними або циклічними зліва.

Можна сформулювати на основі декомпозиції ЕБ ДТФ (3-7) для відповідних косинусних і синусних частин теорему еквівалентності структур матриці показників степені і матриці знаків.

Теорема. Для квадратних матриць показників степені (5) і знаків косинусу (7) порядку {N/2}×{N/2}, обчислених на основі періодичності та симетрії ЕБ ДТФ, а для матриці показників степені (3) і відповідної матриці знаку синуса (6) порядку {N×N}, обчислених на основі періодичності, та переставленні рядків і стовпців за допомогою підстановок з циклічного розкладу, існує еквівалентність структур для довільного обсягу N.

Доведення. Для взаємозв'язку елементів декомпозиції ЕБ ДТФ необхідні аналогічні дії по переставленню рядків і стовпців, у відповідність до матриці Vp, над матрицями (6,7) знаків Z's синуса і Z'c косинуса. Так, як значення величин cos(2πn/N) для n=0(1)N-1 експоненціального базису ДТФ, рівномірно розподілених на відрізку (0,2π), симетричні відносно осі N/2, як для абсолютних величин значень косинуса, так і для їх знаків. Тому достатньо для косинусної частини розглядати елементи декомпозиції Vp та Z'c і одержані їхні властивості для порядку {N/2}×{N/2}, на відміну від антисиметричної функції sin(2πn/N) для n=0(1)N-1 відрізка (0,2π) властивості еквівалентності структур, якої вірні для порядку {N×N}, обчислених на основі періодичності даної функції.

3. ОБЧИСЛЕННЯ ДТФ НА ОСНОВІ ПРОВЕДЕНОЇ ДЕКОМПОЗИЦІЇ

Одержана структура матриць показників степені Vp і відповідних матриць знаків задає процедуру (рис.2) проведення обчислення ДТФ послідовностей довільного обсягу за допомогою циклічних згорток. При цьому необхідно застосовувати швидкі алгоритми згорток [6], що і визначатимуть основні обчислювальні затрати даного трансформування.



Рис.2. Обчислення ДТФ послідовностей довільного обсягу.

На початковому етапі проводиться групування (переіндексація) вхідних даних, що визначається циклічним розкладом підстановки. Одночасно можуть визначатись коефіцієнти косинусних і синусних складових ЕБ ДТФ, що приймають участь в операціях згортки. Для косинусної частини проводиться групування вже об'єднаних вхідних даних:

$$x_c(i) = x(i) \pm x(N-1), \qquad i=1(1)\{N/2\}-1, \quad (8)$$

що може виконуватись в один-три етапи. На основі сформульованої теореми за одержаним циклічним

розкладом підстановки матриця показників степені Vp ЕБ ДТФ послідовностей довільного обсягу і відповідна матриця знаку Z'c, Z's, розбиваються на еквівалентні блочно-матричні структури. Вони можуть визначати додаткові об'єднання даних та обсяг і кількість циклічних згорток необхідних для проведення обчислення. Так, варто зауважити, для парних значень обсягів перетворення ДТФ матриця показників степені Vp для косинусної і синусної частин складатиметься з підматриць Vo, Ve і матиме вид:

$$Vp = \begin{bmatrix} Vo & Ve \\ Ve & Ve \end{bmatrix}, \qquad (9)$$

що вимагає додаткових об'єднань даних. Важливо, що для синусної частини ЕБ ДТФ порядку (N×N) циклічні зліва підматриці містять асиметричну блочно-матричну структуру другого порядку виду (A -блочний елемент):

$$V'p = \begin{bmatrix} A & -A \\ -A & A \end{bmatrix}. \qquad (10)$$

Ці основні особливості структур узагальненого підходу проведення обчислення ДТФ на основі циклічних згорток, звичайно, не перекривають всього розмаїття конкретних матричних структур для ДТФ послідовностей довільного обсягу. Тому ці особливості, за виконаним аналізом для цілих значень обсягів N, можна виділити в підмножинах S1, S2, S3, S4, S5, S6 множини натуральних чисел.

S1={3,5,7,9,11,13,15,17,...}- множина непарних значень обсягів, а множина парних значень поділяється на підмножини:

S2={6,10,18,30,90,...},

S3={14,22,26,34,38,42,46,50,58,...},

S4={12,20,28,36,44,52,60,...}.

S5={2,4,8,16,32,64,128,256,...}.

S6={24,40,48,56,72,80,88,96,104,...}.

Наприклад, для обсягів, що належать підмножині S5, елементи якої дорівнюють цілій степені двійки, легко подати кількість згорток s для косинусної частини на основі даного підходу у вигляді таблиці 1.

Таблиця 1.

N	8	16	32	64	128	...
s	1	4,1 1	8,4,1 4,1 1	16,8,4,1 8,4,1 4,1 1	32,16,8,4,1 16,8,4,1 8,4,1 4,1 1	...

На завершальному етапі на основі одержаної матричної структури, об'єднуємо з попередніх етапів одержані значення згорток і компонуючи дійсні та уявні частини, проводимо формування вихідних значень дискретних коефіцієнтів Фур'є або Хартлі.

ВИСНОВКИ

В роботі показано, як на основі переставлення елементів вхідної послідовності, що визначається на основі перших двох рядків матриці показників (3), обчислюється ДТФ за допомогою алгоритмів згортки. Визначення переставлення не потребує спеціальних обчислень і для обсягів N, N/2 - простих чисел може формуватись елементарним відбором елементів натурального ряду. Використання циклічного розкладу підстановки P(z) приводить до однотипового підходу проведення організації обчислення ДТФ послідовностей довільного обсягу. Окреме проведення обчислень для косинусної та синусної частин ЕБ дозволяє розпаралелювати процес обчислення і визначати не тільки ДТФ, але і дискретне перетворення Хартлі послідовностей довільного обсягу.

ЛІТЕРАТУРА

1. Cooley, J.W., J.W. Tukey, An Algorithm for the Machine Computation of Complex Fourier Series, Math. Comp. 19, 1965, p.297-301.

2. Отнес Р., Эноксон Л. Прикладной анализ временных рядов. Основные методы: Пер. с англ - М.: Мир, 1982. -428 с.

3. Ryszard Stasinski, The techniques of the Generalized Fast Fourier Transform Algorithm IEEE Transactions on Signal Processing, vol.39, NO. 5, MAY 1991, p.1058-1069

4. Сверлик М.Б., Троянский А.В. Обобщенное матричное описание алгоритма БПФ. Радиоэлектроника №7, 1995, с.27-33

5. C.M.Rader, Discrete Fourier transform when the number of data samples is prime. Proc.IEEE 56, 1968, p.1107-1108.

6. Макклеллан Дж. Х., Рейдер Ч. М. Применение теории чисел в цифровой обработке сигналов: Пер. с англ.-М.: Радио и связь, 1983. -264с.

7. S. Winograd, On cumputing the discrete Fourier transform. Math.Comput. 32, 1978, p.175-199.

8. Ш. Зохар, Алгоритм Винограда для дискретного преобразования Фурье. В кн.: Быстрые алгоритмы в цифровой обработке изображений/ Под. ред. Т.С. Хуанга. Пер. с англ. -М.:Радио и связь, 1984. -224 с.

9. Prots'ko I.E. An algorithm for computing the discrete Fourier transform.//Proceedings of LSPIC-90. -Riga, april 22-26, -1990. -P.123-127.

ДВОВИМІРНЕ ПЕРЕТВОРЕННЯ ФУР'Є З КОМПЛЕКСНОЮ ФАЗОЮ

Микола Будник

Інститут кібернетики НАНУ, просп. Акад. Глушкова 40, Київ 03680, Україна
Тел: (044) 266-1267, Факс: 266-3348, Електронна пошта: budnik @ d220.icyb.kiev.ua

ABSTRACT: In general, report is devoted to development of a theory of Fourier Transform (FT) for multidimensional signal domain. In particular, a new approach for the two-dimensional (2D) FT is proposed. Essence of above approach consists in replacing of the single (scalar) phase to two phases which are proportional to Cartesian co-ordinates of the FT garmonics onto the 2D signal plane. It was shown that the proposed FT have hipercomplex nature and the two orthogonal phases one can interpret as the vector-phase so as components of complex number. It was introduced the 22 so-called "derivative" spectra, which permit to analyze 2D imaging in a detail way. Developed results are useful for the spatial filtration of 2D imaging and simulation of appropriate technical systems as a 2D complex spatial filters.

ВСТУП

Багато прикладних задач зводиться до аналізу двовимірних функцій. Сюди відносяться задачі прикладної оптики [1], розпізнавання зображень [2] та оброблення двомірному інформації довільної природи. Відомо, що одним із розповсюджених методів аналізу є інтегральне перетворення Фур'є (ПФ). У найбільш загальній комплексній формі для одно - та багатовимірного випадків ПФ має вигляд [3,4]

$$C(k)=(1/2\pi)\int f(x)\exp[-ikx]dx=A(k)\exp[-i\varphi(k)], \quad (1)$$

$$S(x)=\int C(k)\exp[ikx]dk=\int A(k)\exp[i\{kx-\varphi(k)\}]dk, \quad (2)$$

$$C(K)=(1/2\pi)^N\int f(X)\exp[-i<KX>]dX, \quad (3)$$

$$S(X)=\int C(K)\exp[i<KX>]dK. \quad (4)$$

В операторній формі (1,3) та (2,4) можна записати у вигляді $C=F(f)$ та $S=F^{-1}(C)$, де C, S - фур'є-спектр та відновлений сигнал, F та F^{-1} оператор ПФ та оберненого ПФ (ОПФ). В (3,4) вхідна змінна та хвильове число (частота) є векторами $X=\{x_p\}$, $K=\{k_p\}$, $p=1,...N$; $<KX>$ - їх скалярний добуток, $\int dX(K)$ - N-кратні інтеграли $\int...\int dx_1(k_1)...dx_N(k_N)$. Відмітимо, що (1-4) описують математичний спектр, визначений як для додатних (реальних) частот, так і для від'ємних $k_p\in]-\infty,\infty[$ на нескінченній області визначеності $x_p\in]-\infty,\infty[$. Вирази (1,3) для оберненого ПФ (ОПФ) є розв'язок оберненої задачі. Для її розв'язання, тобто відновлення функції f(x), необхідно знати комплексний спектр C(k). З другого боку ОПФ можна розглядати як оператор, що переводить амплітудний спектр у вихідний сигнал $G^{-1}(\varphi):A\rightarrow f$ і залежить від фази як від параметра. Таким чином, одномірне ОПФ має один параметр. Обгрунтуємо, що у N-мірному випадку оператор повинен залежати від N параметрів.

Для цього необхідно зрозуміти, чому ПФ одновимірної функції має два параметра – амплітуду та фазу. Причина полягає в тому, що графік цієї функції є підмножиною прямого добутку області визначення та множини значень $G(f)\subset X\times Y$. При $X\subset R$, $Y\subset R$, маємо $G\subset R\times R=R^2$, тобто графік є підмножиною двовимірного простору. Тому одновимірне ПФ повинно мати два параметри, а оператор ОПФ, що переводить $A\rightarrow f$ - один параметр. Аналогічно для N-вимірного випадку маємо $X\subset R^N$, $Y\subset R$, та $G\subset R^{N+1}$. Тому ПФ повинно описуватися N+1 параметрами, а оператор ОПФ, що переводить $A\rightarrow f$ – N параметрами.

1. ПОСТАНОВКА ПРОБЛЕМИ

Зрозуміло, що такими додатковими параметрами повинні бути додаткові фази. Проте N-вимірне ПФ має всього лише одну фазу $\varphi(K)$. Її вигляд можна встановити, виділивши явно в комплексному спектрі, аналогічно одновимірному випадку

$$C(K)=A(K)\exp[-i\varphi(K)]. \quad (5)$$

Підстановка (5) в (4) дає вигляд сумарної фази ОПФ

$$S(X)=\int A(K)\exp[i\varphi_\Sigma]dX, \quad \varphi_\Sigma=<KX>-\varphi(K). \quad (6)$$

Якщо x одновимірна і має смисл часу t, то $k=\omega$ та $\varphi=\omega t_0$, де t_0 має смисл часу затримки гармоніки. Тоді у багатовимірному випадку запропонуємо

$$\varphi(K)=<KX_0>=\Sigma\varphi_p, \qquad \varphi_p=k_px_{op}, \quad (7)$$

де φ_p можна інтерпретувати як фазовий спектр, що відповідає p–ій координаті вхідного сигналу. Таким чином, в рамках існуючого багатомірного ПФ, що описується амплітудою та скалярною фазою, у фур'є-області відбувається втрата інформації. Для адекватного опису вхідного сигналу такому ПФ недостає N-1 ступеня вільності, яким відповідають додаткові параметри. Шлях до розробки ПФ, при якому не відбувається втрата інформації, полягає у переході від скалярної фази до векторної. Зрозуміло, що таку

© UkrObraz'2000

векторизацію фази легко провести на основі (7), вважаючи величини φ_p проекціями деякої вектор-фази Ф. Метою роботи є розробка ПФ з такою "векторною" фазою для N=2.

2. СТАНДАРТНЕ ПФ

У випадку періодичних функцій ПФ зводиться до ряду. Функція f(X) задана на *N*-мірному кубі $x_p \in [-L_p, L_p]$. Далі будемо користуватися нормованими величинами $k_{op}x_p$ k_p/k_{op}, де $k_{op}=\pi/L_p$ – характерна частота по *p*–ій координаті. Тоді при p=1,2, $x_1=x$, $x_2=y$, $k_{0x}=k_{01}$, $k_{0y}=k_{02}$ маємо $x=k_{0x}x_1$, $y=k_{0y}x_2$, $k_x=k_1/k_{ox}=n$, $k_y=k_2/k_{oy}=m$, а $(x,y)\in[-\pi,\pi]$. Фізичний спектр $n,m\in[1, \infty[$, в одновимірному випадку має вигляд [4]

$$S(x)=a_0/2+\Sigma A_n\sin\{nx+\varphi_n\},\quad A_n^2=a_n^2+b_n^2,\quad tg\varphi_n=a_n/b_n, \quad (8)$$

$$a_n=(1/\pi)\int f(x)\cos(nx)dx,\quad b_n=(1/\pi)\int f(x)\sin(nx)dx, \quad (9)$$

$$C_n=(1/\pi)\int f(x)\exp[-inx]dx=A_n\exp[-i\varphi_n]=a_n-ib_n, \quad (10)$$

$$S(x)=a_0/2+\Sigma C_n\exp[inx]. \quad (11)$$

У багатовимірному випадку приводиться вираз для математичного спектру при $n,m\in]-\infty,\infty[$ [3]

$$S(x,y)=\Sigma\Sigma C_{nm}F^{-1},\quad F^{-1}=1/F=\exp[i(nx+my)], \quad (12)$$

$$C_{nm}=(1/2\pi)^2\iint f(x,y)F dxdy,\quad F=\exp[-i(nx+my)], \quad (13)$$

На практиці, як правило, розглядають чотири коефіцієнти Фур'є, одержані в результаті розкладу спектру (13) по добутках косинусів і синусів [5].

$$4\pi^2a_{nm}=\iint f(x,y)\,g_{11}dxdy,\ 4\pi^2b_{nm}=\iint f(x,y)\,g_{12}dxdy, \quad (14)$$

$$4\pi^2c_{nm}=\iint f(x,y)\,g_{22}dxdy,\ 4\pi^2d_{nm}=\iint f(x,y)\,g_{22}dxdy, \quad (15)$$

$$S(x,y)=\Sigma\Sigma[a_{nm}g_{11}+b_{nm}g_{12}+c_{nm}g_{21}+d_{nm}g_{22}], \quad (16)$$

$$g_{11}=\cos(nx)\cos(my),\ g_{21}=\cos(nx)\sin(my), \quad (17)$$

$$g_{12}=\sin(nx)\cos(my),\ g_{22}=\sin(nx)\sin(my). \quad (18)$$

При цьому всі амплітудні спектри (a_{nm} - косинус-спектр, d_{nm} – синус-спектр; b_{nm} і c_{nm} - крос-спектри) дійсні, а фаза ігнорується. Такі парціальні спектри не мають прозорого фізичного сенсу, а тому використовується сумарний спектр (19).

$$A_{nm}^2=a_{nm}^2+b_{nm}^2+c_{nm}^2+d_{nm}^2. \quad (19)$$

Враховуючи комплексний характер ПФ (13), $F=g_{11}-g_{22}-i(g_{12}+g_{21})$ одержимо, що спектр дорівнює

$$C_{nm}=\Lambda\exp[-i\gamma],\quad tg\,\gamma=(b_{nm}+c_{nm})/(a_{nm}-d_{nm}), \quad (20)$$

$$\Lambda=[A_{nm}^2+2(b_{nm}c_{nm}-a_{nm}d_{nm})]^{1/2}.$$

Проте такий підхід пов'заний із двома проблемами По-перше, безпосередній фізичний смисл фази γ невідомий, а по-друге – амплітуда Λ, яка відповідає такій фазі, відрізняється від (19). (Причина таких проблем полягає в тому, що в дійсності двовимірне ПФ не може бути адекватно описане комплексними числами (див. П.7.)). Тому автору невідомі приклади використання фази у випадку двовимірного (тим більше багатовимірного) ПФ.

Зідно наших уявлень (7), така фаза повинна бути сумою двох фаз-проекцій

$$C_{nm}=A_{nm}\exp[-i\varphi_{nm}],\qquad \varphi_{nm}=\varphi_n+\varphi_m, \quad (21)$$

$$\varphi_n(x)=k_{0X}x_{0n},\quad \varphi_m(y)=k_{0Y}y_{0m}, \quad (22)$$

де $\varphi_n(x)$ та $\varphi_m(y)$ - відповідні фази-проекції, що описують положення nm-ї гармоніки на площині. Далі будуть одержані вирази (40,41), які показують, що дійсно $\gamma\equiv\varphi_{nm}$ та $\Lambda\equiv A_{nm}$. Але для повного описання двовимірного ПФ однієї амплітуди та однієї фази недостатньо. Таким чином, за незалежні параметри (тобто спектри), що підлягають визначенню і мають фізичний сенс, необхідно взяти амплітуду A_{nm} та ортогональні фази $\varphi_n(x)$, $\varphi_m(y)$ відносно вісей прямокутної декартової системи координат на площині.

3. ВВЕДЕННЯ 2-ВИМІРНОГО ПФ

Спочатку покажемо, що вираз для одновимірного фізичного спектру (8), що використовується на сьогодні, не зовсім коректний. Для цього необхідно перейти до математичного спектру

$$S(x)=\Sigma A'_n\sin\{nx+\varphi_n\},\quad a'_n=a_n/2,\ b'_n=b_n/2. \quad (23)$$

і записати вираз для амплітуди 0-ї гармоніки відновленого сигналу $S_0(x)=A'_0\sin\varphi_0$. Проте з точки зору ПФ постійна складова вихідного сигналу визначається тільки відповідним коефіцієнтом Фур'є, тобто $f_0(x)=A'_0$. Для того, щоб узгодити обидва вирази $S_0(x)=f_0(x)$ необхідно покласти $\varphi_0=\pi/2$. Останнє означає, що нульова гармоніка повина мати фазу (=$\pi/2$), що не має фізичного сенсу. Реально, як це добре видно з (8) або (9), постійна складова є чисто дійсною, тобто їй відповідає нульова фаза. Вихід із цього протиріччя полягає у переході до нової фази ψ, що задовільняє умову $\psi_0=0$. В результаті отримаємо

$$S(x)=\Sigma A'_n\cos\{nx+\psi_n\},\quad \psi_n=\pi/2-\varphi_n,\quad tg\psi_n=b'_n/a'_n. \quad (24)$$

Причина такого протиріччя полягає в тому, що тангенс кута φ (8) є відношенням дійсної частини до уявної частини комплексного спектрального коефіцієнта, отже φ є кутом між останнім і його уявною частиною. Це приводить до того, що рівність (9) $C_n=A_n\exp[-i\varphi_n]=a_n-ib_n$, є формальною тому, що строго маємо $Re\{C_n\}=b_n$, $Im\{C_n\}=-a_n$. Заміна (24) дає кут ψ між комплексним спектром і його дійсною частиною, що відповідає формалізму комплексних чисел

$$C=A\exp(i\psi),\ a=Re\{C\}=A\cos\psi,\ b=Im\{C\}=A\sin\psi. \quad (25)$$

Для подальшого використання знову перейдемо до фізичного спектру, де для уніфікації врахуємо очевидне співвідношення $a_0=A_0$

$$S(x)=A_0/2+\Sigma A_n\cos\{nx+\psi_n\},\quad tg\psi_n=b_n/a_n. \quad (26)$$

Для обчислення двовимірного ПФ подамо двовимірну функцію у вигляді добутку двох одновимірних

$$f(x,y)=f(x)g(y). \quad (27)$$

Можливість такого представлення у загальному випадку грунтується на властивостях матриць. З точки зору теорії матриць, двовимірна функція, задана в дискретних відліках, є матрицею, а одновимірна

- вектором. Відомо [6], що кожна матриця може бути представлена прямим добутком двох векторів. Неважко переконатись, що йому відповідає добуток двох функцій, отже і кожна двовимірна функція може бути подана як добуток одновимірних.

При підстановці (27) у (12,13) ядра ПФ та ОПФ вироджуються, тобто подвійні інтеграл та ряд розщеплюються на добутки одинарних, що дає

$$S(x,y) = S(x)S(y), \quad C_{nm} = C_nC_m. \qquad (28)$$

Щоб отримати явний вираз для S(x,y), підставимо в (28) вирази для одновимірних ОПФ у коректному вигляді (25), тоді формально маємо

$$S(x,y)=S_0+S_X+S_Y+S_{XY}, \quad S_0=A_{0X}A_{0Y}/4, \qquad (29)$$
$$S_X=(A_{0Y}/2)\Sigma A_n\cos\{nx+\psi_n\}, \quad tg\psi_n=b_n/a_n, \qquad (30)$$
$$S_Y=(A_{0X}/2)\Sigma A_m\cos\{my+\psi_m\}, \quad tg\psi_m=b_m/a_m, \qquad (31)$$
$$S_{XY}=\Sigma\Sigma A_nA_m\cos(nx+\psi_n)\cos(my+\psi_m). \qquad (32)$$

З (29-32) випливає, що двовимірний спектр складається з 4-х парціальних компонент: постійної складової S_0, двох одновимірних спектрів S_X, S_Y та власне двовимірного S_{XY}. Важливо відмітити, що парціальні одновимірні спектри не співпадають зі звичайними одновимірними, а власне двовимірний спектр не має смислу кореляційного спектру.

Проте, двовимірний спектр у вигляді (9,29-32) не може нас задовольнити тому, що він описується 4-ма параметрами – двома амплітудами та двома фазами. Така кількість завелика, незалежних параметра повинно бути тільки три, а тому необхідно шукати рівняння зв'язку між ними. Для цього розкладемо (32), використовуючи формули для косинуса від двох аргументів. В результаті одержимо, що власне двовимірний спектр має вигляд (19) з коефіцієнтами

$$a_{nm}=A_nA_m\cos\psi_n\cos\psi_m, \quad b_{nm}=A_nA_m\sin\psi_n\cos\psi_m, \qquad (33)$$
$$c_{nm}=A_nA_m\cos\psi_n\sin\psi_m, \quad d_{nm}=A_nA_m\sin\psi_n\sin\psi_m, \qquad (34)$$

або, враховуючи (25)

$$a_{nm}=a_na_m, \quad b_{nm}=b_na_m, \quad c_{nm}=a_nb_m, \quad d_{nm}=b_nb_m. \qquad (35)$$

Вирази (33-35) можна одержати також в результаті підстановки (25) у (28)

$$C_{nm}=A_nA_m\exp[i(\psi_n+\psi_m)]=a_{nm}-d_{nm}+i(b_{nm}+c_{nm}). \qquad (36)$$

4. ЗАГАЛЬНИЙ ВИРАЗ

Комбінування виразів (33) дають можливість ввести нові двовимірні коефіцієнти Фур'є A, B, C, D, що зв'язані з коефіцієнтами a, b, c, d одновимірного ПФ (далі, для простоти, індекси nm в них опущено)

$$A=(a^2+c^2)^{1/2}=a_nA_m, \quad B=(b^2+d^2)^{1/2}=b_nA_m, \qquad (37)$$
$$C=(a^2+b^2)^{1/2}=A_na_m, \quad D=(c^2+d^2)^{1/2}=A_nb_m, \qquad (38)$$

звідки одержимо шукані рівняння зв'язку

$$A_{nm}^2=A^2+B^2=C^2+D^2=A_n^2A_m^2, \qquad (39)$$
$$tg\psi_n = B/A = d/c = b/a \equiv b_n/a_n, \qquad (40)$$
$$tg\psi_m = D/C = c/a = d/b \equiv b_m/a_m, \qquad (41)$$

що визначають три незалежних параметра – амплітуду A_{nm} та дві фази ψ_n, ψ_m як нелінійні комбінації 4-х амплітуд (14,15) двовимірного ПФ, які повністю описують фур'є-образ та мають безпосередній фізичний смисл. Смисл парціальних фаз – зміщення гармонік на площині відносно центра системи координат, яке, згідно (21) та (24), має вигляд

$$x_{0n} = \psi_n/k_{0x}, \quad y_{0m} = \psi_m/k_{0y}. \qquad (42)$$

Згідно (29-41), фізичний спектр має загальний вигляд

$$S(x,y)=A_{00}/4+\Sigma(A_{n0}/2)\cos\{nx+\psi_n\}+ \qquad (43)$$
$$+\Sigma(A_{0m}/2)\cos\{my+\psi_m\}+\Sigma\Sigma A_{nm}\cos(nx+\psi_n)\cos(my+\psi_m),$$
$$A_{00}=A_{0X}A_{0Y}, \quad A_{n0}=A_nA_{0Y}, \quad A_{0m}=A_{0X}A_m, \qquad (44)$$
$$A_{nm}^2=A^2+B^2, \quad tg\psi_n=B/A, \quad tg\psi_m=D/C, \qquad (45)$$
$$A^2=a^2+c^2, \quad B^2=b^2+d^2, \quad C^2=a^2+b^2, \quad D^2=c^2+d^2, \qquad (46)$$
$$\pi^2a=\iint f(x,y)\,g_{11}dxdy, \quad \pi^2b=\iint f(x,y)\,g_{12}dxdy, \qquad (47)$$
$$\pi^2c=\iint f(x,y)\,g_{21}dxdy, \quad \pi^2d=\iint f(x,y)\,g_{22}dxdy, \qquad (48)$$

де функції g_{ij} описуються виразами (17,18). Коефіцієнти Фур'є для фізичного спектру (47,48) в 4 рази більші від (14,15), згідно з (35).

Замість (45) можна також користуватися формулами, залежними від "малих" коефіцієнтів

$$A_{nm}^2=a^2+b^2+c^2+d^2, \quad tg\psi_n = b/a, \quad tg\psi_m = c/a, \qquad (49)$$

З (49) випливає, що амплітудний спектр описується виразом, аналогічним (19). Відмітимо, що для відшукання парціальних фазових спектрів вирази (40, 41) дають ще можливість $tg\psi_n=d/c$, $tg\psi_m=d/b$, але для визначеності вибрано варіант подібний до одномірного за позначеннями. Вирази (43, 44, 45 або 49, 46-48) описують ПФ у загальному вигляді. Для практичних обчислень можна привести спрощені вирази для $A_{00}=a_{00}$, A_{n0} та A_{0m}

$$A_{00}=\pi^{-2}\iint f(x,y)dxdy, \quad A_{n0}^2=a_{n0}^2+b_{n0}^2, \quad A_{0m}^2=a_{0m}^2+c_{0m}^2, \qquad (50)$$
$$\pi^2a_{n0}[b_{n0}]=\iint f(x,y)\cos(nx)[\sin(nx)]dxdy, \qquad (51)$$
$$\pi^2a_{0m}[c_{0m}]=\iint f(x,y)\cos(mx)[\sin(mx)]dxdy. \qquad (52)$$

5. ДОСЛІДЖЕННЯ ПФ

Для математичного спектру $(n,m\in]-\infty, \infty[)$ замість (43) маємо однорідну формулу

$$S=\Sigma\Sigma A_{nm}G_{nm}(x,y), \quad G_{nm}=\cos(nx+\psi_n)\cos(my+\psi_m), \qquad (53)$$

Тут парціальні спектри обчислюються за допомогою (49), а коефіцієнти Фур'є – (14,15). Набір функцій $\{G_{nm}\}$ є повною ортонормованою системою функцій для двовимірного ПФ. Таким чином, (53) дає загальний вигляд двовимірної тригонометричної функції, справедливий і для фізичного спектру. Тому, більш коректно подати останній у вигляді, подібному (53)

$$S=\Sigma\Sigma E_{nm}G_{nm}(x,y), \quad E_{nm}=A_{nm}, \quad n,m>1, \qquad (54)$$
$$E_{00}=A_{00}/4, \quad E_{n0}=A_{n0}/2, \quad E_{0m}=A_{0m}/2, \qquad (55)$$

Де коефіцієнти в (54) та (55) визначаються, відповідно, виразами (47-49) та (50-52). Також (53) дозволяє узагальнення на континуальний випадок

$$A(k_X,k_Y)=(1/2\pi)^2\iint f(x,y)G(k_X,k_Y,x,y)dxdy, \quad (56)$$
$$S(x,y)=\iint A(k_X,k_Y)G^{-1}(k_X,k_Y,x,y)dk_Xdk_Y, \quad (57)$$
$$G^{-1}(k_X,k_Y,x,y)=\cos[k_Xx+\psi(k_X)]\cos[k_Yy+\psi(k_Y)], \quad (58)$$

У вигляді (56-58) ПФ є інтегральним перетворенням з ядром G. Враховуючи континуальний аналог (42)

$$x_0(k_X)=\psi(k_X)/k_x, \quad y_0(k_Y)=\psi(k_Y)/k_y \quad (59)$$

одержимо, що ядро двомірного ОПФ G^{-1} дорівнює

$$G^{-1}(k_X,k_Y,x,y)=\cos[k_X(x+x_0)]\cos[k_Y(y+y_0)], \quad (60)$$

Приймаючи до уваги аналогію з виразами для ядер ПФ та ОПФ, одержимо, що ядро двомірного прямого ПФ є простим узагальненням одномірного випадку.

$$G(k_X,k_Y,x,y)=\cos[k_X(x-x_0)]\cos[k_Y(y-y_0)]. \quad (61)$$

Звідси легко одержати випадок періодичної функції, коли спектр дискретний, тоді відповідне ПФ має ядро, що є узагальненням (53) та (61)

$$A_{nm}=(1/2\pi)^2\iint f(x,y)H_{nm}(x,y)dxdy \quad (62)$$
$$H_{nm}(x,y)=\cos[(k_{0X}(nx-x_{0n})]\cos(k_{0Y}(my-y_{0n})]. \quad (63)$$

З огляду на те, що фаза є вектор, можна перейти до довжини Ф та кута θ в полярній системі координат, а також аналогічних параметрів ρ, α для зміщень.

$$\Phi^2=\psi_n^2+\psi_m^2, \quad tg\theta=\psi_m/\psi_n, \quad (64)$$
$$\rho^2=x_{0n}^2+y_{0m}^2, \quad tg\alpha=y_{0m}/x_{0n}=(k_{0X}/k_{0Y})tg\theta, \quad (65)$$

Це має сенс при аксіальній симетрії функції $f(\rho,\alpha)=f(\rho)g(\alpha)$, тоді $k_\rho^2=k_X^2+k_Y^2$, $tgk_\alpha=k_Y/k_X$. Таким чином, можна розглядати 5 спектрів – 3 двовимірних та 2 одновимірних: A_{nm} - амплітудний, ψ_n - x-фазовий, ψ_m - y-фазовий, Φ_{mn} – радіально-фазовий, θ_{nm} – аксіально-фазовий. Замість останніх 4-х можна розглядати X-спектр x_{0n}, Y-спектр y_{0m}, радіальний ρ_{mn} та α_{nm} аксіальний спектри. Будемо називати їх первісними тому, що вони описують розподіл у фур'є-області.

6. ПОХІДНІ СПЕКТРИ

Спектри, утворені з первісних спектрів як параметрично заданих функцій, назвемо похідними спектрами. Можна утворити всього 22 різних похідних спектра – 6 двомірних та 16 одномірних. Двовимірні: амплітудно-фазовий $A_{nm}(\psi_n,\psi_m)$ або амплітудно-радіально-аксіально-фазовий $A(\Phi,\theta)$, амплітудно-просторовий $A_{nm}(x_{0n},y_{0m})$ або амплітудно-радіально-аксіальний $A(\rho,\alpha)$, радіально-амплітудно-аксіальний фазовий $\Phi(A,\theta)$, аксіально-амплітудно-радіальний фазовий $\theta(A,\Phi)$, радіально-амплітудно-аксіальний $\rho(A,\alpha)$, аксіально-амплітудно-радіальний $\alpha(A,\rho)$.

Одновимірні: амплітудно-радіально-фазовий $A(\Phi)$, амплітудно-радіальний $A(\rho)$, амплітудно-аксіально-фа-зовий $A(\theta)$, амплітудно-аксіальний $A(\alpha)$, радіально-аксіальний фазовий $\Phi(\theta)$, радіально-аксіальний $\rho(\alpha)$, радіально-фазово-аксіальний $\Phi(\alpha)$, радіально-аксіально-фазовий $\rho(\theta)$ – всього 8 прямих та стільки ж обернених.

Наявність значної кількості спектрів (5 первісних та 22 похідних) дає можливість всебічно аналізувати властивості двовимірного зображення як у фур'є-області (первісні спектри), так і в площині (похідні спектри). Наприклад, амплітудно-просторовий спектр $A(x_0,y_0)$ дає розподіл гармонік в площині, радіально-аксіальний $\rho(\alpha)$ – залежність зміщення гармоніки від напрямку (діаграма направленості) і т.д. Можливо, в прикладних задачах матимуть сенс й інші комбінації наведених спектрів.

7. ГІПЕРКОМПЛЕКСНИЙ СПЕКТР

Так введений спектр має гіперкомплексний (ГК) характер [5], а саме – є тривимірним ГК числом, бо має дві фази. Легко показати, що (53) є дійсною частиною ГК спектру $S=Re\{Z\}$ як інтегрального перетворення, одержаного узагальненням (12,13)

$$4\pi^2C=\iint f(x,y)F(x,y)dxdy, \quad (66)$$
$$C=a+ib+jc+ijd=Aexp[i\psi_n+j\psi_m], \quad (67)$$
$$Z=\Sigma\Sigma CF^{-1}=\Sigma\Sigma A_{nm}exp[i(nx+\psi_n)]exp[j(my+\psi_m)], \quad (68)$$
$$F(x,y)=exp[-(inx+jmy)], \quad F^{-1}=1/F=exp[inx+jmy]. \quad (69)$$

Уявні одиниці i,j мають смисл орт-векторів площини. Спектр (67) можна вважати добутком двох комплексних чисел (КЧ) з різними уявними одиницями. Зображення ГК спектру як КЧ приводить до інтерпретації як КЧ також і фази (Ф та θ даються (64))

$$C=Aexp[i\psi_n+j\psi_m]=Aexp[i\Psi], \Psi=\Phi exp[-k\theta], k=ij. \quad (70)$$

Фаза Ψ описує координати двовимірної фур'є-гармоніки на площині, яка розглядається як площина КЧ.

ВИСНОВКИ

Розвинутий в роботі підхід дозволяє:

1) зберегти всю інформацію у вхідному сигналі, до якого застосовується ПФ;
2) одержати багато похідних спектрів, що дають змогу детально проаналізувати вхідний сигнал;
3) одержати просторовий розподіл амплітуд двовимірних фур'є-гармонік на площині;
4) проводити фазову фільтрацію зображення;
5) досліджувати різноманітні системи реєстрації зображень як двовимірні просторові фільтри.

ЛІТЕРАТУРА

1. Апенко М.И., Дубовик А.С., *Прикладная оптика*. – М:Наука, 1982.
2. Иванюгин В.М., Петухов С.В., Машинное стереозрение / *Зарубежная радиоэлектроника*. - № 7-9, 1993. – с. 56-66.
3. *Математическая энциклопедия*, т.4, /Гл. ред. И.М. Виноградов. - М:Сов. Энциклопедия, 1982.
4. Бронштейн И.Н., Семендяев К.А., *Справочник по математике*. - М:Наука, 1980.
5. Пытьев Ю.П., Чуличков А.И., *ЭВМ анализирует форму изображения*. – М:Знание, 1988.
6. Ланкастер П., *Теория матриц*. – М:Наука, 1978.
7. Кантор И.Л., Солодовников А.С., *Гиперкомплексные числа*. – М:Наука, 1973.

Фільтрація ближнього поля системи взаємодіючих випромінювачів

В.П. Савчук, І.М. Глушко
Одеський Державний Політехнічний Університет
м. Одеса, пр. Щевченка, 1, тел. 22-34-74
E. mail: postmaster@opi.odessa.ua

Розглядається задача фільтрації ближнього поля багатоелементної випромінюючої антени, елементи якої взаємодіють по полю, що випромінюється. Метод фільтрації заснований на корекції амплітудного розподілу коливальних швидкостей елементів антени. Представлений рекурсивний алгоритм, що використовує багатомірну параметричну ідентифікацію оператора взаємодії випромінювачів.

ВСТУП

Випромінюючі антени Тротта дозволяють створити в своєму ближньому полі умови, аналогічні полю плоскої хвилі, що біжить. Такі антени застосовуються як в метрології, так і для рішення задач дефектоскопії і діагностики. Конструктивно антена Тротта являє собою m лінійно розташованих груп випромінювачів, кожна з яких складається з n рознесених на рівну відстань випромінювачів. Кожна група здвинута відносно сусідньої групи на відстань d, рівну відстані між окремими випромінювачами. Амплітудний розподіл в апертурі кожної групи має биноміальний характер. Діаграма спрямованості антени Тротта виражається наступною функцією кута θ:

$$D(\theta) = \frac{\sin(m \cdot \varphi)}{m \cdot \sin(\varphi)} \cdot \cos^n(\varphi), \qquad (1)$$

де $\varphi = \pi \cdot \lambda / d \cdot \sin\theta$, λ — довжина хвиля.

Така форма діаграми дозволяє отримати в ближньому полі антени умови, відповідні плоскій хвилі, що біжить. Однак, якщо випромінювачі взаємодіють по полю, що випромінюється, то амплітудний розподіл на стороні управління не буде співпадати з розподілом на стороні випромінювання. У цьому випадку у діаграми з'являються далекі бічні пелюстки, а ближнє поле стає неоднорідним. Усунути неоднорідність ближнього поля можна за допомогою коректуючого амплітудного розподілу на стороні управління. Корекція розподілу еквівалентна включенню просторового фільтра, що усуває неоднорідність ближнього поля [1].

1. АЛГОРИТМ ФІЛЬТРАЦІЇ

Зв'язок нормальної складаючої коливальної швидкості v_n випромінювача з номером n і напруги u_m, що підводиться до випромінювача з номером m, визначається як:

$$u_m = \sum_{n=1}^{N} z_{mn} \cdot v_n, \quad m = 1,\ldots,N, \qquad (2)$$

де N — число елементів решітки; z_{mn} — взаємний опір випромінювання перетворювачів з номерами m и n (при $m \neq n$) або власний, включаючи механічний, опір випромінювання перетворювача з номером m (при $m = n$).

Будемо розглядати сукупність швидкостей v_n як вектор $\mathbf{v}$ величин, що спостерігаються з деякою похибкою, а сукупність напружень u_m — як вектор керуючих впливів. Тоді рішення системи рівнянь (2) відносно швидкостей можна записати у вигляді:

$$\mathbf{v} = \mathbf{B} \cdot \mathbf{u} + \boldsymbol{\varepsilon}, \qquad (3)$$

де $\mathbf{B} = \mathbf{Z}^{-1}$ — матриця коефіцієнтів взаємного впливу елементів решітки; $\mathbf{Z}$ матриця власних і взаємних опорів; $\boldsymbol{\varepsilon}$ — вектор, характеризуючий похибку вимірювання швидкості.

Якщо опори випромінювання не відомі, то співвідношення (3) являє собою систему багатомірний рівнянь регресії вектора швидкостей, що спостерігаються по вектору $\mathbf{u}$ вхідних впливів з матрицею $\mathbf{B}$ невідомих параметрів [2]. Матрицю $\mathbf{B}$ можна ідентифікувати, якщо виміряти J векторів $\mathbf{v}(j)$ (j — номер поточного вимірювання) для заданих впливів $\mathbf{u}(j)$. В тому випадку, коли компоненти вектора некоррельовані і розподілені згідно з нормальному законом, критерій якості оцінки $\mathbf{B}$ у відповідність з методом найменших квадратів можна представити у вигляді:

$$Q_n(J) = \sum_{j=1}^{J} q_j \cdot [v_n(j) - \tilde{\mathbf{b}}_n(j) \cdot \mathbf{u}(j)]^2, \qquad (4)$$

де n — номер випромінювача; $v_n(j)$ — його коливальна швидкість; q_j — довільний ваговий коефіцієнт; $\tilde{\mathbf{b}}_n(j)$ — оцінка рядка з номером n матриці $\mathbf{B}$ на кроці j.

© UkrObraz'2000

Рішення системи нормальних рівнянь

$$\frac{\partial Q_n}{\partial \tilde{\mathbf{b}}_n} = 0\,,\ n = 1,\ldots,N$$

відносно $\tilde{\mathbf{b}}_n(j)$ мінімізує критерій (4) і дає оптимальну оцінку $\tilde{\mathbf{B}}$ матриці $\mathbf{B}$ на кожному кроці j. Це рішення можна представити у вигляді процесу послідовних ітерацій [3]:

$$\tilde{\mathbf{b}}_n(j) = \tilde{\mathbf{b}}_n(j-1) + q_j \cdot \mathbf{\Gamma}(j) \cdot \mathbf{u}(j) \cdot \Delta v_n, \quad (5)$$

де

$$\Delta v_n = [v_n(j) - \tilde{\mathbf{b}}_n(j-1) \cdot \mathbf{u}(j)],$$

$$\mathbf{\Gamma}(j) = \mathbf{\Gamma}(j-1) - q_j \cdot \Delta\mathbf{\Gamma}(j-1),$$

$$\Delta\mathbf{\Gamma}(j-1) = \frac{\mathbf{\Gamma}(j-1) \cdot \mathbf{u}(j) \cdot \mathbf{u}^{\mathrm{T}}(j) \cdot \mathbf{\Gamma}(j-1)}{1 + q_j \cdot \mathbf{u}^{\mathrm{T}}(j) \cdot \mathbf{\Gamma}(j-1) \cdot \mathbf{u}(j)},$$

T — знак транспонування.

Для побудови просторового фільтра ближнього поля антени до процедури ідентифікації, заданої співвідношенням (5), необхідно приєднати рівняння послідовного управління амплитудным розподілом коливальних швидкостей в апертурі антени:

$$\mathbf{u}(j) = \tilde{\mathbf{B}}^{-1}(j) \cdot \mathbf{w}\,, \quad (6)$$

де $\mathbf{w}$ — амплітудний розподіл швидкостей в апертурі антени, що забезпечує необхідні власти-вості її ближнього поля.

Процес формування амплітудного розподілу складається з двох етапів. На першому етапі по заданому вектору вхідних впливів на електричній стороні випромінювачів і вектору швидкостей, що спостерігається на їх акустичній стороні, оцінюються коефіцієнти регресії відгуку по вектору впливів. На другому етап будується вектор збудження, що є вхідним для наступного кроку ідентифікації.

Початкові умови для розглянутого процесу можна записати у вигляді:

$$\mathbf{B}(0) = \mathbf{I}\,,\ \mathbf{u}(0) = \beta \cdot \mathbf{w}\,,\ \Gamma(0) = \alpha \cdot \mathbf{I}\,, \quad (7)$$

де $\mathbf{I}$ — єдинична матриця, $\alpha, \beta = const$, $\alpha >> 1$.

Процес завершується при умові:

$$\|\mathbf{w} - \mathbf{v}(j)\| < \mu\,, \quad (8)$$

де $\|\circ\|$ — норма вектора, μ — константа, яка характеризує точність формування амплітудного розподілу.

2. Результати моделювання

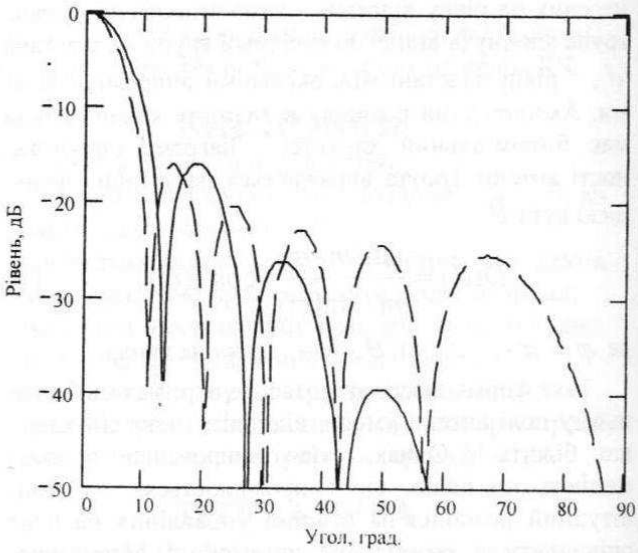
У цьому розділі приведені результати моделювання процесу формування ближнього поля для одновимірних антен Тротта при різних видах взаємного зв'язку їх елементів (мал. 1).



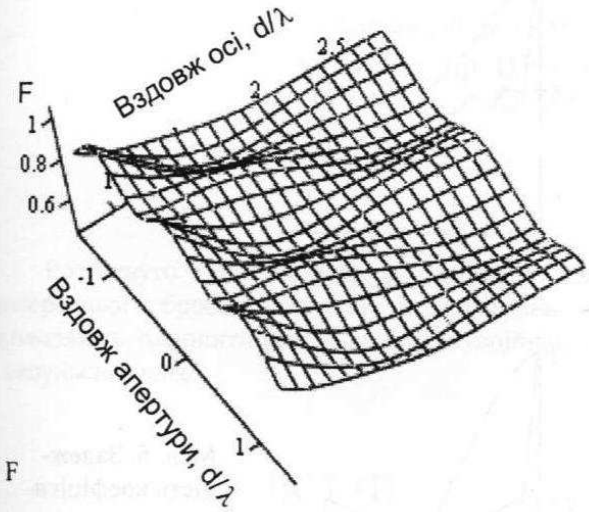
Мал.1. Залежність коефіцієнтів взаємного впливу від різниці номерів елементів для різної міри зв'язку:
—— сильний;
--- середній;
-·-·- слабкий.

На мал. 2 показані діаграми спрямованості лінійної антени з сильно зв'язаними елементами. Число елементів антени $N = 15$ $(n = 5,\ m = 11)$, відстань між ними $d = 0.4 \cdot \lambda$, λ — довжина хвилі, що випромінюється. Для формування рівномірного розподілу тиску в ближньому полі використано п'ять кроків корекції. На мал. 3 представлена структура ближнього поля, а на мал. 4 приведені лінії постійного рівня відносного тиску в ближньому полі. Ці малюнки показують, що для отримання однорідного ближнього поля необхідно усунути в діаграмі спрямованості далекі по куту бічні пелюс-тки.

Мал.2. Діаграма спрямованості:
--- до корекції ;
—— після корекції .

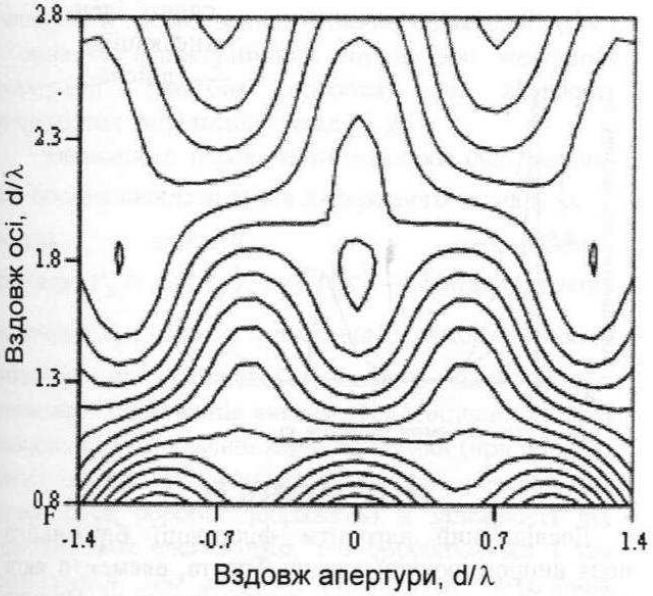


На мал. 5а приведений амплітудний розподіл Тротта на стороні випромінювання, який забезпечує необхідну структуру ближнього поля. На мал. 5б показаний розподіл, який виходить через взаємний вплив випромінювачів на стороні випромінювання, якщо на стороні збудження введений розподіл Тротта. На мал. 5в показаний отриманий внаслідок корекції розподіл на стороні збудження, який забезпечує розподіл Трота на стороні випромінювання при даному взаємозв'язку елементів.



а)

Мал. 3. Ближнє поле:
а) --- до корекції ; б) —— після корекції .

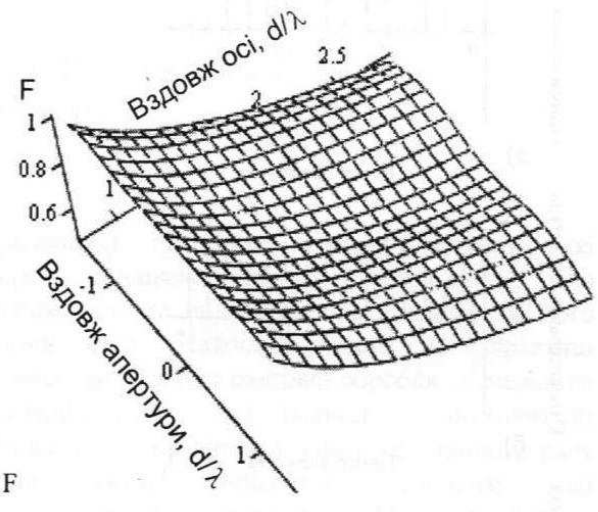


а)

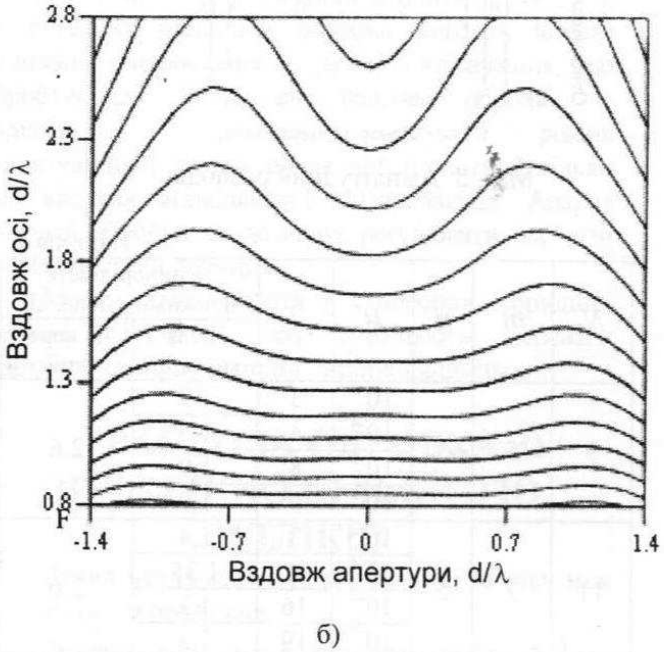
Мал. 4.. Лінії рівня відносного тиску в ближньому полі:
а) --- до корекції ; б) —— після корекції .

Процес формування однорідного ближнього поля завершується за число кроків, менше ніж число елементів антени. Але ідентифікація матриці взаємозв'язку завершується не повністю. Це говорить про те, що має місце неоднозначний зв'язок між зміною структури ближнього поля і коефіцієнтами взаємного впливу випромінювачів антени.

На мал. 6 показана залежність дійсних коефіцієнтів взаємного впливу і коефіцієнтів, отриманих в процес ідентифікації, від відстані між взаємодіючими елементами. Аналогічні результати отримані для інших антен Тротта.



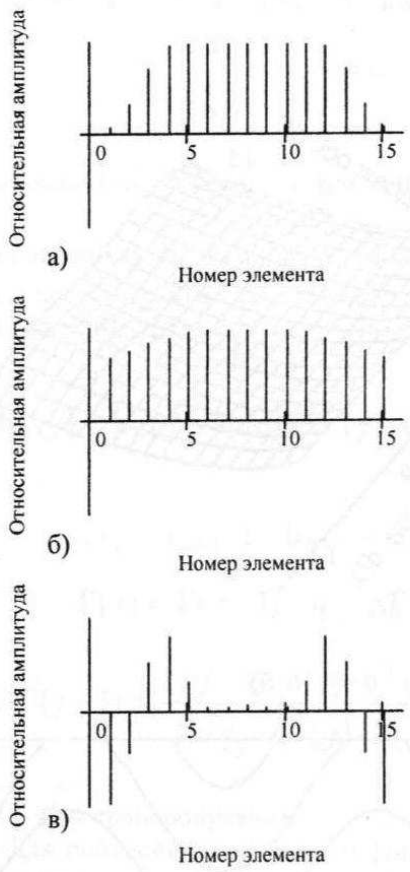
б)



б)

Результати моделювання представлені в таблиці 1, де визначено: N — число випромінювачів; m — число груп; n — число в групі; $\mu = \frac{|\mathbf{w}-\mathbf{v}|^2}{|\mathbf{w}|\cdot|\mathbf{v}|}$ — похибка корекції; J — число кроків корекції.

Для аналізу впливу похибки вимірювання коливальної швидкості на стороні випромінювання проведено моделювання процесу фільтрації ближнього поля методом Монте-Карло. Результати моделювання показують, що із зростанням похибки вимірювання швидкості при фіксованому значенні точності корекції зростає як середнє число ітерацій, так і його варіація. Наприклад, для 29-елементної антени при точності корекції $\mu = 0.01$ дані моделювання представлені в таблиці 2.



Мал. 5. Амплітудний розподіл

Таблиця 1.

N	m	n	μ	J	Неоднорідність ближнього поля, дБ	
					Після корекції	До корекції
9	5	5	10^{-2}	3	1.9	2.6
			10^{-4}	4	1.9	
			10^{-7}	8	1.9	
			10^{-8}	9	1.9	
11	7	5	10^{-2}	3	1.4	2.6
			10^{-4}	4	1.35	
			10^{-7}	16	1.3	
			10^{-8}	19	1.3	
15	11	5	10^{-2}	3	2.0	3.2
			10^{-4}	5	1.8	
			10^{-7}	9	1.8	
			10^{-8}	36	1.8	
29	23	7	10^{-2}	4	1.8	3.5
			10^{-4}	6	1.6	
			10^{-7}	12	1.6	
			10^{-8}	14	1.6	

Таблиця 2.

Пхибка вимірювання (СКО)	Число ітерацій			
	Середне	Мин.	Макс.	СКО
0.3	6	5	8	0.6
0.4	7	6	9	0.65
0.5	8	6	11	1.1
0.6	9	6	16	1.5



Мал. 6. Залежність коефіцієнтів взаємного впливу від різниці номерів елементів:
—— отримана внаслідок ідентифікації;
--- дійсна.

ВИСНОВКИ

Досліджений алгоритм фільтрації ближнього поля випромінюючої антени Тротта, елементи якої взаємодіють по полю, що випромінюється. Процес корекції амплітудного розподілу в апертурі антени сходиться до бажаного розподілу значно швидше, ніж ідентифікація матриці коефіцієнтів взаємного впливу. Це говорить про те, що має місце неоднозначний зв'язок між зміною структури ближнього поля і коефіцієнтами взаємного впливу випромінювачів антени. Із зростанням похибки вимірювання швидкості при фіксованій точності корекції зростає число кроків корекції.

ЛІТЕРАТУРА

1. Зверев В. А. "Радиооптика". — М.: Радио и связь, 1975. —304 с.
2. Гроп Д. "Методы идентификации систем". — М.: Мир, 1979. —302 с.
3. Изерман Р. "Цифровые системы управления". — М.: Мир, 1984. —542 с.

ПОПЕРЕДНЯ ОБРОБКА СИГНАЛІВ ТА ЗОБРАЖЕНЬ

С. Г. Антощук, В. М. Крилов, В. О. Давидов
Одеський державний політехнічний університет
м. Одеса, пр. Шевченко 1. Факс (0482)600-83E-mail
Krylov@enf.opu.odessa.ua

Розглянуто гибридні методи лінійно-рангової попередньої обробки сигналів та зображень, що дозволяють зменшити рівень як флюктуаційної, так і імпульсної завад.

ВСТУП

При розробці систем розпізнання та обробки сигналів і зображень, які працюють в масштабі реального часу, основною проблемою, що обмежує область їхнього застосування, є шуми і завади різноманітного походження. Вплив останніх можна зменьшити застосуванням лінійної фільтрації (ЛФ) у випадку флюктуаційних шумів або медіанної фільтрації (рангова обробка) за наявності розрізнених імпульсних завад [1, 2].

Визначимо переваги та недоліки цих засобів. ЛФ послідовності відліків дискретного сигналу $\{x_k\}$ можна описати як перетворення вигляду: $y_k = L_k(x_k)$, де $L_k(x)$ — лінійна функція величини x_k, що є зваженою сумою відліків сигналу, які складають заданий окіл даного елементу. Зважування виконується у відповідності зі значеннями імпульсної характеристики (при обробці сигналів) або зі значеннями функції розсіювання точки (при обробці зображень) в залежності від відстані між елементами, що обробляються і що зважуються. ЛФ дозволяє виконувати виділення необхідної для подальшої обробки області спектрів сигналів і зображень, здійснювати різноманітні лінійні операції (диференціювання, інтегрування та інш.), проводити оптимальні вінеровську та погоджену обробки. До переваг цієї групи засобів відноситься наявність швидких алгоритмів і розвинених засобів синтезу фільтрів. До недоліків засобів ЛФ відносяться чутливість до імпульсної завади (в одномірному випадку) або до завади типу «яскрава точка» (у двомірному випадку). Поява значних по інтенсивності завад такого типу впливає на результати ЛФ у всьому фрагменті сигналу (зображення), що обробляється.

Рангові алгоритми здійснюють нелінійне перетворення сигналу $y_k = \Phi_k(x_k)$, де $\Phi_k(x_k)$ — функція, вигляд якої визначається деякою заданою підмножиною так званих рангів або рангових (порядкових) статистик вибірки відліків, які утворені відліком сигналу (зображення), що обробляється, та відліками, які складають його заданий окіл. Найбільш часто на практиці використовується вид рангової обробки — медіанна фільтрація (МФ), яка полягає у знаходженні центрального (медіанного) члена варіаційного ряду та в заміні інтенсивності елементу, що обробляється, цією величиною. Цей вид обробки дозволяє істотно знизити рівень імпульсних завад і завад типу «яскрава точка», однак не виявляє значного впливу на флюктуаційний шум. Крім того, застосування медіанної обробки вносить істотні нелінейні викривлення у сигнал (зображення, що обробляється). Інший вид рангової обробки — «зрізка» — дозволяє знижувати рівень флюктуаційної завади, однак він дієздатний тільки при високих відношеннях сигнал/завада. Апарат рангової обробки не дозволяє регулювати частотні характеристики фільтрів.

Метою даної роботи є створення гібридних методів фільтрації, які суміщають переваги рангових методів і методів лінійної фільтрації.

РАНГОВА ОБРОБКА ЗІ ЗВАЖЕНОЮ КІЛЬКІСТЮ ВІДЛІКІВ

Даний метод враховує кореляційні зв'язки між елементами зображення.

Введемо позначки: S-окіл елементу (x_i, y_i) — задана певним чином в просторовій області множина елементів зображення, що оточує центральний елемент (x_i, y_i); N_s - об'єм околу, тобто кількість елементів, що її складають; $V_s(r)$ - варіаційний ряд з елементів S-околу, $r = 0,1,\ldots,N_{s-1}$; $r_s(v)$ — ранг елементу в варіаційному ряді; M-окіл — підмножина S-околу, визначена за деяким правилом.

Запропонований гібридний метод обробки передбачає такі етапи:

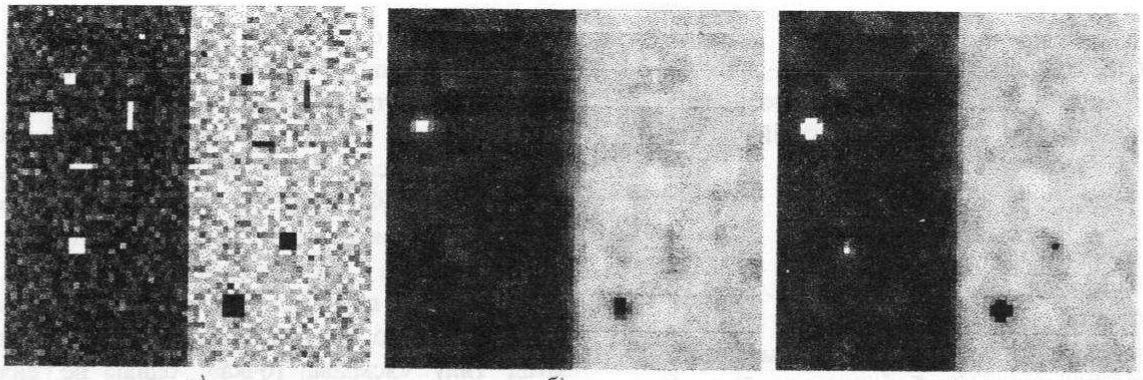
– S-околи елементу, що обробляється, створюють M-околи, які розрізняються за відстанню до елемента, що обробляється. Кожному M-околу присвоюється свій цілочисельний ваговий коефіцієнт;

© UkrObraz'2000

– складається варіаційний ряд з елементів М-околів. У відповідності з вибраним цілочисельним ваговим коефіцієнтом кількість членів варіаційного ряду М-околу збільшується;

– вибирається центральний (медіанний) член варіаційного ряду $V_s(r)$ з рангом $E[r_s(v)/2]+1$. Тут Е — операція цілочисельного ділення. Елемент, що обробляється, замінюється на вибране медіанне значення $V_s(r)$.

Результати обробки тестового зображення, ушкодженного як завадою типу «яскрава точка», так і флюктуаційною подані на рис. 1.



а) *б)* *в)*

Рис. 1. Приклад обробки зображення ранговими фільтрами: а) вхідне зображення; б) медіанний фільтр; в) медіанний фільтр зі зваженою кількістю відліків.

З рис. 1. видно, що частотні характеристики розробленого гібридного фільтру по відношенню до медіанного змінилися, а рівень нелінійних викривлень малорозмірних об'єктів зменшився. Проведені експерименти показали, що запропонований вид обробки дозволяє зменшувати рівень імпульсних завад та завад типу «яскрава точка», зменшувати рівень флюктуаційних завад, змінювати частотні характеристики розроблених фільтрів шляхом зміни вагових коефіцієнтів. Нелінійні викривлення, які вносяться при такій обробці, у порівнянні з медіанною обробкою зменшуються, тому що результат обробки зміщується у варіаційному ряді до значення інтенсивності елементу, що обробляється. У відсутності флюктуаційної завади різниця між результами обробки за допомогою розробленого методу і з використанням медіанної фільтрації незначна. До недоліків методу слід віднести деяке зниження швидкодії по відношенню до медіанної фільтрації.

ЛІНІЙНО-РАНГОВА ФІЛЬТРАЦІЯ

Іншим методом попередньої обробки зображень є лінійно-рангова обробка (ЛРФ) з медіанним вибором середнього. В цьому випадку обробка полягає у наступному:

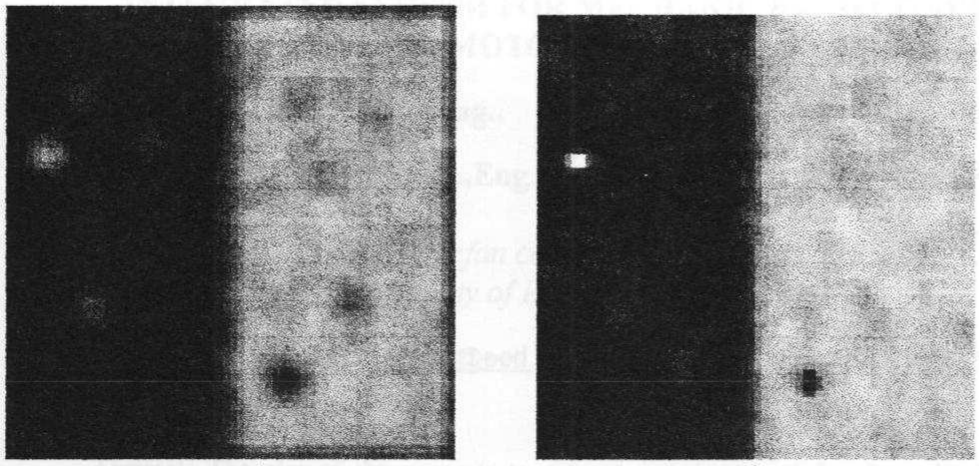
– окіл елементу, що обробляється, зважується у відповідності з імпульсною характеристикою або функцією розсіювання точки фільтру;

– з результатів зважування складається варіаційний ряд $V_s(r)$. Вибирається центральний (медіанний) член варіаційного ряду $V_s(r)$ з рангом $E[r_s(v)/2]+1$. Елемент, що обробляється, замінюється на вибране медіанне значення $V_s(r)$.

Члени варіаційного ряду, що визначаються імпульсною завадою, розміщені на периферії варіаційного ряду і на результати обробки не впливають. Тому запропонований метод дозволяє, на відміну від ЛФ, усунути вплив імпульсної завади на результати обробки.

У відповідності з законом великих чисел у формі Чебишева, середнє в околі елементу, що обробляється, і медіана варіаційного ряду є незміщенними і ефективними оцінками математичного очикування. Тому можна стверджувати, що при необмеженому збільшенні фрагменту, що обробляється, результати ЛРФ і ЛФ збігаються. Такий же збіг спостерігається при флюктуаційному характері завади. Тому частотні характеристики ЛРФ незначно відрізняються від частотних характеристик відповідних ЛФ. Внаслідок цього при реалізації ЛРФ можуть бути використані відомі методи синтезу ЛФ.

Результати обробки тестового зображення (див. рис. 1, а), яке ушкоджено як завадою типу «яскрава точка», так і флюктуаційною завадою, подані на рис. 2.



а) *б)*

Рис. 2. Приклад обробки зображення лінійними фільтрами: а) лінійний фільтр; б) лінійний фільтр з медіанним вибором середнього.

ПРОСТОРОВЕ АВТОМАТИЧНЕ РЕГУЛЮВАННЯ ПІДСИЛЕННЯ

Перспективним методом поліпшення якості зображень при наявності мультиплікативної флюктуаційної завади є просторове автоматичне регулювання підсилення (ПАРП). Система ПАРП працює таким чином: на вхід регулюючого елементу подається сигнал строки зображення. Коефіцієнт підсилення змінюється у відповідності із значенням управляючої дії. Опорне значення інтенсивності визначається за головним максімумом гістограми зображення, тобто за глобальною оцінкою фона. В системі визначається середнє значення інтенсивності в деякій заданій області елементу, що оброблюється, яке відповідає локальній оцінці фону. У відповідності з різницею між локальною та глобальною оцінками фона змінюється коефіцієнт підсилення регулюючого елемента. В наслідок цього фон кожного локального фрагменту наближається до глобальної оцінки фона. Дослідження реальних сигналів показали, що даний метод дозволяє усунути в процесі обробки завади типу "яскраве чорне або біле пляма". Для цього можна використовувати операцію рангової зрізки для усіх об'єктів, яскравість яких виходить за рамки наперед заданого діапазону (яскравість цих об'єктів замінюється глобальною оцінкою фона).

ВИСНОВКИ

Проведені експерименти підтвердили, що запропоновані види обробки дозволяють водночас зменшувати рівень завад типу «яскрава точка» і рівень флюктуаційних завад. До недоліків методів слід віднести зниження швидкодії по відношенню як до медіанної, так і до лінійної фільтрації.

Розроблені методи рекомендуються для використання у випадку обробки сигналів і зображень, ушкодженних як імпульсною так і флюктуаційною завадами. В цьому випадку традиційну спільну обробку ранговим і лінійним фільтрами доцільно замінити обробкою одним з запропонованих гібридних методів. Швидкодія гібридних методів менша швидкодії лінійної і рангової фільтрації окремо, але значно вища швидкодії спільної рангової і лінійної обробки. Практичне використання алгоритмів попередньої обробки показало, що у разі наявності мультиплікативної флюктуаційної завади, система просторового автоматичного регулювання підсилення є досить ефективною. Рівень нелінійних викривлень при використанні розроблених методів зменшується і з'являється можливість синтезувати гібридні фільтри з заданими частотними характеристиками.

Автори використовували наведані методи під час вирішення ряду практичних задач для розпізнавання об'єктів в умовах високого рівня завад.

Література

1. Прэтт У. Цифровая обработка изображений.—Т. 2 – М. Мир.—1982.
2. Крылов В. Н., Антощук С. Г., Бодалевский А. А. Ранговая обработка двумерных изображений в пространстве оценок и решений //Труды Одес. политехн. ун-та. — Одесса, 1998. - с. 105-108.

INTEGRATED SYSTEM FOR MECHANIC BAND CONVEYORS WITH HYBRID MOTOR, TV CAMERA AND ROBOT

Radu PENTIUC, Ph.D.,Eng., Stefan-Gheorghe PENTIUC, Ph.D.,Eng.,

Adrian GRAUR Ph.D.,Eng., Mariana MILICI Lecturer Eng.

University "Stefan cel Mare" of Suceava, Romania
Faculty of Electrical Engineering

radup@eed.usv.ro, pentiuc@eed.usv.ro adrian@eed.usv.ro dam@eed.usv.ro

Abstract: The aim of the paper is to present a system used for the unloading of the mechanic band with low speed and carrying heavy materials. The movement is assured by an hybrid electrical induction motor able to develop a sufficient force according to the speed of the band and the mass of the transported blocks. The blocks are uncharged and pushed into a stack by a robot, in fact a manipulator with 6 liberty degrees controlled by a computer. Two TV cameras are connected to the computer. Based on the image analysis the computers synthesises the commands for the robot.

Keywords: hybrid electrical induction motor, digital image processing, robot, automation

INTRODUCTION

The mechanic conveyor bands are widely used in constructions and in the industry of construction materials.

Constructions band conveyors can work in the environment, in a tuff climate activating against the electric and mechanic elements. The sand, the dust, the frost and the humidity have a negative influence on the equipment's working and affects especially the function of the reduction gears of revolution.

The main inconvenient of the low speed mechanic band conveyors, trained up with electric motors, usually asynchronous, is that the mechanic reduction gears of revolution, coupled with the activating engines, present a high incidence of the defects by reducing the period of working of the technological equipment because of the frequent interventions for reparations.

This paper describes some main aspects of an integrated system for unloading a conveyor band (see figure 3), presenting the control of this operation by a robot guided by computers processing the images captured by two cameras and scanning the outputs of an automate controlling the hybrid electrical induction motors that trains up the conveyor band.

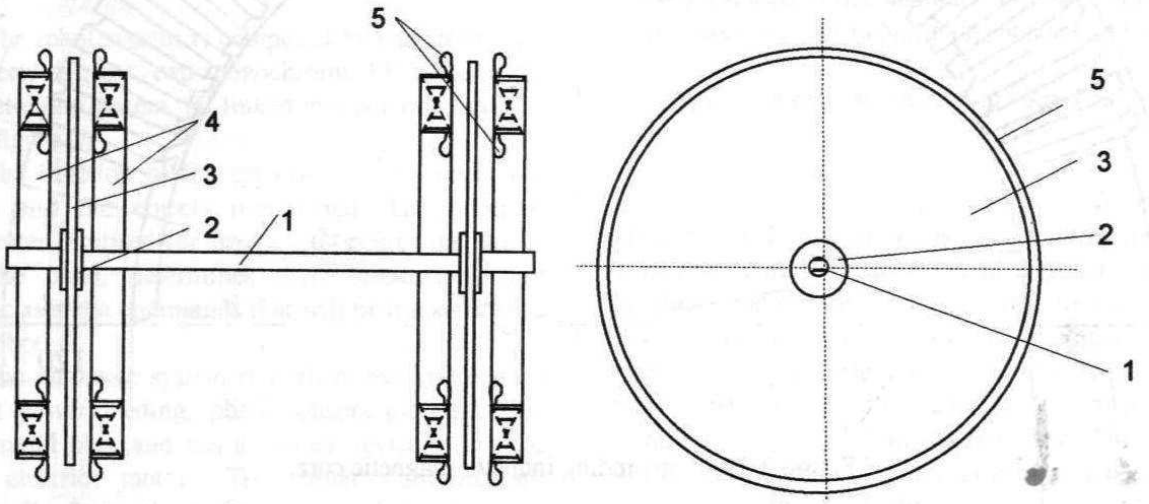


Figure1. Activating system at low speed with EIHM for mechanic band conveyors.

1. ELECTRIC ACTIVATING SYSTEM AT LOW SPEED WITH HYBRID MOTOR FOR MECHANIC BAND CONVEYORS

It is introduced a new modern activating system at low speed, equipped with an electric induction hybrid asynchronous motor (EIHM) [3] for the mechanic band conveyors used in constructions.

1.1. Particularities of an activating system with EIHM

The technical problem that is resolved by our installation consists of the expulsion of the reduction gear of revolution from the action system by directly training the band conveyor with low speed EIHM, this way resulting the reducing of the off-times for the repairing of the defects of the reduction gears, reducing the costs of the conveyor equipment, the weight of the equipment, diminishing extra expenses for maintaining, when the movement from the motor to the band is done directly, without using a reduction gear, the power on the installation goes higher, the transporting installation can be also built in a modulate way with doubled groups of EIHM, at every drum made in order to activate the band. The conveying band trained directly by the EIHM, in the new proposed variant moves away the up mentioned disadvantages because the activating system is realised with EIHM, which develop a low speed revolution motion [5]; this fact permits the exclusion of the cinematic chain of the mechanic reduction gear of revolution.

An EIHM is recommended to be built with $p \geq 4$ poles number. After the radius R dependent, pole pitch τ, the tangential synchronous velocity v_1 is computed using the primary suply tension frequency f_1:

$$v_{1(R)} = 2f_1\tau_{(R)} \quad (1)$$

where $v_{1(R)}$ and $\tau_{(R)}$ are values dependent segment circle inductor radium.

In the base of the maximum radius of inductor R_{max} and the angular speed ω, the tangential synchronous velocity $v_{1(Rmax)}$ is calculated.

1.2. Compounds parts of an EIHM with bilateral toroidal inductors

The EIHM has its primary winding on the toroidal inductors placed bilateral from the disc - made of conductive material - together with the activating drummer of the conveying band, this disk is also the inducted with the secondary winding of the hybrid motor..

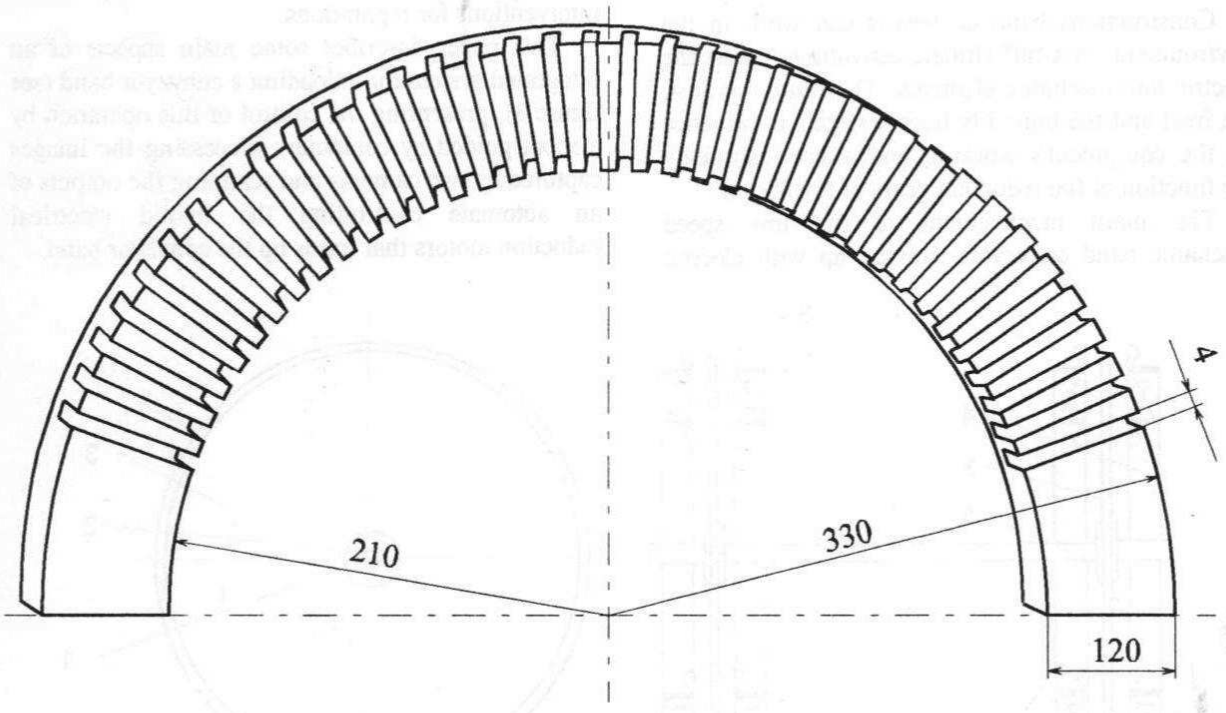


Figure 2. Built unwinding inductor magnetic core.

In fig.1 is presented an activating system with EIHM for conveying bands used in constructions, mining exploitations, in quarries, in the metallurgical enterprises, in port fitting outs, for cereals and food business, for rolling stairs or other systems of transport a 1 drummer on which the rolling band of

the conveyer is held, by using the elastic couple 2, there are linked on the lateral sides two disks made of electric secondary winding of the hybrid motors. On both sides of every disk 3, there are two fixed toroidal inductors 4, on which is placed the distributed primary winding 5, or in the form of a ring.

In order to improve the performances of this electric machine, the disk of conductive material can be played of the middle, a sandwich-like structure, by using super magnetic material (silicosis sheet iron or other iron -magnetically material).

The toroidal inductor of the EIHM may have an opening segment as a semi-toroidal (semicircle), the construction of the primary magnetic core without winding being represented by fig.2, or as a complete semicircle opened at 360^0, case in which the engine becomes an asynchronous one with a rolling disc.

The core in fig 2 has been projected, built and tested by the author in lab by using a testing.

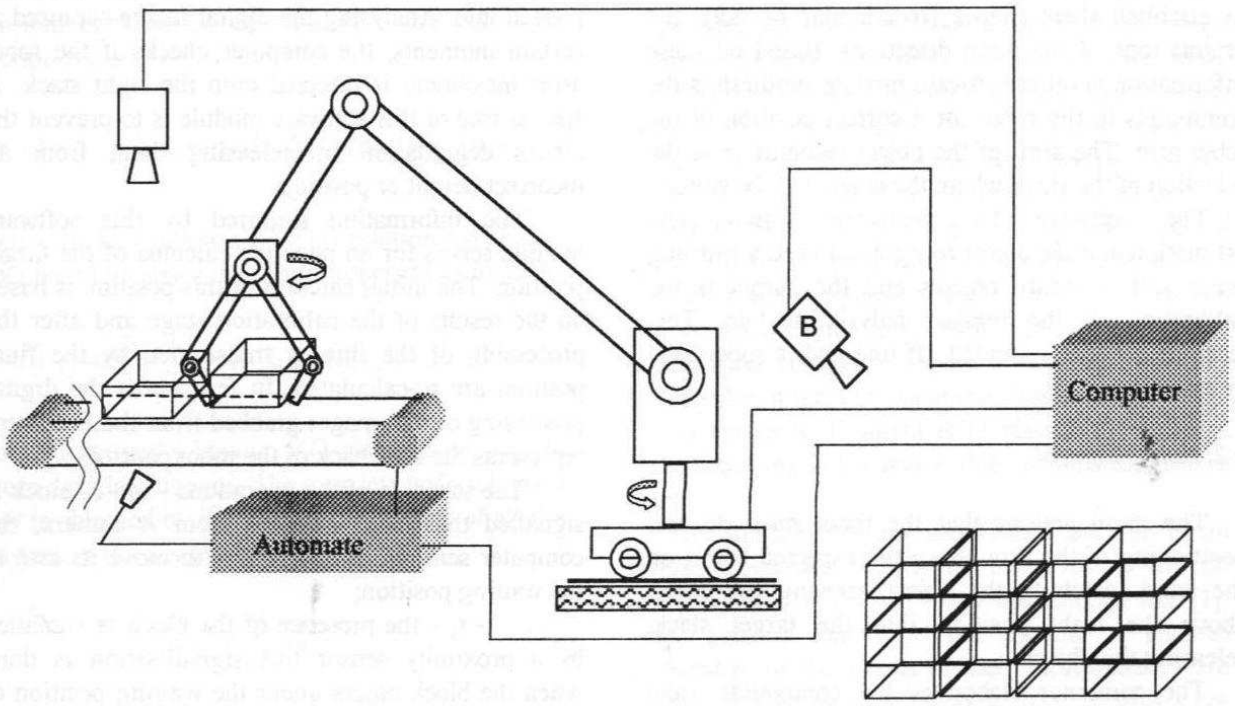


Figure 3. Integrated system for conveyor band unloading.

2. ROBOT CONTROL SYSTEM

The robot system is composed by one robot with 6 liberty degrees, two monochrome TV camera and an automate system, all linked to a computer system (see figure 3).

The function of the camera A is to survey the band and the objects transported. The program computer digitised the image and locates the blocks on the band, determines their orientations and synthesises the commands that will be transmitted to the robot.

The automate system centralises the information about power feeding, photo sensors placed in the horizontal plan and has a special device to control the electric motor. The binary information, especially from photo-sensor are redirected to the computer.

The B camera surveys the stack made by the robots. The monochrome image is grabbed and digitised by the computers. The results of the image analysis are used in the elaboration of the commands for the robot in order to build a right stack of blocks.

2.1. Acquisition and analysis of the band surface images

The surface of the band is surveyed by the monochrome TV camera named A in the figure 3. It is assumed that exists a sufficient contrast between the blocks and the surface of the band. In most cases this contrast is assured by the black colour of the band and the grey colour of the blocks (BCA). The area surveyed by the camera is artificially illuminated with constant parameters. For these reasons, the use of a monochrome camera is perfectly justified, but this is not an imperative requirement. The main stages of the image processing are: digital image acquisition, binarisation, object labelling, filtering, object

analysis (shape, orientation, geometrical parameters estimation).

In order to have a certain robustness face to the variations of band colours or blocks, illumination conditions, the binarisation is made with an algorithm based on pattern recognition technique [2]. The binarysed image permits the execution of a quickly procedure for labelling objects and detecting a number of ROIs (Region of interest) containing a single object. These ROI's may be analysed one by one, leading to a high speed of processing. Some ROI that are not eliminated in the filtering process contain relevant object which are analysed in order to establish their shapes (rectangular or not), the orientations of the main directions. Based on these information another software module synthesises the commands to the robot for a correct position of the robot arm. The aim of the object recognition is the selection of the stack where the object will be stored.

The software for geometrical parameters estimation and the object recognition need a learning stage with standard objects and the output is the calibration of the image analysis module. This learning stage is executed off-line and is supervised by an operator.

2.2. Robot Control

The main actions that the robot must do are: positioning of the arm above the expected object on the band, catching the object, moving the object above the right position into the target stack, releasing the object.

The computer elaborates the commands and transmits them by a serial link to the robot. The fields of a robot command must contain polar co-ordinates in the robot's own co-ordinating system. In the frame of the whole control system of the robot, including TV camera, there may be distinguished 7 co-ordinates systems:

- object surveying system (S1) - measuring units: cm
- graphic system for the A TV camera (SA) - measuring units: pixels
- object catching area system (S2) - measuring units: cm
- robot co-ordinates system - polar co-ordinate for each arm (S3), measuring units CAN
- stacks systems (S4) and (S5) - measuring units: cm (one for the stack with standard objects and one for the non-standard ones)
- graphic system for the B TV camera (SB) - measuring units: pixels

These systems assure the flexibility of the operation. The co-ordinates transformation relations is based on coefficients computed during the learning stages (calibration). For example, in order to have a correct transformation between the co-ordinates in S2 and S3, the arm of the robot is moved, by a manual command of the operator, in the origin of the S2 and in a point placed on the Ox axis of the S2. As answer to a special command emitted by the computer the robot transmits the co-ordinates (in S3) of its arms.

2.3. Stack building

The arrangement of the blocks into the stacks is supervised by the computers. The B TV camera surveys the stack where the standard objects are pushed into. Analysing the digital image captured at certain moments, the computer checks if the robot arms movement is directed onto the right stack. A special role of this software module is to prevent the blocks degradation by releasing them from an incorrect height or position.

The information acquired by this software module serves for an adaptive calculus of the target position. The initial calculus of this position is based on the results of the calibration stage and after the processing of the images transmitted by the final position are re-calculated. In few words the digital processing of the images grabbed from the B camera represents the feed-back of the robot control.

The schedule of the operations - to - a. block is signalled the image analyses from A camera; the computer send an order to robot to move its arm in the waiting position;

- t_1 - the presence of the block is signalled by a proximity sensor this signallisation is done when the block enters under the waiting position of the robot arm; the commands to catch the block are transmitted to the robot
- t_2 - the robot begins to the block to the target stack.
- t_3 - another series of commands based on the image analysed the B camera is launched to the robot
- t_4 - the robot release the block on the stack.

At this final point a new block enters in the area surveyed by the A camera and a new cycle begins. The computer control the speed of the band by estimating the frequency of the objects on the band.

3. IMAGE PROCESSING AND RECOGNITION

3.1. Image processing related aspects

During the acquisition of monochrome images different noise sources may diminish the quality of digitised image. It may be distinguished two main types of noise sources: deterministic (e.g. optical system imperfections) and stochastic. The restoration

of a monochrome digitised image is based on the following model of the image degradation and image restoration processes.

The original image is represented by the function *o(x, y)* giving the brightness value in the point of co-ordinates *x* and *y*. We assume that the degradation was done by a filter with the transfer function *H(u, v)* and an additive noise *z(x, y)*. For the restoration is responsible a filter with the transfer function *R(u, v)*. The restored image function, *f(x, y)*, must approximate the function *o(x, y)* representing the original image. The degraded image function *d(x, y)* is assumed that is a sum of the result of a space invariant convolution operator *h( )* and the noise function *z(x, y)*:

$$d(x,y)=\int_{-c}^{c}\int_{-c}^{c} h(x,y;\alpha,\beta)o(\alpha,\beta)d\alpha d\beta+z(x,y) \quad (1)$$

where *c* is the window size. This relation may be re-wrote based on space invariant properties as follows

$$d(i,j)=\sum_{k=-c}^{c}\sum_{l=-c}^{c} o(i-k,j-l)h(k,l)+z(i,j) \quad (2)$$

A neural networks having *N.G* neurones is used to restore dgradate images. The grey level of each pixel *r(i, j)* is obtained by the sum of the outputs of all the neurones in the assigned group.

$$r(i,j)=\sum_{k=1}^{G} v_{ij}^{k} \quad (3)$$

where v_{ij}^{k} is the output of *k*-th neurone in the group of *G* neurones assigned to the pixel *(i, j)*. This output is obtained by applying a threshold function *g(u)* to the total input of the neurone. All the neurones are assumed to be interconnected by symmetric links.

We consider a monochrome image digitised in *G* grey levels and $N = x_{max} \cdot y_{max}$ pixels. After the binarisation follows the phase of objects isolation that is based on a convex compounds labelling method. It must be underlined that every connecting area of pixel is interpreted as an object marked by the same label.

After the identification of the object representing a block it must be determined its orientation. The main axes of the object are computed using an analogy to the samples points in an stochastic process, and were approximate by a regression line

$$y = a x + b \quad (4)$$

where the coefficients *a* and *b* are determined by the least square method using the pixels points in the object waiting area image. Their values minimize the function F(a,b), computed with the formula:

$$F(a,b)=\sum_{(xi,yi)\in Object}(y_i-ax_i-b)^2 \quad (5)$$

where the *Object* stands for the set of the pixels in the object image representing the carried block.

Having the coefficients *a* and *b* the main orientation of the block is determined.

3.2 Recognition of standard objects

The significant objects identified on the conveyor band may be standard object (whole blocks), pieces of a fragmented block etc. the standard objects recognition is necessary to send to the robot the order that will move the objects in the standard objects stack, and the other objects in the other stack. The recognition process is based on a the concept of parallel separated patterns [4]. A pattern is described by an array of features or a pattern may be viewed as a point in a n-dimensional space. The similar patterns can be grouped into a finite number of sets. We suppose that these sets are disjoints and we name them classes; their representation in the pattern space consists in a number of regions that may be separable or not. For example a class contains

To recognize a pattern is to classify them into a class, according to the region that contains the pattern point.

The similarity measure used in this case may be an inverse dependence from a distance measure between two patterns $x, y \in E$. In many applications like [1], arises patterns that have not all the *n* features measurable, or the results of their measurements are strongly affected by noise. Suppose there are only *k* out of *n* features measurable for all the patterns. Then we will work in a projection of pattern space. More difficult is the situation that the patterns have not the same features incomplete measured. In this situation the usual solutions can be:

a) isolating incomplete patterns by:

- elimination from the set *E*;
- classification of the incomplete patterns in a special class [1].

b) restoring the missing values by:

- substitution of the missing features by the means of the known values of the features of the other ones patterns;
- computation of missing features by the aid of a regression function determined from complete patterns.

A major disadvantage of these methods is the requirement of a large mount of data to make possible the elimination or restoring the missing data. But in many cases it is not necessarily to compute the missing features. More important is to gain the final objective of the pattern recognition problem. In the supervised learning context the system receives as inputs data a set composed of prototypes with known classification composed of prototypes with known classification:

$$T = \{ (x, C_x) / C_x \subseteq E \text{ and } x \in C_x \} \quad (1)$$

where T is called training set, with C_x a class in the partition P over E. Considering T, the recognition system must assign a class of P to any new pattern y. This assignment may be done directly by a search in T, (pattern matching) or by a classifier that functions as follows:

- for K classes are K functions g_i;
- it classifies the new pattern y in C_i class iff

$$g_i(y) \geq g_j(y) \quad (\forall)\, i,j \leq K \quad (2)$$

The functions $g_i(x)$ may be linearly or non-linearly.

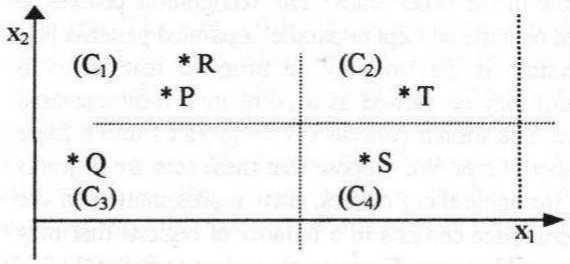


Figure 4. Case of four patterns classes inseparable in an one dimensional feature space (x_1 or x_2).

The quickly classifier is a "third order classifier" type. Clearly, a classifier obtained for complete patterns is inapplicable for patterns with missing features. An expansive solution is to obtain several classifier g_{ij}. For n features it needs a number of N_c classifiers

$$N_c = C_n^1 + C_n^2 + \ldots + C_n^n \quad (3)$$

It is important to observe that the separability of the classes may be not maintained in the projection space (figure 4).

Here we had supposed that the pattern space is a bi-dimensional one and we have four classes labelled by $C_1, \ldots, C_4$. In this case if a feature is missed, it is not possible to separate the pattern in four classes, but only in two (e.g. if x_2 is missing, it may observe a class containing the patterns from $C_1 \cup C_3$ and a class corresponding to $C_2 \cup C_4$). The separabillity analysis is accomplished during the determination of classifiers, when the cases of uncertain separability are rejected for automate recognition. The corresponding objects may be falses or non-standard objects.

The classifiers determined in the learning phase are algorithms with polynomial time complexity very appropriate to the decision making process running in the time that the object is moved from the A TV camera visual area to the area where it will be picked up by the robot arm.

CONCLUSIONS

The installation presents the following advantages:

- the direct transmission of the movement between bilateral toroidal inductors which represent the main fitting of the action motor and the inducting solidary disk by mechanically coupling with the action motor and the inducting solidary disc by mechanically coupling with the activating drummer of the conveying band;
- the expulsion of the reduction gears of revolution which diminishes the global efficiency of the equipment;
- easy and cheaper maintenance;
- protection against shock and trepidations of the mechanic reduction gears of revolution;

The system assured also the increase of the reliability and liability.

The correctness of the robot movements are depenedent of the calibration stages. These stages must be done only if one component of the system (conveyor, robot, cameras, stacks) changed its position.

REFERENCES

[1] Celeux G., Diday E. et all *Dclassification automatique des donnees.* Dunod, Paris, 1989 .

[2] St.Gh.Pentiuc,R.Prodan,R.Haliuc,L.Iurescu. "Monochrome Image Binarization Method Based on a Pattern Recognition Technique". *The II International Conference on Microelectronics and Computer Science,* Vol.II, pp. 203-206. Technique Publishing House, Chişinău Moldavia, 1997.

[3] R. Pentiuc, L. Cantemir. "Consideraţii asupra alterării câmpului magnetic datorită efectului de bord la motoarele hibride cu inductoare toroidale." *First International Conference SIELMEC'97*, Vol. I, pp. 302-305. Technique Publishing House, Chişinău Moldavia. 1997.

[4] St. Gh. Pentiuc. *Aplcatii ale recunoasterii formelor in diagnosticul automat*, Ed. Tehnică, Bucureşti. 1997.

[5] R. Pentiuc. "Le calcul des reactances de magnetisation des enroulements de tipe anneau pour les moteurs hybrides", The *Annals of the "Ştefan cel Mare" University of Suceava,* Electrical Section, No.8, pp. 50-55. 1998.

Towards a Natural User Interface for Comprehensive Support of Conceptual Shape Design

György Kuczogi
Zoltán Rusák
Imre Horváth
Faculty of Design, Engineering and Production
Delft University of Technology
Jaffalaan 9, 2628 BX, Delft,
The Netherlands
TELEPHONE: +31 15 278 1196
FAX: +31 15 278 1839
E-MAIL: g.kuczogi@io.tudelft.nl

ABSTRACT

The lack of possibility of creative communication in the conceptualization phase of the design process significantly decreases the effectiveness. The aim of this research is to find effective non-traditional shape conceptualization techniques.
The paper presents an analysis of the activities in the conceptualization process and the man–machine communication. It also shows how to evaluate and select the most efficient shape conceptualization techniques.

1. INTRODUCTION

1.1. Problem statement

It is known for long time that the lack of creative communication between humans and CAD systems is one of the most significant problems that are hindering efficiency. An adequate interface should allow designers to concentrate only on their specific design task, without paying attention to how communicate with the computer. There exist solutions supporting specific activities of native human communication, but we cannot find any system on the market yet, which wraps up all issues relating the conceptualization design tasks.

1.2. Hypothesis

In our opinion, the ideal computer aided conceptual design (CACD) system would be similar to the one bounded by the dashed rectangle in Figure 1. The ideal system should have an intelligent user interface, which understands the high level natural communication of the human being, corrects the small mistakes without notifying the user, and advises and proposes solutions.

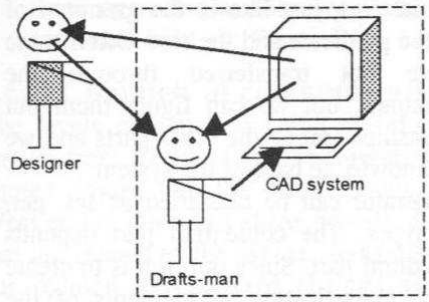


Figure 1 The ideal CACD system should behave like a smart drafts-man and one of the recent CAD software together

1.3. Aims of the research

Our long-term aim is to create a design system, which supports the complete shape conceptualization process by integrating technologies based on the specific needs of the application field.
The objective of the paper is to (a) analyze the shape conceptualization process, (b) discover the significant elements and influencing factors of the communication in the conceptualization process, and (c) show how can we arrive at more effective shape communication techniques.

1.4. Research methodology

Research into creative communication needs a specific methodology. This paper covers those steps that are indicated in italic of our research methodology described below.

- *Definition of application field*
- *Analysis of the activities of computer aided conceptualization*
- *Characterization of the input*
  - *Contents of communication*
  - *Channels of communication*
  - *Level of abstraction in communication*
- *Selection of appropriate techniques for user interface in particular application field*
  - *Elaboration of the aspects of evaluation*
  - Evaluation/ranking of the categories of techniques
  - Selection of the techniques from the top-ranked categories
- Composing a framework of an effective system interface for a particular application field by integrating the top-ranked technologies.
  - Proposing algorithms for the selected technologies
- Implementing the technologies
- Evaluation and testing

© UkrObraz'2000

2. INVESTIGATION OF THE CONCEPTUALIZATION PROCESS

As a starting point of the analysis of the activities related to the application field, we investigate the conceptualization process in general. Afterwards, we look at the specific activities of shape exteriorization.

2.1. The spiral of the conceptualization process

The conceptualization process is a loop formed by steps for creation/modification and evaluation of concepts. In the case of computer supported conceptualization, we can describe the loop by three partial models:

- the mental model of design concepts,
- the computer model of design concepts, and
- the mental model of the model of the computer (we might call it the designer's view on the computer model).

These three models crate a kind of loop based on their contextual dependencies. By introducing one more dimension of investigation for the conceptualization loop, we can represent the evolution of concepts and models as a spiral (Figure 2). The diameter of the spiral shows how close the three models are to each other semantically. During the conceptualization process, each model usually changes and gets closer to the others. When a given computer model resembles the design concept so much that the designer is satisfied with it, the process is finished.

Let us have a closer look at the steps of a general conceptualization process. Consider the following scenario:

1. The three models are given. (The computer model may be an empty model at the beginning.)
2. The designer evaluates the models and the design concept:
   - He compares his view about the computer model to the original mental model to judge how accurately is the concept represented by the computer.
   - Based on his view about the computer model, he investigates the design concept, whether the concept really fulfills the requirements.
3. The designer communicates the needed changes to the computer.
4. The computer modifies its digital model.
5. The computer communicates the modified model to the designer. (The loop continuos with the step 2.)

2.2. Activities of the conceptualization process

There are three activities in this scenario mentioned above, where the computer has important role. Although it needs much deeper investigation to see how can we speed up the conceptualization process, in this subchapter we limit ourselves to the essence of the activities.

The first activity is when the designer communicates the required changes of the model to the computer. The

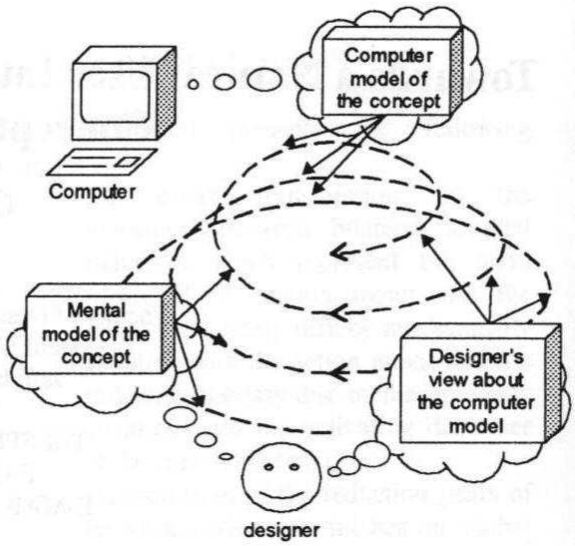


Figure 2 Evolution of models in the conceptualization process

computer must interpret and react on the communicated contents. We will go into a more detailed analysis of it in the next chapter.

The second activity is modification of the computer model. The effectiveness and opportunities of modification are determined mainly by the applied representation. The computer has to be capable to handle the incomplete and vague concepts, which is one of the most dominant characters of conceptualization. We have decided to use discrete model representation. The paper of Rusák et al [1] discusses it in detail.

The third activity is communication of the computer model to the designer. There are several existing output systems on the market. We can find numerous solutions in the field of scientific visualization, for example, displaying the model in virtual reality, producing photo-realistic pictures etc. According to our observations, the output systems tend not to be the bottleneck of the computer support of the conceptualization process.

3. INVESTIGATION OF THE INPUT COMMUNICATION

3.1. Structure of communicated instructions

The communicated instructions have a communication independent structure. Each of them has a procedural part, and a contextual part, just like in the grammar of natural languages: the predicate and the rest. Often some of the parts are not transferred through the communication channels, but we can figure them out based on the relationships among the other parts and we need to rely on the knowledge base of the system.

Each procedural operator can be classified as set, get, create, or destroy types. The contextual part depends largely on the procedural part. Since our aim is to create a shape model, the classification of elements of the contextual part of communication is exactly the same as with the regular geometric models of CAD systems. We

distinguish geometry (morphology and scale), topology, identification information, location, and attributes. Depending on the category of a particular element of a user instruction, it may have a most effective device to communicate it.

3.2. Dependence of the aspects of communication

Referring to the first activity in the earlier mentioned scenario, if the designer is not satisfied with the digital shape model, he wants to change it by communicating one or more instructions to the computer. He can do it in several ways. He may choose from the available modeling operators, communicate his choice on different abstraction levels, and select different kinds of communication channels. We can conclude that the three aspects of communication, i.e. (a) content of communication, (b) the channel of the communication, and (c) the level of abstraction, are independent from each other.

3.3. Contents of communication

In order to describe a domain of the desired shape or modify its attributes, location etc., we use operators. The operators are determined by the their types (procedural part) and parameters (contextual part). We have operators to create, modify, destroy, or get any of the parts of the geometric model. The most influencing operators are the shape modification operators, therefore we present the categories of them based on the theory they apply.

- The create operator is very simple. It does not determine the boundary of the object on a procedural way, but with a large set of discrete points. E.g. this can be applied by using data gloves or 3D scanner.
- The copy operator makes a copy of the boundary or a part of the boundary of a particular object. It facilitates not only the regular copy/paste type of operations, but to apply metaphors, as well.
- We have a number of volumetric operators, which use the set theory, e.g., the "mix" operator [3] uses the union. They are used by all cutting type of modifications.
- Weight operators can modify the shape like parameterized features by weight functions.
- Physical based operators [3] use the theory of finite element methods. We can modify our shapes with these operators as they were physical objects, we can bend, twist etc. them.

3.4. Channels of communication

The levels of communication of information are well known from the literature. The number of levels is still argued issue, but the main points are established. Weaver [2] discusses three aspects of communication that are related to the syntax, semantics, and effects. In our opinion, the selected channel and coding of the communication concerns the physical, statistical, and syntactical levels of communication. The channel can be voice, keyboard, mouse, hand gesture etc., which form the physical layer of communication. We can apply different coding, e.g., in the case of voice based information transfer, we can use natural languages. The coding covers the statistical and the syntactical levels. What we get at the end of the processing of this input is supposed to be syntactically correct information. No matter what sort of channel or coding we choose, the same semantic elements (the type and parameters of the modeling operators) have to be transferred through. Therefore, it has to be possible to convert the input to a channel independent syntax. We can express exactly the same by hand gestures through a sign language as by talking, although both the channel and the coding are different.

3.5. Level of abstraction

We can express the same intent on the same channel several different ways. The basic difference appears in the level of abstraction. This has also something to do with the levels of information communication.

For instance, we can use the lowest syntactic level. In this case, the syntactic description that we get from the processing the input signals is executable without interpretation. This means that the user must follow a prescribed syntax, and each syntactically complete instruction can directly be connected to a modeling operator.

A more advanced level is the semantic level, where we can attach semantic meanings to the content. In our case, the semantics refer to the modeling operators. This level of the communication also allows the user to apply casual syntax, for instance, he may even skip the required parameters. However, the semantic meaning of the communication must be unambiguous. The reason of it is that the computer has to figure out, which modeling operator was intended to be used. Applying the semantic level of communication may make the use of the shape conceptualization system more understandable for the user.

The third level of abstraction is the pragmatic level, when the user does not need to specify at all, the appropriate modeling operator to achieve the desired shape. He can declare requirements for the concept, and the computer has to come up with proposals, how to fulfill. For instance, the user can declare that he wants a more streamlined shape, and the computer is supposed to propose changing the curvatures of specific regions of the shape.

The fourth level is the apobetical level. It relates much more to the human being than to the shape conceptualization. We consider here feelings, motivations, jokes, etc.

4. SELECTION OF INPUT TECHNIQUES

A particular combination of the communication channel, level of communication, and a clearly defined set of modeling methods is called here shape conceptualization

Table 1 The needed knowledge categories for a given communication level

Level	Knowledge required
Physical	Nothing
Statistical	Coding rules
Syntactical	Type information of the categories of the elements of the geometric models Type information and algorithm of operators Relationships between the types of operators and categories of elements
Semantic	Predefined objects, including semantic parts of the objects (e.g.,front, top), what the user can refer to Operators with predefined conditions, which has pictorial meaning for the user Predefined constant attributes, e.g., "red" instead of numbers
Pragmatic	Relationships between the functional requirements and theories of shape modifications
Apobetical	Human feelings, motivations, intentions

technique. The primary issue is to find the most efficient input techniques for our shape conceptualization system.
In order to achieve our goal, we can create a three-dimensional matrix, one aspect of communication represented on each axis. We can indicate several techniques at each cells of the matrix. Obviously the techniques that we put the same cell of the matrix, use the same theory of their operators, the same channel and the same level of communication. (For instance, saying "Take the intersection of the two objects" or "Take the union of the two objects".) There can also be minor differences, e.g., in the expression mode, or some operators may require different number of parameters etc. Although, this may influence the efficiency, we can assume that the efficiency of the techniques that share a cell is nearly the same. Therefore, we can conclude that it is not necessary to name and define all the techniques before evaluating them. It is sufficient to evaluate only the cells of the matrix, and we can define the actual techniques later.
The evaluation procedure is twofold. On the one hand, the theory of the operators and the level of the abstraction determine the actual syntax to be transferred. On the other hand, a particular channel of the communication can be more suitable for transferring the given syntax than other channels. Let us elaborate more about the two issues:
In Chapter 3.3., we have introduced the five theories of shape definition/modification that we want to use. Therefore, we have to choose, which level of abstraction is the most appropriate for the particular theories. Although, we have to send significantly less data to the system by increasing the level of communication, it is not always the case that the higher level we communicate the faster the result is achieved. We have to consider that using high level communication accompanied by uncertainty. The level of communication is proportional the knowledge the system requires (Table 1). If the knowledge of the system about a particular task is not exactly the same as the knowledge of the user then the result can be different that the user expects. We can conclude that the higher level of abstraction we choose to communicate, the less data we have to transfer, and the less predictable the result is.
The evaluation of possible channels of communication (i.e. the device and the coding the channel uses) for transferring a particular syntax can include many aspects e.g., device costs, or the needed time and effort learning how to use the device etc. However, the main point of the efficiency of the transfer is the speed.

5. FURTHER RESEARCH

The research will go on by a comprehensive evaluation of the categories of the shape conceptualization techniques. We are going to use the abstract prototyping tool of Opiyo [4]. Based on the ranked categories, we will define the actual techniques to use. Since our aim is not to produce a fully functioning software prototype but only a proof of idea implementation, we will filter out several techniques, which is not necessary for the proofing and takes too much effort to implement.

REFERENCES

[1] Rusák, Z., Horváth, I., Vergeest, J.S.M., Kuczogi, G., Jansson, J.,"Discrete Domain Representation for Shape Conceptualization", Proceedings of the 4th International Conference of Engineering Design and Automation-EDA 2000
[2] Shannon, C. E., Weaver, W., "The Mathematical Theory of Communication" The Board of Trustees of the University Illinois 1949
[3] Horváth I., Rusák Z., Vergeest J. S. M., Kuczogi, G., "Vague Modeling for Conceptual Design", TMCE 2000 conference, pp. 131-144
[4] Opiyo, E. Z., Horváth, and Vergeest, J. S. M., " Software Tools For Abstract Prototyping Of Design Support Tools", Proceedings of the 20th Computers and Information in Engineering (CIE) Conference: 2000 ASME DETC Conferences, September 10-13, 2000, Baltimore, Maryland, USA, in CD-ROM.

ВИЗНАЧЕННЯ ПЕРІОДУ ПАПІЛЯРНОГО ВІДБИТКА НА ЗОБРАЖЕННІ НА ОСНОВІ ПОБУДОВИ ФУНКЦІОНАЛІВ МАТЕМАТИЧНОГО СПОДІВАННЯ

Ростислав Косаревич, Ігор Ісаєв, Богдан Русин
Фізико-механічний інститут ім Г.В. Карпенка НАН України, 79601, м. Львів, вул. Наукова 5,
тел.(0322) 65-45-30, e-mail: dep32@ah.ipm.lviv.ua

Представлено метод визначення періоду відбитка папілярних ліній на зображенні за допомогою побудови функціоналів математичного сподівання по випадковій реалізації. Розроблено новий підхід для визначення центра відбитка на зображенні.

1. ВСТУП

Сліди папілярних відбитків (ПВ) за своїм ідентифікаційним значенням займають перше місце серед біометричних характеристик особистості таких як обличчя, голос, підпис та ін. Візерунки, які утворюють папілярні лінії (ПЛ) нігтьових фаланг рук не змінюються протягом цілого життя. Якщо у однієї особи відібрати відбитки всіх десяти пальців, то вірогідність збігу із візерунками іншої людини буде рівна 1 з 64 млрд. Взявши до уваги загальну чисельність населення земної кулі, можна вважати, що збіг відбитків пальців двох людей практично неможливий [1].

На даний час широко поширені як автоматичні так і автоматизовані системи для ідентифікації особистості за допомогою візерунків папілярних ліній. Для роботи вони використовують або безпосереднє сканування поверхні нігтьових фаланг, або сканування відбитків візерунків папілярних ліній. В системах ідентифікації особистості, як правило, використовується другий спосіб оскільки він більш зручний. Результатом сканування є зображення розміром 512×512 , що має 256 градацій яскравості сірого кольору (Рис.1).

Для обробки зображень ПВ та ідентифікації особистості на їх основі, важливе значення має величина так званого періоду ПВ. Під ним розуміють суму величин ПЛ та міжпапілярного проміжку (МП) в одиницях виміру зображення. Ефективне покращення зображень ПВ, як правило, полягає у підвищенні контрасту між ПЛ та МП, а також в усуненні шумів на зображеннях. Таке покращення зводиться до застосування певного фільтру із заданими розмірами [2]. Ці розміри мають бути спільномірними з періодом ПВ, крім того коефіцієнти фільтра та їх розташування залежать від діапазону в якому змінюються величини ПЛ та величини МП. Вважається, що товщина ПЛ є величина постійна, але при утворенні ПВ вона змінюється, оскільки поверхня пальця випукла, а слід переноситься на плоску поверхню. Крім того, при утворенні цифрового зображення слід квантується, що також вносить певні спотворення у величини ПЛ та МП, причому ці спотворення не є однаковим для різних ділянок ПВ. Коефіцієнти фільтра розташовуються так, щоб посилити ПЛ на зображенні та послабити МП. Таким чином для побудови фільтра необхідно знати діапазон зміни величини ПЛ та МП.

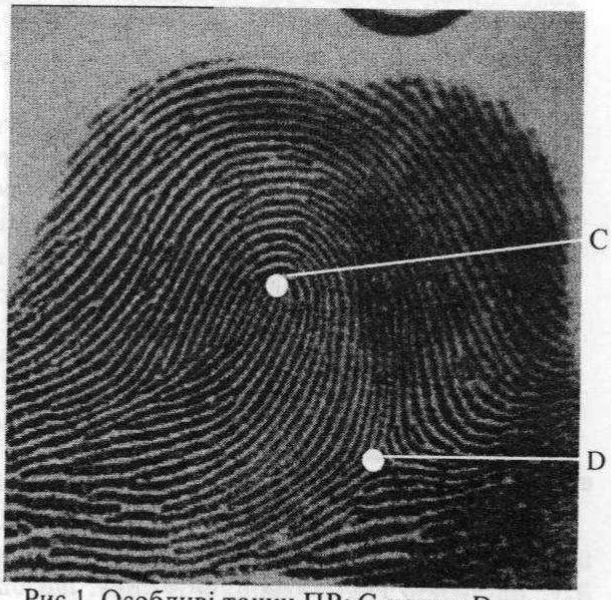


Рис.1. Особливі точки ПВ: C-центр; D-дельта.

В [3] запропонований метод оцінки періоду ПВ на основі профілів рядків та стовпців зображення. Цей метод дає можливість тільки приблизно оцінити межі зміни періоду ПВ. З метою більш точної оцінки пропонується для визначення періоду скористатися елементами теорії ПКВП.

Метод визначення діапазону зміни періоду ПВ за допомогою профілів має своїм недоліком те, що він не враховує розташування папілярних ліній на зображенні. У випадку, якщо ПЛ розташована під кутом відмінним від

© UkrObraz'2000

90° величина періоду буде спотворена і це спотворення тим більше чим менший кут між прямою,що відповідає ряду або стовпцю зображення і папілярною лінією. Для більш точного визначення періоду ПВ та діапазону його зміни запропоновано новий метод, який полягає у скануванні зображення ПВ з його центра у декількох напрямках. На основі отриманих даних за допомогою методики ПКВП обчислюється період ПВ.

2. ЗНАХОДЖЕННЯ ЦЕНТРА ВІДБИТКА

2.1. Позначення і означення

Нехай D так званий "directional image" - поле напрямків напівтонового зображення G розміром $g \times g$, яке у випадку зображення ПВ містить ПЛ та МП. Напрямок у кожній точці представляє локальну орієнтацію ПЛ в цій точці. Напрямок $D_{(i,j)}$ в точці (i,j) зображення – це напрямок d, для якого S_d – сума різниць значень яскравості на зображенні вздовж напрямку d є мінімальною. Ця сума визначається як

$$S_d = \sum_{m=0}^{M-1} \left| G(i_m, j_m) - G(i, j) \right|, \text{ для d =1,...,N,}$$

де $G(i,j)$ та $G(i_m,j_m)$ – значення яскравості в точках (i,j) та (i_m,j_m) відповідно. Точка, яка відповідає (i_m,j_m) є m–ою точкою вздовж напрямку d від (i,j), M–кількість точок вибрана для обчислення , N – кількість напрямків.

На основі поля напрямків D отримується поле напрямків блоків B (Рис.2) за допомогою знаходження домінуючого напрямку в блоці точок розміром $b_s \times b_s$. Це також зменшує розмірність з $(g \times g)$ до $(g/b_s \times g/b_s)$. Операції виконуються за допомогою обчислення гістограми напрямків H, яка формується на основі поля напрямків блоків B. Гістограма напрямків H – послідовність $\{H(0),...,H(N-1)\}$, де $H(j)$ – частота напрямку j в деякому околі.

В наших дослідженнях ми використовували $g=512$, $N=8$, $M=8$, $b_s=16$. Вибір величини g пов'язаний із вимогою якнайкращого представлення ПЛ на зображенні. Величина блоків b_s вибирається із припущення, що один блок містить принаймні одну ПЛ та МП. Величини кількості напрямків N та кількості точок M необхідних для обчислення напрямку пов'язані між собою. Більше число напрямків N вимагатиме збільшення кількості точок M. В наших експериментах, а також в інших дослідженнях відомих з літератури [4,5] вибір величини $N=8$ дав добрі результати. Більше N дає більш точніше представлення напрямків ПЛ на зображенні але також спричинює значне зростання обчислень для кожної точки зображення.

Гістограма напрямків H може бути наступних видів[4]:

а) пік гістограми зліва – коли напрямки менші або рівні ніж 90° домінуючі, іншими словами, сума частот напрямків які менші ніж 90° перевищує суму частот напрямків більших ніж 90°;

б) пік гістограми справа – коли напрямки більші або рівні ніж 90° домінуючі;

в) піки гістограми на краях – коли домінуючими є напрямки, які не належать діапазону 45°–135°.

Особливими точками зображення ПВ є точки центрів та точки дельт (Рис.1). За означенням точкою центру ПВ є точка, яка знаходиться на вершині найбільш внутрішньою ПЛ центрального потоку ПВ. Для точки центру ПВ область, що розташована зліва від неї буде мати гістограму, пік у якої зліва, а для області, що розташована справа – гістограма буде мати пік справа. Точкою дельти є точка в якій с три потоки ПЛ зустрічаються один з одним. Тому для такої точки додатково до критерію точки центру повинна виконуватись умова, що гістограма області, яка розташована нижче дельти, матиме піки на краях.

2.2. Знаходження особливих точок ПВ.

Оскільки поле напрямків блоків B зберігає характеристики напрямків ПЛ на багатоградаційному зображенні, то подальші дії по виділенню і знаходженню особливих точок ПВ доцільно проводити на ньому.

2.2.1. Виділення.

Якщо елементам B поставити у відповідність числа від 1 до 8 за наступним правилом: 0°,180°→1, 22°30',→8, 45°→7, 67°3'→6, 90°→5, 112°3'→4, 135°→3, 157°3'→2, то як показали дослідження необхідними умовами для віднесення точки до розряду особливих є наступні вирази:

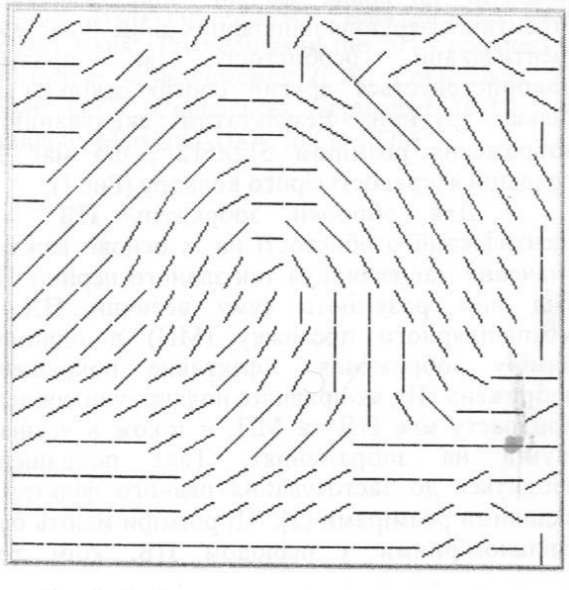


Рис.2. Зображення поля напрямків блоків B.

$$B[j,i]=5 \wedge B[j,i+1] \in \{3,4\},$$
$$B[j,i]=6 \wedge B[j,i+1] \in \{2,3,4,5\}, \quad (1)$$
$$B[j,i]=7 \wedge B[j,i+1] \in \{3,4,5\},$$
$$B[j,i]=8 \wedge B[j,i+1]=2.$$

При виконанні хоча б однієї умови (1) $B[j,i]$ відноситься до особливих точок ПВ. Про те спотворення вихідного зображення G можуть спричинити виявлення хибних особливих точок ПВ. Для того щоб цього не було запропоновано для кожної точки, що задовольняє вирази (1) перевіряти виконання наступної умови. Якщо позначити точки околу $B[j,i]$розміром 3×3 через $c_1, c_2, \ldots, c_8$, то точка $B[j,i]$ не береться до уваги, якщо для неї не виконується умова $\sum_{i=1}^{7} |c_i - c_{i+1}| \le p$. Невиконання такої умови означає, що окіл точки $B[j,i]$ є однорідним, а в самій точці є локальне збурення напрямку, що можливе при спотвореннях зображення G.

2.2.2. Кластеризація.

Наступним кроком знаходження особливих точок ПВ є кластеризація. Вона необхідна оскільки за означенням центра ПВ це є верхня точка найбільш внутрішньої ПЛ. Тому умовам (1) можуть задовольняти точки, які розташовані поряд на сусідніх ПЛ. Для зменшення їх кількості використовується алгоритм знаходження центрів кластерів. Далі особливими точками ПВ вважаються точки, що відповідають центам виявлених кластерів. Для виявлення центрів кластерів використовується проста евристична процедура наведена у [6]. Її перевага в невеликих обчислювальних затратах, оскільки для виділення центрів кластерів достатньо однократного перегляду даних.

2.2.3. Відокремлення центра.

Після проведених операцій множина знайдених особливих точок ПВ мітить точку, що відповідає центру ПВ та точки, які відповідають точкам дельт ПВ (якщо є). Для того щоб виділити єдину точку центра пропонується наступна процедура.

Спочатку розглядається окіл розміром 5×5 навколо кожної особливої точки. Цей окіл розбивається на чотири частини $R1$, $R2$, $R3$, $R4$ (Рис. 3).

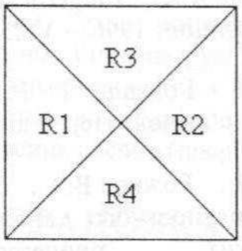


Рис. 3. Частини околу особливої точки.

Далі для кожної з частин $R1$, $R2$, $R3$, $R4$ формуються гістограми напрямків $H1$, $H2$, $H3$, $H4$ і в кожній гістограмі вибирається напрямок h_1, h_2, h_3, h_4, що має максимальну частоту. Якщо для особливої точки ПВ величин h_1, h_2, h_3, h_4 задовольняють хоча б одній з наступних умов

$$(h_1=4)\wedge(h_2\in\{5,6\})\wedge(h_3\in\{1,2,7,8\})\wedge(h_4\in\{3,4,5,6,7\}),$$
$$(h_1=5)\wedge(h_2\in\{4,5,6\})\wedge(h_3\in\{1,2,7,8\})\wedge(h_4\in\{3,4,5,6,7\}),$$
$$(h_1=6)\wedge(h_2\in\{3,4,5\})\wedge(h_3\in\{1,2,7,8\})\wedge(h_4\in\{3,4,5,6,7\}),$$
$$(h_1=7)\wedge(h_2\in\{4,5,6\})\wedge(h_3\in\{1,2,7,8\})\wedge(h_4\in\{3,4,5,6,7\}),$$
$$(h_1=8)\wedge(h_2\in\{3,4,5,6\})\wedge(h_3\in\{1,2,7,8\})\wedge(h_4\in\{3,4,5,6,7\}),$$

то вважаємо, що ця особлива точка – шуканий центр відбитка. Наведені умови відображають наведене в розділі 2 означення центра ПВ.

3. ФОРМУВАННЯ ВИПАДКОВОЇ РЕАЛІЗАЦІЇ

Після того як знайдено центр ПВ необхідно сформувати дані для обчислення періоду ПВ. Для цього з точки центра ПВ вздовж восьми ліній до перетину з краєм зображення G зчитуємо значення елементів (Рис. 4).

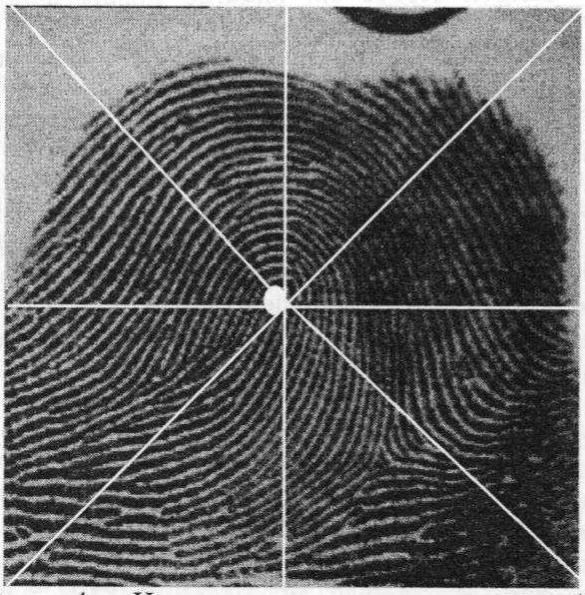


Рис. 4. Напрямки зчитування елементів зображення

Вибір напрямків зчитування пов'язаний із тим, що при правильній орієнтації ПВ проведені лінії в більшості випадків утворюють кути близькі до 90° із ПЛ на зображенні. Це дасть змогу більш точно оцінити період ПВ.

4. ЗНАХОДЖЕННЯ ПЕРІОДУ ПАПІЛЯРНОГО ВІДБИТКА

Оскільки у випадковій реалізації, сформованій по ПВ, присутні періодичні коливання, модульовані випадковими збуреннями, найбільш адекватною моделлю для її дослідження і пошуку періоду, є модель у вигляді періодично корельованих випадкових

процесів (ПКВП). Ця модель є узагальненням детерміністичного і стаціонарного підходів до обробки сигналів. ПКВП визначаються як періодично нестаціонарні випадкові процеси з інваріантними відносно періоду корельованості математичним сподіванням і кореляційною функцією

$$m(t) = E\xi(t) = m(t+T),$$

$$b(t,u) = E\overset{\circ}{\xi}(t)\overset{\circ}{\xi}(t+u) = b(t+T,u),$$

$\overset{\circ}{\xi}(t) = \xi(t) - m(t)$. T - період корельованості процесу [7].

На основі теорії ПКВП у ФМІ НАН України були розроблені методи пошуку прихованих періодичностей, які попередньо грунтувалися на одній з перших історичних схем для виявлення періодичностей - схемі Бюй-Балло. Ця схема має певні селективні властивості і при обробці реалізацій ПКВП. Після дослідження властивостей різних методів оцінки періоду, було показано, що метод максимальної правдоподібності дає асимптотично незміщену і асимптотично ефективну оцінку періоду корельованпості ПКВП. Проте він вимагає знання детальної структури кореляційної функції. Були розроблені методи, які й програють по збіжності порівняно з методом максимальної правдоподібності, проте є більш простими по своїй реалізації, і менш вибагливими до апріорним відомостям про характеристики процесу. Ці методи базуються на властивостях функціоналів, побудованих подібно до виразів для когерентних і компонентних статистик ПКВП, з тою відмінністю, що замість істинного значення періоду, в них використовується деяка пробна величина.

Оскільки властивістю періодичності володіють математичне сподівання, дисперсія - на рівні перших моментів, кореляційна функція, спектральна густина і їх компоненти - на рівні других моментів, то побудова функціоналів для оцінювання періоду можлива по всіх цих характеристиках випадкового процесу. В даній роботі обмежимось дослідженням функціоналів, що побудовані на властивостях математичного сподівання.

Аналогічно до когерентного методу оцінювання ПКВП, побудуємо когерентний функціонал по математичному сподіванню для оцінювання періоду корельованості ПВ:

$$\hat{m}(t_0, \tilde{T}) = \frac{1}{2N+1} \sum_{n=-N}^{N} \xi(t + n\tilde{T})$$

Когерентний функціонал при $\tilde{T} = kT$, k - ціле число, $\tilde{T}$ - пробний період, приймає максимальні значення, що дозволяє уточнити межі, в яких знаходиться значення основного періоду ПЛ. Оскільки особливістю когерентного функціоналу є відбирання значень через пробний період, то за його допомогою можливе знаходження тільки цілих значень.

Для точнішого визначення періоду використаємо схему, основану на інтегральних перетвореннях Фур'є. На рівні математичного сподівання розглянемо компонентний функціонал

$$\hat{m}_k(\tilde{T}) = \frac{1}{\theta} \int_{-\theta}^{\theta} \xi(t)\cos(\tilde{\Lambda}_k t)dt\,,\ \tilde{\Lambda}_k = \frac{2k\pi}{\tilde{T}}\,.$$

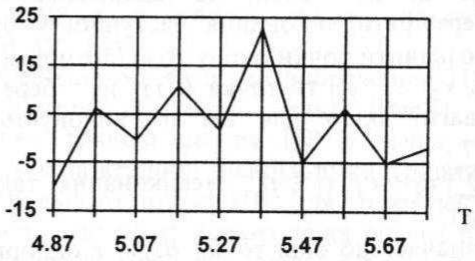


Рис. 5. Результат обчислення функціоналу 2-го компоненту математичного сподівання.

ВИСНОВКИ

Запропонований підхід до визначення особливих точок ПВ дозволяє ефективно проводити виділення центра відбитка.

По сформованій реалізації можливе оцінювання періоду ПВ за допомогою компонентних функціоналів математичного сподівання. Більш точне значення періоду досягається при використанні функціоналів по дисперсії та кореляційній функції.

ЛІТЕРАТУРА

1. Дубовий О.П., Лукашенко В.Я., Рибалко Я.В. та ін. "Криміналістичне дослідження слідів рук": Науково-практичний посібник. / За ред. Я.Ю. Кондратьєва. - К.: Атіка, 2000. - 152 с.: іл.
2. O`Gorman L., Nickerson J.V. "An approach to fingerprint filter design" // Pattern recognition.- 1989.- Vol. 22.- №. 1.- P. 29-38.
3. Косаревич Р.Я. "Метод визначення діапазону зміни періоду папілярного відбитка" / Матеріали XV відкритої науково-технічної конференції молодих науковців і спеціалістів ФМІ ім. Г.В. Карпенка НАН України "КМН - 2000", Львів, 2000. - С. 110-111.
4. Srinivasan V.S., Murthy N.N. "Detection of singular points in fingerprint images" // Pattern recognition.- 1992.- Vol. 25.- №. 2.- P. 139-153.
5. Karu K., Jain A.K., "Fingerprint classification" // Pattern Recognition, 1996. - Vol. 29. - №. 3. - P. 389-404.
6. Ту Дж., Гонсалес Р. Принципы распознавания образов: Пер с англ. - М.: Мир, 1978.- 411 с.
7. Драган Я.П., Рожков В.А., Яворский И.Н. "Методы вероятностного анализа ритмики океанологических процессов".- Л.: Гидрометеоиздат, 1989. - 319с.

АЛГОРИТМ КЛАСИФІКАЦІЇ РОСЛИННИХ ОБ'ЄКТІВ З МОЖЛИВІСТЮ ШВИДКОЇ АДАПТАЦІЇ

Д. М. Піза, С. В. Морщавка
Запорізький державний технічний університет
м. Запоріжжя, вул. Жуковського 63
тел.: (0612)644662, факс.: (0612)642774 e-mail: svmorsh@zstu.edu.ua

У статті розглядаються питання розподілу об'єктів на два класи за допомогою алгоритму, який формує лінійні вирішальні межі по мінімуму метрики Махаланобіса. Вирішується задача скорочення часу навчання класифікатору. Пропонується алгоритм, який дозволяє, навіть при використанні великої кількості ознак, проводити навчання у режимі реального часу.

ВСТУП

У сучасних дослідженнях все частіше використовується багатоканальна зйомка у широкому оптичному діапазоні, від ультрафіолетової області до довгохвильової частини інфрачервоного діапазону. Спектральна інформація, що отримується при зйомці багатоканальними датчиками, часто служить для виділення одних об'єктів на фоні других. При цьому у якості ознак для такого розподілу служать коефіцієнти спектральної яскравості на різних довжинах хвиль, що відповідають різним діапазонам.

Прикладом таких багатоканальних систем може бути розпізнавання різноманітних об'єктів з космосу. З другого боку, широке використання флуоресцентної мікроскопії і спектроскопії для дослідження біологічних об'єктів можна також віднести до подібних задач. Ще однією важливою галуззю використання багатоканальної інформації про оптичні властивості об'єктів є сільське господарство. Така інформація може бути потрібна не тільки для визначення врожайності або засміченості посівів з космосу [1], але й для безпосереднього управління сільськогосподарською технікою в залежності від стану або виду рослини, яка обробляється у цей час.

Така технологія була запропонована в ЗДТУ [2]. В ній, зокрема, коефіцієнти спектральної яскравості використовуються для розділення культурних рослин та бур'янів. Рішення про клас рослини, що обробляється, використовується в подальшому для вибіркового знищення бур'янів (прополки) або часткового корисних рослин (проріджування).

1. ПОСТАНОВКА ЗАДАЧІ

Як показують дослідження, що були проведені як у ході розробки означеної технології, так і іншими вченими [1, 3] існує декілька особливостей при розробці методів розпізнавання по багатоканальних даних.

По-перше, для отримання більшого обсягу вхідної інформації і, відповідно, більш якісного розпізнавання, необхідно використовувати як можна більше ознак. Наприклад, програмний комплекс MultiSpec, який був розроблений в університеті Пурдю (Індіана, США), використовує у якості вхідних даних знімки поверхні Землі, які зроблені літаковим спектрометром AVIRIS на 220 довжинах хвиль [3]. Нове покоління спектрометрів може дати ще більше інформації – до декількох тисяч спектральних інтервалів.

Друга особливість – все більша кількість задач потребує, щоб класифікація об'єктів проходила у режимі реального часу. Очевидно, що чим більше спектральних інтервалів (ознак) використовується, тим важче цього досягти.

Ще однією особливістю є необхідність використання алгоритмів з можливістю швидкого навчання або адаптивних алгоритмів для більш точного урахування змінних у часі оптичних властивостей об'єктів.

Всі ці особливості чітко прослідкувалися і у ході вирішення задачі розпізнавання видів рослин для використання у новітніх сільськогосподарських технологіях. Як показали дослідження, що були проведені у 97-99 роках [4], для розпізнавання краще усього використовувати не одну-дві, а декілька ознак. Було також встановлено, що оптичні властивості рослин змінюються від поля до поля та залежать від періоду вегетації. Однак при цьому все рівно зберігаються відмінності культурних рослин від бур'янів. Таким чином, якщо проводити параметричну адаптацію класифікатора, то можна досягти практично стабільної якості розділення рослин.

Для класифікації виявилося можливим використовувати оптимальні алгоритми, що приймають рішення по максимуму правдоподібності. Така можливість є завдяки наявності нормального розподілу оптичних властивостей рослин [1]. Найбільш оптимальним по співвідношенню швидкості та якості розпізнавання було визначено алгоритм, який застосовується у лінійному дискримінантному аналізі. Він використовує

© UkrObraz'2000

класифікацію по мінімуму метрики Махаланобіса при одній коваріаційній матриці для обох класів (модель Фішера) [4]. Схема відповідного класифікатора зображена на рис.1.

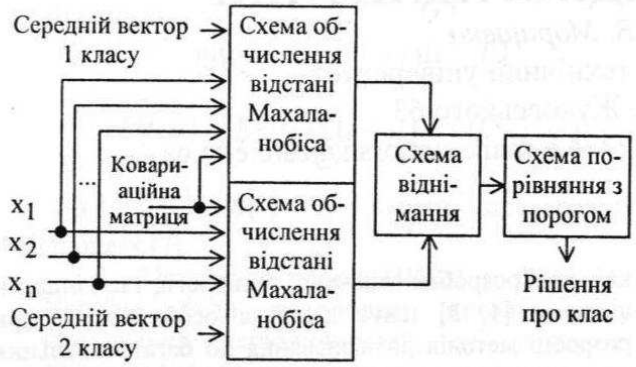


Рис. 1. Схема класифікатора.

При цьому, якщо вважати вірогідності появи і штрафи за переплутування класів однаковими, то вирішальне правило виглядає як [5]:

$$\left(\mathbf{X_k} - \frac{\mathbf{U_1}+\mathbf{U_2}}{2}\right)^t \mathbf{V}^{-1}(\mathbf{U_2}-\mathbf{U_1}) \underset{H_2}{\overset{H_1}{\lessgtr}} 0, \qquad (1)$$

де: $\mathbf{X_k}$ – $\mathbf{k}$-й тестовий вектор; $\mathbf{U_i}$ – середній вектор для класу i $(i=1,2)$, якого отримано по навчальних вибірках; ^{t} – знак транспонування; H_i – рішення, яке приймається при дотриманні одного з відповідних нерівенств $(i=1,2)$; $\mathbf{V^{-1}}$ – обернена коваріаційна матриця, що відшукується за формулою:

$$\mathbf{V}^{-1} = \left(\frac{\mathbf{V_1}+\mathbf{V_2}}{2}\right)^{-1}, \qquad (2)$$

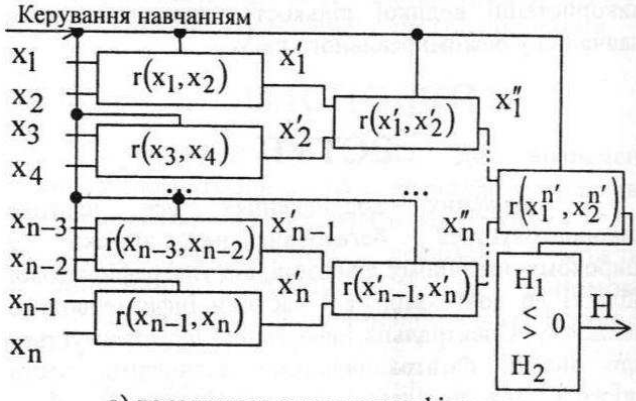
где: $\mathbf{V_i}$ – коваріаційна матриця для класу i $(i=1,2)$, яка отримується по навчальних вибірках.

Такий алгоритм дозволяє реалізувати процедуру розпізнавання з мінімальним використанням обчислювальних ресурсів. Ще однією перевагою такого алгоритму, як і всіх параметричних, слід вважати достатньо просту процедуру його адаптації до змінних властивостей класів рослин. Для такого навчання достатньо тільки заново, або за допомогою оцінки, що ковзає, отримати нові вектор середніх та коваріаційні матриці, що описують властивості класів рослин.

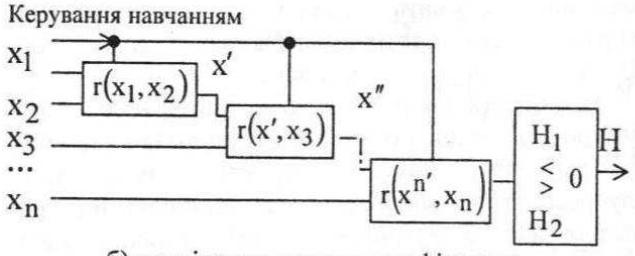
Однак при безпосередній реалізації навчання або адаптації класифікатора рослин по багатоканальних даних виникає одна суттєва проблема. Вона полягає у необхідності усякий раз обертати коваріаційну матрицю, що входить у (1). При цьому у разі збільшення кількості спектральних інтервалів (ознак) – ситуація значно ускладнюється. Тільки кількість операцій множення та розділення, які виконуються найбільш довго, росте пропорційно п'ятому ступеню розмірності коваріаційної матриці. Таке виходить, якщо знаходити мінори через приведення матриць до трикутного вигляду. Якщо ж використовувати прямий алгоритм для знаходження відповідних детермінантів, то кількість операцій зростає пропорційно факторіалу розмірності. Це значно знижує швидкість адаптації, і тому не дозволяє проводити навчання у реальному часі. Спосіб вирішення цієї проблеми пропонується у наступному розділі.

2. МЕТОДИ ВИРІШЕННЯ

Для зменшення витрат часу, необхідного для адаптації класифікатора, пропонується замінити традиційну схему з одноетапним пониженням розмірності інформаційного вектора (рис.1), на кількоетапну (рис.2).



а) паралельна схема класифікатора;



б) послідовна схема класифікатора.

Рис. 2. Схеми класифікаторів з поступовим зниженням розмірності.

При цьому виграш буде досягатися за рахунок невеликої кількості операцій у комірках пониження розмірності, и того, що кількість таких комірок зростає набагато повільніше п'ятого ступеню розмірності.

Очевидно, що найбільший виграш за швидкістю буде при використанні схеми з комірками, що мають два інформаційні входи и один керуючий (рис. 2). Для того щоб з'ясувати принцип функціонування таких комірок слід розглянути процедуру класифікації двовимірних векторів до одного з двох класів. Відповідна геометрична інтерпретація для цієї ситуації при застосуванні моделі Фішера наведена на рис.3.

Задача класифікації в такому випадку полягає у тому, щоб вирішити по яку сторону від межі

знаходиться поточна точка (x_{km}, x_{kh}). Для того, щоб визначити це можна використати відстань r від цієї точки до межі, що вирішує.



Рис. 3. Геометрична інтерпретація класифікації в двовимірному просторі ознак.

Враховуючи, що межа лінійна, відстань від точки до неї можна знайти як:

$$r = \frac{1}{\sqrt{a^2 + b^2}}(a \cdot x_{km} + b \cdot x_{kh} + c), \qquad (2)$$

де a,b,c – постійні коефіцієнти, які входять в рівняння лінійної межі. Їх можна знайти з рівняння межі, яке виводиться з (1). Так, рівнянням межі, що вирішує, буде:

$$\left(\mathbf{X_k} - \frac{\mathbf{U_1} + \mathbf{U_2}}{2}\right)^t \mathbf{V}^{-1}(\mathbf{U_2} - \mathbf{U_1}) = 0, \qquad (3)$$

або, якщо розписати для двовимірного простору ознак:

$$\left[x_{km} - \frac{u_{1m} + u_{2m}}{2} \quad x_{kh} - \frac{u_{1h} + u_{2h}}{2}\right] \cdot \rightarrow$$
$$\rightarrow \cdot \begin{bmatrix} v_{mm} & v_{hm} \\ v_{hm} & v_{hh} \end{bmatrix}^{-1} \cdot \begin{bmatrix} u_{2m} - u_{1m} \\ u_{2h} - u_{1h} \end{bmatrix} = 0 \qquad (4)$$

де: (x_{km}, x_{kh}) – координати точки в двовимірному просторі ознак на виході m-го і h-го спектральних каналів; v_{mm} і v_{hh} – варіації ознак m-го і h-го каналів, відповідно; v_{mh} – коваріація ознак між m-м і h-м каналами; (u_{1m}, u_{1h}) і (u_{2m}, u_{2h}) – координати середніх векторів першого та другого класів, відповідно.

Для зручності розрахунків можна помножити ліву і праву частини цього рівняння на детермінант коваріаційної матриці, вважаючи, що він не нульовий. В протилежному випадку, якщо детермінант дорівнює нулю, то межа, що вирішує, вироджується у точку на прямій, уздовж якої розмістилися всі навчальні вибірки (одновимірна задача класифікації), і тоді формула (1) стає непридатною для обчислень. Проводячи прості перетворення можна знайти коефіцієнти:

$$a = v_{hm} \cdot (u_{2h} - u_{1h}) - v_{hh} \cdot (u_{2m} - u_{1m}), \qquad (5)$$

$$b = v_{hm} \cdot (u_{2m} - u_{1m}) - v_{mm} \cdot (u_{2h} - u_{1h}), \qquad (6)$$

$$c = \tfrac{1}{2} \cdot \left(v_{hh} \cdot \left(u_{2m}^2 - u_{1m}^2\right) + \rightarrow\right.$$
$$\rightarrow + v_{mm} \cdot \left(u_{2h}^2 - u_{1h}^2\right)\Big) - \rightarrow \qquad (7)$$
$$\rightarrow - v_{hm} \cdot (u_{2m} \cdot u_{2h} - u_{1m} \cdot u_{1h})$$

На практиці, для прискорення класифікації поточної вибірки, зручніше користуватися не формулою (4), а більш простим співвідношенням:

$$r' = \frac{1}{|a|}(a \cdot x_{km} + b \cdot x_{kh} + c). \qquad (8)$$

При цьому:

$$r'(x_{km}, x_{kh}) = K \cdot r(x_{km}, x_{kh}), \qquad (9)$$

де K – коефіцієнт пропорційності. Можливість такого підходу пояснюється тим, що абсолютне значення величини r не грає великого значення. Для класифікації вибірки в такому просторі важливо знати її знак. При обчисленні відстані (8) з використанням коефіцієнтів (5)...(7), якщо її знак буде позитивним, то, згідно (1), рішення приймається на користь другого класу, та навпаки.

Пристрій, що реалізує (8), можна представити як комірку пониження розмірності від двовимірного простору ознак до одновимірного. Стан виходу такого пристрою у свою чергу можна вважати ознакою, що містить у собі об'єднану інформацію з двох вхідних каналів.

Таким чином, при використання комірок пониження розмірності з двома інформаційними входами, можна запропонувати варіанти побудови класифікаторів: с паралельною структурою (рис. 2а), с послідовною структурою (рис. 2б) або їх комбінацію.

Паралельний і послідовний варіанти досліджувалися з метою перевірки їх роботоспроможності.

3. ЕКСПЕРИМЕНТАЛЬНІ ДОСЛІДЖЕННЯ

Для досліджень використовувалися данні, що були отримані у ході вивчення спектрів відбиття від різних культурних рослин і бур'янів. У сукупності даних були представлені наступні рослини: культурні – соняшник та кукурудза (по 52 вибірки даних), бур'яни – щириця, лобода, молочай, мишій, осот, пирій (по 24 вибірки даних). В кожен набір даних входили відносні (с точністю до постійного множника) коефіцієнти відбиття від рослин або їх надземних частин. Такі коефіцієнти були отримані

на 55 довжинах хвиль від 300 до 850 нм з кроком 10 нм і смугою випромінювання 10 нм. З них безпосередньо використовувалися 32 з найбільш інформативних частин спектру – від 420 до 650 нм та від 720 до 790 нм включно.

Експеримент проводився наступним чином. Вибірки корисних рослин розділялися на дві частини за видами рослин. Відповідно, експеримент також поділявся на дві частини: навчання і розділення кукурудзи та всіх видів бур'янів на два класи; перенавчання і розділення соняшників та всіх видів бур'янів на два класи. Це відповідає переходу від поля з однією культурою до поля з другою. Очевидно, що задача переходу з поля на поле з одним і тим же видом рослин, але з декілька відмінними оптичними характеристиками буде вирішуватися ще простіше.

Безпосереднє навчання та перенавчання розподілу рослин на два класи проходило наступним чином. В кожній частині експерименту випадково обиралася одна вибірка, та, маючи апріорну інформацію про її клас, змінювалися відповідні середній вектор і коефіцієнти варіації та коваріації в усіх комірках послідовно по шляху розповсюдження сигналу. Після цього оцінювалася середня по усім вибіркам вірогідність вірного розпізнавання без зміни параметрів. Така послідовність процедур повторювалася багато разів. В результаті такого навчання добре видно підвищення середньої по двом класам вірогідності вірного розпізнавання до 85% (кроки 1…50 на рис. 4).

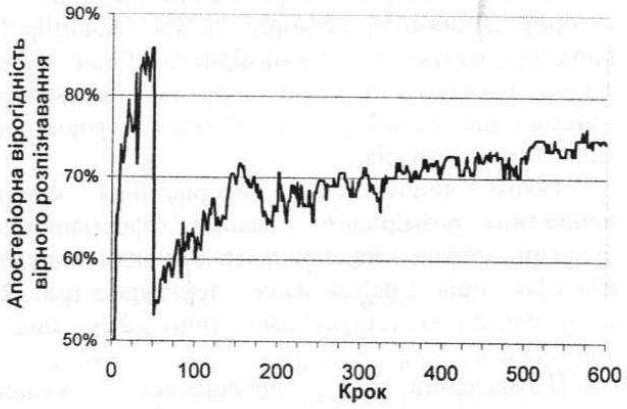


Рис. 4. Результати експериментів.

Після того, як вірогідність правильного розпізнавання переставала рости, набір вибірок замінювався на інший. При цьому вірогідність вірного розпізнавання різко зменшувалася (крок 51). И тільки після перенавчання вірогідність знов виростала до значень біля 75%. Слід відмітити, що, як і очікувалося, на початкове навчання потрібно набагато менше часу, ніж на перенавчання. В ході експериментів, для отримання середніх значень и коваріацій використовувалися згладжені оцінки.

Рис. 4 ілюструє результати, отриманні при дослідження паралельної схеми побудови класифікатора. У випадку використання послідовної схеми досягти таких результатів виявилося неможливим. Проблема полягає в тому, що при використанні поточної навчальної вибірки необхідно оцінювати середні значення та варіації і коваріації кожної комірки першого ступеню. Після цього знаходиться їх реакція на цю вибірку. И тільки потім оцінюються параметри у комірках наступного, по ходу розповсюдження сигналу, ступеню. Таким чином, для послідовної схеми шлях навчального сигналу зі входу на вихід виявляється довшим, ніж для паралельної схеми. Тому, внаслідок випадкового характеру навчальних вибірок, накопичуються помилки. В результаті такий класифікатор гірше навчається.

ВИСНОВКИ

Як показали проведені дослідження, запропонований алгоритм класифікації по багатоканальним даним дозволяє виконувати навчання і перенавчання у реальному часі. Кількість ознак при цьому може бути скільки завгодно великою. В той же час, зберігаються всі переваги алгоритмів, які класифікують по максимуму правдоподібності.

З двох запропонованих схем класифікаторів була обрана паралельна тому, що вона забезпечує більшу швидкість навчання. Послідовна схема простіше з точки зору програмної реалізації, але вона не забезпечує достатню точність и якість навчання при великій кількості ознак.

Ще однією важливою перевагою запропонованого алгоритму є можливість розпаралелення роботи комірок класифікатора, що визначається самою його структурою. Це дозволяє, навіть у випадку використання дуже великої кількості ознак, відносно просто забезпечити роботу у режимі реального часу.

ЛІТЕРАТУРА

1. *Рачкулик В.И.,Ситникова М.В. Отражательные свойства и состояние растительного покрова.-Л: Гидрометеоиздат, 1981.-287с.*

2. *Заявка на патент №94010291 (Украина). МКИ А01 В 39/18. Способ прополки пропашных культур и устройство для его осуществления/ Пиза Д.М., Пиза А.Д., Татарчук И.Н.–Заявл. 17.12.92 г.*

3. *Hoffbeck, Joseph P. and David A. Landgrebe, Classification of Remote Sensing Images having High Spectral Resolution, Remote Sensing of Environment, Vol. 57, No. 3, pp 119-126, September1996.*

4. *Пиза Д.М., Морщавка С.В., Скоробогатов Ю.В. Выбор эффективного метода распознавания растений по коэффициентам спектральной яркости., «Радиоэлектроника, информатика, управление», №1, Запорожье:ЗГТУ,1999р., с.85-88.*

5. *Айвазян С.А., Бухштабер В.М., Енюков И.С., Мешалкин Л.Д.,Прикладная статистика. Классификация и снижение размерности.Справочное издание., М.: "Финансы и статистика", 1989,– 530 с.*

ІНТЕЛЕКТУАЛІЗАЦІЯ ВІДЕОСИСТЕМ КОЛОВОГО ОГЛЯДУ

Боюн В. П., Довгань В. В.

Інститут кібернетики ім. В.М. Глушкова НАН України, Київ-187, пр. Академіка Глушкова, 40,
Факс 266-33-48, тел. 266-21-49, E-mail: vboyun@boyun.icfcst.kiev.ua

Запропоновано спосіб 2-х етапного перетворення відеопослідовності, отриманої з поворотної відеокамери, для підвищення "інтелектуальності" систем колового огляду. Перший етап – перетворення відеопослідовності у статичне панорамне зображення. Другий етап – виділення і реєстрація різницевого зображення між двома обертами відеокамери. При цьому значно покращуються умови роботи оператора, який здійснює контроль колового огляду, а також створюються умови для автоматичного виділення рухомих об'єктів або об'єктів, що змінюються, забезпечується значна компресія зображення. Алгоритм може служити основою створення інтелектуальних систем колового огляду.

ВСТУП

В нинішній час спостерігається широкий розвиток систем відеоспостереження (охоронні системи колового огляду, системи слідкування, системи спеціального призначення і таке інше). З'являються системи видеоспостереження за виробничими процесами, об'єктами. Великі обсяги відеоинформації та складність аналізу рухомого на екрані зображення створюють дуже важкі умови для роботи оператора. Виходом з цього положення може бути підвищення інтелектуальності систем відеспостереження, які можуть автоматично виділяти переміщення об'єктів, що відбуваються у зоні спостереження, або зміни геометричних розмірів, яскравості, кольору.

Найбільш розповсюдженими засобами компресії динамічних зображень, що враховують зміни у відеопослідовності, є стандарти MPEG-1, MPEG-2, які являють собою комбінацію ряду засобів компресії. В них використовуються засоби як просторової (всередині кадра), так і міжкадрової компресії. До перших відносяться засоби дискретно-косинусного перетворення і кодування по засобу RLE. Міжкадрова компресія базується на компенсації переміщення (оцінці вектору переміщення для блоку пікселів) і двунаправленої інтерполяції кадру по двом сусіднім [1,2].

Ці засоби є універсальними і забезпечують достатньо високий ступінь компресії динамічних видезображень. Однак вони вимагають виконання значних обсягів обчислень і складної апаратної підтримки для їхньої реалізації у реальному часі, крім того вони орієнтовані на компресію рухомих зображень у системах реєстрації і передачі інформації і погано пристосовані для систем аналізу видеозображень в реальному часі.

Більш повне врахування умов колового огляду, особливостей таких систем і вимог, що пред'являються до них, дозволяють сподіватись на створення більш спеціалізованих і значно більш ефективних засобів компресії і виділення динаміки у відеопослідовності.

В доповіді розглянуто один з варіантів підвищення інтелектуальності систем колового огляду. Передумовами для вирішення поставленої задачі є:

– скорочення надлишковості у відеопослідовності;
– виключення з розгляду і аналізу постійного фону.

ПОБУДОВА ПАНОРАМНОГО ЗОБРАЖЕННЯ

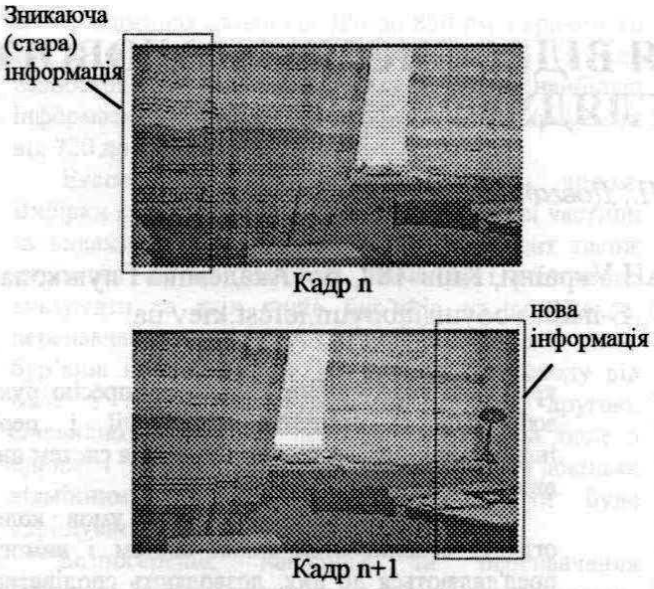
Аналіз сусідніх кадрів видеопослідовності з камери колового огляду показує, що наступний кадр видеопослідовності містить значну частину попереднього, зсунуту в сторону, протилежну напряму обертання камери. При цьому деяка частина зображення виходить з кадру і така ж частина нового зображення входить у кадр (див. мал. 1).

Якщо тепер з усієї відеопослідовності виділити тільки нові неповторювані частини зображень і "зклеїти" їх, то можна отримати широкоформатне панорамне статичне зображення одного повороту камери. Це панорамне зображення може бути представленим на екрані дісплею або у вигляді однієї широкоформатної смуги, або у вигляді 2-х, 4-х смуг по половині або чверті оберту відеокамери.

Основними проблемами отримання панорамного зображення є визначення границь зістиковки кадрів та згладження місця зістиковки для кращого візуального сприйняття.

При постійній швидкості обертання камери або при змінній швидкості із застосуванням датчика швидкості обертання зона нової частини зображення може бути визначена в залежності від швидкості обертання і від куту захвату видеокамери. Обсяг обчислювальної роботи для уточнення границі області не є великим.

© UkrObraz'2000



Мал. 1. Кадри відеопослідовності при обертанні камери праворуч. Прямокутником виділена інформація, що з'являється у кадрі n+1 у порівнянні з попереднім кадром.

Ступінь стиснення при постійній швидкості визначається як відношення об'єму інформації у кадрі відеопослідовності до об'єму інформації у новій частині зображення або як співвідношення ширини кадру B і ширини нової частини зображення b.

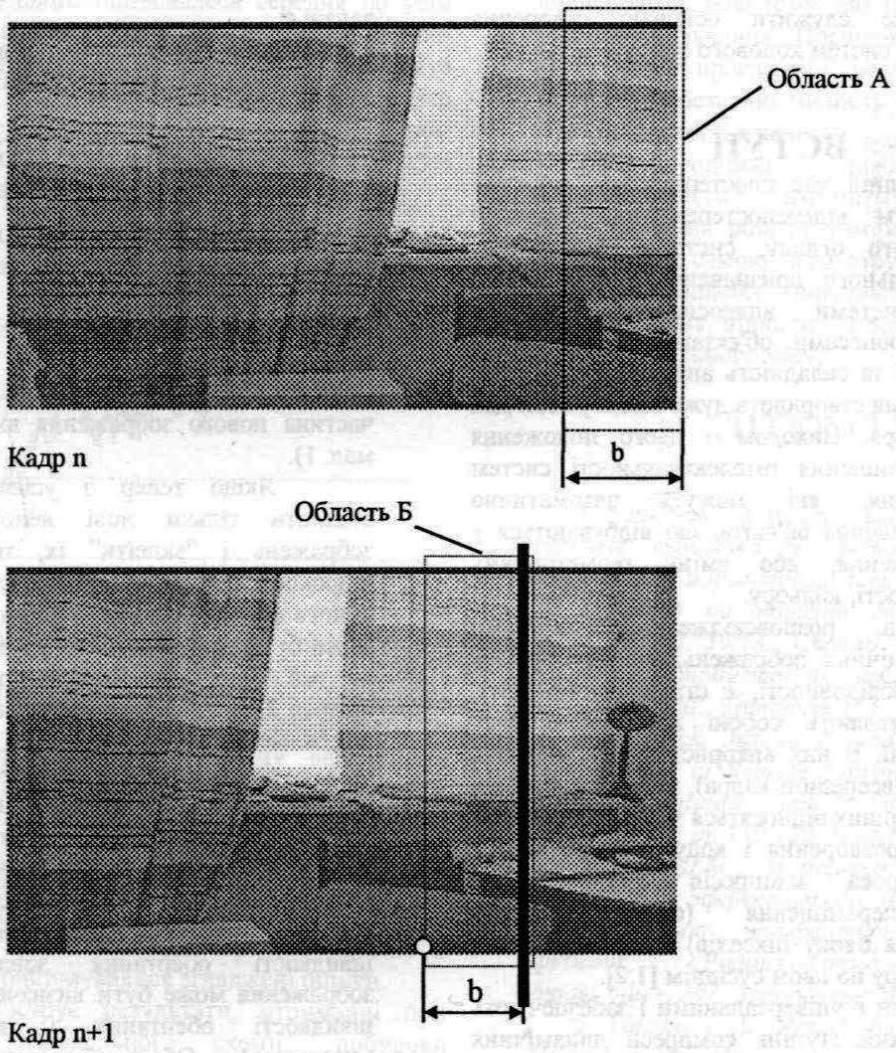
При змінній швидкості обертання видеокамери і відсутності датчика швидкості пошук границі попереднього кадру в наступному здійснюється шляхом сканування смуги краю зображення кадру n (область А) шириною b у зоні пошуку B-b у кадрі n+1, починаючи з лівої сторони кадру.

Ступінь відповідності обраної смуги з кадру n в кадрі n+1 на кожному кроці сканування визначається, наприклад, по мінімуму функції [3.4]

$$F = \sum_{(i,j \in R)} |A(i,j) - B(i,j)|,$$

де A (i, j) - значення яскравості елементів смуги краю зображення кадру n;

B (i, j) - значення яскравості елементів смуги, що обирається у зоні пошуку кадру (n+1).



Мал. 2. Пошук границі появи нової інформації у кадрі в порівнянні з попереднім. Жирною лінією показана границя входження нової частини відеозображення.

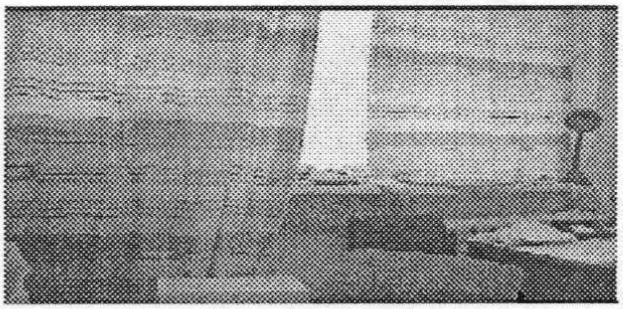
Мінімум функції F буде спостерігатися при визначенні ступеня подібності областей А та Б (див. мал. 2).

Складність процедури пошуку краю зображення при нерівномірній швидкості обертання може бути визначена слідуючим чином.

Для однієї операції порівняння зі смугою b·h пікселів необхідно виконати ≈2bh операцій типу додавання. Якщо пошук здійснюється по зоні сканування шириною (B-b), то загальна кількість операцій визначиться як 2(B-b)·bh.

Якщо визначити ширину контрольованої смуги краю зображення у відносних одиницях від ширини кадру як b=k·B, де k≤1, тоді кількість операцій для реалізації процедури пошуку визначиться як $2kB^2bh(1-k)$.

Для поліпшення зорового сприймання панорамного зображення доцільно виконувати зістиковку нової частини зображення зі згладжуванням границі зістиковки [5] (зображення повинні плавно переходити друг в друга) – див. мал. 3, 4 а, б.



Мал. 3. Панорамне зображення, що є сумою кадрів, зображених на мал. 2.

Описаний спосіб панорамування відеопослідовності може бути використаний при проведенні зйомки з об'єкту, що рухається прямолінійно (наприклад, з поїзду, автомобіля). В результаті буде одержана широка панорама і значна компресія. Цей підхід може бути ефективно використаний при побудові систем віртуальної реальності [6].

Середній ступінь компресії відеопослідовності за рахунок панорамування сягає 10.

Представлення відеопослідовності у вигляді панорамного зображення значно полегшує умови контролю оператором зони спостереження.

Отримана послідовність широкоформатних відеозображень може додатково підлягати стисканню будь-яким алгоритмом компресії динамічних відеозображень [7], в тому числі і MPEG. При цьому забезпечується значно вищий ступінь компресії порівняно з компресією вихідної послідовності кадрів (ступінь компресії залежить від швидкості колового руху відеокамери, частоти кадрів зйомки, а також від динаміки зміни оточення).

Вхідне динамічне зображення може бути повністю відновлене з панорамного шляхом послідовного вирізування з панорамного зображення окремих областей. Для цього при складанні панорамного зображення слід зберігати ширину області, що додається до панорамного зображення від кадру до кадру (це необхідно тільки в випадку змінної кутової швидкості відеокамери). Динамічна послідовність відеокадрів може бути відновлена з панорамного зображення у темпі зйомки, або в довільному (уповільненому або прискореному) темпі.

ПОБУДОВА РІЗНИЦЕВОГО ЗОБРАЖЕННЯ

Наступним етапом зменшення надлишковості інформації, що представляється оператору, може бути вилучення фону з зображення і виділення нової інформації між обертами відеокамери. Це може бути реалізоване шляхом виділення різницевого зображення між панорамними зображеннями двох сусідніх обертів відеокамери.

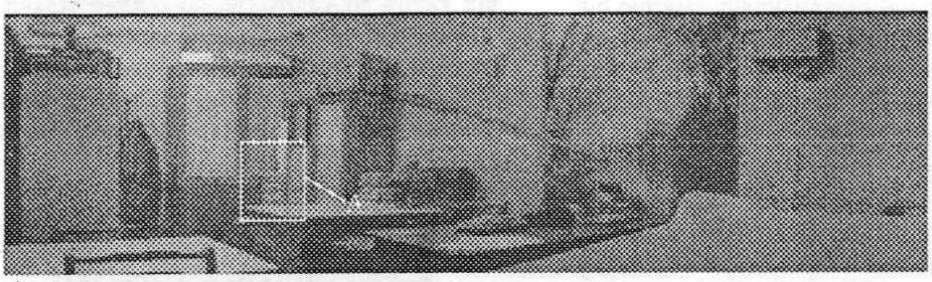
На мал. 4 а, б показана панорама огляду відеокамери відповідно для першого та другого оберту камери. Як видно, зображення фону не підлягає змінам, розбіжність між цими двома зображеннями полягає в переміщенні предмету з положення А у положення Б за час оберту камери. На підставі цих двох зображень побудоване третє, кожний піксел якого є різницею значень яскравості пікселів цих двох зображень (див. мал. 4 в). Якщо обмежити це зображення по порогу, та можна автоматично визначити не тільки наявність змін в обстановці, але і локалізувати місце, де відбуваються зміни. У результаті автоматичного визначення отримаємо дві області підвищеної уваги (на мал. 4 в вони виділені прямокутниками). Таким чином можливо робити автоматичну реєстрацію і сигналізацію наявності переміщення або локальної зміни яскравості (наприклад, при загорянні) в полі зору відеокамери колового огляду.

Похибки побудови панорамних зображень з ряду смуг, що складаються, границя між якими не завжди може бути визначена з точністю до піксела, а також наявність похиб ки від механіки системи колового огляду, призводять до появи деякого "шуму контурів" на різницевому зображенні. Однак такий шум у більшості випадків усувається високочастотною фільтрацією [5,8]. В той же час, зниклі, рухомі предмети або предмети що змінили геометричну форму, розміри, яскравість або колір будуть виділені практично повністю (якщо вони відрізняються від фону).

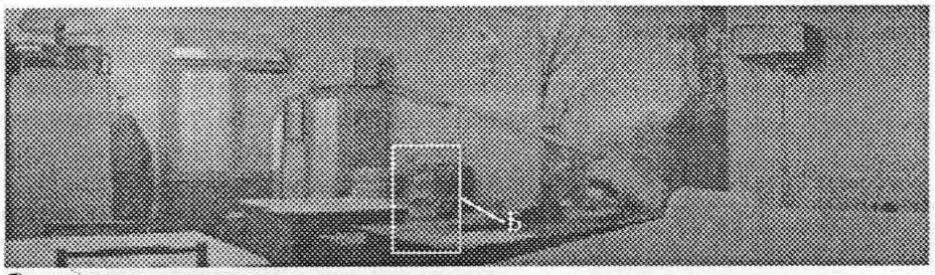
Використання різницевого панорамного зображення значно полегшує роботу оператора і скорочує обсяг інформації, що реєструється. В залежності від частоти і масштабності змін, що відбуваються ступінь компресії може досягати сотен

тисяч і більше разів. Якщо виконувалась постійна реєстрація (без фільтрації) різницевих зображень, то панорамні зображення, а при необхідності, і вхідна відеопослідовність можуть бути відновлені у реальному або довільному темпі.

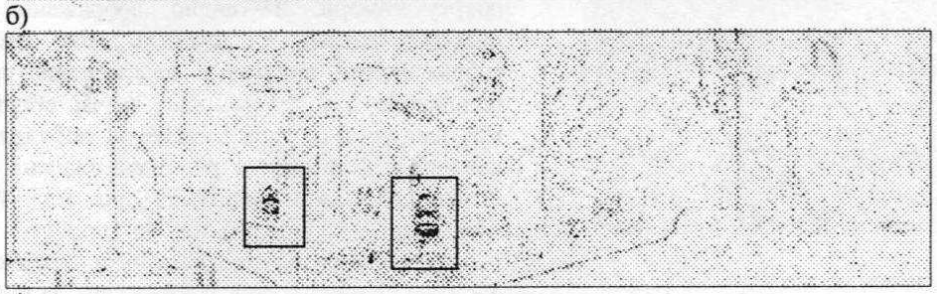
Крім того, отримане різницеве зображення створює передумови для автоматичного викриття змін, що відбуваються, тобто з'являється реальна можливість створення інтелектуальних систем колового огляду, що видають сигнал тривоги і реєструють виниклі зміни, або звертають увагу оператора на зміни, що мали місце.



а)



б)



в)

Мал. 4. Приклад отримання різницевого панорамного зображення, побудованого з 160 кадрів відеопослідовності.
а) панорамне зображення, зняте на 1-ому оберті відеокамери
б) те ж на 2-ому оберті відеокамери
в) різниця зображень а) та б). Значення різниці посилені у 4 рази, зображення інвертовано.

ВИСНОВКИ

Запропонований спосіб панорамування відеопослідовності з камери колового огляду і отримання різницевого зображення, що виділяє зміни між двома обертами відеокамери, значно полегшує роботу оператора і скорочує обсяг інформації, що реєструється.

Крім того, забезпечується можливість створення інтелектуальних систем колового огляду, які автоматично виявляють зміни в коловій панорамі.

ЛІТЕРАТУРА

1. Зелов С. Цифровое кодирование изображений. Компьютер – Пресс №3,4 1997.
2. Сидоренко Ю. "MPEG"–Путь в цифровое завтра. Компьютерное обозрение. № 17(90) 1997.
3. В.В. Старовойтов. Локальные геометрические методы цифровой обработки и анализа изображений. – Минск: Институт технической кибернетики НАН Беларуси, 1997.-284 с.
4. Липанов А. В. , Путятин Е. П. Исследование алгоритмов обнаружения объектов на основе методов корреляционного распознавания и алгоритма параллельной нормализации. – Харьков: Радиоэлектроника и информатика №3, 1998.
5. Прэтт У. Цифровая обработка изображений. – М.: Мир. 1982.Т.2.
6. Хирозе М. Создание виртуальных миров на основе реальных изображений. Компьютеры+Программы №9(42), 1997.
7. М. Г. Георгиев, И.Г. Таранцев. Алгоритм сжатия последовательности изображений в системах визуального контроля редко меняющейся обстановки. – Новосибирск: Автометрия №2, 1999.
8. Абламейко С. В., Лагуновский Д. М. Обработка изображений: технология, методы, применение. – Минск: Институт технической кибернетики НАН Беларуси, 1999.-300 с.

ПОМІХОСТІЙКІ ШТРИХОВІ КОДИ

М. Ф. Козак, В. В. Ніколенко
УжДУ, каф.кібернетики та прикладної математики,Ужгород, вул.Підгірна,46
тел.(03122)43325, jvtech@mail.uzhgorod.ua

ВСТУП

Штриховому кодуванню інформації, що надзвичайно широко використовується в країнах заходу, в нашій країні приділяється дуже мало уваги. Такий стан речей є незаслуженим – штрихові коди покликані полегшити та прискорити рутинну роботу людини в найрізноманітніших галузях людської діяльності. Широка сфера застосування штрихових кодів спричинила велику кількість найрізноманітніших типів штрихкодів. Нижче буде описано процес створення нового поміхостійкого типу штрихового коду Hem-8. Потреба в створенні цього коду викликана комп'ютеризацією контролю за переміщенням персоналу на підприємстві. Для цього кожного співробітника потрібно забезпечити пластиковою карткою-перепусткою, на якій окрім фото та звичайної інформації про робітника, повинен був бути штриховий код для автоматичного зчитування інформації про робітника сканерами, що підключені до комп'ютера. Сформульовано вимоги до штрихового коду: простота та економічність у виготовленні, можливість кодувати до 8 чисел, і висока стійкість до помилок. Остання вимога найважливіша – при зчитуванні сканером штрихкоду, часто з'являються помилки. Причинами цього є ряд факторів: простота скануючих пристроїв, рух картки через сканер із прискоренням (робітник сам повинен протягувати її через сканер, а швидкість руху руки може бути нерівномірною), пошкодження штрихового коду у виді забруднення, подряпин, підтертостей. Помилки, що виникають внаслідок цього, є причинами того, що штриховий код не вдається розкодувати, або, що ще гірше, він розкодовується неправильно, вони не дають змоги використовувати штрихові коди вже існуючих типів і виникає потреба в створенні нового поміхостійкого типу штрихкоду, основною особливістю якого є висока стійкість до помилок.

В ході роботи по створенню нового типу штрихового коду було вивчено та порівняно основні типи існуючих штрихкодів [3,4,5,6], на їх основі описано основні методи побудови нових типів штрихкодів, отримано математичний апарат, що дає змогу оцінювати стійкість коду до помилок, його надлишковість [1], та на основі цього апарату створено програмний пакет для дослідження стійкості до помилок різних типів штрихових кодів. Результатом цієї роботи став штриховий код Hem-8 основною відмінністю якого, від вже існуючих типів штрихових кодів, є його висока стійкість до помилок, яка була досягнута використанням в штриховому кодуванні методів корекції помилок, що базуються на алгоритмі Хеммінга.

1. Способи побудови штрихових кодів

У штриховому кодуванні існує два способи графічного представлення бінарних значень '0' та '1'. Перший спосіб: значення '0' та '1' кодуються відповідно двома кольорами - білим та чорним. Наприклад: бітова послідовність 10110011101100011 буде мати наступне штрихове представлення (мал. 1). В цьому способі штрихи, що відповідають '0' та '1' мають однакову ширину. В разі якщо в бінарній послідовності йдуть одне за одним кілька однакових n значень '0' чи '1' їм буде відповідати білий чи чорний штрих n-кратної ширини.



Мал.1

Другий спосіб представлення бінарних значень '0' та '1' у штриховому коді є спосіб, коли біти задані не різними кольорами, а різними значеннями ширини штрихів. Тобто є чотири атомарні графічні символи: два вузькі штрихи та два широкі, білого та чорного кольору. В такому штриховому коді білі та чорні штрихи чергуються, а значенням '0' та '1' відповідають відповідно широкі та вузькі штрихи. В цьому разі наведена вище бінарна послідовність буде мати вигляд (мал.2).



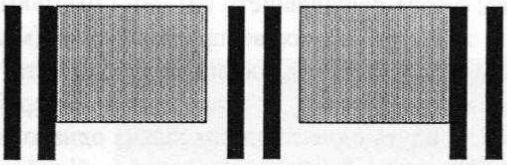
Мал. 2

Кожен із цих варіантів має як переваги, так і недоліки. В першому варіанті штриховий код буде коротшим, в наслідок того, що всі біти кодуються однаковими по ширині штрихами. По цій же причині в першому варіанті штриховий код бінарної послідовності зі сталим числом бітів буде мати сталий розмір, в той час як у другому варіанті розмір штрихового коду буде залежати від співвідношення нулів та одиниць. По цій причині перший варіант використовується якщо розмір інформації, що кодується, сталий, а другий варіант, коли кількість інформації, що кодується, змінна.

Розглянемо інші особливості побудови штрихових кодів, які також використовуються для

© UkrObraz'2000

класифікації штрихових кодів. Однією з них є наявність чи відсутність контрольних штрихів (бітів). В випадку відсутності контрольних штрихів, при нерівномірній швидкості зчитування штрихкоду, є ймовірність, що цей код буде інтерпретовано неправильно. Щоб цьому запобігти, на початку та в кінці коду розміщується набір із принаймні двох контрольних штрихів. Після зчитування сканером цього коду, обчислювальна машина може судити про зміну швидкості сканування штрихового коду і відповідно корегувати процес декодування. Прикладом застосування контрольних штрихів може бути штриховий код типу EAN-13. В ньому контрольні штрихи наявні не тільки на початку та в кінці, а і в середині коду (мал. 3).



Мал.3

Ще однією особливістю при побудові штрихового коду є наявність чи відсутність контрольної суми. Для гарантування правильності декодування штрихового коду в деяких типах штрихового коду до інформації, що кодується, додається деяка контрольна сума, яка функціонально залежить від кодованої інформації. Ця контрольна сума кодується в штриховий код разом з основною інформацією, а при декодуванні ЕОМ знову вираховує контрольну суму цього коду і порівнює із заданим. Зрозуміло, що неспівпадання цих двох контрольних сум свідчить про неправильно зісканований штрихкод. Поряд із контрольними числами, які кодуються разом з інформацією, нерідко в бітовій послідовності використовуються також біти контролю парності. Всі вищезазначені методи, покликані мінімізувати ймовірність неправильного інтерпретування штрихового коду, що сканується. В цьому штрихові коди чимось нагадують бітові послідовності сигналів, що передаються через цифрові канали зв'язку, тощо. Але штрихові коди мають одну дуже важливу відмінність, бітова послідовність, яка графічно представлена штрихкодом, не повинна містити велику кількість підряд розміщених однакових бітів, тому що графічне представлення такої послідовності, утвореної з одних тільки "0", перетвориться на тривалу темну смугу на штрихкоді, виміряти довжину якої для сканера буде складно. Тому, як правило, на штрихкоди накладається обмеження на кількість підряд розміщених однакових бітів, найчастіше це не більш ніж 3 одинакові елементи. Все вище описане приводить до збільшення надлишковості інформації у штриховому коді, але в той самий час збільшує надійність коду. Надлишковість інформації присутня практично в кожному типі штрихового коду в більшій чи меншій мірі, але у випадку штрихового кодування міру надлишковості можна умовно вважати мірою надійності (стійкості до помилок) штрихового коду.

2. Типи помилок, що виникають при штриховому кодуванні

В ході дослідження та створення штрихових кодів особливу увагу приділяють типам помилок, що виникають, та факторам, які стають причинами цих помилок. Найчастіше - це нерівномірна швидкість сканування штрихового коду, яка може бути наслідком багатьох причин: а) ручне сканування, дуже часто рука людини, що тримає сканер чи носій штрихкоду, рухається з прискоренням; б) траєкторія руху штрихкоду відносно сканера нелінійна; в) короткі зупинки штрихкоду під час сканування. Запобігти таким помилкам можна ще на стадії створення нового типу штрихового коду, ввівши в нього додатково контрольні біти-штрихи (мал.3), чи розмістивши поряд зі штрихкодом шкалу синхронізації сканера (мал. 4).


Мал.4

Нерівномірна швидкість сканування спричиняє помилкові ситуації, які досить важко виправити – порушуються пропорції в ширині штрихів і, як наслідок, в бінарному коді після сканування та декодування з'являються зайві біти, виявити котрі дуже складно.

Ще одним фактором, що спричиняє помилки при декодуванні штрихового коду, є забруднення та пошкодження штрихкоду. Такі проблеми виникають значно рідше ніж попередні, передусім, тому що з'являються, як правило, тільки при тривалому та частому використанні носія зі штрихкодом (посвідчення особи, бібліотечна справа, і т. п.). В таких випадках виникають три типи помилок: а) написи на штрихкоді, подряпини, перегини; б) плями; в) підтертості та іншого роду зменшення контрасту штрихового коду. В першому випадку пошкодження спричиняють заміну 1-2 бітів при декодування на протилежні. У другому випадку в бітовій послідовності з'являється тривала послідовність однакових бітів. В третьому випадку з'являється висока ймовірність неправильного декодування штрихового коду.

Усі ці три випадки формулюють специфічні задача по розпізнаванню:

1. Локалізація та видалення написів із штрихової поверхні (мал.5).
2. Оцінка розмірів, локалізація та видалення плям (у разі якщо пляма не покриває повністю штрихкод). (мал.6).

3. Очистка штрихкоду та збільшення його контрастності (мал.7).



Мал.5



Мал. 6



Мал. 7

3. Нет-8

Нет-8 штриховий код створений для використання на посвідчені особи в межах великого підприємства. Цей штрихкод дозволяє кодувати 8-ми значні порядкові номера. Головною особливістю цього типу штрихкоду є використання методу Хеммінга для виявлення та корекції помилок. Серед існуючих типів штрихових кодів використовуються найрізноманітніші методи контролю за помилками, але методи корекції помилок, або не використовуються, або слабо реалізовані. Нижче описується процес створення штрихового коду Нет8.

Вимоги до штрихкоду, що були сформульовані до його створення:

Вимога 1. Мінімальна кількість символів, яку повинен містити штрихкод – 6. Можливо збільшити це число до 7 або 8, ввівши резервне або контрольне число.

Вимога 2. Алфавіт нашого штрихового коду складається тільки з 10 цифр 0,1, ...,9. Тому на кодування одного символу мінімально потрібно 4 біти $2^4=16$ ($2^3<10$) (насправді ця кількість буде дещо більшою, за рахунок уведення контрольних бітів та накладення деяких додаткових умов. Однією з таких умов буде наступна).

Вимога 3. Для запобігання злиття двох сусідніх бітових рядів, необхідно щоб кожен із таких рядів розпочинався з '0' та закінчувався '1'-цею.

Вимога 4. Велика кількість підряд розміщених однакових символів збільшує ймовірність помилки. Тому накладається ще одна умова: в бітовому ряді не повинно бути підряд більш ніж 3 однакових бітів.

Вимога 5. Як зазначалося у вимозі 2, в наслідок вимог 3-4, на кодування одного символу потрібно 6 біт.

Штриховий код Нет-8 забезпечено властивістю самовиправлення помилок, яка базується на методі Хеммінга. Для цього необхідно з'ясувати, яку кількість бітів потрібно відвести під перевірочні. Виходячи з відомої формули $2^k>n+1$, де n - загальна кількість бітів, k - кількість перевірочних бітів, m=n-k - кількість інформаційних бітів, отримуємо наступну таблицю:

Табл. 1.

K	N	m=k-n	K	N	m=k-n
1	1	0	**5**	31	26
2	3	1	**6**	63	57
3	7	4	**7**	127	120
4	15	11	**8**	255	247

Мінімальне число бітів у бітовому ряді 6, якщо під інформаційні відвести 26 біт, то можна було б кодувати 26/6=4 символи. Така кількість є недостатньою, тому виділимо

Під інформаційні виділено 57 бітів (при 26 можна кодувати тільки 26/6=4 символи), таким чином можна кодувати до 57/6=9 символів, що є навіть надлишково для сформульованих вимог.

В таблиці 2 вказано всі можливі 6-ти значні бітові ряди, що задовольняють нашим умовам.

Табл. 2.

1	000101	5	001101	9	010111
2	000111	6	010001	10	011001
3	001001	7	010011	11	011011
4	001011	8	010101	12	011101

Цієї кількості бітових рядів достатньо, щоб кожному символові алфавіту штрихового коду було поставлено у відповідність по одному з них. Але розгляд матриці кодових відстаней цих бітових рядів, вказує на деякі недоліки набору. В разі виникнення однієї помилки в штриховому коді, вона буде виправлена за допомогою алгоритму Хеммінга. В разі виникнення двох помилок, метод Хеммінга перестає бути ефективним, в такій ситуації принаймні повинно бути виявлено наявність помилки. Перевірку на наявність подвійних помилок забезпечується вводом восьмого контрольного числа. Воно функціонально залежить від перших семи, і ця залежність вибирається так, щоб при зміні одного з перших семи чисел, восьме контрольне також обов'язково змінилось. Якщо при виникненні двох помилок, вони обидві знаходяться в одному бітовому ряді, то наявність такої помилки буде виявлено контрольним числом, але якщо ці помилки з'являться в різних бітових рядах, тобто будуть змінені два числа із семи, то можливі такі комбінації чисел, коли помилка не буде виявлена взагалі. Для запобігання таких помилок набір бітових рядів потрібно підібрати так, щоб матриця кодових відстаней цього набору не містила одиниць. Тоді при виникненні двох помилок у різних бітових рядах, ці новоутворені бітові ряди не будуть належати нашому наборові, що буде свідчити про помилку. В разі, якщо матриця кодових відстаней не містить одиниць,

то, для того щоб помилковий код не виявлявся, необхідно, щоб помилки виникли по дві в двох різних бітових рядах. Тобто вивід із матриці кодових відстаней одиниць, значно покращує стійкість коду до можливих помилок. Але для того щоб вивести з матриці кодових відстаней одиниці, потрібно скоротити набір бінарних рядів, але тоді цього набору буде замало щоб поставити один бітовий ряд у відповідність одному символу нашого алфавіту. З'являється необхідність у збільшення довжини бітових рядів з 6-ти до 7-ми бітів. При 7-ми бітових рядах можна кодувати 57/7=8 символів. Це задовольняє вимогам. В таблиці 3 вказано всі можливі бітові ряди, що задовольняють сформульованим умовам.

Табл. 3.

1	0001001	9	0011011	17	0110001
2	0001011	10	0011101	18	0110011
3	0001101	11	0100011	19	0110101
4	0010001	12	0100101	20	0110111
5	0010011	13	0100111	21	0111001
6	0010101	14	0101001	22	0111011
7	0010111	15	0101011		
8	0011001	16	0101101		

Матриця кодових відстаней новоутворених бітових рядів також містить одиниці. Після відбору з 22-х бітових рядів десять таких, щоб їх матриця кодових відстаней не містила одиниць отримаємо таблицю 4.

Табл. 4.

0	0001001	4	0011101	8	0110011
1	0010001	5	0100111	9	0110101
2	0010111	6	0101011		
3	0011011	7	0101101		

При використанні цих бітових рядів Hem-8 повинен мати наступну структуру:
0xxxxx10xxxxx10xxxxx10xxxxx10xxxxx10xxxxx10xx
xxx10xxxxx1nnnnnn , де n – контрольні біти.

Ця структура містить поки що один невеликий недолік - останні шість бітів можуть бути причиною деяких помилок. Наприклад: а) вище ми накладали обмеження на кількість підряд розміщених однакових бітів, на комбінацію nnnnnn в даному випадку таке обмеження не поширюється; б) в разі, якщо останній байт чи байти будуть рівні "1", то при графічному представленні цього коду вони будуть втрачені, оскільки їм буде відповідати білий колір і вони зіллються з фоном штрихкоду. Щоб запобігти цим небажаним явищам, вводяться додаткові нульові біти в кінець коду. Кінцева структура коду:
0xxxxx10xxxxx10xxxxx10xxxxx10xxxxx10xxxxx10xx
xxx10xxxxx101nn01nn01nn010

ВИСНОВКИ

Стійкість до помилок штрихкоду Hem-8 було перевірено за допомогою спеціально розробленої прикладної програми для ПК, яка генерувала помилки та намагалася їх виявити чи виправити. Результати отримані за допомогою комп'ютера, підтвердили наші теоретичні розрахунки: при появі одного помилкового біта він виправляється завжди, при появі двох таких бітів будь-де в коді в 100% випадків фіксується наявність помилки та в 59% випадків виправляється. Помилки, які не виявляються, можуть з'явитись тільки при появі 4 помилкових бітів. Усе це свідчить про високу стійкість створеного штрихового коду.

Для порівняння штрихкоду Hem-8 з існуючими типами підрахуємо його ентропію. Hem-8 складається з 71 біта, кількість комбінацій чисел, які можна утворити 10 мільйонів (10^7), оскільки восьме число залежить тільки від перших семи, ми його не враховуємо. Кількість інформації, що несе в собі штрихкод Hem-8 рівна I(a)=log(10^7). В [3] наведена формула для знаходження величини надлишковості коду:

$$r = 1 - \frac{I(a)}{n \cdot \log L} \qquad (1)$$

алфавіт двозначний L=2, кількість бітів n=71, тому

$$r = 1 - \frac{\log 10^7}{71 \cdot \log 2} = 0.67 \quad (2)$$

Коефіцієнт надлишковості штрихкоду EAN13:

$$r = 1 - \frac{\log 10^{12}}{95 \cdot \log 2} = 0.58 \quad (3)$$

Надлишковість в штриховому коді не є критичною, завдяки цій надлишковості коди забезпечуються більшою стійкістю до помилок, що при їх застосуванні більш важливо ніж надлишковість.

ЛІТЕРАТУРА

1. Шинон К. Работы по теории информации и кибернетике / Пер. с англ. под ред. Добрушина Р.Л.,Лупанова О.Б. -М.: Издательство иностранной литературы, 1963. 830 с.
2. Котов П.А. Повышение достоверности передачи цифровой информации. -М.: Связь, 1966. 184 с.
3. Боховский В. А. Кибернетические системі в медицине. –М.: Наука, 1973. 328 с.
4. Cracking as an art, [Barcodes][Instant Access] / Lesson C(1). www.annon.penet.fi
5. Handbuch zum barcode-lesegerat / ABP-Programierhandbuch fur Einbauscanner.
6. Barcodes FAQ. : www.azalea.com
7. Barcodes. : www.incom.ru/barcodes/s1.html

New Generation Video Surveillance System Based on Image Processing and FPGA Processor towards Autonomous System

A. Benkhalil, S. Ipson and W. Booth
Department of Electronic & Electrical Engineering
University of Bradford, Bradford, West Yorkshire, DB8 9EU, UK
akbenkha@bradford.ac.uk, *+44-1274-234075*

A stand-alone two-camera automatic security system is described. The system utilises a single Field Programmable Gate Array (FPGA) chip which performs a variety of video-rate arithmetic and logical operations on images, grabs and stores image frames and generates all necessary control signals. Results illustrating the highly successful operation of the system are included.

1. INTRODUCTION

Modern digital computing, semiconductor technologies and communication technologies have introduced a complete change of perspective in the design of the architectures of surveillance systems. Recently, several research groups have devoted a great deal of effort to devising prototype video surveillance systems based on robust methods of dynamic scene interpretation. Most of this work relies on complex and expensive hardware such as multiple DSP boards or on high-speed host computer platforms in order to achieve near real-time performance[1], [2], [3].

The research reported in this paper was directed at developing an image processing based surveillance system that does not need operator intervention, or a host computer or expensive hardware. The major aim of the research was to develop a stand-alone automatic security camera system (ASCS) capable of grabbing Comité Consultatif International des Radiocommunications (CCIR) format images, detecting motion and providing high resolution images of moving objects. A further aim was to take advantage of the high logic density of recently introduced FPGAs to implement motion detection and tracking algorithms and control functions in a single chip [4].

The paper is organized as follows. In section 2 an overview of the system is presented. In section 3 algorithms developed to detect motion, to update reference images and to carry out adaptive thresholding are described. Section 4 deals with the functionality of the FPGA including the implementation of the algorithm used for target classification. Section 5 presents some results to illustrate the performance of the system under different operating conditions and finally section 6 draws conclusions and indicates possible future developments.

2. SYSTEM OVERVIEW

A block diagram of our prototype ASCS, which employs two cameras working in a master slave combination, is shown in Fig. 1. The master camera points in a fixed direction and incorporates a 16mm lens to provide a field of view 22° wide in the horizontal direction. It provides greyscale images that are analyzed by the FPGA device every 40ms in order to detect target motion. The slave camera, which is mounted on a robotic arm for mobility, has a 150mm focal length lens to provide a high-resolution view of the target. It is moved to point at a target by the control logic hardware on the basis of information provided by the master camera. In the system currently implemented, the slave camera is driven to view one of the 100 sub-regions into which the field of view of the master camera is divided. The direction chosen is the one that is closest to the direction of the detected target.

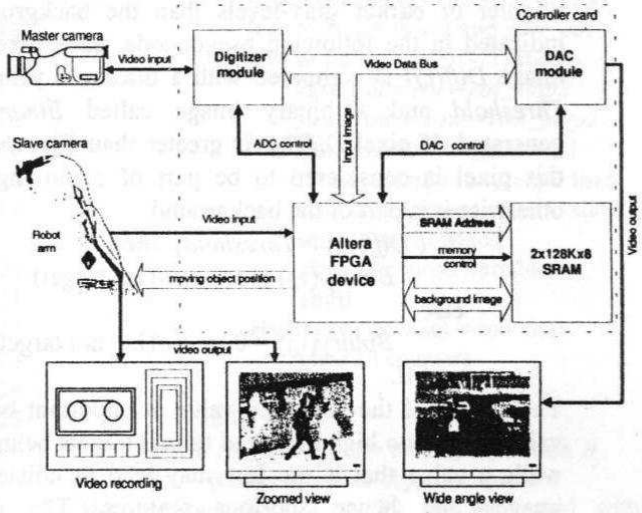


Fig. 1. Block diagram of the ASCS

The FPGA device shown in Fig. 1 is responsible for controlling the digitization and display of each incoming live-video image and the storage of a continually updated background image in the external SRAM. It also performs a variety of video-rate arithmetic and logic operations associated with moving target detection prior to transferring the co-ordinates of a detected target to the slave camera via the robot arm. A detailed description of the functionality of the FPGA is presented later in section 4. The digitizer module shown in Fig. 1 was implemented using a conventional flash Analogue-to-

© UkrObraz'2000

Digital Converter (ADC) for digitization, a sync separator IC for timing control and a PLL (Phase Locked Loop) device to produce the system's 10MHz global clock from the 15.625 kHz CCIR line frequency signal. The Digital-to-Analogue Converter (DAC) module shown in Fig.1 uses an off-the-shelf DAC integrated circuit that automatically generates all the pedestal signals necessary to produce a standard video signal (CCIR). The system is designed to grab a stream of gray scale images with a resolution of 512(H) × 512(V) × 8 bits.

3. MOTION DETECTION

Since the system was intended to operate at video rates, a simple but efficient motion detection algorithm was chosen for implementation in the FPGA:

$$Diff(i,j) = \left| Live_im(i,j) - Bg(i,j) \right| \quad (1)$$

where $Live_im\{(i,j) \in R\}$ is the current image, $Bg\{(i,j) \in R\}$ is the background image and R={(i,j): i=1...M (512), j=1...N (512)}. This technique of subtracting the background image *Bg(i,j)* from the input image *I(i,j)* to produce the difference image *Diff(i,j)*. is called background differencing and it still forms the cornerstone of many of today's more complex systems [1], [2], [5]. The background image should, ideally, not contain moving objects and the system can use either a stored single live image or, more usefully, a continually updated stored image that is computed using a temporal filter as described at the end of this section. The absolute value of *Diff(i,j)* is used in order to detect moving objects which may have brighter or darker gray-levels than the background. As indicated in the following pseudocode, each pixel of the image *Diff(i,j)* is compared with a threshold value called *Threshold* and a binary image called *Binary(i,j)* is generated. If pixel *Diff(i,j)* is greater than *Threshold* then this pixel is considered to be part of a moving object, otherwise it is part of the background.

```
if ( Diff(i,j) > Threshold) then
        Binary(i,j) = 1        //(i.e. target)
else
        Binary(i,j) = 0        //(i.e. not target)
end if.
```

The choice of the threshold value is important because a value that is too high may lead to real targets being missed while a value that is too low may lead to noisier binary images and hence spurious features. The optimum threshold value also varies with the instantaneous illumination level and a suitable compromise has been achieved by continuously adapting the threshold value

$$Threshold = f \times \sum_{k=1}^{m \times n} \frac{Live_im\ (k)}{(m \times n)} \quad (2)$$

Here, $m \times n$ is the total number of pixels and f is a multiplying factor whose value can be set by the user. A value of f = 1 was found to be appropriate for the surveillance scenes used for testing so the threshold values actually used in these cases were the average gray levels of the live images. Fig. 2 demonstrates the detection of moving features in a sequence of images captured using the master camera.

The background image is continually updated using the temporal low-pass filter [5] defined by:

$$Bg_{k+1}(i,j) = Bg_k(i,j) + G \times \left(Live_im_k(i,j) - Bg_k(i,j) \right) \quad (3)$$

Here k labels the current frame and G is a gain factor defined by the user. This feature causes the system to adjust the background automatically to take account of changes in the scene caused, for example, by changes in the illumination or by an object moving into the scene but then stopping. In either case the change of scene will become part of the updated background within a time-scale related to the value of G. Image sequence simulations carried out using MATLAB indicated that G = 0.125 is an appropriate value to choose. Then the background image takes about 20 frame periods to adjust to a step change in the scene, which enables new static features to become part of the background in a reasonably short time whilst ensuring that no moving objects are added to the background.

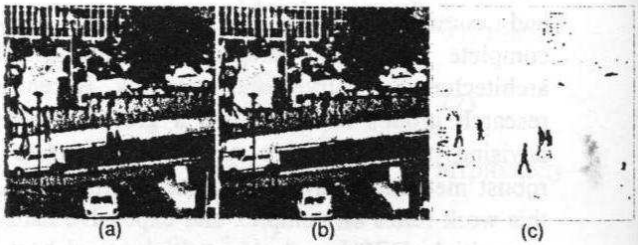


Fig. 2. The detection of moving objects. Here (a) is one of the images in a sequence of images, (b) is the current background image and (c) is the binary difference image after applying the current threshold.

4. FUNCTIONALITY OF THE FPGA

One FPGA chip forms the heart of the single board surveillance system. The functionality of this logic device is shown in Fig.3 by means of a schematic diagram. It can be seen that this chip performs a number of tasks, implemented as individual modules, in order to generate control signals for the slave camera from an analysis of incoming video images. The diagram also shows the data, address and control signal pathways connecting the different modules. Two of the FPGA's modules, frame-grabber control and memory control, provide all the control signals required by the external components in order to synchronize the image processing hardware to the CCIR composite video signal. These modules are implemented using 56 and 29 logic cells respectively, where 1 logic cell is equivalent to about 20 gates. The robot arm control module generates the signals required to move the slave camera and is implemented using only about 10 logic cells. Three image-processing modules are included in the FPGA. The first two of these, background image update and threshold selection, implement the algorithms described in the previous section within 17 and 68 logic cells respectively. Both algorithms require either a multiply or

divide operation (implemented by bit shifting) but the latter also requires an accumulator (implemented as a 20 bit adder) which accounts for the difference in the amount of resources used. The third image-processing module, target detection, is the most complex of all the internal modules and its implementation requires 286 logic cells. It decides whether a connected group of changing pixels is a target of interest or not. Because of its major influence on the workings of the system, the remainder of this section concentrates on describing the algorithm used to achieve the desired functionality.

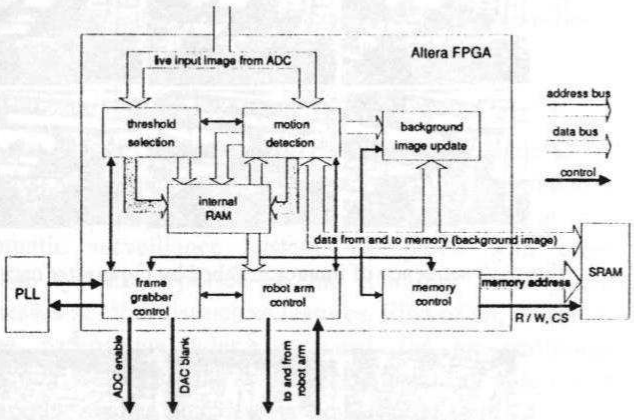


Fig. 3. Schematic diagram of the FPGA part of the overall system design.

The target detection module receives the live and background image data at video rates, pixel by pixel on a row by row basis, and applies differencing and thresholding as described earlier. The result of this process is illustrated in Fig. 4, which shows a small region cropped from Fig. 2(c). A potential target group of pixels, whose rate of change with time is faster than that set by the time constant of the temporal filter, is shown in black in Fig. 4 while background pixels are indicated in white. Moving along each row, a 4-bit counter *col_count* counts the number of successive pixels whose gray levels exceed *Threshold*. If the length of this run of black pixels is less than a user defined width limit *col_udtp1* (appropriate to people) the counter is cleared. If however, this run length is greater than or equal to *col_udtp1* another 4-bit counter, *row_count* checks the height of the identified object and compares it with a height limit *row_udtp1*. The parameter *row_count* counts the number of successive rows in which runs of black pixels, starting at about the same column position, exceed the width limit. This situation is illustrated by position *b* in Fig. 4 for the case where *col_udtp1* is chosen to be four. If *row_count* equals the height limit, then the co-ordinates of the target are saved in internal memory and also sent to the slave camera controller. This situation is illustrated by position *c* in Fig. 4 for the case where *row_udtp* is chosen to be 6.

The logic implemented in the target detection module is defined more precisely by the following pseudo-code. As indicated in this code, a section of the logic is repeated, the first time to detect large objects such as cars and the second time to detect smaller objects such as people. The temporary variables and saved data indicated in the pseudocode are stored in the internal RAM of the FPGA.

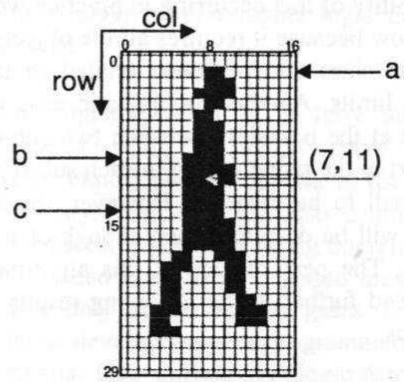


Fig. 4. A small region of Fig. 2(c) showing a detected object.

```
//within each of the 100 sub-regions do the following
1   Clear all counter variables to zero
2   for row =1 to end of current sub-region
3   for col =1 to end of current sub-region
4       if pixel at location [col,row] > Threshold then
5           col_count= col_count + 1
6       else
7           if col_count > or = to col_udtp2 then // (this
              could be a large object i.e. car)
8             col_count = 0
9             if row_count zero then // (first row of object)
10                save_col = col – col_udtp2
11                save_row = row – row_udtp2
12                row_count = row_count + 1
13            else if row_count not zero then
                  // (subsequent rows of same object)
14                if save_row = row-1 and save_col =
                  (col or col + 1 or col – 1) then
15                  save_col = col – col_udtp2
16                  save_row = row – row_udtp2
17                  row_count = row_count + 1
18                  if row_count = row_udtp2 then
19                  send target co-ordinates to slave
                  camera controller and
20                  clear all counter variables
21                  endif
22              else if save_row not = row then
23                  clear all counters
24              endif
25          endif
26      else if col_count > or = to col_udtp1 then //
            (this is a smaller object i.e. person)
    The previous lines 8 to 25 are duplicated with
    col_udtp2 changed to col_udtp1, row_udtp2 changed
    to row_udtp1.
45      end if
46    endif
47  end for
48  end for
```

This relatively simple algorithm detects only one potential target object within each sub-region and ignores any other potential target to the right of the first. This is of no concern in the current system because, as long as one target is detected, a view of that sub-region will be captured showing all targets that are present. However, it is conceivable that the presence of a number of objects,

smaller than the size threshold, to the left of a real target could prevent the detection of the real target. The probability of this occurring in practice would be expected to be low because it requires all the objects present to have specific sizes and positions related to target width and height limits. Another problematic case is when a target occurs at the boundary between two sub-regions, because the part of the target falling in each sub-region may then be too small to be detected. However, the target is moving and it will be detected when enough of it is within a sub-region. The performance of this algorithm in practice is discussed further in the following results and conclusions sections.

5. RESULTS

All the functions described in previous sections have been encapsulated within a single Altera EPF10K10LC84 device using the design and simulation tools provided by the Altera MAX+PLUS II development system. Two Altera library megafunctions, namely a fast 16-bit adder/subtractor and an allocator for the FPGA's internal RAM (EABs) were each used more than once in the design to shorten the design cycle time. It was found appropriate to create each of the six sub-designs shown in Fig. 4 as text files using Altera Hardware Description Language (AHDL). The complete design consists of about 1600 lines of AHDL source code. It uses most of the logic capacity and about 70% of the memory capacity of the specified device.

The complete system was tested by mounting the two cameras side by side, about 11m above ground level, and setting them to view a thoroughfare in our university campus that is used by both pedestrians and road vehicles. In order to illustrate the system performance, two synchronized frame grabbers were used to capture simultaneously image sequences from the master and slave cameras. Fig. 5 shows a typical short sequence of 12 images captured from the master camera over a 6s period on a rainy day. In this sequence one person (indicated by an enclosing black rectangle for ease of identification) can be seen walking down the sidewalk.

Fig. 6 shows higher resolution views of the moving person captured by the slave camera during the same period as it follows him. These images show the direction of the slave camera remaining fixed while the target moves through one of the 100 sub-regions, e.g. B1, B2 and B3, then changing to the next sub-region, e.g. C1, as the target progresses further. Because the slave field of view contains 512 by 512 pixels, there is sufficient resolution to recognize the person if the images are appropriately magnified. In this particular sequence, only one moving person is present and the slave camera accurately tracks the movement. When multiple moving targets are present, however, the system no longer tracks an individual target. It is designed to cause the slave camera to view them in order of their positions starting from the top left and moving from left to right and top to bottom through the list of target coordinates stored in the FPGA's internal memory. The list of targets is updated every frame period.

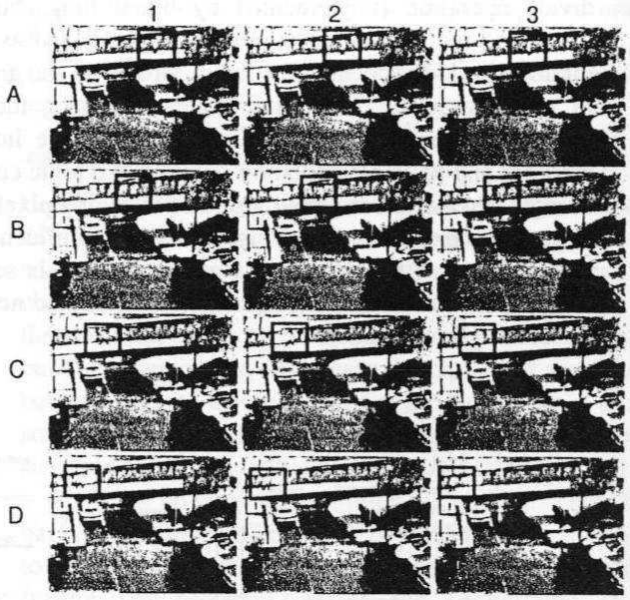


Fig. 5. A sequence of images grabbed by the master camera.

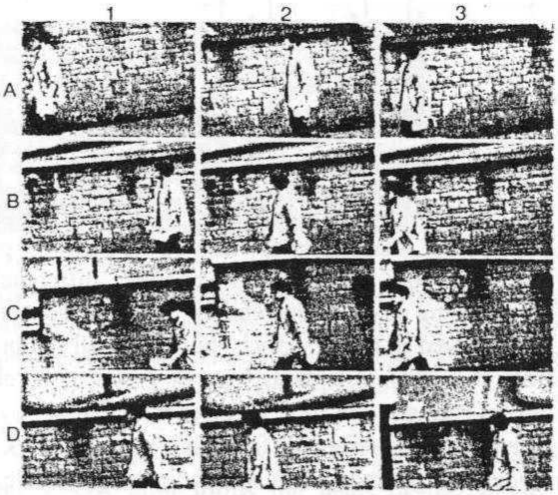


Fig. 6. A sequence of images grabbed by the slave camera.

Results on the detection and tracking of people recorded in one hour under different environmental conditions are shown in Fig. 7. The results demonstrate the ability of the system to work well in different situations. In bad atmospheric conditions, such as strong winds and heavy rain, the system generates a few false alarms but this is considered to be preferable to missing genuine targets. These false alarms arise from the movements of objects similar in size to people such as the branches of trees. On sunny days only a very small percentage of target people are missed, due entirely to their clothing being indistinguishable in gray level from the background. The possibility alluded to in the previous section of targets being missed due to the presence of other smaller objects was not observed in the trials and is much less important than the lack of contrast between some targets and the background.

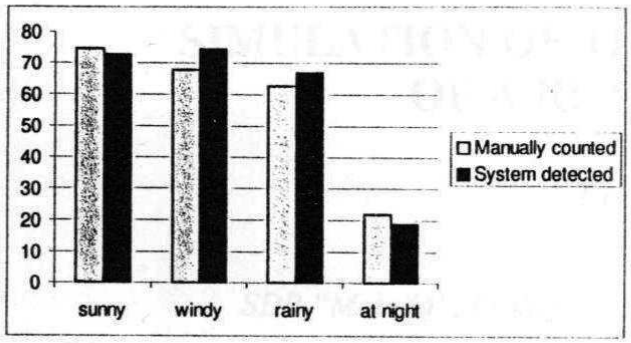


Fig. 7. The number of people detected over a one hour period under different environmental conditions.

6. CONCLUSIONS

All the major objectives of the research, namely the creation of a low cost stand-alone system incorporating a single FPGA controller, have been achieved in our automatic surveillance system. Compared with other similar systems reported in the literature, our system possesses several distinctive features. First of all, its stand-alone nature precludes the need for an additional computer. Secondly, the system operates at full video rates and grabs images and detects /tracks objects at 25 or 30 frames/s depending on the video standard used. Most of the alternative systems are significantly slower than this. Thirdly, the components on the board cost only about $100 which is much less than all the other systems. Fourthly, unlike the other systems, it is specifically designed to provide several high-resolution views of all targets of interest within the scene. Like all of the systems mentioned in the introduction, our system is capable of working effectively under varying illumination and weather conditions.

Improvements to the current system could be made on two fronts. Firstly, the relatively slow speed action of the pan and tilt movement, which is currently provided by a robot arm, must be improved to take full advantage of the very fast detection and tracking performance of our system. This would involve developing a lighter and more accurate opto-mechanical means of changing the direction of the optical axis of the slave camera. By amending the control logic, the slave camera could then be directed to point at the center of each target of interest rather than at a sub-region. This would permit even higher optical magnification to be employed by the slave camera. The second improvement relates to image processing functionality. Although the current system works, it has less sophistication in terms of image segmentation and target classification than computer-based systems. Adding to the existing design an algorithm that fully segments the whole binary image in each frame period would enable all moving objects in the field of view to be located and improved target classification to be added. This would require extra internal memory to store one or two lines of run-length encoded live video data, temporary variables, arrays and the bounds of each segment comprising a group of connected pixels. Additional logic resources would also be necessary. These modifications could be implemented in a larger programmable logic device at very little extra cost.

The successful implementation of a fully automatic surveillance system using a master slave camera combination is an important development in the area of vision-based security. The work has also demonstrated that a number of useful image-processing operations can be performed at video rates using a single inexpensive FPGA containing only 30000 useable gates. It is very clear that the latest developments in programmable logic devices, in particular their million gate logic density and their ability to perform several operations in parallel, afford new opportunities for researching and prototyping intelligent autonomous surveillance systems.

7. REFERENCES

[1] G. L. Foresti, "A real-time system for video surveillance of unattended outdoor environments," *IEEE Trans. Circuits Sys. Video Technol.*, vol.8, pp. 697-704, Oct. 1998.

[2] A. Iketani, A. Nagai, Y. Kuno, and Y. Shira, "Detecting persons on changing background," *IEEE 14th Int. Conf. on Pattern Recognition*, Queensland, Australia, 1998, pp. 74-76.

[3] B. A. Boghossian, and S. A. Velastin, "Real-time motion detection of crowds in video signals," *Inst. Elec. Eng. Colloquium High Performance Architectures for Real-time Image Processing*, IEE, London, 1998, pp. 12/1-12/6.

[4] D.W Trainor, D.J. Ridge, and R. Hamill, "Generic megafunction for high performance processing on FPGAs/PLDs," *IEEE Asilomar Conf. Signals, Systems & Computers*, Pacific Grove, CA, 1997, pp. 1047-1051.

[5] A. Makarov, "Comparison of background extraction based intrusion detection algorithms," *Proc. IEEE Int. Conf. Image Processing*, Lausanne, Switzerland, 1996,vol. 1, pp. 521-524.

SIMULATION OF THE RADIATION CREEP OF A REACTOR FUEL

Tarasov V.A.

SDB "Molnija", Prohorivskaja St.45, 65005, Odessa, Ukraine

Phone: (048) 7323164 Fax: (048) 7324906

E–Mail: kpmo@farlep.net

Are submitted the computer system for radiation creep simulation, developed on the basis of the offered model of a radiation creep and the method of dynamic simulation of a dislocation movement, and results of computer experiment on radiation creep research on an example of technical uranium and its alloys with the small additions of molybdenum (0,9 % ÷ 1,3 %).

1. INTRODUCTION

The special interest represents study of a radiation creep of materials. The majority of experimental data on radiation creep concerns to already used in working reactors materials, which were tested in conditions close to working modes of reactors. The experimental researches of radiation damage of materials and change of their properties are connected to the large expenses of work, means and time, and also frequently with significant difficulties of technical and technological character. The large variety of theoretical models of radiation creep is united by one common parcel - presence in an irradiated material of high concentration of point defects (PD) and sinks for them. The account of this parcel with attraction of dislocation representations allows to explain features of radiation creep of materials under a loadings by result of gliding and climbing of dislocations in a field of the distributed obstacles. The theoretical models, constructed on the basis of these representations and using the analytical description, allow to receive analytical expressions for speed of radiation creep, allowing qualitatively to judge character of influence of some factors on radiation creep rate. However quantitative revealing of features of radiation creep on the basis of model of dislocation movement in a field of the distributed obstacles is possible only with attraction of methods of mathematical simulation on a computer. Therefore special value is represented by development of computer systems for radiation creep simulation, allowing to receive estimations of radiation creep speed in significant ranges of change of such parameters, as temperature, loading and density of a radiation flow.

The computer simulation of nuclear structures of defects in crystals and their reorganizations under influence of loadings, thermal fluctuations and irradiation became the important method of research. The authors [1] consider, that the time comes nearer, when the statement anyone (especially expensive) experiment in radiation physics will be preceded by its computer simulation. Many models of computer simulation of radiation defects allowing are nowadays known to find in this or that approximation types of defective structures. Clearly, however, that the finding of defective structures, that is finding of distributions in space of a various sort of defects of a crystal lattice, is only first step of research. Its following step should consist in study of influence of all set of defects on change of those or other physical properties of metals at an irradiation. The development of computer simulation of radiation damages, as was emphasized in [2], is connected to development of complex computer simulation systems, which will unite in themselves various more private algorithms of simulation of nuclear structures of defects in crystals and their reorganization. Preliminary structure of such system submitted in a fig.1 there was discussed.

The development of ideal computer system for radiation creep simulation demands development of complex simulation system, which structure will be coordinated to structure submitted on a fig. 1. However development of such system while remains by affair of the future. Today efforts are concentrated on development simplified simulation radiation creep system, which main component is the system of dislocation movement simulation, for which work all other necessary making (fig. 1) are set as entrance parameters, for with the help of such systems it is possible to receive estimations of radiation creep speed. In the future the developed simulating systems of dislocation movement will enter by a key part in ideal computer system simulating radiation creep of materials.

© UkrObraz'2000

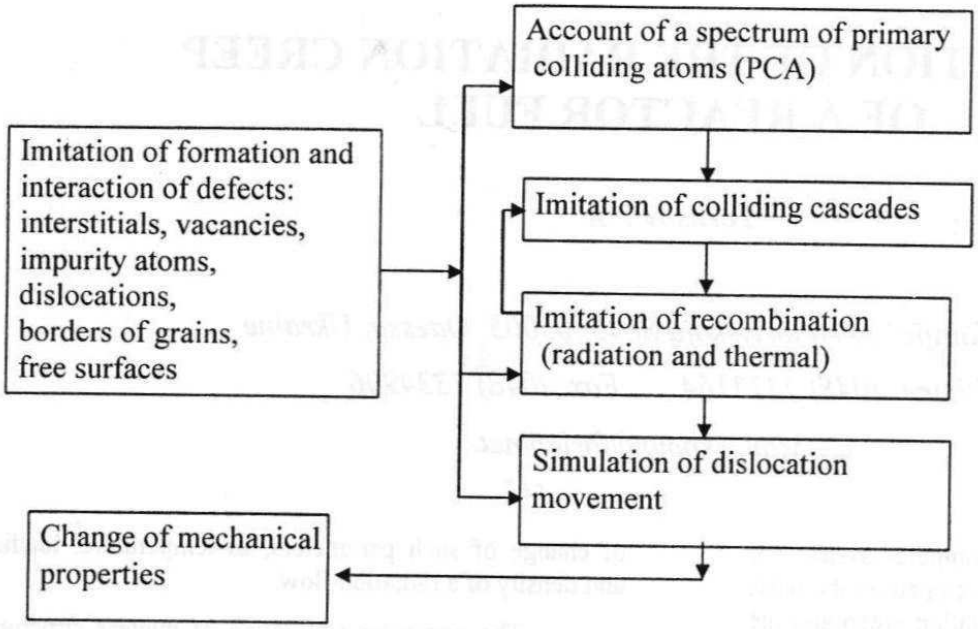


Fig. 1. The structure of system simulating influence of an irradiation on mechanical properties.

2. PHYSICAL MODEL OF RADIATION CREEP

The physical model accepted by the author, assumes presence in an irradiated material active to movement dislocations, obstacles for driving dislocations, as which the spherical centers of expansion of various capacity and dislocation loops of a edge type were considered, and PD (interstitials and the vacancies), generated by an irradiation (in divided materials PD are generated also by process of braking of splinters of division), and the established concentration PD considerably exceeds thermoequilibrium. The model of a radiation creep is explained within the framework of the mechanism of gliding and climbing dislocations based on the conception of a dislocation as not ideal sink for PD [3]. The model is efficient for installed concentration PD, considerably exceeding thermally steady state concentration. The gliding of dislocation are describing as due to moving dislocation kinks in Peierl's relief. The climbing of dislocation are describing as due to moving dislocation jogs.

3. METHOD OF DYNAMIC SIMULATION OF FLEXIBLE DISLOCATIONS MOVING

The mathematical model of dislocation movement is based on model of dislocation gliding in a Peierl's relief, carried out by means of movement of kinks, and on model of dislocation climbing carried out by means of movement of dislocation jogs being sinks for PD [3]. The obstacles are described by the fields of stress, for example. voids, spherical congestions of point defects, spherical allocation in alloys and, under certain conditions, admixture atoms in solid solutions of replacement, and also the introduced splinters of divided materials are described by a stress field of a spherical-symmetric center of expansion.

According to a method of dynamic simulation of dislocations movement mathematical simulation on the computer of gliding and climbing of flexible dislocation was carried out as follows. On a dislocation line gets out N of basic points (each point there corresponds an elementary segment of a dislocation line), the movement by each of which pays off according to the equations of the following system:

$$(\vec{V},\vec{\xi})=0,$$

$$(\vec{V},\vec{n}_{gl})=A_1\left|d\vec{F}_{el}\right|\left(\text{sign}\left\{d\vec{F}_{fr}-(d\vec{F}_{el},\vec{n}_{gl})\right\}-1\right), \quad (1)$$

$$(\vec{V},\vec{n}_{cl})=A_2^I(d\vec{F}_{el}+\vec{F}^I_{con},\vec{n}_{cl})-A_2^V(d\vec{F}_{el}+\vec{F}^V_{con},\vec{n}_{cl}),$$

where $\vec{V}$ - speed of movement of a basic dislocation point, $\vec{\xi}$ - unit tangent vector of a dislocation line, $\vec{n}_{gl}$ - unit vector determining a direction of gliding, $\vec{n}_{cl}$ - unit vector determining a direction of climbing, $d\vec{F}_{el}$ - force caused by elastic stress fields (a stress of external loading, a stress of obstacles, a stress of dislocation self-action), $d\vec{F}_{fr}$ - force of friction determined by a starting stress of dislocation, A_1-coefficient, the meaning for which can be received from expression for speed of gliding dislocation, A_2^α - the coefficients, the meanings for which can be received from expression for speed of climbing dislocation, $\vec{F}^\alpha_{con}$ - concentration forces caused by superfluous concentration of PD [3].

The equation system (1) is adhered to local system of coordinates, formed by vectors $\vec{\xi}, \vec{n}_{gl}, \vec{n}_{cl}$, and at movement of flexible dislocation these vectors can change the direction, therefore in common system of coordinates the making speeds of movement of a basic point are defined by the following expression

$$V_i' = \ell_{ij} V_j , \qquad i, j = 1,2,3, \qquad (2)$$

where V_j - are calculated on the equations (1), ℓ_{ij} - cosines of angles between axes of local and common systems of coordinates. The location of a basic point "n" in common system of coordinates through k- steps of account is defined by expression

$$(x_i^n)^k = (x_i^n)^{k-1} + (V_i'^n)^k \cdot \tau , \text{ where } i = 1,2,3. \quad (3)$$

The location of a dislocation line is defined by means of approximation of the locations of basic points. At realization of simulating accounts it was considered, that at the initial moment of time dislocation is rectilinear and has edge character, that the boundary points of a dislocation line are so removed from obstacles, that the direction of a dislocation line in these points remains constant.

4. COMPUTER SYSTEM, SIMULATING OF A RADIATION CREEP, AND RESULTS OF SIMULATION

The computer system simulating of radiation creep of the block of a model crystal, taking place under action of stretching loading, is developed, according to above described theoretical model of radiation creep and method of dynamic simulation of gliding and climbing dislocations. It is considered, that in the block of a model crystal a planes of active dislocation gliding are focused in relation to the enclosed external loading according to angle $\lambda = 45^\circ$ and active dislocations, which movement is simulated and also which movement causes creep, rely originally rectilinear and parallel to an axis OX_1.

Agrees [4] creep rate of the simulated block of a crystal is estimated according to expression

$$\dot{\varepsilon} = b\rho u^* , \qquad (4)$$

where $$u^* = \frac{1}{\Delta\tau \rho V} \sum_i^N \Delta S_i , \quad (5)$$

ρ - density of mobile dislocations, b - module of a Burger's vector, ΔS_i - area, covered i dislocation in time $\Delta\tau$, V - volume of a crystal, N - number of dislocations in volume V.

As obstacles were considered spherical centers of expansion and edge dislocation loops (simulates of a prismatic dislocation loop both interstitial and vacancy type formed in a plenty in uranium at an irradiation).

The account of stress (external stress, stress of the spherical center of expansion, stress of a dislocation loop, stress of dislocation self-action) is carried out in elastic continuum approach.

At realization the computer experiments on radiation creep research for the block of a modeling crystal the parameter of a crystal lattice, size of atom, coefficients of diffusion, characteristics and other parameters, which task is necessary for modeling were set appropriate for uranium. Let's result some of them: $\mu = 7{,}05 \cdot 10^{10}$ N/m^2, $\nu = 0{,}35$, radius of atom $r_a = 1{,}41 \cdot 10^{-10}$ m, Burger's vector of dislocation $b = 2{,}5 \cdot 10^{-10}$m, dislocation density $\rho \sim 10^{13}$ m^{-2}, formation energies of vacancy and interstitial accordingly - $F_V \approx 2 \cdot 10^{-19}$ J and $F_I \approx 5 \cdot 10^{-19}$ J, formation energy of a dislocation kink $F_K \approx 1 \cdot 10^{-19}$ J, formation energy of $F_{jog} \approx 1 \cdot 10^{-19}$ J. In a wide range of temperatures active dislocation, ensuring deformation of uranium, are dislocations, having a Burger's vector in direction [110] and system of sliding (010) [110].

At realization the computer experiments on radiation creep research of technical uranium and its alloys with the small additions of molybdenum (from 0,9 % up to 1,3 %) the following results are received. The received dependence of the established radiation creep rate of technical uranium from external loading has linear character and is in the good consent with experimental data [5]. The dependence of the established radiation creep rate of technical uranium from interstitial concentration is received. As the interstitial concentration is proportional to density of a neutrons flow, the comparative analysis of the submitted curve and experimental dependence of established radiation creep rate of technical uranium from density of a flow of neutrons is competent [5]. The comparative analysis shows that settlement and experimental dependence qualitatively and quantitatively are coordinated among themselves [5]. Comparison of the received results and experimental data also allows to make a conclusion that the established interstitial concentration in uranium at density of a flow $\sim 10^{12}$ n/cm^{2}s is in an interval $10^{-11} \div 10^{-13}$.

In the fig.2 the curve of dependence of the established radiation creep rate from temperature for the technical uranium received as a result of simulating accounts at loading $\sim 1 \cdot 10^7$ N/m^2 and established interstitial concentration$\sim 10^{-12}$ is submitted. In the fig.2 the curve of dependence of the established radiation creep rate from temperature for alloys of uranium with the small additions of molybdenum (0,9%÷1,3 %), received as a result of simulating accounts at loading

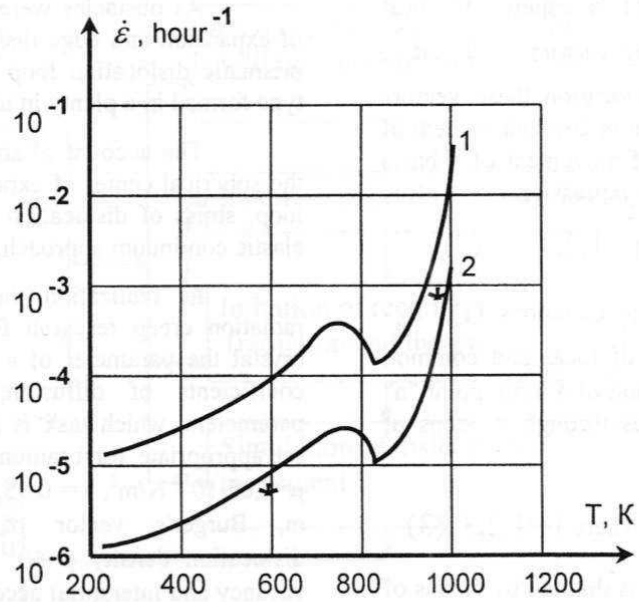


Fig. 2. The temperature dependencies of the established radiation creep rate $\dot{\varepsilon}$ for the technical uranium (curve 1) and for alloys of uranium with the small additions of molybdenum (0,9%÷1,3 %) (curve 2) at neutron flow density $\sim 10^{12}$ n/cm^{2}s and loading $\sim 1 \cdot 10^7$ N/m^2, where T - temperature (* - experimental meanings [6, 7]).

$\sim 1 \cdot 10^7$ N/m^2 and established interstitial concentration $\sim 10^{-13}$, is also submitted. At construction of a curve the experimental data given in works [6, 7] were taken into account and marked on figure by asterisks.

5. CONCLUSION

The computer system for radiation creep simulation, developed on the basis of radiation creep model, considering gliding and climbing dislocations as a not ideal sink for PD, and method of dynamic modeling of dislocations movement, is submitted.

The results of computer experiment on radiation creep of technical uranium and its alloys with the small additions of molybdenum (from 0,9 % up to 1,3 %) are received.

Are submitted the curve of dependence of the established radiation creep rate from temperature for the technical uranium and the curve of dependence of the established radiation creep rate from temperature for alloys of uranium with the small additions of molybdenum (0,9%÷1,3 %), received as a result of simulating accounts. The submitted curves allow to predict temperature behavior of uranium and its alloys and having practical importance for successful designing of reactor elements, their safe operation and forecasting of their behavior in emergencies.

The received results show that the developed computer system is effective addition to direct experimental researches.

REFERENCES

1. Кирсанов В.В., Орлов А.Н. Моделирование на ЭВМ атомных конфигураций дефектов в металлах. // Успехи физических наук.–1984.–Т. 142, вып.2.–С. 219-264.
2. Агранович В.М., Кирсанов В.В. Проблемы моделирования радиационных повреждений в кристаллах. // УФН.–1976.–Т. 118, вып.1.–С. 3-51.
3. Тарасов В.А. Математическое моделирование радиационной ползучести реакторного топлива на примере урана и его сплавов. // Ядерная и радиационная безопасность.–1999.– Т. 2, вып. 1.– С. 77–107.
4. Предводителев А.А. Анализ возможностей построения теории упрочнения из опытов по дислокациям.–К.: Наукова думка, 1972. –С. 74 –94.
5. Айтхожин Э.С., Ибрагимов Ш.Ш., Кусаинов С.К., Пятилетов Ю.С. Исследование радиационной ползучести альфа-урана. // Известия АН Каз. ССР. Серия физико-математическая, 1981, N6, с. 42-46.
6. Дашковский А.И, Маскалец В.Н., Тарасов В.А. Исследование физико-механических свойств металлов и сплавов: Отчет о НИР / МИФИ. –№ 81020398,0282. 0075503. –М.: 1981. –34с.
7. Дашковский А.И., Маскалец В.Н., Тарасов В.А. Исследование физико-механических свойств металлов и сплавов: Отчет о НИР / МИФИ. –№ 81020398, 0283.0035655. –М.: 1982. –18с.

ПРО ОДИН МЕТОД АДАПТИВНОЇ АПРОКСИМАЦІЇ НЕПЕРЕРВНИХ СИГНАЛІВ АПРОКСИМУЮЧИМИ ІМПУЛЬСНИМИ СПЕКТРАМИ

Воронова О.С.

Відділення гібридних моделюючих та керуючих систем в енергетиці ІПМЕ НАН України,
м. Київ, вул. ген. Наумова 15, т. 444-98-35

Викладено метод адаптації сітки для неперевних сигналів з довільною кількістю особливостей, який базується на апроксимуючих імпульсних спектрах (АІС). Наведено приклади застосування методу. Отримані результати порівнюються з результататми апроксимації по рівномірним сіткам.

При обробці інформації часто виникає потреба апроксимації сигналів, що є комбінацією швидко та повільно змінюючихся компонент. У випадку апроксимації таких сигналів по системі локальних базисних функцій виникає питання про вибір частоти дискретизації або, іншими словами, кроку сітки. Використання рівномірної сітки прзводить або до надлишкового представлення сигналу на ділянках, де переважають компоненти, що повільно змінюються (якщо вибір кроку визначається точністю апроксимації швидких компонент), або до великих похибок в апроксимації швидко змінюючихся компонент сигналу (див. мал.1,б).

В роботі пропонується один з алгоритмів адаптації сітки та апроксимації такого роду сигналів. Даний алгоритм спирається на особливості локально-імпульсних систем базисних функцій, а саме, системи базисних функцій, що складається з блочно-імпульсних та кусково-лінійних базисних функций, яка є модифікацією системи запропонованої в [1, 2]:

$$v_i(t) = \sigma\left(t-(i-1)h-a\right)-\sigma\left(t-ih-a\right), \qquad (1)$$

$$w_i(t) = \left(\frac{2(t-a)}{h}+1-2i\right) v_i(t), \quad i = 1, \ldots, m. \qquad (2)$$

Зручність такої системи базисних функцій в тому, що по-перше, ця система є ортогональною, а по-друге, в кожній підсистемі проміжки визначення базисних функцій не перетинаються. Отже ми можемо розглядати апроксимацію сигналу на кожному інтервалі $[a+(i-1)h, a+ih]$ незалежно від решти області визначення сигналу. Дійсно, при використанні локально-імпульсних функцій (1)-(2) ми отримуємо кусково-поліноміальную апроксимацію сигналу:

$$x_a(t) = \sum_{i=1}^{m} \left(X_0(i)v_i(t) + X_1(i)w_i(t)\right), \qquad (3)$$

де X_0 та X_1 – компоненти вектора $\tilde{X}$ спектра сигналу $x(t)$ [2]:

$$X_0(i) = \frac{1}{h}\int_{a+(i-1)h}^{a+ih} x(t)v_i(t)dt = \frac{1}{h}\int_{a+(i-1)h}^{a+ih} x(t)dt, \qquad (4)$$

$$X_1(i) = \frac{3}{h}\int_{a+(i-1)h}^{a+ih} x(t)w_i(t)dt = \frac{3}{h}\int_{a+(i-1)h}^{a+ih} x(t)\left(\frac{2(t-a)}{h}+1-2i\right)dt. \qquad (5)$$

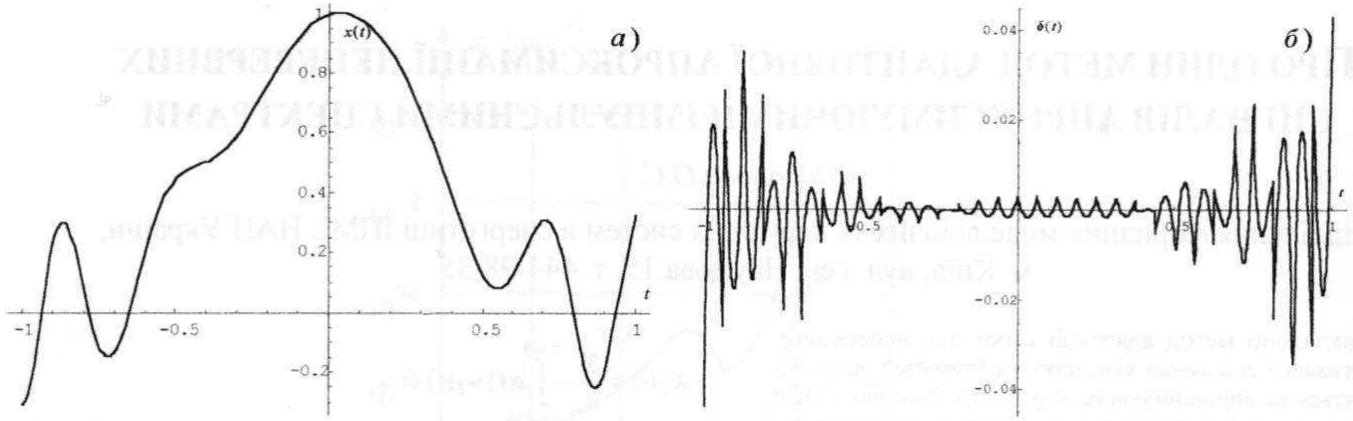
Як можна помітити з (3), апроксимуючий поліном $x_a(t)$ фактично скаладається з m полиномів першої степені, кожен з яких визначено на проміжку $[a+(i-1)h, a+ih]$. Не важко переконатися, що відхилення апроксимації $x_a(t)$ від заданого сигналу на всьому інтервалі $[a, b]$ є сумою відхилень від сигналу поліномів вигляду: $X_0(i)v_i(t)+X_1(i)w_i(t)$, де $i = 1, m$ – номер підінтервалу.

Нехай розглядається сигнал $x(t)$ на проміжку зміни аргументу $t:[a,b]$. На цьому інтервалі введемо рівномірну сітку з кроком h_0: $\bar{\omega}_0 = \{t_i = a+ih_0, i = 1, m_0\}$. Припустимо, що середньоквадратичка похибка повинна розподілятися рівномірно, тоді можна перейти від оцінки похибки на всій області визначення до оцінки середньоквадратичної похибки на i-тому підінтервалі. При цьому на кожному інтервалі $[a+(i-1)h_0, a+ih_0]$ інтеграл середньоквадратичної похибки повинен складати ϵ/m_0, де ϵ – задана точність. Проте, при апроксимації сигналу по рівномірній сітці $\bar{\omega}_0$ зустрічатимуться підінтервали, на яких похибки будуть значно більше (менше) заданого порогового значення (див. мал.1,б).

Оберемо один з підінтервалів, на якому похибка більше заданого значення. Нехай для визначенності це буде проміжок $[a, a+h_0]$. Розглядаючи його як нову область визначення сигналу, що досліджується, можна ввести на ньому рівномірну сітку $\bar{\omega}_1$ з меншим кроком h_1 (скажімо, $h_1 = h_0/2$) та отримати на ній більш точну апроксимацію. Оцінивши відхилення на кожному підінтервалі сітки $\bar{\omega}_1$, можна знову виділити ділянки, на яких похибка перевищує задану точність та застосувати до них запропоновану процедуру.

Таким чином, ми отримуємо ітераційний процес покращання сітки. Якщо на кожній ітерації ми будемо модифіковувати сітку, замінюючи ті ділянки сітки $\bar{\omega}_{j-1}$, на яких похибка перевищує порогове значення, більш точними сітками $\bar{\omega}_j$, то отримаємо нерівномірну сітку $\bar{\omega}$, на кожному підінтервалі якої апроксимація відхиляється від сигналу у припустимих межах. Більш того, якщо в процесі модифікації сітки ми будемо зберігати спектр сигналу на тих ділянках,

© UkrObraz'2000



Мал. 1. Приклад аппроксимації сигнала по рівномірній сітці. (а) Сигнал, що апроксимується, та (б) похибка апроксимації цього сигналу по рівномірній сітці з кроком $h = 0.0625$.

де похибка апроксимації лежить у припустимих межах, можна отримати спектр та апроксимацію сигнала по сітці $\bar{\omega}$.

Оскільки система базисних функцій (1)-(2) безпосередньо залежить від вибору сітки, то можна говорити про ітераційну зміну системи базисних функцій. Можна також зауважити, що в запропонованій процедурі використовуються елементи теорії мультирезолюційного аналізу [3,4], оскільки отриману таким чином систему базисних функцій, визначену на нерівномірній сітці, можна розглядати як систему, що породжується з "материнської функції" (в даному випадку двох "материнських функцій") шляхом масштабування та зсуву. Користуючись тією ж термінологією про вихідну систему (1)-(2) можна сказати, що вона породжується з материнських функцій лише за допомогою зсуву.

Для наочності продемонструємо запропонований алгоритм апроксимації та відновлення сигналу на прикладі. Розглянемо сигнал (мал.1,а):

$$x_1(t) = \frac{t}{\pi}\cos(4\pi t^2) + e^{-5t^2} \qquad (6)$$

на інтервалі зміни аргументу $[-1, 1]$ та задамо порядок точності $\epsilon = 10^{-3}$. Введемо вихідну рівномірну сітку $\bar{\omega}_0$, в якої кількість інтервалів дорівнює $m = 8$. Далі будуємо за формулами (3)-(5) апроксимацію сигнала (мал.2,а) та оцінимо інтеграл середньоквадратичної похибки на кожному підінтервалі сітки:

$$\epsilon_h = \{\, 0.00745,\ 0.00036,\ 0.00005,\ 0.00005,\ 0.00011,\ 0.00002,\ 0.00029,\ 0.00804 \,\}$$

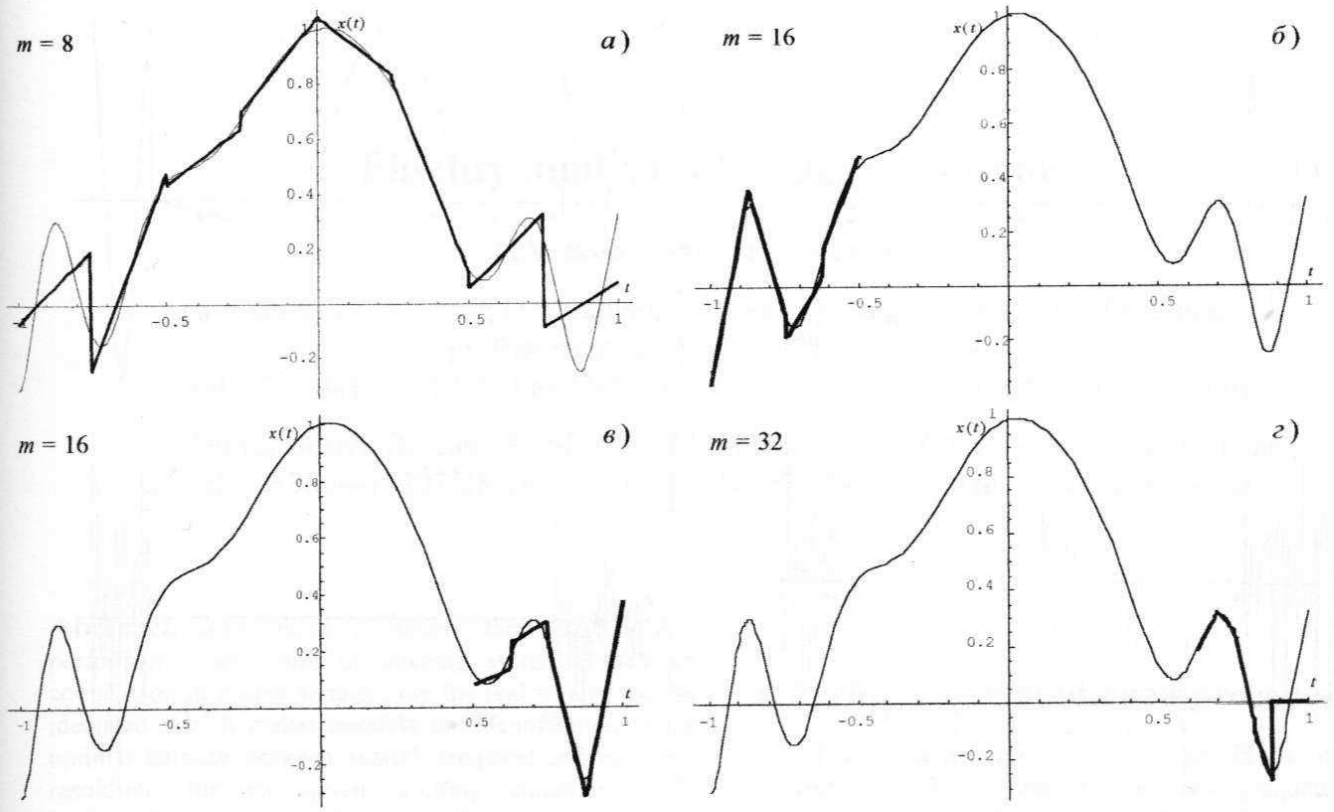
Якщо врахувати припущення про рівномірний розподіл похибки, то отримаємо, що на кожному підінтервалі похибка не повинна перевищувати $\epsilon/m = 0{,}125 \cdot 10^{-3}$. В нашому випадку це виконується скрізь за винятком 1, 2, 7 та 8 підінтервалів $\bar{\omega}_0$. Модифікуємо AI-спектр функції $x_1(t)$, записавши нулі замість значень спектру на перших двох та на останніх двох підінтервалах та збережемо отриманий спектр. Для спрощення процедури відновлення сигналу необхідно також зберегти інформацію про поточну сіттку. В якості параметрів, що описують сітку були обрані величина кроку, кількість врахованих підінтервалів та ліва межа області, на якій визначено сітку. Таким чином, частина спектру функції $x_1(t)$, яка відповідає $m = 8$, матиме вигляд:

$$\left\{\{0,0,0.53575,0.86909,0.94073,0.46411,0,0,\ 0,0,0.10302,0.16958,-0.10692,-0.34964,0,0\},\ \frac{1}{4},8,-1\right\}$$

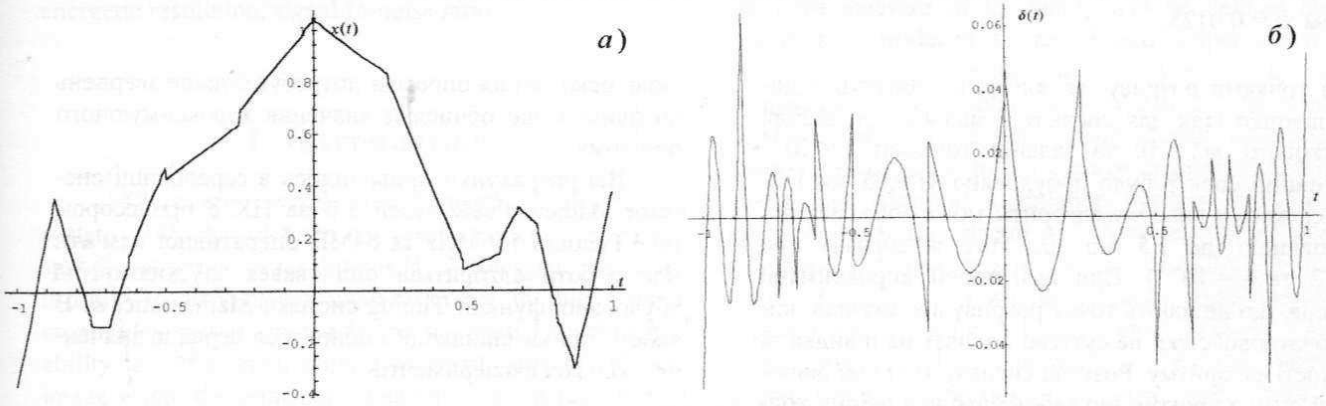
Згідно запропонованої процедури, розглянемо дві області: $I_1 = [-1, -0.5]$ та $I_2 = [0.5, 1]$. Кожна з цих областей складається з двох інтервалів сітки $\bar{\omega}_0$. Збільшимо (для визначеності подвоїмо) кількість інтервалів розбиття для кожної з областей I_1 та I_2. Іншими словами, ми переходимо до нової рівномірної сітки $\bar{\omega}_1$ з $m = 16$, але розглядатимемо сингал лише на перетині $\bar{\omega}_1$ з вказаними областями Оскільки з точки зору мультирезолюційного аналізу параметр m можна розглядати як параметр масштабу, то можна говорити про представлення сигналу на новому масштабі. Отже перейдемо на новий масштаб та послідовно повторимо для областей I_1 та I_2 вищезгадану процедуру. Будемо повторювати цю процедуру для все більш точних масштабів представлення сигналу доти, доки на кожному підінтервалі отриманої нерівномірної сітки похибка не буде перевищувати ϵ/m_j, де m_j – масштабний параметр відповідної сітки.

В нашому прикладі процес уточнення сітки завершиться після трьох ітерацій (мал.2): $m = 8$, $m = 16$, $m = 32$, і ми отримаємо спетр функції $x_1(t)$:

$$\hat{X} = \left\{\left\{\{0,0,0.53575,0.86909,0.94074,0.46411,0,0,\ 0,0,0.10302,0.16958,-0.10692,-0.34964,0,0\},\ \frac{1}{4},8,-1\right\}\right.$$
$$\left\{\{-0.00679,0.12961,0,0.30781,\ 0.35818,-0.23501,0,0.17337\},\ \frac{1}{8},4,-1\right\},$$
$$\left\{\{0.10906,0,0,0.03274,\ 0.02875,0,0,0.34338\},\ \frac{1}{8},4,\frac{1}{2}\right\}$$
$$\left\{\{-0.13549,-0.01376,-0.00099,0.10951\},\ \frac{1}{16},2,-\frac{3}{4}\right\},$$
$$\left.\left\{\{0.24720,0.28762,0.08756,-0.19398,\ 0.06183,-0.03499,-0.15465-0.09257\},\ \frac{1}{16},4,\frac{5}{8}\right\}\right\}.$$



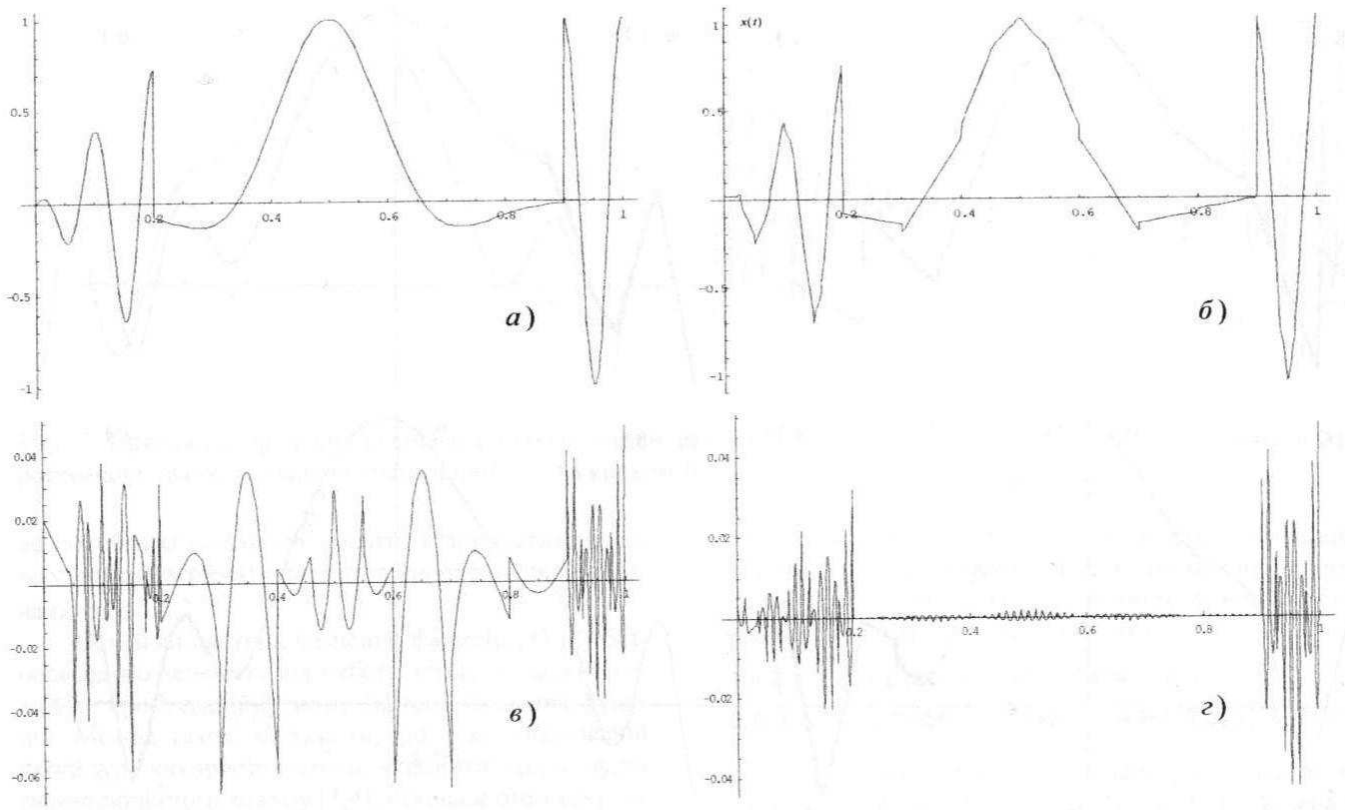
Мал. 2. Кілька ітерацій процесу адаптації сітки. Жирною кривою показані апроксимуючі поліноми на рівномірних сітках, що введені на підінтервалах, а тонкою кривою – сигнал, що апроксимується.



Мал. 3. (а) Апроксимація функції $x_1(t)$ по нерівномірній сітці та (б) графік функції похибки.

Відновити сигнал можна наступним чином. Для кожної рівномірної підсітки за формулою (3) будуємо апроксимуючий поліном, при цьому всю необхідну інформацію (елементи вектора спектра, границі сумування, розмір кроку, межі інтервала апроксимації) можна отримати з відповідної компоненти вектора $\hat{X}$. Сума отриманих апроксимуючих поліномів і дає апроксимацію сигнала на всій області визначення $[a, b]$ (мал.3,а). На мал.3,б показано графік похибки апроксимації $x_1(t)$ по адаптованій сітці. При цьому інтеграл середньоквадратичної похибки на проміжку $[-1, 1]$ склав $0.522786 \cdot 10^{-3}$. Для порівняння наведемо результати апроксимації даного сигналу по рівномірній сітці. Максимальне значення масштабного параметра, яке використовувалось під час побудови адаптованої сітки дорівнює $m = 32$. Цьому масштабу відповідає рівномірна сітка з кроком $h = 0.0625$. Графік похибки апроксимації по цій сітці продемонстровано на мал.1,а; при цьому інтеграл середньоквадратичної похибки склав $0.114275 \cdot 10^{-3}$.

Запропонований алгоритм було реалізовано у вигляді пакету в середовищі системи Mathematica 3.0 [5] та було протестовано на сигналах різних типів (наприклад, мал.1,а та мал.4,а). При проведенні чисельних розрахунків також оцінювалась швидкодія алгоритму у порівнянні з алгоритмами апроксимації сигналів по рівномірній сітці AI-спектрами другого порядку. Наведемо деякі отримані результати. При побудові адаптивного спектру сигнала, що має точки розриву (мал.4,а), збіжність алгоритму в значній мірі залежить від вибору початкової рівномірної сітки. Якщо вихідну сітку обрано так, що її вузли співпада-



Мал. 4. Адаптивна апроксимація сигналу, що має точки розриву 1-го роду. (а) Сигнал, який складається з двох сінусоїд та так званого "мексиканського капелюха", (б) апроксимація цього сигналу по адаптованій сітці , (в) графік похибки апроксимації по адаптованій сітці та (г) відхилення від сигналу апроксимації по рівномірній сітці з кроком $h = 0.0125$.

ють з точками розриву, то алгоритм збігається досить швидко (так для сигналу з мал.4,а при виборі початкового $m_0 = 10$ та заданій точності $\epsilon = 10^{-3}$ адаптивний спектр було побудовано за 9,65 сек.). У протилежному випадку алгоритм може збігатися дуже повільно (до 1,5 год. для того ж сигналу при $m_0 = 2$ та $\epsilon = 10^{-3}$). При адаптивній апроксимації сигналів, що не мають точок розриву або скачків, вибір початкової сітки не суттєво впливає на швидкість збіжності алгоритму. Розклад сигналу $x_1(t)$ по рівномірній сітці з кроком, що забезпечує відповідну точність, виконується (в залежності від вигляду сигнала) в 3-10 разів швидше.

Проте, оскільки при адаптивній апроксимації відбувається суттєва компресія інформації про сигнал (до 45% в залежності від типу сигнала) у порівнянні з представленням за допомогою рівномірної сітки, що забезпечує відповідну точність, при відновленні сигналу по його адаптивному спектру ми отримуємо значний виграш у швидкодії. Для ілюстрації, порівняємо час, витрачений на побудову графіку функції похибки: $\delta(t) = x(t) - x_a(t)$, де $x(t)$ – сигнал, а $x_a(t)$ – апроксимуючий поліном. Так, на побудову графіків з мал.1,а та мал.4,г знадобилося по 14.2сек. та 39.9сек., відповідно, а на побудову графіків з мал.3,б та мал.4,в – 4,25сек. та 9.6сек., відповідно. При обчисленні інтегралів середньоквадратичної похибки для апроксимації по рівномірній та адаптованій сіткам різниця у витратах часу стає ще більшою, оскільки ця опреація потребує більше звернень до функції, що обчислює значення апроксимуючого поліному.

Всі розрахунки проводились в середовищі системи Mathematica® версії 3.0 на ПК с процесором Intel Pentium 166MHz та 64MB оперативної пам'яті. Час роботи алгоритмів оцінювався за допомогою вбудованої функції Timing системи Mathematica®. В якості оцінки швидкодії вибиралося середне значення з кількох експериментів.

Література

1. Симак Л.А. Аппроксимирующие импульсные спектры в приложении к дробно-дифференциальному анализу. - Киев, 1989. - 56с. (Препр./ АН УССР. Ин-т проблем моделирования в энергетике; 89-8).
2. Васильев В.В., Симак Л.А., Чечь В.В., Воронова О.С., Косова А.М., Пилипенко Н.Н. Методы аппроксимации и цифровой обработки сигналов мониторинга. - Киев, 1988. - 72с. (Препр./ НАНУ. ОМУСЭ ИПМЭ; №02/98).
3. Васильев В.В., Симак Л.А., Воронова О.С., Аппроксимация в спектральных методах моделирования динамических систем. Сравнительный анализ. - Киев, 1988. - 70с. (Препр./ НАНУ. ОГМУСЭ ИПМЭ; №04/98).
4. Mallat S.G. A theory for multiresolution decomposition: The wavelet representation. // IEEE Trans. on Pattren. Anal. and Machine Intell., – 1989, vol.11. – p.674-693.
5. Wolfram S. The Mathematica Book: Wolfram Media / Cambridge University Press .– 1996, – 1403p.

Fidelity analysis of imaging systems.

V.N. Borovytsky [a], *V.V. Fesenko* [b]

[a] Optical Engineering Department, National Technical University of Ukraine,
pr. Peremogy 37, Kyiv 252 056, Ukraine
Tel. (380-44) 441-10-22 Fax (380-44) 274-59-32 E-mail borovytsky@spie.org.ua

[b] Central Design Bureau "ARSENAL" Moskovskaya Str. 8, Kyiv-10, 03 601 Ukraine
Tel. (+380-44) 293-28-19 Fax. (+380-44) 573-94-29 E-mail vfes@ukrpack.net

Abstract. The paper represents the technique for performance prediction of imaging systems based on comparison of output signals from the real system and the idealized one. It makes possible the identification of the optimal balance between spatial, temporal and energetic resolution for the given working conditions. The optimization procedures based on fidelity maximization are also discussed.

Key words: fidelity, imaging systems, spatial resolution energetic resolution, signal-to-noise ratio.

1. Introduction

There are several requirements to an imaging system (IS) that defines the performance - energetic, spatial and temporal resolutions [1,2]. Energetic resolution is ability to distinguish a small difference in output signal amplitudes that is limited by noise. Spatial resolution is ability to differ small objects or small parts of object's image where the principal limitation factor is reduction of high spatial frequencies caused by optics and detector dimensions. Temporal resolution is ability to differ fast changing of output signal amplitude in time when object moves in a field of view. Its limitation factor is reduction of high temporal frequencies in detector and read-out electronics. Generally the energetic and spatial resolutions are the most important requirements in IS design.

The key problem of IS design is to maximize the integrate parameter of IS quality called performance [1,2]. It is obvious that the performance depends on the mentioned requirements. Thus the results of design procedures will be defined by the technique of performance prediction. Summarizing said above it is clear how important is to investigate the objective techniques for performance prediction that consider all principal effects of analog-digital and analog-digital-analog IS with or without human observer [3,4].

2. Performance prediction procedure.

The principal application of an IS is image representation. For example, an analog-digital IS represents the input optical signal in form of a matrix of digital values, analog-digital-analog IS represents it in a form of a definite realization of an analog signal. Thus the objective measure of IS quality will be defined by the distortions introduced by an IS into output signal. To calculate these distortions it is convenient to create mathematical abstraction as an idealized IS. The idealized IS is a model of an IS with infinity spatial, temporal and energetic resolutions. Infinity spatial resolutions mean that spatial and temporal modulation transfer functions are equal to one and a detector and electronic do not produce any noise. Mathematically we can calculate the output signal of this IS. Now we can introduce performance as a measure of similarity between output signals from an investigated IS and an idealized one. Of course, there are known some possibilities to calculate signal difference. The least-square error shows the absolute value of the difference, coefficient of correlation clears the difference between linear combinations of the signal and the losses of information represents the difference of amplitude distributions [4,6].

3. Fidelity evaluation technique.

We consider the fidelity using the proposed abstraction - an idealized IS. This approach is different from the techniques mentioned in [4,5]. The fidelity will be an absolute measure of difference between the output signals of investigated and idealized IS in sense of normalized least-square error [4,5]. It helps as to calculate the losses in spatial and energetic resolutions. We consider digital representation of output signals. Thus the fidelity can be written in the following form [4]:

© UkrObraz'2000

$$F = 1 - \frac{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}\left(U(i,j) - U_C(i,j)\right)^2}{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U(i,j)^2} \quad (1)$$

where i, j - the coordinates of pixels,

N_X, N_Y - the dimensions of the output signal matrix,

U(i,j), U_C(i,j) - the output signals of the idealized IS and the real one that are two-dimensional arrays of pixels, respectively. The output signal of the real IS contains two parts: the signal losses after spatial transformations including spatial filtering, sampling, digital processing - U_{SL}(i,j) and noise U_N(i,j) [4,5]:

$$U_C(i,j) = U(i,j) - U_{SL}(i,j) + U_N(i,j)$$

We accept an assumption about absence of correlation between noise and the losses U_{SL}(i,j) and noise U_N(i,j). Now we can rewrite (1) in the following form doing some transformations:

$$F = 1 - \frac{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}\left(U(i,j) - U_C(i,j)\right)^2}{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U(i,j)^2} = \quad (2)$$

$$= 1 - \left(\frac{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U_{SL}(i,j)^2}{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U(i,j)^2} + \frac{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U_N(i,j)^2}{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U(i,j)^2}\right) =$$

$$= 1 - \left(NSL^2 + NEL^2\right)$$

where NSL – the normalized spatial losses coefficient that characterizes the relationship between the signal losses and energy of an output signal of an idealized IS:

$$NSL^2 = \frac{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U_{SL}(i,j)^2}{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U(i,j)^2}$$

NEL – the normalized energetic losses coefficient that characterizes the relationship between variance of noise and energy of the output signal of an idealized IS. It is connected with the signal-to-noise ratio (SNR):

$$NEL^2 = \frac{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U_N(i,j)^2}{\sum_{i=1}^{N_X}\sum_{j=1}^{N_Y}U(i,j)^2} = \frac{1}{SNR^2}$$

It is obviously that each system has the definite ratio between NSL and NEL. Other words, an IS represents a definite balance between spatial and energetic resolutions for the given working conditions. To get the better IS it is necessary to find the better balance between spatial and energetic resolutions.

4. Parameter optimization.

It is possible to apply the expression (2) for optimization considering NSL and NEL as functions of IS parameters including the coefficients of an embedded digital filter:

$$F = 1 - \left(NSL^2 + NEL^2\right) = \quad (3)$$

$$1 - \left(NSL^2(p_1, p_2 .. p_M, c_1, c_2 .. c_N) + \right.$$

$$\left. + NEL^2(p_1, p_2 .. p_M, c_1, c_2 .. c_N)\right) \rightarrow \max$$

where $p_1, p_2 .. p_M$ – the parameters of an IS,

$c_1, c_2 ... c_N$ – the coefficients of a digital filter.

According to (2) optimization changes a ratio between NSL and NEL to reach an optimum balance between energetic and spatial resolutions to guarantee the highest similarity of the output signals.

Note, NSL and NEL can be found for any IS without human observer. But by adding characteristics of human vision into the process of calculation of NSL and NEL it is possible to get fidelity F for an IS with human observer [1,2]. It is obvious that F≡1 when an IS do not introduce any distortions into the output signal, other word when the IS acts as an idealized one (3). The case F=0 represents the situation when energy of the total distortions is equal to one of the output signal of idealized IS and the output signal of an IS is almost corrupted (2,3).

The following results illustrates the application of the proposed approach for performance evaluation. We consider the IS with typical values of parameters: it has diffraction limited optics with aperture diameter 120 mm, focal length 200 mm, a scanning system shifts image with the linear velocity 1.2 m/sec, a linear focal plane array contains CTH photodiodes working in 8 – 12 mkm optical range and electronics that does signal integration during 20 mks. The object is a standard 4 bars test with background temperature Tb=293 K, temperature difference "target-background" ΔT, angular period of a bar α_X [1,2]. The IS computer simulation software calculated the characteristics F, NSL and NEL as functions of the dimension of a photosensitive elements Xe for the following working conditions: case 1 - α_X = 1', ΔT=0.1 K; case 2 - α_X = 2', ΔT=0.2 K; case 3 - α_X = 2', ΔT=0.05 K (Pic.1,2,3). When the dimension Xe is getting bigger the IS will have higher energetic resolution and the lower spatial one (Pic.2). On the contrary, the smaller dimension increases spatial resolution but it decreases energetic one (Pic.3). Thus for

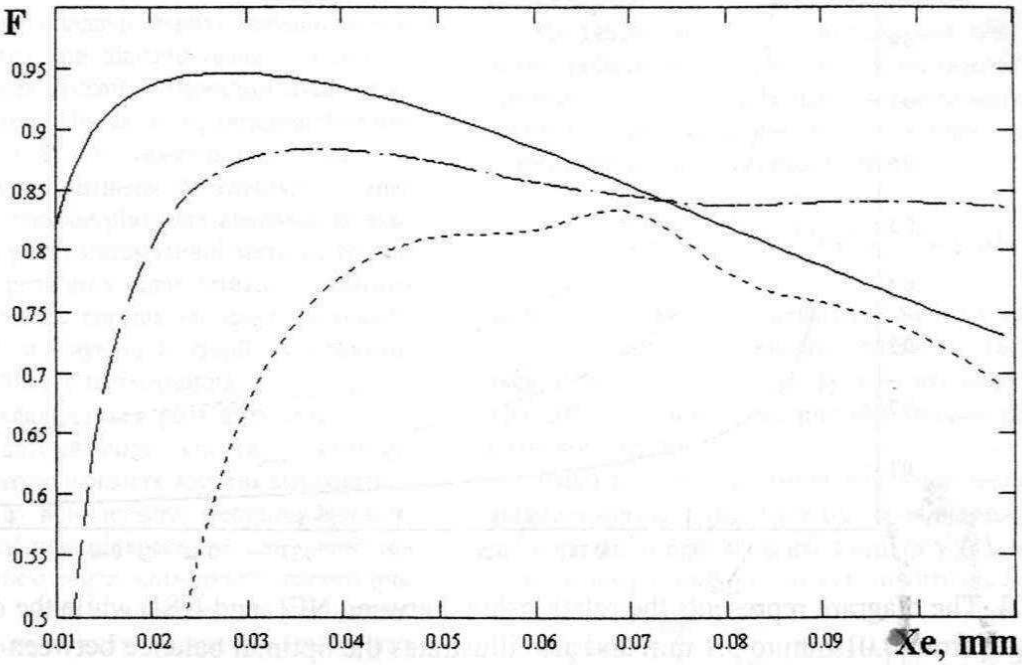
the given working conditions it exists the optimal values Xe which guarantees the highest performance. The results of computer simulation confirm this fact (Pic.1). The characteristics F = F(Xe) has one global maximum that corresponds the highest performance (Pic.1). The optimal values Xe are 0.02 – 0.03 mm for the critical working conditions (case1 – case 3) and they are dramatically depends on test-object characteristics (Pic.1). These results correspond to engineer's practice where the dimensions X_P, Y_P are in the range 0.025 – 0.05 mm (Pic.1,3). These results also show that for performance maximization it is necessary to reduce NSL because the influence of spatial harmonic reduction in case of a small test-object has much more negative influence than noise. Thus the optimal balance between energetic and spatial resolution will be represented by the ratio NEL/NSL that is equal 0.05 - 0.1 for the mentioned above working conditions (Pic.2, 3).

5.Conclusions.

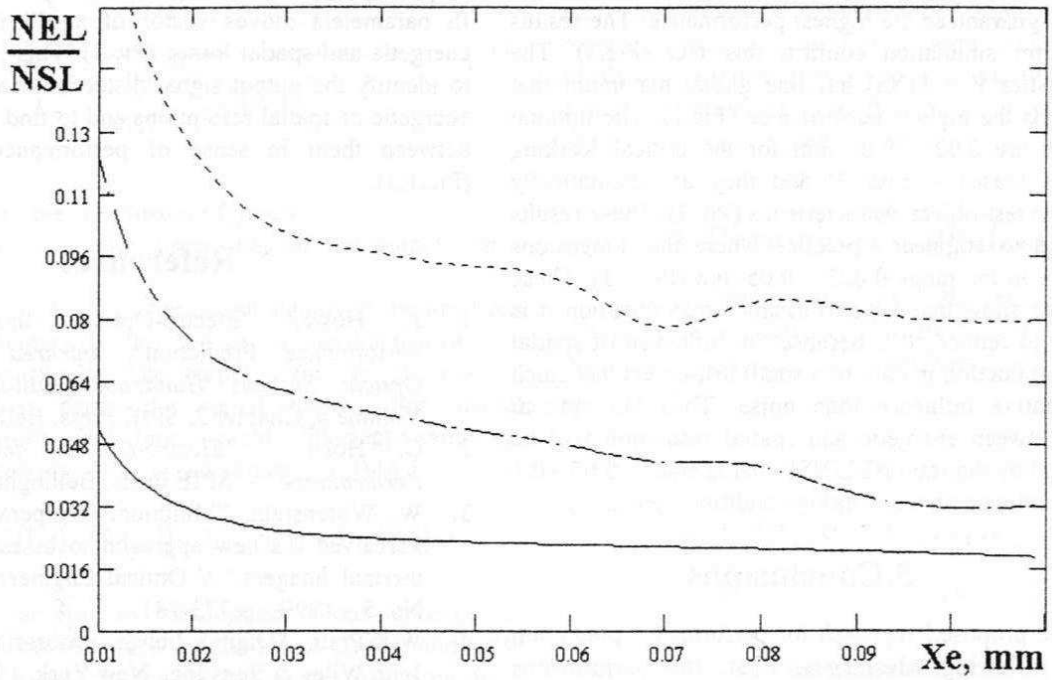
The proposed approach for performance prediction has the following advantages. First, the performance becomes an objective measure of output signal distortions introduced by an IS for given working conditions. The proposed mathematical apparatus takes into account all principal factors that define efficiency of an IS – characteristics of input signal, IS parameters, characteristics of output signal. Second, the performance is specified by a fidelity F. It makes possible parametric optimization of an IS by maximization of a merit function (Pic.1). Third, there is a convenient and convincing graphical illustration of optimization process: variation of IS parameters moves vector of an IS in space of the energetic and spatial losses (Pic.3). This illustration helps to identify the output signal distortions caused by limited energetic or spatial resolutions and to find optimal balance between them in sense of performance maximization (Pic.1,3).

References

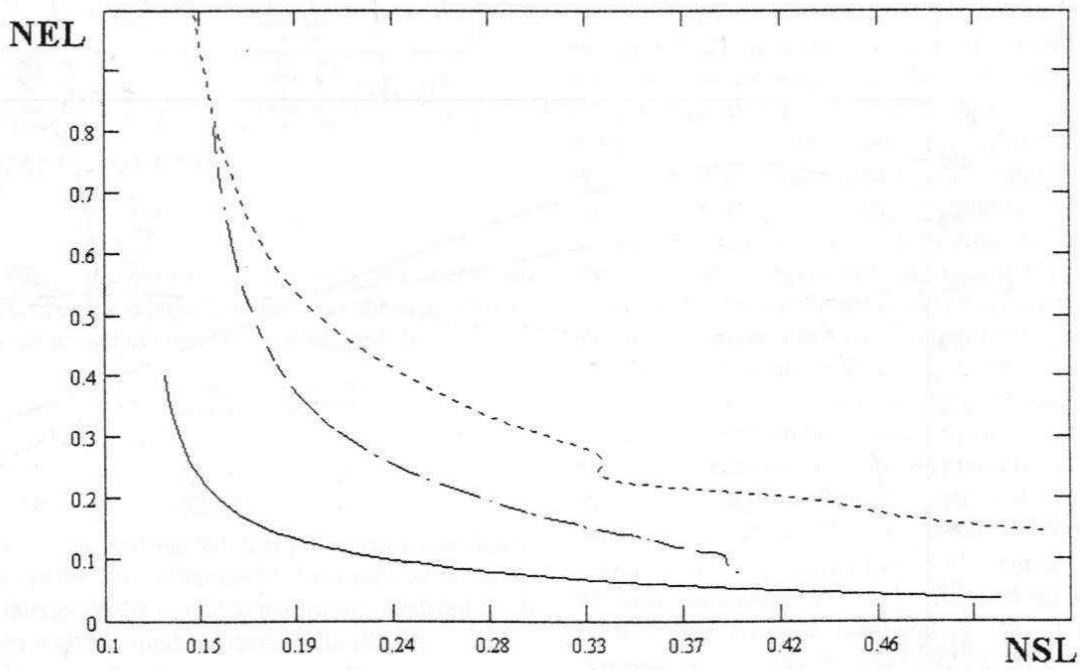
1. J. Howe, "Electro-Optical Imaging System Performance Prediction", *Infrared and Electro-Optical Systems Handbook.* Editor Dudzik M., Volume 4, Chapter 2, SPIE Press, Bellingham, 1993.
2. G. Holst , *"Electro-Optical Imaging System Performance."* - SPIE Press, Bellingham, 1995.
3. W. Wittenstein, "Minimum Temperature Difference Perceived - a new approach to assess undersampled thermal imagers." // Optical Engineering – Vol.38 – No. 5 – 1999 – p.773-781.
4. W.K.Pratt, "Digital Image Processing", p.685-686, John Wiley & Sons Inc., New York, 1991.
5. S.K. Park, Z. Rahman, "Fidelity analysis of sampled imaging systems." // Optical Engineering – Vol.38 – No. 5 – 1999 – p.786-800.
6. F.O. Huck, R.A. Gartenberg, S.K. Park, Z. Rahman, "Informatic-theoretic assessment of sampled imaging systems." // Optical Engineering – Vol.38 – No. 5 – 1999 – p.742-762.



Pic.1. The diagram represents the fidelity F as a merit function F(Xe) for the parametric optimization: solid line - α_X = 1', ΔT=0.1 K; dash line - α_X = 2', ΔT=0.2 K; dot line - α_X = 2', ΔT=0.05 K.



Pic.2. The diagram represents the ratio NEL/NSL as a function of the dimension Xe to illustrate the optimal balance between the spatial and energetic resolutions :
solid line - α_X = 1', ΔT=0.1 K; dash line - α_X = 2', ΔT=0.2 K; dot line - α_X = 2', ΔT=0.05 K.



Pic.3. The diagram represents the relationship between NEL and NSL while the dimensions Xe changes from 0.01 mm to 0.1 mm and also illustrates the optimal balance between the resolutions:
solid line - α_X = 1', ΔT=0.1 K; dash line - α_X = 2', ΔT=0.2 K; dot line - α_X = 2', ΔT=0.05 K.

АЛГОРИТМ І ПРИСТРІЙ ОЦІНКИ СТІЙКОСТІ НЕЛІНІЙНИХ НЕСТАЦІОНАРНИХ СИСТЕМ, ЯКІ ДОПУСКАЮТЬ ЛІНЕАРИЗАЦІЮ

СЕРГІЙ ЮХИМЧУК

Вінницький державний технічний університет, Україна, 21021, м. Вінниця, Хмельницьке шосе, 95.Тел.(0432) 446848. Факс: (0432) 433375. E-mail: uhim @ vstu. vinnica.ua

Наводиться алгоритм і функціональна схема пристрою оцінки стійкості нелінійних нестаціонарних систем, які можливо лінеаризувати. За рахунок потактної роботи пристрою, на який отримано патент України, він дозволяє розв'язати задачу автоматизації визначення стійкості класу систем, що розглядаються, у частотній області.

ВСТУП

При проектуванні любих автоматичних систем необхідно забезпечити їх працездатність, тобто при функціонуванні вони повинні досягти поставленої мети із заданою точністю та якістю. Зрозуміло, що успішно розв'язати такі задачі можливо лише при забезпеченні нормального їх функціонування та нечутливості до неминучих сторонніх збурень, від яких у принципі неможливо позбавитися при експлуатації будь -яких автоматичних систем контролю та управління технічних і технологічних об'єктів та процесів. Тобто, автоматичні системи повинні стійко працювати, незважаючи на дію на них різноманітних збурень.

Саме тому, перша задача аналізу поведінки, яка завжди розв'язується при діагностуванні таких систем - задача оцінки їх стійкості. Причому, якщо стійкість лінійних систем, для яких справедливий принцип суперпозиції, є їх властивістю (тобто вимога стійкості таких систем визначає її структуру та значення відповідних параметрів), для визначення якої існують відносно прості математичні методи, то для нелінійних систем розв'язок задач стійкості набагато складніший. Для таких систем не існує загальних критеріїв стійкості, а існуючи критерії, як правило, мають обмежену область застосування.

Ще більше ускладнюється розв'язок задач стійкості нелінійних автоматичних систем у випадку, коли під дією неконтрольованих збурень змінюються їх параметри, тобто вони стають нестаціонарними. При цьому слід особливо підкреслити, що, в принципі, дуже важко забезпечити незмінність параметрів систем любого класу при їх використанні у реальних промислових умовах, в яких практично неможливо позбавитися від таких неконтрольованих збурень як температура, вологість, запиленість, вібрації, електромагнітні завади, тощо.

В інженерній практиці широко застосовуються такі методи розрахунку нелінійних систем, які дозволяють, після деяких спрощень, застосувати математичний апарат, який використовується для дослідження лінійних систем. Найпоширенішим засобом, який дозволяє здійснити таке спрощення, є лінеаризація нелінійних систем Крім того, коли параметри систем, що можуть бути лінеаризовані, змінюються не більше ніж на 20% на деякому інтервалі часу, то, якщо використати метод заморожуваних коефіцієнтів на цьому інтервалі, то лінеаризовану нелінійну систему зі змінними параметрами можливо розглядати як лінійну із постійними параметрами, аналіз поведінки якої можливо здійснювати добре відомими методами [1, 2].

1.МАТЕМАТИЧНА МОДЕЛЬ ОЦІНКИ СТІЙКОСТІ

Якщо врахувати особливості Фур'є - інтегрального методу ідентифікації ФІМІ [3], то метод, що пропонується, можливо використати для аналізу стійкості систем при випадкових їх вхідних та вихідних сигналах.

Нагадаємо основні співвідношення ФІМІ, виходячи з роботи [3]. Відомо, що для лінійних або лінеаризованих систем зв'язок між їх характеристиками та характеристиками вхідного та вихідного сигналів задається рівнянням Вінера-Хопфа:

$$R_{yx}(\tau) = \int_0^{\infty} R_x(\tau - t)g(t)dt, \qquad (1)$$

де $R_{yx}(\tau)$ — взаємно кореляційна функція вихідного y(t) та вхідного x(t) сигналів систем, що досліджуються; $R_x(\tau)$ — кореляційна функція вхідного сигналу x(t); g(t) — вагова характеристика лінійної або лінеаризуємої системи.

Якщо t_0 — час взаємної кореляції вихідного та вхідного сигналів ($R_{yx}(t_0) \approx 0$), то на відрізку $[-t_0, t_0]$ дві кореляційні функції, які входять у (1), можливо розкласти у ряди Фур'є, які будуть мати вигляд:

$$R_x(\tau) = \frac{a_0}{2} + \sum_{l=1}^{\infty} a_l \cos l\omega_1 \tau, \qquad (2)$$

© UkrObraz'2000

$$R_{yx}(\tau)=\frac{a_0}{2}+\sum_{l=1}^{\infty}b_l\cos l\omega_1\tau+\sum_{l=1}^{\infty}c_l\sin l\omega_1\tau, \quad (3)$$

де $\omega_1=\frac{\pi}{t_0}$, а

$$\begin{cases} a_l=\frac{1}{t_0}\int_{-t_0}^{t_0}R_x(\tau)\cos l\omega_1\tau d\tau, \\ b_l=\frac{1}{t_0}\int_{-t_0}^{t_0}R_{yx}(\tau)\cos l\omega_1\tau d\tau, \\ c_l=\frac{1}{t_0}\int_{-t_0}^{t_0}R_{yx}(\tau)\sin l\omega_1\tau d\tau. \end{cases} \quad (4)$$

Якщо підставити (2), (3) в (1), то після нескладних перетворень, здійснених в роботі [3], можна отримати, що

$$\frac{b_0}{2}+\sum_{l=1}^{\infty}b_l\cos l\omega_1\tau+\sum_{l=1}^{\infty}c_l\sin l\omega_1\tau=\frac{a_0}{2}P_0+ \sum_{l=1}^{\infty}a_lP(l\omega_1)\cos l\omega_1\tau-\sum_{l=1}^{\infty}a_lQ(l\omega_1)\sin l\omega_1\tau \quad (5)$$

Порівняння коефіцієнтів при однакових гармонійних складових у співвідношенні (5) приводить до системи рівнянь:

$$\begin{cases} P(0)=\frac{b_0}{a_0}, \\ b_l=a_lP(l\omega_1), \\ c_l=-a_lQ(l\omega_1). \end{cases} \quad (6)$$

А з (6) витікає, що:

$$\begin{cases} P(l\omega_1)=\frac{b_l}{a_l}, \quad l=0,\ 1,\ 2,...,\ n; \\ Q(l\omega_1)=-\frac{c_l}{a_l}, \ l=0,\ 1,\ 2,...,\ n, \end{cases} \quad (7)$$

де $P(l\omega_1)$ та $Q(l\omega_1)$ —відповідно, значення дійсної та уявної частотних характеристик систем, що досліджуються, на частотах $\omega=l\omega_1$ (l=1, 2, 3...n) , а P(0) — значення дійсної частотної характеристики систем на частоті $\omega=0$ (відмітимо, що $Q(0)=0$).

Якщо використати формулювання частотного критерію Найквіста для стійкої розімкненої системи, то замкнена система буде стійкою, якщо виконується наступна умова:

$P(l\omega_1)> -1$ при любому $Q(l\omega_1)$. (8)

Таким чином, якщо використати умову (8), а відповідні значення частотних характеристик знайти за допомогою ФІМІ, то можна побудувати пристрій для автоматичного визначення стійкості замкнених нелінійних систем, які можливо лінеаризувати, по частотних характеристиках розімкнених систем. При цьому підкреслимо, що якщо організувати потактну роботу пристрою, що пропонується, таким чином, щоб за такт його роботи параметри практично не змінювались, то такий пристрій можливо використати для оцінки стійкості нестаціонарних лінеаризованих систем. З викладеного зрозуміло, що таких умов можливо досягти, коли параметри не змінюються (або зміни незначні) за час t_0. Тобто, при виконанні вищезгаданих умов можливо розв'язати задачу робастної стійкості нелінійних нестаціонарних систем, динамічні процеси в яких можливо описати лінійними співвідношеннями за рахунок лінеаризації.

2.АЛГОРИТМ ОЦІНКИ СТІЙКОСТІ

Якщо підсумувати вище наведене, то при умові, що за час t_0 параметри лінеаризованих систем не змінюються, або змінюються незначно, то для автоматизації процесу визначення стійкості систем, що розглядаються, необхідно:

— підрахувати значення функцій $R_{yx}(\tau)$ та $R_x(\tau)$, які входять у (1);

— визначити значення коефіцієнтів їх розкладу у ряди Фур'є за співвідношенням (6.1.4.) та за допомогою цих коефіцієнтів визначити значення дійсної $P(l\omega_1)$ та уявної $Q(l\omega_1)$ частотних характеристик систем на частотах $\omega=l\omega_1$ (l=0, 1,...,n) за допомогою виразів (6);

— провести перевірку виконання умов (8) та забезпечити індикацію виконання чи невиконання цих умов.

3.ОПИС РОБОТИ ПРИСТРОЮ ОЦІНКИ СТІЙКОСТІ

Пристрій працює таким чином. Вхідний x(t) та вихідний y(t) сигнали системи, стійкість якої необхідно визначити (СК), поступають на входи корелятору, який підраховує значення кореляційних функцій $R_x(\tau)$ та $R_{yx}(\tau)$. Значення $R_x(\tau)$ та $R_{yx}(\tau)$ поступають, відповідно, на перший та другий інформаційні входи першого блоку запису та зчитування. Одночасно з початком подачі сигналів із виходів корелятора подається сигнал з виходу генератора тактових імпульсів на вхід першого лічильного тригеру, а з його інверсного виходу - на вхід управління першого блоку запису та зчитування (БЗС 1). Цей сигнал дозволяє запис кодованих значень $R_x(\tau)$ та $R_{yx}(\tau)$ у БЗС 1. Тривалість запису сигналів у БЗС 1 визначається частотою імпульсів, які виробляються генератором і залежить від області застосування пристрою та характеристик систем, що досліджуються. Для успішної роботи пристрою, що пропонується, в телемеханічних системах необхідно вибирати частоту імпульсів за співвідношенням:

$$f=\frac{1}{2t_u}, \quad (9)$$

де t_u - тривалість вимірювання сигналів x(t) та y(t).

Враховуючи відмічені раніше особливості методу, на якому базується робота пристрою, необхідно, щоб

$$t_u \approx t_0. \quad (10)$$

Зрозуміло, що частоту імпульсів, які виробляє генератор, можливо змінювати в залежності від умов використання даного пристрою та особливостей зміни параметрів систем, стійкість яких необхідно визначити. При цьому зміна частоти не впливає на роботу пристрою, робота якого розглядається.

Після закінчення часу t_u сигнал з інверсного виходу першого лічильного тригеру дозволяє зчитування кодованих значень кореляційних функцій у БЗС 1, з якого вони поступають на входи першого та другого перетворювачів Фур'є, які здійснюють розклад $R_x(\tau)$ та $R_{yx}(\tau)$ у ряди Фур'є з частотою першої гармоніки ω_1 за співвідношеннями (4).

Сигнали, які відповідають значенням коефіцієнтів a_l $(l = \overline{0, n})$, поступають на інформаційний вхід першого комутатора, а значення коефіцієнтів b_l та c_l $(l = \overline{0, n})$ - відповідно, на перший та другий інформаційні входи другого комутатора. В цей момент часу подається сигнал запуску на вхід третього тригеру.

Сигнал з виходу елементу АБО дозволяє проходження сигналів, які відповідають значенням a_l, на вхід першого блоку порівняння, а значення коефіцієнтів b_l та c_l, відповідно, на входи першого функціонального блоку та блоку зміни знаку.

У першому блоці порівняння реалізується операція перевірки умови $a_l=0$. Якщо ця умова виконується, то сигнал з одного із виходів даного блоку порівняння поступить на третій вхід індикатору та на одиничний вхід третього тригеру, сигнал з прямого виходу якого подається на вхід елементу АБО та забороняє передачу сигналів з першого та другого комутаторів до того часу, поки не поступить сигнал запуску пристрою на нульовий вхід третього тригеру. Відмітимо, що умова $a_l=0$ є умовою припинення роботи пристрою, що пропонується.

Якщо ж $a_l \neq 0$, то сигнал з першого або третього виходів першого блока порівняння подається на синхронізуючий вхід регістра пам'яті, на інформаційний вхід якого подається сигнал, відповідний значенню коефіцієнту $a_l \neq 0$.

При наявності сигналу на вході синхронізації регістра, значення коефіцієнту a_l подаються на перший та другий функціональні блоки, на виході яких формуються сигнали, відповідні значенням $P(l\omega_1)$, $Q(l\omega_1)$ у відповідності з (7).

Ці значення АФЧХ записуються у другий блок запису та зчитування (БЗС 2) при поступленні на вхід управління БЗС 2 сигналу з виходу одновібратора. Сигнал з одновібратора подається на нульовий вхід другого тригеру з прямим динамічним входом. На одиничний вхід другого тригеру подається сигнал з виходу елементу АБО. Сигнал з цього тригеру подається на вхід блоку затримки сигналу, причому час затримки $\tau_з$ дорівнює тривалості виконання операцій, необхідних для отримання значень $P(l\omega_1)$ та $Q(l\omega_1)$ у відповідності з (7).

Після закінчення часу $\tau_з$ із виходу одновібратора подається імпульс, який дозволяє запис значень $P(l\omega_1)$ та $Q(l\omega_1)$ у БЗС 2.

Одночасно з моментом подачі імпульсу з одновібратора на вхід БЗС 2, подається сигнал, код якого дорівнює одиниці, на вхід другого блоку зміни знаку та сигнал нуля - на вхід третього блоку порівняння. На другі входи другого та третього блоків порівняння подаються, відповідно, значення $P(l\omega_1)$ та $Q(l\omega_1)$.

При виконанні умов (8) для кожного значення $P(l\omega_1)$ та $Q(l\omega_1)$ з'явиться сигнал на виході другого елементу І. Зрозуміло, що для того, щоб прийняти рішення про стійкість системи, яка досліджується, умови (8) повинні виконуватися для всіх $P(l\omega_1)$ та $Q(l\omega_1)$ $(l = \overline{0, n})$. Якщо умови (8) виконуються для всіх значень l $(l = \overline{0, n})$, то на виході другого елементу І сигнал з'явиться n+1 разів.

Подальша робота даного пристрою присвячена перевірці виконання умов (8) для всіх значень гармонічних складових (5). Для того, щоб здійснити таку перевірку, сигнал з виходу другого елементу І подається на лічильний вхід лічильника, який здійснює підрахунок кількості імпульсів, які поступають з другого елементу І. Сигнал, відповідний коду значень l $(l = \overline{0, n})$ з лічильника, подається на другий вхід четвертого блоку порівняння, на перший вхід якого подається код числа гармонічних складових L=n+1 рядів (2) та (3). Якщо умови (8) будуть виконанні для всіх знайдених значень $P(l\omega_1)$ та $Q(l\omega_1)$, то з'явиться сигнал на другому виході четвертого блоку порівняння, який подається на перший вхід блоку індикації, що засвідчує стійкість системи, що контролюється (СК). Поява такого сигналу є ознакою припинення роботи пристрою.

Якщо ж умови (8) не виконуються хоча б для однієї гармонічної складової, то з'являться сигнали на інших виходах другого та третього блоків порівнянь, які будуть подані на входи третього елементу І, з якого сигнал поступить на другий вхід індикатору. Поява такого сигналу є ознакою того, що система, яка досліджується, є нестійкою та також є ознакою припинення роботи пристрою.

Окремо відзначимо, що короткий імпульс з виходу другого елементу І подається також на вхід інвертора, на виході якого з'явиться сигнал, відповідний рівню логічного нуля, який подається на вхід першого елементу І, на виході якого буде рівень логічного нуля, який подається на вхід елементу АБО та забороняє зчитування сигналів з комутаторів. Після того, як імпульс з другого елементу І запишеться у лічильник, на вхід інвертора буде поданий сигнал, відповідний рівню логічної одиниці, а на його виході з'явиться сигнал, відповідний рівню логічного нуля.

Якщо ж число гармонічних складових L=l+1 менше числа, яке підраховане в лічильнику та яке подається на вхід четвертого блоку порівняння, то з'явиться сигнал на першому виході четвертого блоку порівняння. Цей сигнал буде поданий на другій вхід першого елементу І, з виходу якого буде поданий на вхід елементу АБО, що дозволить подальше зчитування сигналів з комутаторів та перевірку умов (8).

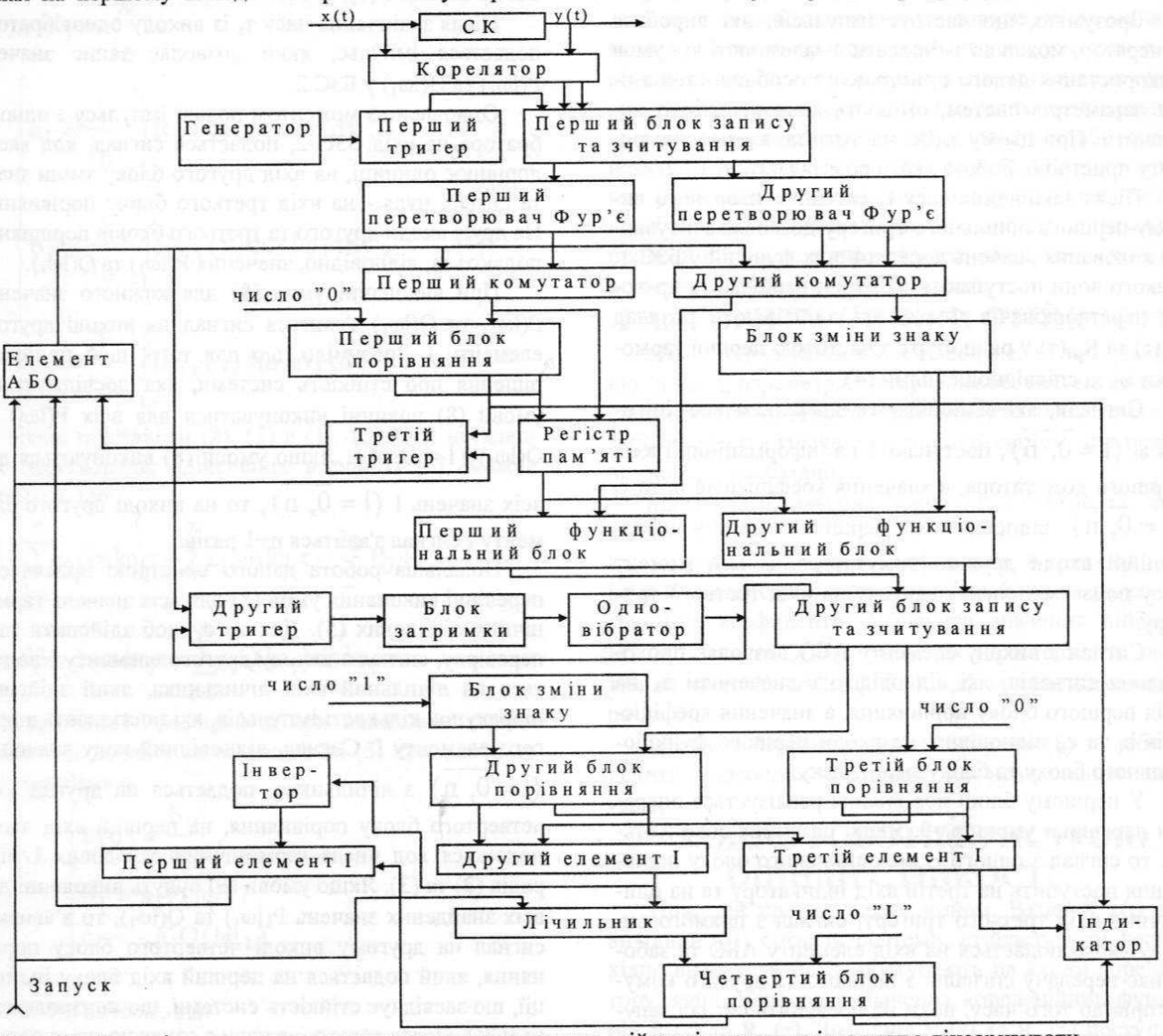


Рис.6.1. Функціональна схема пристрою для визначення стійкості систем, які можливо лінеаризувати.

ВИСНОВКИ

Таким чином, запропонований пристрій працює потактно до моменту появи сигналу на одному із виходів індикатору. Новий цикл перевірки на стійкість починається після подачі сигналу запуску, який приводить лічильник та третій тригер у початковий стан. При цьому слід підкреслити, що момент подачі сигналу запуску необхідно узгоджувати із тривалістю імпульсів, які генеруються генератором тактових імпульсів.

На завершення статті відзначимо, що пристрій, який запропоновано, розв'язує задачу визначення стійкості систем, що досліджуються, у ході їх нормальної експлуатації, без подачі тестових сигналів, не порушуючи їх функціонування. Важливим є те, що використання коефіцієнтів рядів Фур'є для розв'язку рівняння (1), як відзначено в роботі [3], дозволяє підвищити точність розв'язку в зв'язку з тим, що при використанні таких коефіцієнтів завжди здійснюється фільтрація адитивної перешкоди з нульовим середнім, тобто вони володіють регуляризаційними властивостями у більшій ступені, ніж будь-яка інша система ортогональних функцій.

ЛІТЕРАТУРА

1. Д'Анжело Г. "Линейные системы с переменными параметрами". - М.: Мир, 1974. - 288 с.
2. Заде А., Дезоер Ч. "Теория линейных систем": Пер. с англ. -М.: Наука, 1970.- 703с.
3. Мокин Б.И., Корбич Ю. "Математические модели контроля и управления в энергетике" - К.: Техника, 1990, Зелена Гура: Изд-во Высшей инженерной школы, 1990. - 192 с.
4. Патент 29994 України, МКІ G 05 B 23/02. "Пристрій для визначення стійкості систем управління"/С.В.Юхимчук, Х.М.Батиха.-№93006169; Заявл. 16.07.93; Опубл. 25.12.98; Бюл. №6.- 16с.

ІНТЕРАКТИВНА ІМІТАЦІЯ ТА АНАЛІЗ РОЗКРАШЕНОЇ СІТКИ ПЕТРИ З ВИКОРИСТАННЯМ АЛГЕБРИ ПРЕДИКАТНИХ ОПЕРАЦІЙ

Н.С. Кравець

Харківський державний технічний університет радіоелектроніки, 61166 Харків, пр. Леніна14, телефон (0572) 40-94-46, факс (0572) 40-91-13

АБСТРАКТ

У роботі розглядаються базові елементи та операції фундаментальної алгебри предикатних операцій та їх застосування при розробці імітаційної моделі ординарної розкрашеної сітки Петрі. Розроблено метод визначення множини тупікових становищ системи з використанням представлення логічної структури розкрашеної сітки Петрі в термінах фундаментальної алгебри предикатних операцій. Запропонована програмна система ISICPN, яка дозволяє розробляти і виконувати імітаційні моделі систем на мові розкрашених сіток Петрі без контролю часу, корегувати структуру сітки та аналізувати деякі її властивості, що базується на апараті фундаментальної алгебри предикатних операцій.

Ключові слова: предикати, алгебра предикатних операцій, розкрашені сітки Петрі, імітаційне моделювання, множина тупікових становищ,

ВСТУП

Алгебра предикатів та алгебра предикатних операцій належать до області логичної математики, яка виникла у зв'язку з необхідністю формального опису об'єктів і процессів.

Алгебра предикатів є узагальненням апарату булевих функцій (перехід від двійчастих знаків до букв) та апарату багатозначної логіки (перехід від функцій до відношень) і являє собою універсальний математичний апарат, який використовується теорією інтелекту для формального опису відношень[1].

Алгебра предикатних операцій являє собою надбудування над алгеброю предикатів, її поширення. Алгебра предикатів охоплюється алгеброю предикатних операцій.

Предикатні операції використовуються у тих випадках, коли виникає необхідність проводити дії над предикатами або описувати з'вязки між ними [2]. Алгебри предикатних операцій можна використовувати під час проектування систем обробки інформації, різних інформаційних структур та їх електронних схем. Алгебри предикатних операцій доцільно використовувати для формульного запису властивостей інтелектуальної поведінки того, що випробується, під час вивчення механізмів інтелекту людини. Сама ж ця поведінка вдало описується мовою алгебри предикатів.

Алгебра предикатів та алгебра предикатних операцій дозволяють формалізувати відношення, будь вони реакцією людини на зовнішні діяння або відношення між буквами в слові та словами в реченні, але при моделюванні динамічних об'єктів виникають деякі труднощі.

Існують інші інструменти дослідження систем, більш придатні для наведення динамічної поведінки системи що моделюється, наприклад сітки Петрі. На цей час існує безліч модифікацій сіток Петрі, запропонованих для того, щоб більш адекватно та зручно виражати у термінах сіток особливості функціонування реальних дискретних систем. Ці модифікації приводять до класів сіток, більш потужних, ніж клас сіток Петрі. Наприклад, імітаційні сітки Петрі, регулярні сітки Петрі, чергові сітки Петрі, регулярні стохастичні сітки Петрі, розкрашені або позначені сітки Петрі [3], алгебраїчні сітки Петрі[4].

При моделюванні сітками Петрі дискретних систем фішки часто відповідають об'єктам, що передаються від компоненту до компоненту системи (даним в інформаційних системах, деталям, ресурсам і т.і.). частіше ці об'єкти мають додаткові атрибути, які дозволяють розпізнавати їх та використовувати ці різниці для управління функціонуванням системи. Однак, фішки в сітках Петрі "безособові" і не відображають такі різниці.

Розкрашені сітки Петрі являють собою модифікацію сіток Петрі, в якій фішкам приписані деякі ознаки, наприклад, різні кольори, а умови зпрацьовування переходів та правила зміни розмітки сітки задаються спеціальною таблицею, що враховує кольори фішок.

Розкрашені сітки Петрі - це мова моделювання развинена для систем, в яких розподіл ресурсів, комунікацій та синхронізація грають важливу роль. Розкрашені сітки Петрі об'єднують в собі силу ординарних сіток Петрі з силою мови програмування високого рівня.

Для моделювання деякіх систем може виникати потреба у визначенні додаткових умов зпрацьовування переходів, а також більш загальний математичний опис функціонування сітки. Це тягне за собою розробку нових модифікацій - алгебраїчних сіток та алгебраїчних сіток Петрі високого рівня.

© UkrObraz'2000

1. АЛГЕБРИ ПРЕДИКАТНИХ ОПЕРАЦІЙ

У термінах алгебри предикатів [1] відношення визначаються таким чином. Вхідні сигнали, які пред'являють тому, хто випробується в експериментах, що відносяться до теорії Т, називаються предметами теорії Т. Непорожня множина U всіляких предметів теорії Т називається універсумом теорії Т. Нехай $x_1, x_2,..., x_m$ - всілякі місця предметів, що використовуються в експериментах теорії Т. Вони називаються предметними змінними теорії Т. Якщо предмет а знаходиться на місці x_i ($i = \overline{1, m}$), то говорять, що змінна x_i приймає значення а і пишуть $x_i = a$.

Якщо $a_1, a_2,..., a_m \in U$ та $x_1 = a_1, x_2 = a_2, ..., x_m = a_m$, то говорять, що вектор (або набір) $(a_1, a_2,..., a_m)$ належить до простору предметів U^m. Простір U^m теорії Т утворено із всіляких векторів $(x_1, x_2,..., x_m)$, які складені з предметів $x_1, x_2,..., x_m$ в теорії Т. Число m називається розмірністю простору U^m. Воно повинно збігатися із числом місць, що використовується в експериментах теорії Т.

Будь-яка множина Р простору U^m називається відношенням, що задане на U^m. Будь-яка функція, $P(x_1, x_2,..., x_m)$, що відображує множину U^m у множину $\Sigma = \{0, 1\}$, називається предикатом, що заданий на U^m, або просто предикатом на U^m. Елементи множини Σ називаються логічними константами. Елемент 0 називається хибою, елемент 1 - істиною. Предикат Р називається кінцевим, якщо множина U кінцева, та безкінцевим - у протилежному випадку. Відмінні $x_1, x_2,..., x_m$ називаються аргументами предиката $P(x_1, x_2,..., x_m)$.

Таким чином, алгебра предикатів дозволяє описувати у вигляді рівнянь тільки детерміновані та дискретні процеси. Таке визначення виключає з розгляду безперервні інформаційні процеси.

Нехай U - універсум предметів; $X_1, X_2, ..., X_m$ - предметні відмінні. М - множина всіх предикатів Р $(X_1, X_2, ..., X_m)$ на предметному просторі U^m. Множина М називається універсумом предикатів. Відмінні, що визначені на множині М, називаються предикатними відмінними. Їх значення - предикати, що задані на U^m. Множина M^n називається предикатним простором розмірністю n над предикатним простором U^m. Елементи множини M^n (набори предикатів) називаються предикатними векторами.

Будь-яка функція

$$F(X_1, X_2, ..., X_n) = Y, \qquad (1.1)$$

що відображує множину M^n у множину М називається предикатною операцією. R - множина усіх предикатних операцій. Алгеброю предикатних операцій над R називається будь-яка алгебра, що задана на носії R.

У [2] визначені булева алгебра предикатних операцій, алгебра предикатних операцій із константами та відмінними, диз'юнктивно-кон'юнктивна алгебра предикатних операцій і фундаментальна алгебра предикатних операцій.

Фундаментальною алгеброю називається алгебра предикатних операцій, в якій базисними операціями служать диз'юнкція та кон'юнкція, а базисними елементами - предикат 0 та всілякі предикати пізнавання предметів x_i^a ($i = \overline{1, m}$, $a \in U$) та предикати пізнавання предикатів X_j^P ($j = \overline{1, n}$, $P \in M$).

Теорема щодо повноти фундаментальної алгебри. Фундаментальна алгебра для будь-яких U, m и n повна. Доведення теореми наведено в [2].

2. ПРЕДСТАВЛЕННЯ ЛОГИЧНОЇ СТРУКТУРИ РОЗКРАШЕНОЇ СІТКИ ПЕТРІ У ТЕРМІНАХ ФУНДАМЕНТАЛЬНОЇ АЛГЕБРИ ПРЕДИКАТНИХ ОПЕРАЦІЙ

Представимо у термінах фундаментальної алгебри предикатних операцій розкрашену сітку Петрі. Розкрашена сітка Петрі визначається такою п'ятіркою: $N = (P, m, F, H, Cond, V, M_0)$,

де Р - скінченна непорожня множина символів $\{P_1, P_2, ... P_n\}$, які називаються місцями. Представимо їх як скінченну непорожню множину предикатів виду

$$P_i(x_1, x_2, ... x_m) = x_1^{a1} \wedge x_2^{a2} \wedge ... \wedge x_j^{aj} \wedge ... \wedge x_m^{am}, \quad (2.1)$$

де a_j - кількість фішок j - го типу в місці P_i, $a_j \in \{0, 1, 2, ...\}$;

n - кількість місць;

m - кількість типів фішок (кольорів).

Будемо вважати, що в розкрашеній сітці в кожне місце можуть поцілити фішки тільки визначеного типу і ніякі інші. Таким чином, якщо деяке a_j при будь-якій розмітці дорівнює нулю, то x_j^{aj} - можно уважати несуттєвою відмінною.

F - функція інцидентності, що вказує на наявність дуг, які з'єднують місця з переходами. Представимо її як предикат $F(p_1, p_2, ... p_n)$, де $p_i \in \{1, 2,... k\}$. Якщо існує дуга із P_1 в V_1, то $p_1^1 = 1$. Таким чином, функція інцидентності для сітки Петрі з рис.1 може бути записана в такому вигляді: $F: = p_1^2 \wedge p_2^3 \wedge \underline{p_3^4 \wedge p_4^5} \wedge p_5^6 \wedge p_6^2 \wedge p_7^3 = 1$, де для p_i^j $i = \overline{1, n}$; $j = \overline{1, k}$; $i = \overline{1, 7}$; $j = \overline{1, 6}$.

H - функція інцидентності, що вказує на наявність дуг, які з'єднують переходи з місцями, визначається аналогічно функції F, як предикат $H(v_1, v_2, ... v_k)$. Якщо існує дуга із V_1 в P_2, то $v_1^2 = 1$, де $v_i \in \{1, 2,... n\}$. Таким чином, таблиця інцидентності для сітки Петрі з рис.1 може бути записана у такому вигляді: $H = v_1^1 \wedge v_2^2 \wedge \underline{v_3^3 \wedge v_4^4} \wedge \underline{v_5^5 \wedge v_5^6} \wedge v_5^7 = 1$, де для v_i^j $i = \overline{1, k}$; $j = \overline{1, n}$; $i = \overline{1, 5}$; $j = \overline{1, 7}$.

Cond - множина умов зпрацьовування переходів $\{C_1, C_2,... C_k\}$, де к - кількість переходів сітки. Для розкрашеної сітки Петрі Cond буде мати вигляд

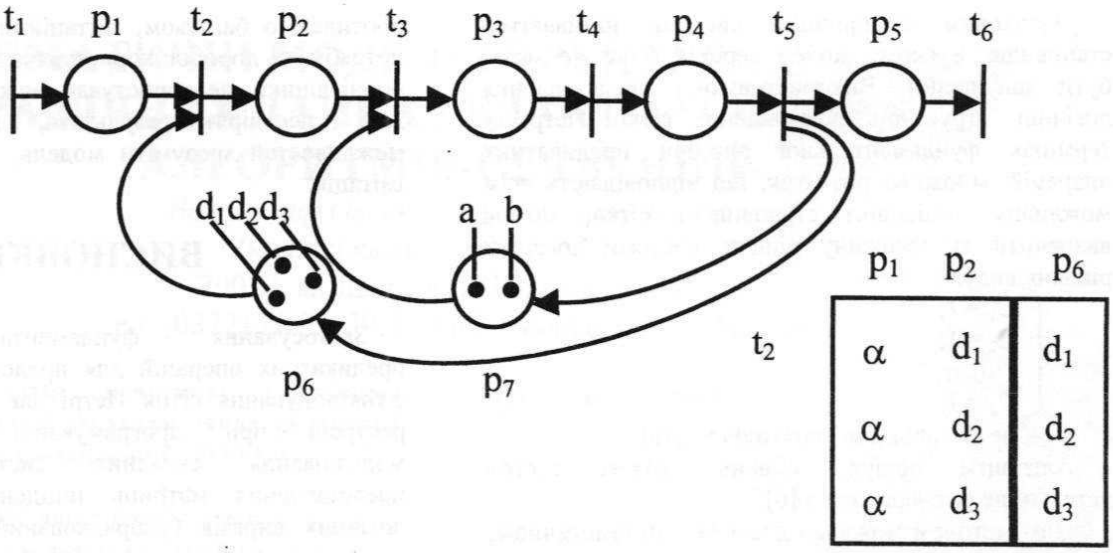


Рисунок 2.1 Розкрашена сітка Петрі та умови зпрацьовуваання переходу t_2.

кон'юнкції логічних виразів, так для сітки з рис.1 умова C_2 для переходу V_2 буде мати вигляд:

$$C_2 = (x_6 / 0\ P_1\ (x_1,\dots, x_{10}) \vee x_1 / 0\ P_6\ (x_1,\dots, x_{10})) \wedge$$
$$\wedge (x_6 / 0\ P_1\ (x_1,\dots, x_{10}) \vee x_2 / 0\ P_6\ (x_1,\dots, x_{10})) \wedge$$
$$\wedge (x_6 / 0\ P_1\ (x_1,\dots, x_{10}) \vee x_3 / 0\ P_6\ (x_1,\dots, x_{10})) =$$
$$= 0. \qquad (2.2)$$

Зручніше записати умови у вигляді системи рівнянь:

$$C_{21} = x_6 / 0\ P_1\ (x_1,\dots, x_{10}) \vee x_1 / 0\ P_6\ (x_1,\dots, x_{10}) = 0,$$
$$C_{22} = x_6 / 0\ P_1\ (x_1,\dots, x_{10}) \vee x_2 / 0\ P_6\ (x_1,\dots, x_{10}) = 0,$$
$$C_{23} = x_6 / 0\ P_1\ (x_1,\dots, x_{10}) \vee x_3 / 0\ P_6\ (x_1,\dots, x_{10}) = 0.$$

Для того щоб визначити загальний вигляд умови зпрацьовування переходу для розкрашеної сітки Петрі введемо предикатну операцію G:

$$G_i\ (P, X) = x_i / X\ P\ (x_1,\dots, x_m), \qquad (2.3)$$

де X може приймати значення 0; x_i^{ai-1}; x_i^{ai+1}.

Тоді умова (2.2) буде мати вигляд :

$$C_{21} = G_6\ (P_1, X) \vee G_1\ (P_6, X) = 0,$$
$$C_{22} = G_6\ (P_1, X) \vee G_2\ (P_6, X) = 0,$$
$$C_{23} = G_6\ (P_1, X) \vee G_3\ (P_6, X) = 0, \qquad (2.4)$$

де X = 0.

Використовуючи (2.3), умови зміни розмітки вихідних місць переходу V_2 сітки с рис.1 запишемо у вигляді:

$$R_{21} = G_1\ (P_2, X),$$
$$R_{22} = G_2\ (P_2, X),$$
$$R_{23} = G_3\ (P_2, X), \qquad (2.5)$$

где $X = x_i^{ai+1}$.

V - скінченна непорожня множина предикатних операцій $\{V_1, V_2,\dots V_k\}$, що зовуться переходами. Зпрацьовування перехода є множиною операцій над предикатами P_i та, отже, предикатом розмітки M які мають вигляд:

$$V_i = C_{il}\ (G_s\ (P_j, X)) \vee R_{il}\ (G_s\ (P_j, Y)), \qquad (2.6)$$

де $l = \overline{1, u}$; $i = \overline{1,k}$; $j = \overline{1, n}$;

$s \in \{1, m\}$;

$X = x_s^{as-1}$; $Y = x_s^{as+1}$.

M_0 - початкова розмітка. Розмітка являє собою предикат виду

$$M = P_1 \vee P_2 \vee \dots \vee P_n. \qquad (2.7)$$

Визначення обмеженої ординарної сітки Петрі в термінах алгебри предикатних операцій наведено в [5].

2.1 ВИЗНАЧЕННЯ МНОЖИНИ ТУПІКОВИХ СТАНОВИЩ

Сітки Петрі були розроблені і використовуються для моделювання. Однак само моделювання малокорисне. Необхідно провести аналіз системи, що моделюється. Цей аналіз, можливо сподіватися, приведе до глибокого проникнення у поведінку системи, що моделюється. Декілька методів анализу вже розроблено, однак в області аналізу сіток Петрі існує ще багато проблем.

Сітки Петрі можна аналізувати трьома різними методами. Перший метод - імітація. Це дуже схоже на відладку та виконання програми. Це означає, що ми можемо виконувати модель сітки Петрі для отримання статистичних результатів про поведінку системи, що моделюється.

Другий метод аналізу - входження графів (або простір становищ, або досягаємість графів). Основною ідеєю після входження графів є побудова орієнтованого графу, який має вузол для кожного можливого становища системи і дугу для кожної можливої зміни становища. Очевидно, такий граф може бути дуже великим навіть для маленької сітки Петрі. Однак, він може бути побудований і проаналізований повністю автоматично. Існують техніки, які роблять можливим таку обробку ущільнених входжень графів. Ці техніки основані на класах еквівалентності.

Третій метод аналізу - інваріанти місць. Цей метод дуже схожий з використанням інваріантів у простій веріфікації програми. Користувач конструює множину рівнянь, які задовольняють кожному становищу системи яке можна досягти. Рівняння використовуються для доведення властивостей системи, що моделюється, тобто відсутність мертвих точок (тупікових становищ).

Тупіковим становищем системи називається становище, в якому жоден перехід сітки не може бути запущений. Використовуючи представлення логічної структури розкрашеної сітки Петрі у термінах фундаментальної алгебри предикатних операцій, множина розміток, що відповідають всім можливим тупіковим становищам сітки, можна визначити як множину рішень системи логічних рівнянь виду:

$$\begin{cases} C_1 = 1; \\ C_2 = 1; \\ \ldots\ldots\ldots \\ C_k = 1, \end{cases}$$

де к - кількість переходів у сітці.

Алгоритм пошуку рішень схожих систем детальніше розглядається в [6].

Цим методом можливо довести у математичному розумінні, що система має певну множину властивостей поведінки. Однак, промислові системи часто такі складні, що не можливо або, при наймні, дорого зробить повне доведення їхньої коректності. Тому, формальні методи веріфікації повинні розглядатися як доповнення до більш неформальних результатів, що одержані за допомогою засобів імітації. Використання формальної веріфікації часто виділяє більш важливі підсистеми або більш важливі аспекти всієї системи.

3. СИСТЕМА ISICPN ІНТЕРАКТИВНОЇ ІМІТАЦІЇ РОЗКРАШЕНОЇ СІТКИ ПЕТРІ

Система ISICPN - інструментальний засіб для роботи з розкрашеними сітками Петрі без контролю часу, правила зпрацьовування переходів якої реалізуються за допомогою фундаментальної алгебри предикатних операцій. Ця система обладнана зручним інтерфейсом і призначена для інтерактивної імітації розкрашеної сітки Петрі (РСП), з наданням можливості внесення змін у структуру сітки та перевірки властивостей збереження і обмеженості. Програма реалізована на мові Delphi.

Вхідними даними є параметри та структура РСП-сітки що вводяться користувачем, вихідні дані - структура сітки (текстовий файл), значення властивостей для даної структури.

РСП-моделі без контролю часу звичайно використовуються для перевірки функціонально-логичної коректності системи. Під час інтерактивної імітації користувач керує всім процесом. Можливо побачити ефект окремих кроків безпосередньо на графічному зображенні розкрашеної сітки Петрі. Це означає, що користувач може досліджувати різні становища та обирати між можливими переходами. Інтерактивна імітація схожа на покрокову відладку. Це дає можливість при наскрізному контролі РСП-моделі досліджувати різні сценарії та перевіряти, чи відповідає робота моделі очікуванням. Це протилежно багатьом, імітаційним пакетам, що не потребують доробок користувача, які часто діють, як чорні ящики, де користувач може визначати вхідні дані і перевіряти результати, але має дуже мало можливостей зрозуміти модель, на якій зоснована імітація.

ВИСНОВКИ

Застосування фундаментальної алгебри предикатних операцій для представлення та опису функціонування сіток Петрі дає суттєву економію ресурсів при програмуванні, особливо для моделювання складних систем, тому що представлення матриць інцідентності у вигляді логічних виразів (упорядкований рядок символів) значно скорочує час пошуку. Дане твердження несправедливе тільки в одному випадку, якщо в сітці кожне місце буде виявлятися вхідним та вихідним місцем для кожного переходу, що малоймовірно. Фундаментальна алгебра предикатних операцій дозволяє формалізувати такі перетворення сіток Петрі, як приєднання, вилучення, накладення і т.і. сіток Петрі у алгебрі регулярних сіток Петрі [3]. Використання можливостей розкрашених сіток Петрі в поєднанні з апаратом фундаментальної алгебри предикатних операцій дозволяє успішно моделювати функції інтелекту людини. Система ISICPN може бути використана як інструментальний засіб під час розробки інтелектуальних агентів. Однак, розв'язання в даних термінах задач досяжності разметки значно підвищить ефективність моделювання.

ЛІТЕРАТУРА

1. Ю.П. Шабанов-Кушнаренко *Теория интеллекта. Математические средства.* - Х.: Вища шк. Изд-во при Харьк. ун-те, 1984.-144с.
2. З. В. Дударь, Н. С. Кравец, Ю. П. Шабанов-Кушнаренко. *О фундаментальной алгебре предикатных операций.* \\ Проблемы бионики, 1998. Вып. 49.с.3-13.
3. З. В. Е. Котов *Сети Петри.* - М.:Наука, 1984 - 160с.
4. C.Dimitrovici, U.Hummert, L. Petrucci *Semantics, Compositions and Net Properties of Algebraic High-level Nets*, 1991.
5. Н.С. Кравец *Моделирование структуры сетей Петри с помощью алгебры предикатных операций* Вестник ХГПУ Новые решения в современных технологиях, 1999. Вып. 44.с.13-14.
6. Д.Э. Ситников *Методы решения уравнений алгебры конечных предикатов с параметрами и их применение в информационных системах:* Дисс. на соиск. уч. степ. канд. тех. наук /05.25.05/.-Харьков, 1991.-177с.+приложения.

МАТЕМАТИЧНІ МОДЕЛІ СТРУКТУР ТЕКСТОВО-ГРАФІЧНИХ ДАНИХ РЕДАКТОРА ФОРМУЛ АЛГОРИТМІВ-СЕКВЕНЦІЙ

Володимир Овсяк, Василь Бритковський
Українська академія друкарства
79020, м. Львів, вул. Підголоско, 19
тел. (0322) 28-24-29, E-mail: ovsiak@polinet.lviv.ua

У статті подані початкові і кінцеві математичні моделі структур даних редактора формул алгоритмів-секвенцій МОДАЛ

ВСТУП. В алгебрі алгоритмів-секвенцій [1, 2] є елементи тексту і сім знаків операцій і кожен з них має специфічну графічну структуру, яка утворюється при їх наборі в створеному текстово-графічному редакторі, названому МОДАЛ [3, 4], і називається початковою структурою даних. З метою їх збереження в однотипному файлі розроблена кінцева структура даних знаків операцій і елементів тексту.

1. ПОЧАТКОВІ СТРУКТУРИ ДАНИХ

Вони описуються такими формулами алгебри алгоритмів секвенцій [1]:

1) елемент "текст"

$$\not\subset d_1 \quad \overline{\vdash D_1, *, u_1 - ?} ,$$

де

$$D_1 = \overbrace{x_1, y_1, x_2, y_2, \dots, t, f, s, c, p}^{\dots} ;$$

2) знак операції секвентування

$$\not\subset d_2 \quad \overline{\vdash D_2, *, u_2 - ?} ,$$

де

$$D_2 = \overbrace{x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, s, c, p}^{\dots} ;$$

3) знак операції елімінування

$$\not\subset d_3 \quad \overline{\vdash D_3, *, u_3 - ?} ,$$

де

$$D_3 = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, s, c, p}^{\dots} ;$$

4) знак операції інвертування

$$\not\subset d_4 \quad \overline{\vdash D_4, *, u_4 - ?} ,$$

де

$$D_4 = \overbrace{x_1, y_1, x_2, y_2, s, c, p}^{\dots} ;$$

5) знак операції циклічного секвентування

$$\not\subset d_5 \quad \overline{\vdash D_5, *, u_5 - ?} ,$$

де

$$D_5 = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, s, c, p}^{\dots} ,$$

6) знак операції циклічного елімінування

$$\not\subset d_6 \quad \overline{\vdash D_6, *, u_6 - ?} ,$$

де

$$D_6 = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, s, c, p}^{\dots} ;$$

7) знак операції паралелення

$$\not\subset d_7 \quad \overline{\vdash D_7, *, u_7 - ?} ,$$

де

$$D_7 = \overbrace{x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, s, c, p}^{\dots} ;$$

8) знак операції циклічного паралелення

$$\not\subset d_8 \quad \overline{\vdash D_8, *, u_8 - ?} ,$$

де

$$D_8 = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, s, c, p}^{\dots} .$$

2. КІНЦЕВА СТРУКТУРА ДАНИХ

Вона описується такою формулою (1):

$$\left(\begin{array}{l} \not\subset d_1 \quad \overline{\vdash D_1', *, u_1 - ?} \\ , \not\subset d_2 \quad \overline{\vdash D_2', *, u_2 - ?} \\ , \not\subset d_3 \quad \overline{\vdash D_3', *, u_3 - ?} \\ , \not\subset d_4 \quad \overline{\vdash D_4', *, u_4 - ?} \\ , \not\subset d_5 \quad \overline{\vdash D_5', *, u_5 - ?} \\ , \not\subset d_6 \quad \overline{\vdash D_6', *, u_6 - ?} \\ , \not\subset d_7 \quad \overline{\vdash D_7', *, u_7 - ?} \\ , \not\subset d_8 \quad \overline{\vdash D_8', *, u_8 - ?} \end{array}\right) , \qquad (1)$$

де $u_1, u_2, \dots, u_8$ – умови, $D_1', D_2', \dots, D_8'$ - елементи кінцевої структури даних .

Елементи алгоритму описуються формулами:

© UkrObraz'2000

$$D_1' = \overbrace{x_1, y_1, x_2, y_2, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, t, f, s, c, p}\ ,$$

$$D_2' = \overbrace{x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, c_0, c_0, c_0, c_0, c_0, c_0, s, c, p}\ ,$$

$$D_3' = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0, s, c, p}\ ,$$

$$D_4' = \overbrace{x_1, y_1, x_2, y_2, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, s, c, p}\ .$$

$$D_5' = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0, s, c, p}\ .$$

$$D_6' = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0, s, c, p}\ ,$$

$$D_7' = \overbrace{x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, c_0, c_0, c_0, c_0, c_0, c_0, s, c, p}\ .$$

$$D_8' = \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0,\ s, c, p}\ .$$

де $x_1, x_2, x_3, x_4, x_5, x_6, y_1, y_2, y_3, y_4, y_5, y_6, t, f, s, c$ – змінні, c_0 – константа.

3. МОДЕЛЬ КІНЦЕВОЇ СТРУКТУРИ ДАНИХ

Теорема. Кінцева структура даних має модель.

Доведення виконується заданням секвентних областей значень змінних:

$x_1, x_2, x_3, x_4, x_5, x_6 \in Q_1 = \overbrace{0,1,2, \dots, x_{max}}$;

$y_1, y_2, y_3, y_4, y_5, y_6 \in Q_2 = \overbrace{0,1,2, \dots, y_{max}}$;

$t \in Q_3 = \overbrace{t_1, t_2, t_3, \dots, t_n}$; $f \in Q_4 = \overbrace{f_1, f_2, f_3, \dots, f_n}$;

$s \in Q_5 = \overbrace{s_1, s_2, s_3, \dots, s_n}$; $c \in Q_6 = \overbrace{c_1, c_2, c_3, \dots, c_n}$;

$p \in Q_6 = \overbrace{1, 2, 3, \dots, 8}$.

Модель кінцевої структури даних (2):

$$\begin{aligned}
&\mathcal{L}\, d_1 \overbrace{x_1, y_1, x_2, y_2, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, t, f, s, c, p}, *, u_1 - ?\\
&\mathcal{L}\, d_2 \overbrace{x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, c_0, c_0, c_0, c_0, c_0, c_0, s, c, p}, *, u_2 - ?\\
&\mathcal{L}\, d_3 \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0, s, c, p}, *, u_3 - ?\\
&\mathcal{L}\, d_4 \overbrace{x_1, y_1, x_2, y_2, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, c_0, s, c, p}, *, u_4 - ?\\
&\mathcal{L}\, d_5 \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0, s, c, p}, *, u_5 - ?\\
&\mathcal{L}\, d_6 \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0, s, c, p}, *, u_6 - ?\\
&\mathcal{L}\, d_7 \overbrace{x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, c_0, c_0, c_0, c_0, c_0, c_0, s, c, p}, *, u_7 - ?\\
&\mathcal{L}\, d_8 \overbrace{x_1, y_1, x_2, y_2, \dots, x_6, y_6, c_0, c_0,\ s, c, p}, *, u_8 - ?
\end{aligned} \qquad (2)$$

ВИСНОВКИ. 1) Засобами теорії алгоритмів секвенцій описані формалізовані початкові і кінцеві структури текстово-графічних даних редактора формул алгоритмів-секвенцій МОДАЛ. 2) Побудована математична модель розроблених структур даних.

ЛІТЕРАТУРА

1. В.К. Овсяк. Засоби еквівалентних перетворень алгоритмів інформаційно-технічних систем. // Доповіді НАН України, № 9, 1996. – с.83-89.

2. В. Овсяк. Алгебра алгоритмів-секвенцій. // Комп'ютерні технології друкарства, №3, 1999. с.3-12.

3. В. Бритковський. Модель алгоритму розпізнавання об'єктів знаків операцій алгебри алгоритмів та його програмування. // Вісник ДУ "Львівська політехніка", №387, Львів, 2000. с. 470-473.

4. В. Бритковський. Моделі алгоритмів інструментальних засобів системи моделювання алгоритмів та їх програмування. // Комп'ютерні технології друкарства, № 3, 1999. с. 48-49.

АЛФАВІТНИЙ ПОКАЖЧИК. AUTHOR INDEX

Абрамов Сергій 59
Анагностопулос К. 145
Антонов О.В. 223
Антощук С. Г. 261
Арунрунгрусмі С. 125
Астола Йаакко Т. 55, 59
Ахметшин О. М. 117, 153, 161, 173, 213, 231
Бабаєв Т. А. 67
Бабак О. В. 67
Базилевич Роман 87, 91
Байбак Ю. В. 187
Баклан І. В. 67
Бар Б. 45
Бенхаліл А. 291
Боот В. 291
Боровицький В. М. 305
Боюн В.П. 283
Бритковський Василь 317
Буда Антоніна 191, 221
Буда Сергій 221
Будник Микола 253
Бухало О. П. 217
Васильєв В.І. 71, 75
Васильєва Тетяна 191
Вейга Альваро 21
Вергадос Д. 145
Вертеест Жоріс С.М. 29
Вінцюк Тарас 9, 95
Воробель Роман 209
Воронова О.С. 301
Востров Г. М. 241
Гасан С. 145
Гасанов А. С. 67
Гвоздак Андрій 117
Гвоздєва І. М. 193
Герасімов Ю. М. 223
Герцій А. О. 187
Глушко І. М. 257
Горбань І. І. 107
Ґабрані Марія 17
Ґарсіа-Креспі Ф. 129
Ґраур Адріан 265
Давидов В. О. 261
Довгань В.В. 283
Довнар Д. В. 235
Доля Олександр 103
Драбич Олег 121
Драбич Петро 121
Еґіазарян Карен 55
Журавель Ігор 209
Заговора С. О. 37
Ібарра-Піко Ф. 129
Інсон С. 291
Ісаєв Ігор 113, 275
Кайафас Е. 145
Калмиков Володимир 157
Калмиков Дмитро 157
Кийко В. 165
Кириленко Н. Л. 149
Кириченко М. Ф. 83
Кириченко О. П. 231
Кисель А. Г. 137
Клим Б. П. 217
Кнаак У. 145
Кнофель Р. 165
Кобець Н. М. 71
Коваль Олексій 205
Ковешников С. М. 137
Кожем'яко Андрій 191
Козак М. Ф. 287
Козлова Віра 191
Койвісто Перті Т. 201
Конґраттанапрасерт В. 125
Коржов В.І. 227
Король Ольга 221
Корольов В.Ю. 237
Косаревич Ростислав 275
Кошовий В. В. 133, 245
Кравець Н. С. 313
Краснопоясовський А. С. 37, 183
Кривін Є. В. 133
Крилов В. М. 261
Кропачов Д.О. 227
Куенка-Асенсі С. А. 129
Курій В. П. 133
Кутаєв Ю. Ф. 187
Куц Роман 83
Кучоґі Ґйорґі 29, 271
Кхан М. Х. 149
Ланге Т.І. 71
Левицький О. Є. 133
Левко О.В. 169
Лепеха М. П. 83
Лєвий С. В. 107
Лисенко Генадій 221
Лисікова О. М. 235
Лінардіс Панаґіотіс 103
Лозинський А. Б. 245
Лозинський Б. А. 245
Лоренцо-Квінтанілла Й.Й. 129
Лотоцький Р. В. 63
Лукін Володимир 55, 59, 201
Лумос В. 145
Любченко В. В. 241
Магера В. М. 107
Мартинюк Тетяна 191, 221
Машталір В. П. 41
Мельник Володимир П. 201
Микитин Г. В. 217
Милейко Ю. Ю. 79
Миргород В. Ф. 137
Михайлишин Віктор 51
Михальов Д. О. 173
Мілічі Маріана 265
Моралес-Бенавенте Й.Л. 129
Морщавка С. В. 279
Нассар С. 145
Ніколенко В. В. 287
Нюнькін К.М. 169
Овсяк Володимир 317
Окуда М. 125
Онищенко І. В. 177
Опир Наталія 209
Павленко В. Д. 141
Паєвський А. С. 193
Пентюк Раду 265
Пентюк Стефан-Ґеорґ 265
Піза Д. М. 279
Політіс Діонісіос 103
Полякова М. В. 241
Пономаренко Микола 55
Потапова Р. К. 99
Почапський Є. П. 217
Предко К. Г. 235
Процько І. О. 249
Пунтсірі Б. 125
Путятин Є. П. 41
Радомський В. А. 249
Романенко Тетяна 157
Романишин І. М. 133
Рубльов Б. В. 79
Руженцев Н. В. 223
Русак Золтан 29, 271
Русин Богдан 205, 275
Рябоконь Д. І. 179
Рябцев В. Г. 149
Савчинський Б.Д. 25
Савчук В. П. 257
Сажок Микола 111
Сакер А.-Р. 145
Саричева Людмила 33
Семенюк Михайло 197
Сіроджа І. Б. 45
Сіроджа І. І. 45
Скаковська А. М. 183
Собакін А.М. 99
Стассінопулос Г. 145
Струк С.П. 75
Тарасов В.О. 227, 297
Тимченко Л. І. 187
Томаз Карлос Е. 21
Третяк Олег 17
Трохим Георгій 113
Файст Х. 165
Федоренко А. Є. 161
Федорів Р. Ф. 217
Фейтоса Рауль К. 21
Фенога Д. А. 213
Фесенко В. В. 305
Флах Б. 177
Фомін О. О. 141
Хорват Імре 29, 271
Цимбал Олег В. 201
Чамнонтай К. 125
Швейкі Н. О. 187
Шибаєв В. О. 141
Шлезінгер М. І. 177
Штанке Ґ. 165
Шутін Д.В. 153
Юхимчук Сергій 309
Яворський Ігор 51, 113

Abramov Sergey K. 59
Akhmetshin A. M. 117, 153, 161, 173, 213, 231
Anagnostopoulos C. 145
Antonov A. V. 223
Antoshchuk S.H. 261
Arunrungrusmi S. 125
Astola Jaakko T. 55, 59
Babak O. V. 67
Babayev T. A. 67
Backlan J. V. 67
Bar B. 45
Baybak Y. V. 187
Bazylevych Roman 87, 91
Benkhalil A. 291
Bojun V.P. 283
Booth W. 291
Borovytsky V. N. 305
Brytkovs'kyj Vasyl' 317
Buda Antonina 191, 221
Buda Serhij 221
Budnyk Mykola 253
Bukhalo O. P. 217
Chamnongthai K. 125
Cuenca-Asensi S.A. 129
Davydov V. O. 261
Dolia Alexander 103
Dovhan' V.V. 283
Dovnar D. V. 235
Drabych Oleh 121
Drabych Petro 121
Egiazarian Karen 55
Fedorenko A. E. 161
Fedoriv R. F. 217
Feist Ch. 165
Feitosa Raul Queiros 21
Fenoga D. A. 213
Fesenko V. V. 305
Flach B. 177
Fomin O. O. 141
Gabrani Maria 17
Garcia-Crespi F. 129
Gasanov A. S. 67
Gerasimov Yu. M. 223
Gertsiy A. A. 187
Graur Adrian 265
Gvozdeva I. M. 193
Hasan S. 145
Hlushko I. M. 257
Horban' I. I. 107
Horvath Imre 29, 271
Hvozdak Andrij 117
Ibarra-Pico F. 129
Ipson S. 291
Isajev Ihor 113, 275
Javors'kyj Ihor 51, 113
Jukhymchuk Serhij 309
Kalmykov Dmytro 157
Kalmykov Volodymyr 157
Kayafas E. 145
Khan M. H. 149
Kiiko V. 165
Kirichenko O. P. 231
Klym B. P. 217
Knaak U. 145
Knofel R. 165
Kobets' N.M. 71
Koivisto Pertti T. 201
Kongrattanaprasert W. 125
Korol' Olha 221
Koroliov V.Yu. 237
Korzhov V.I. 227
Kosarevych Rostyslav 275
Koshovyj V. V. 133, 245
Koval' Oleksij 205
Koveshnykov S. M. 137
Kozak M. F. 287
Kozhemiako Andrij 191
Kozlova Vira 191
Krasnopojasovs'kyj A. S. 37, 183
Kravets' N. S. 313
Kropachov D. O. 227
Krylov V. M. 261
Kryvin Je. V. 133
Kuc Roman 83
Kuczogi Gyorgy 29, 271
Kurij V. P. 133
Kutaev Y. F. 187
Kyrychenko M. F. 83
Kyrylenko N. L. 149
Kysel' A. H. 137
Lange T.I. 71
Lepekha M. P. 83
Levko O.V. 169
Levyts'kyj O. Je. 133
Lievyj S. V. 107
Linardis Panagiotis 103
Liubchenko V. V. 241
Lorenzo-Quintanilla J.J. 129
Lotots'kyj R. V. 63
Loumos V. 145
Lozyns'kyj A. B. 245
Lozyns'kyj B. A. 245
Lukin Vladimir V. 55, 59, 201
Lysenko Henadij 221
Lysikova O. N. 235
Mahera V. M. 107
Martyniuk Tetiana 191, 221
Mashtalir V. P. 41
Melnik Vladimir P. 201
Mikhalev D. A. 173
Milici Mariana 265
Morales-Benavente J.L. 129
Morshchavka S. V. 279
Mykhajlyshyn Viktor 51
Mykytyn H. V. 217
Mylejko Ju. Ju. 79
Myrhorod V. F. 137
Nassar S. 145
Nikolenko V. V. 287
Niun'kin K.M. 169
Okuda M. 125
Onyshchenko I. V. 177
Opyr Natalia 209
Ovsiak Volodymyr 317
Paevskiy A. S. 193
Pavlenko V. D. 141
Pentiuc Radu 265
Pentiuc Stefan-Gheorghe 265
Piza D. M. 279
Pochaps'kyj Je. P. 217
Poliakova M. V. 241
Politis Dionysios 103
Ponomarenko Nikolay N. 55
Potapova R.K. 99
Predko K. G. 235
Prots'ko I. O. 249
Pungsiri B. 125
Putyatin E. P. 41
Radoms'kyj V. A. 249
Riabokon' D. I. 179
Riabtsev V. H. 149
Romanenko Tetiana 157
Romanyshyn I. M. 133
Rubliov B. V. 79
Rusak Zoltan 29, 271
Rusyn Bohdan 205, 275
Ruzhentsev N. V. 223
Saqer Abdel-Rahim 145
Sarycheva Ljudmila 33
Savchuk V. P. 257
Savchyns'kyj B.D. 25
Sazhok Mykola 111
Schlesinger M. I. 177
Semeniuk Mykhajlo 197
Shutin D.V. 153
Shweiki N. O. 187
Shybajev V. O. 141
Sirodzha I. B. 45
Sirodzha I. I. 45
Skakovs'ka A. M. 183
Sobakin A.N. 99
Stanke G. 165
Stassinopoulos G. 145
Struk S.P. 75
Tarasov V. O. 227, 297
Thomaz Carlos Eduardo 21
Timchenko L. I. 187
Tretiak Oleh J. 17
Trokhym Heorhij 113
Tsymbal Oleg V. 201
Vasilyeva Tetiana 191
Vasyl'jev V.I. 71,75
Veiga Alvaro 21
Vergados D. 145
Vergeest Joris S. M. 29
Vintsiuk Taras K. 9, 95
Vorobel' Roman 209
Voronova O.S. 301
Vostrov G. M. 241
Zahovora S. O. 37
Zhuravel' Ihor 209