

# 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, крес-

лення, 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.

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 місцезнаходження.

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 напрямками:

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 "Авто-референт".

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 радіоактивного 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. Сміслова, текстова та усномовна інтерпретація зображень — перетворювач "зображення—смісл—текст—усна мова"
- 6.3. Комплексне сприйняття та опис аудіовідеоінформації — перетворювач "аудіовідео—смісл-текст"
- 6.4. Комплексна генерація аудіовідеоінформації за заданим "смісл-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 радіоактивного 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, ґрунтуючись на загально визначених 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 вимовлення 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.

Отже, 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 чим 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, структурувати 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 направленного 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-скопичних 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-

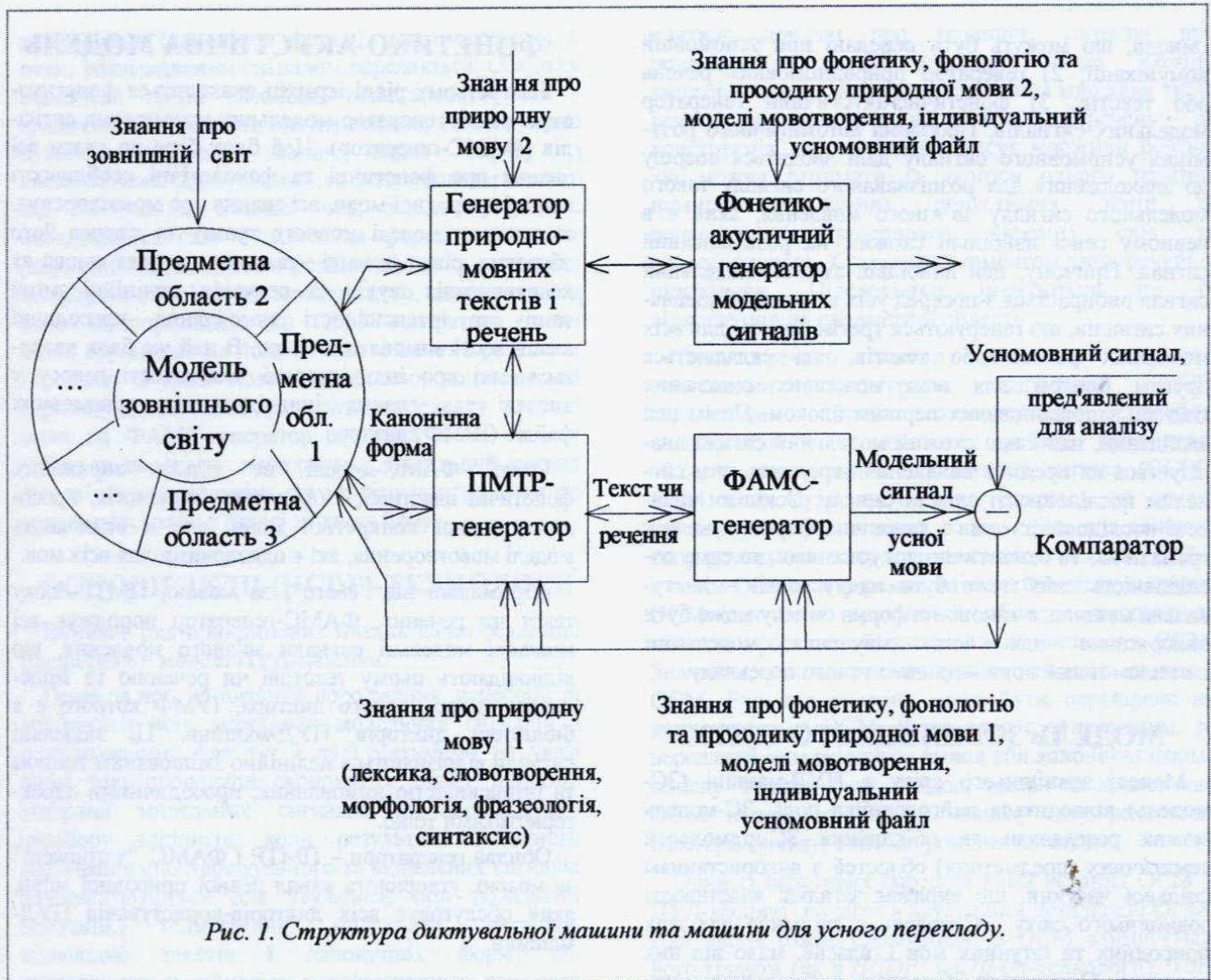


Рис. 1. Структура диктувальної машини та машини для усного перекладу.

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 загально-визнаним 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, концеп-

ції 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

сенси, 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 ґрунтується 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 прийнятою 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 редуція, 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 впливає 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 (НСДР)-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 направленою 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(НСДР)-технологією автоматичного розпізнавання, розуміння та синтезу усномовних сигналів [1-6]. Перша вимога задовольняється використанням ієрархічно (I (H – *Hierarchy*)) структурованих автоматних породжувальних граматик, які синтезують (композиція (K (C – *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 направленою перебору варіантів – зворотніми зв'язками, що йдуть від каналного 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 замінів 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 знайдені 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 *НСДР*- чи *НММ*-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].

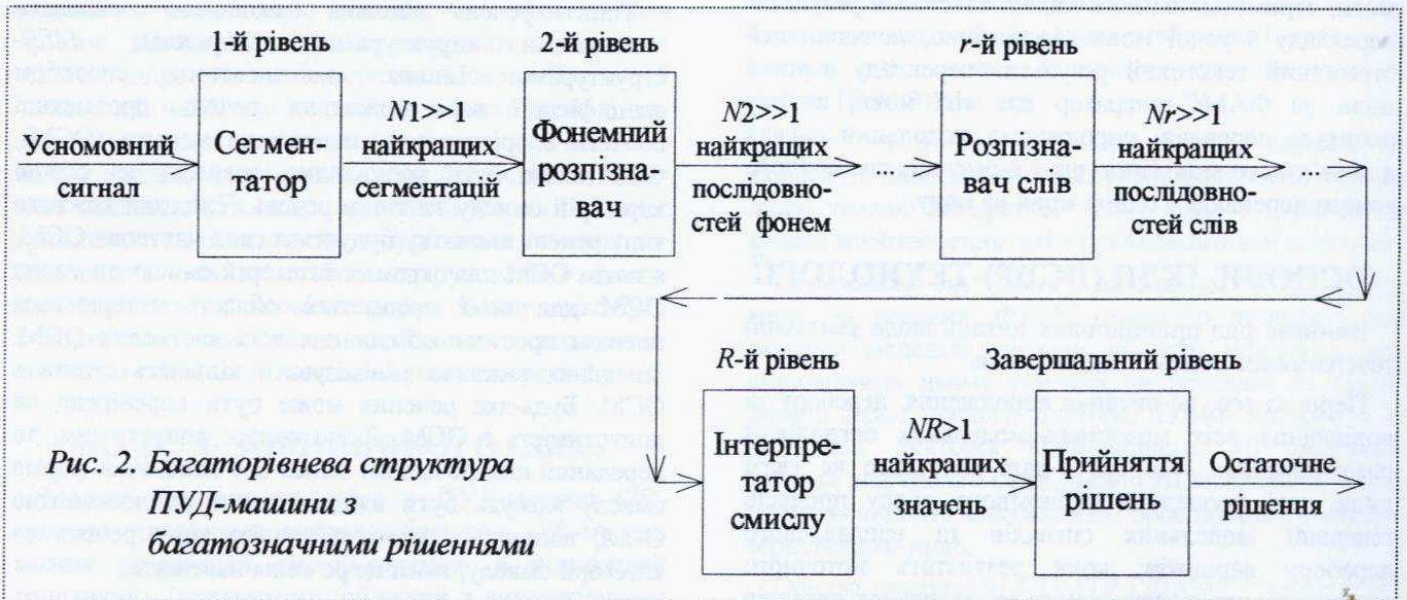
## ГЕНЕРАТИВНА МОДЕЛЬ ОК

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  $N1 \gg 1$  найкращих розбивок усномовного сигналу на сегменти, що

елементам. Далі, на другому рівні, ця багатозначна сегментація відображається в  $N2 \gg 1$  найкращих композицій з більш крупних елементів (літер, окремих деталей, дрібних предметів, елементів фону тощо). На третьому рівні ієрархії знаходить  $N3 \gg 1$  найкращих сегментацій зображення, що відповідають можливим композиціям з більш крупних елементів (надписів, предметів, людей). На вищих рівнях ієрархії виконується опис та смисловий аналіз зображення і сцени в цілому.



відповідають фонемам. Далі, на другому рівні, ця багатозначна сегментація в умовах, наприклад, вільного порядку слідування фонем, відображається в  $N2 \gg 1$  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  $N3 \gg 1$  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  $Nr \gg 1$  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  $N1 \gg 1$  найкращих розбивок зображення на сегменти, що відповідають дугам, відрізням, дрібним

## 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] Т.К. Vintsiuk, “Element-Wise Recognition of Continuous Speech Composed of Words from Specified Vocabulary”, *Kibernetika*, 1971, No. 2, pp 133-143.
- [2] Т.К. 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] Т.К. 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 Винцюк, “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  $f_1'(x) = f_1(T(x))$  in which homologous points in  $f_1(x)$  and  $f_2(x)$  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:

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

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

$$f(\phi(\bullet)) = k \exp(-\int [w * (\phi - \mu)](\mathbf{x})^2 dx) \quad (2)$$

where  $w*$  represents the application of an operator described below (for example,  $w*$  may be equal to  $\Delta$ , 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

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

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

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

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  $\phi$  and  $\mathbf{B}$  by the principle of maximum likelihood. For notational convenience, we

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

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

We seek a minimum over  $\phi$  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,

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

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

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

The functions  $\phi(\cdot)$  for which the  $Q$  is zero (the kernel of  $Q$ ) is the linear space of polynomials of degree less than  $m$ . A minimum over  $\phi$  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

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

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

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

where  $\mathbf{H}(\cdot)$  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)

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

$k = 1, \dots, N$  and the orthogonality constraints

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

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

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

$Q(\mathbf{A}, \mathbf{B})$  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  $\phi$  is a random walk, and the variance of  $\phi$  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

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

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

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

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  $\mu\text{m}$  in the alignment of brain surfaces, and an error of

about voxels 88  $\mu\text{m}$  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 through 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  $\mu\text{m}$ , and the thickness of each section is approximately 40  $\mu\text{m}$ .

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  $\mu\text{m}$ .

	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

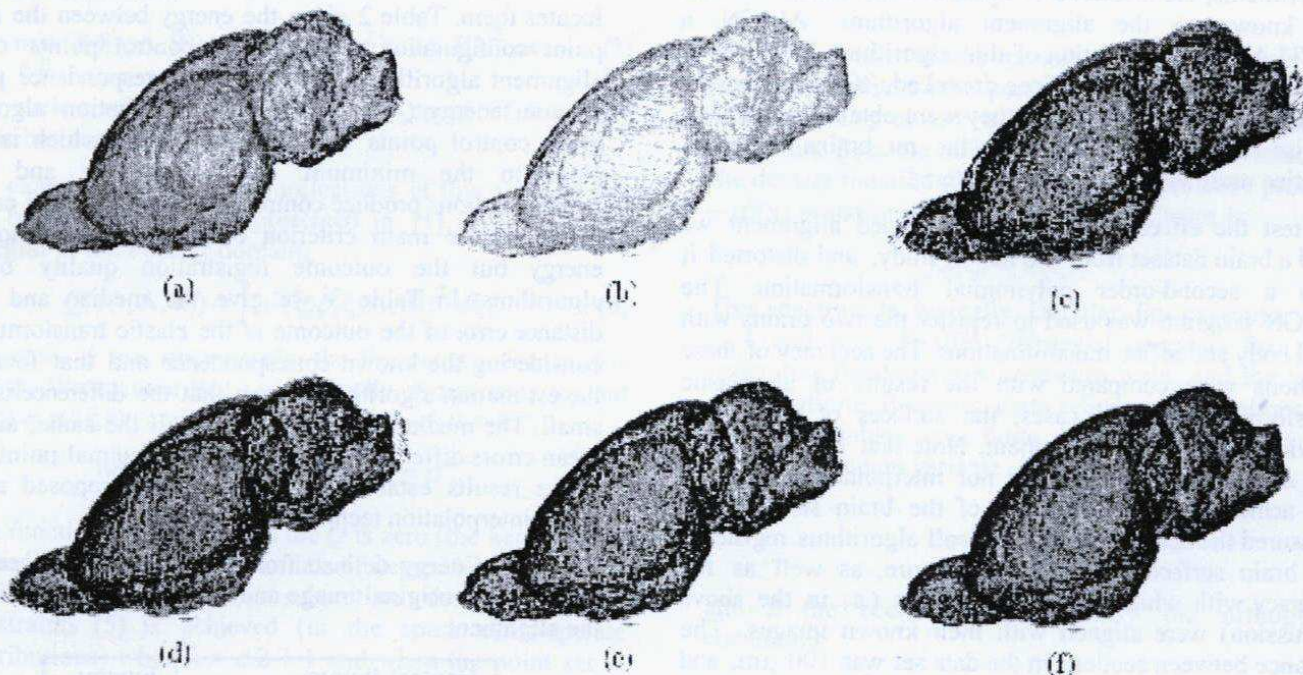
Control Points	Energy
Known Correspondence	0.0309
Elastic Estimation	0.0357
Affine	0.0829
Rigid Body	1.0454

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

Error	Known		Estimated	
	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  $E(\bullet)$  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  $\mathbf{A}$ , 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 principe 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 Kovacic, 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<sup>1</sup>, Raul Queiroz Feitosa<sup>2</sup>, Álvaro Veiga<sup>3</sup>*

<sup>1,2,3</sup>Catholic University of Rio de Janeiro  
Department of Electrical Engineering  
r. Marquês de São Vicente 225, 22453-900, Rio de Janeiro, Brazil

<sup>2</sup>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  $d_1(\mathbf{x}), d_2(\mathbf{x}), \dots, d_g(\mathbf{x})$ , ( $g$  is the number of groups) with the property that if the  $n$ -dimensional pattern vector  $\mathbf{x}$  belongs to the group  $\pi_i$  ( $1 \leq i \leq g$ ) then  $d_i(\mathbf{x}) \geq d_j(\mathbf{x})$ , for all  $i \neq j$ ,  $1 \leq i \leq g$ .

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  $1 \leq i \leq g$ , where  $\mu_i$  and  $\Sigma_i$  are respectively the mean and the covariance of group  $\pi_i$ . 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  $\Sigma_i$  is the Sample Group Covariance matrix defined by:

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

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

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

The matrix  $S_i$  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

$$S_{\text{pooled}} = \frac{(k_1 - 1)S_1 + (k_2 - 1)S_2 + \dots + (k_g - 1)S_g}{k_1 + k_2 + \dots + k_g - g}. \quad (4)$$

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

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

Since more observations are taken to calculate  $S_{\text{pooled}}$ , it will potentially have a higher rank than  $S_i$  and will be eventually full rank. Although the pooled estimate does provide a solution for the algebraic 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  $S_{\text{pooled}}$  and  $S_i$ . It is given by

$$S_{\text{mix } i} = aS_{\text{pooled}} + (1-a)S_i, \text{ where } 0 < a < 1. \quad (6)$$

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 corre-

sponding to the true covariance is placed somewhere between both ellipsoids, as shown by the solid black lines.

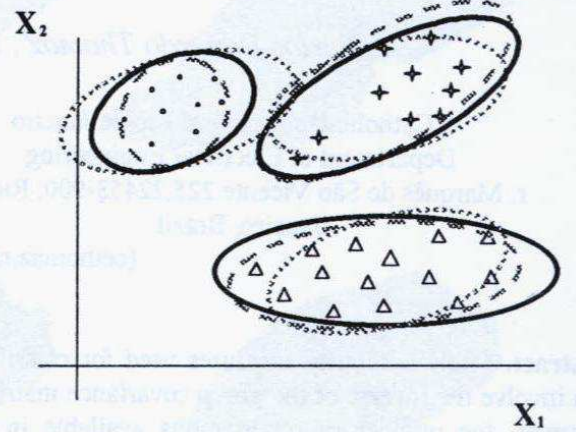


Figure 1: Geometric interpretation of  $S_i$ ,  $S_{\text{pooled}}$  and  $S_{\text{mix}}$ .

Each  $S_{\text{mix } i}$  matrix has the important property of admitting an inverse if  $S_{\text{pooled}}$  does. Let the dimension of the feature space  $n$  be such that  $S_{\text{pooled}}$  is full rank (invertible) and  $S_i$  is not full rank (non invertible). Thus  $S_{\text{pooled}}$  and  $S_i$  are respectively positive definite and positive semi-definite matrices. Since  $0 < a < 1$ ,  $a$  and  $(1-a)$  are both positive numbers. Therefore,  $aS_{\text{pooled}}$  and  $(1-a)S_i$  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]

$$\det(A + B) \geq \det(A) \quad (7)$$

Hence,

$$\det(S_{\text{mix } i}) = \det[aS_{\text{pooled}} + (1-a)S_i] \geq a \det(S_{\text{pooled}}) > 0 \quad (8)$$

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

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

## 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  $C = W^{-1}B$ , where  $W$  is the sample within groups covariance matrix, and  $B$  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  $W$  is singular. As Swets and Weng [3] observed, this problem can be circumvented 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  $W$  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 ( $S_i$ ,  $S_{pooled}$ ,  $S_{mix}$ ). 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  $S_{pooled}$  estimate against the five  $S_{mix}$  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  $S_{mix}$  produced a better performance than the  $S_{pooled}$  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  $S_{mix}$  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,  $S_{mix}$  may have a better or worse performance than  $S_{pooled}$ . The best recognition rate - 96.52% - for all the MDF components considered is reached by the  $S_{mix}$  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  $S_{group}$ ,  $S_{pooled}$ , and the five  $S_{mix}$  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  $S_{group}$  becomes singular, the  $S_{mix}$  estimate reaches better recognition rates than the  $S_{pooled}$  estimate for all values of  $a$  considered in the experiment. The best recognition rate of these experiments - 85.19% - was obtained by the  $S_{mix}$  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.

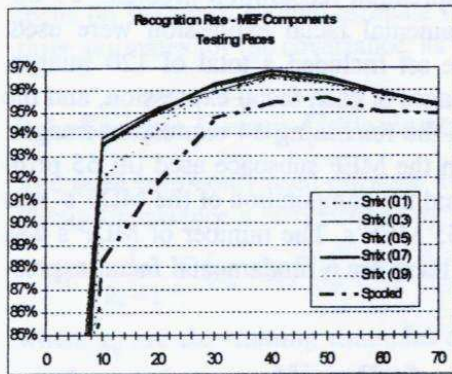


Figure 2: Face Recognition

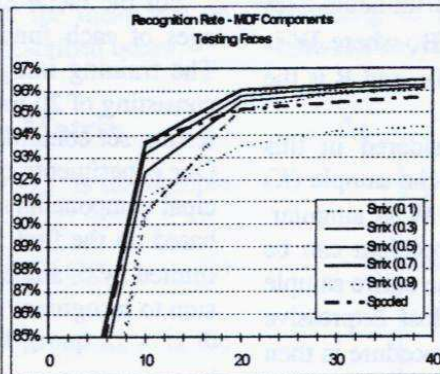


Figure 3: Face Recognition

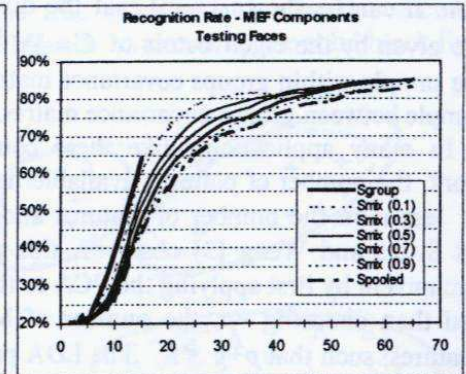


Figure 4: Facial Expression Recognition

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 Retrieval," IEEE Trans. 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 Differential Calculus with Applications in Statistics and Econometrics", Wiley and Sons Limited, pp. 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/МПП  
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 піксели, або **точки**. **Зображенням** називатимемо функцію:

$$x : T \rightarrow R.$$

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

Також розглянемо клас  $\Phi$  **монотонних функцій** вигляду  $\phi : T \rightarrow T$ , тобто:

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

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  $\phi$  точки  $t$  та  $\phi(t)$  називатимемо **відповідними**. Для зручності введемо клас  $\mathcal{P}$  функцій вигляду  $\rho: T \rightarrow T$ , 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 функцією паралакса.

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

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

Позначимо через  $N_i(t)$ ,  $0 \leq i \leq (I-1)$  множину тих значень, які може приймати функція  $\rho$  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 апіорний розподіл  $p(\rho)$  всіх можливих функцій паралакса є рівномірним, а умовний розподіл ознак об'єкта при заданому його стані визначається як:

$$\begin{aligned} p((x_l, x_r)/\rho) &= A \cdot \prod_{t \in T} \exp[-\delta(x_l(t), x_r(\phi(t)))] = \\ &= A \cdot \prod_{i=0}^I \exp[-\delta(x_l(i), x_r(i - \rho(i)))] = \\ &= A \cdot \exp[-\sum_{i=0}^I \delta(x_l(i), x_r(i - \rho(i)))] , \end{aligned}$$

де  $A$  - нормувальний множник.

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

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

де  $C$  та  $B$  - нормувальні множники.

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

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

Розглянемо кілька різних функцій втрат:

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

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

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

Почнемо з (3). Згідно з (2)

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

$$\begin{aligned} &= \arg \min_{q \in \mathcal{P}} (1 - p(q/(x_l, x_r))) = \\ &= \arg \max_{q \in \mathcal{P}} p(q/(x_l, x_r)) = \quad (7) \end{aligned}$$

$$\begin{aligned} &= \arg \max_{q \in \mathcal{P}} \exp[-\sum_{i=0}^I \delta(x_l(i), x_r(i - q(i)))] = \\ &= \arg \min_{q \in \mathcal{P}} \sum_{i=0}^I \delta(x_l(i), x_r(i - q(i))). \quad (8) \end{aligned}$$

Отже, функції (3) відповідає оцінка максимальної правдоподібності (див. (7)) функції  $\rho$ . 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 десятих 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 повернімося до розв'язку задачі (8). Введемо позначення:

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

Між цими величинами існує рекурентне співвідношення:

$$\begin{aligned} F_k(\rho(0), \rho(1), \dots, \rho(k)) &= \\ &= F_{k-1}(\rho(0), \rho(1), \dots, \rho(k-1)) + \\ &+ \delta(x_l(k), x_r(i - \rho(k))), 1 \leq k \leq I, \end{aligned} \quad (9)$$

а, отже:

$$\begin{aligned} \min_{\rho \in \mathcal{P}} F_k(\rho(0), \rho(1), \dots, \rho(k)) &= \\ &= \min_{\rho \in \mathcal{P}} F_{k-1}(\rho(0), \rho(1), \dots, \rho(k-1)) + \\ &+ \min_{\rho(k) \in N_{k-1}(\rho(k-1))} \delta(x_l(k), x_r(i - \rho(k))), 1 \leq k \leq I, \end{aligned} \quad (10)$$

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

Оскільки:

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

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

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

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

Функції втрат вигляду (4) відповідає

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

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):

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

$$\text{де } p(\rho(i)/(x_l, x_r)) = \sum_{\rho \in \mathcal{P}} p(\rho/(x_l, x_r), \rho(i))$$

- сумарна ймовірність всіх функцій паралакса  $\rho$  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:

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

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

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

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

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

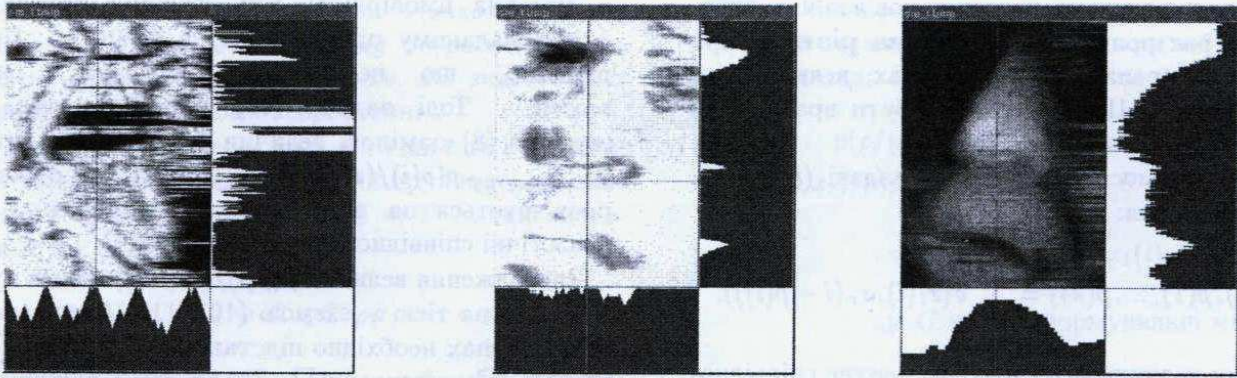


Рис. 1 Результати розв'язання задачі стереозору для трьох описаних випадків. Малюнок зліва відповідає пошуку найімовірнішого рельєфу (див.(8)), по центру - розв'язку байєсівської задачі з "точковим" штрафом (4). І, нарешті, малюнок справа отриманий з розв'язку байєсівської задачі з "квадратичним" штрафом (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  $k = i; i - 1; \dots; 0$ .

### 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. Найлівише 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  $\Phi$ ) 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 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.
- [3] Schlesinger M.I., V.Hlavač. "Deset prednasek z teorie statistického a strukturniho rozpoznani". 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 continuous 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.

## 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

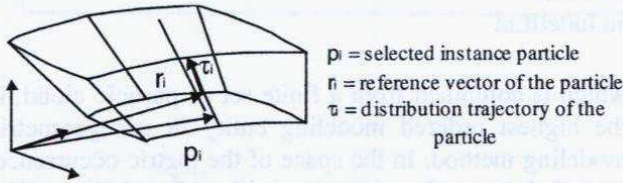


Figure 1: Variables in instance selection

instance generation.

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  $\alpha$ , and the position  $p_i$  of the particle is computed as follows:

$$p_i = r_i + \alpha \cdot \tau_i \quad (1)$$

The condition is that the instance function has to be continuous 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  $\Pi$ , we implemented a method, by which the user can interactively manipulate the particles of a region with user-defined, free-form 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.

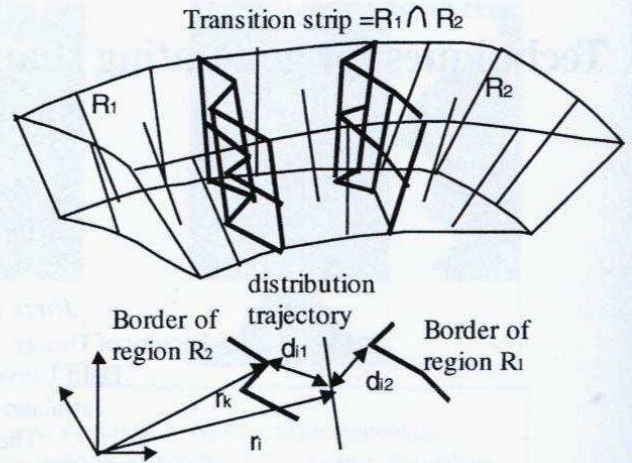


Figure 2: Specification of weights of the selection functions

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 selected to be 8 for  $C^1$  and 32 for  $C^2$ . Then for all particles  $\tau_i$  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:

$$w_{ik} = \frac{d_{ik}}{\sum_{k=1}^N d_{ik}} \quad (2)$$

Figure 2 shows the example of specification of the member values in the transition strip. Afterwards, the position  $p_i$  of the instance particles can be calculated as follows:

$$p_i = r_i + \tau_i \cdot \frac{\sum_{k=1}^N w_{ik} \alpha_k}{\sum_{k=1}^N w_{ik}} \quad (3)$$

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.

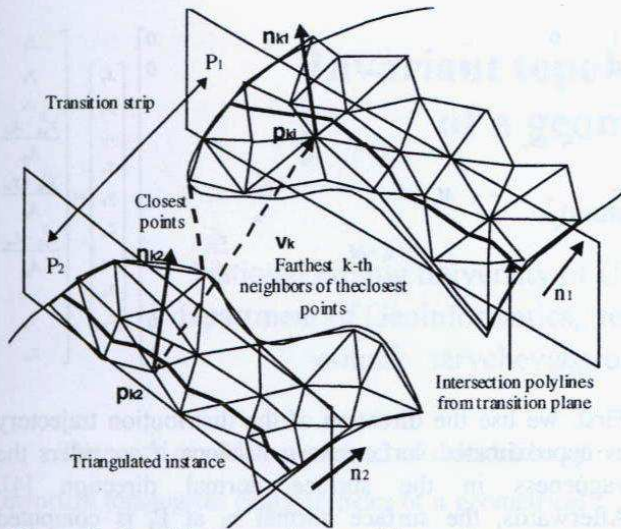


Figure 3: Specification of a transition strip

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

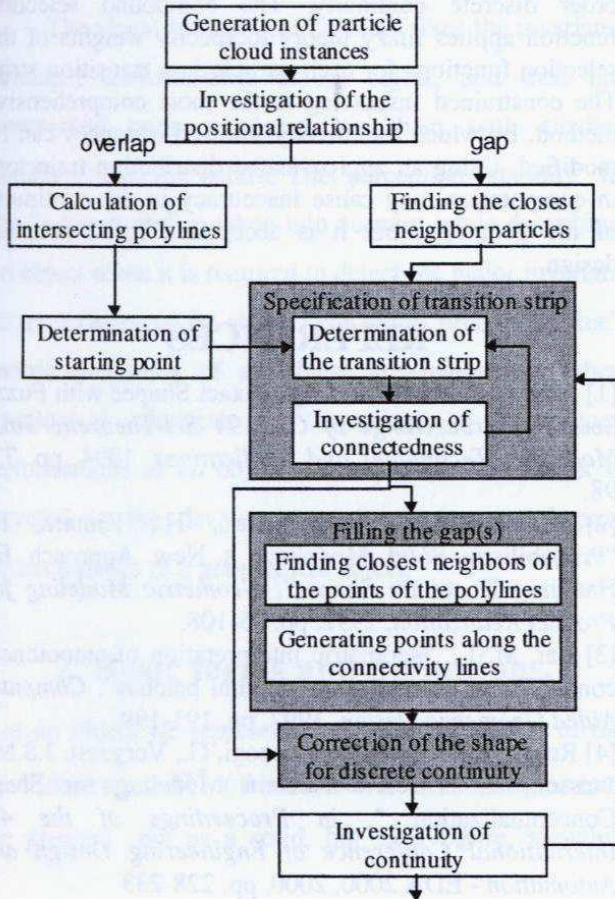


Figure 4: Constrained Instance generation

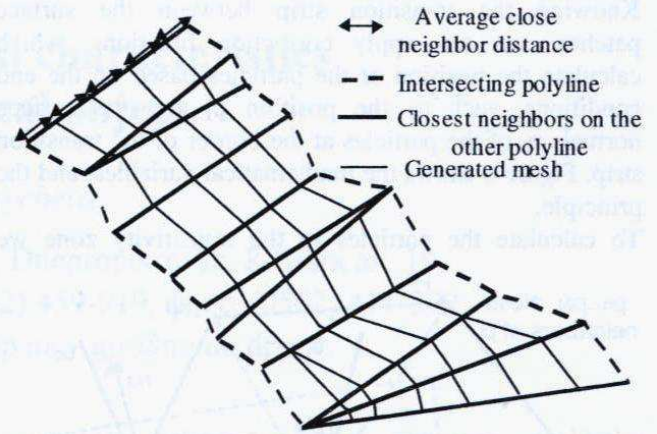


Figure 5: Particle generation in the transition strip

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

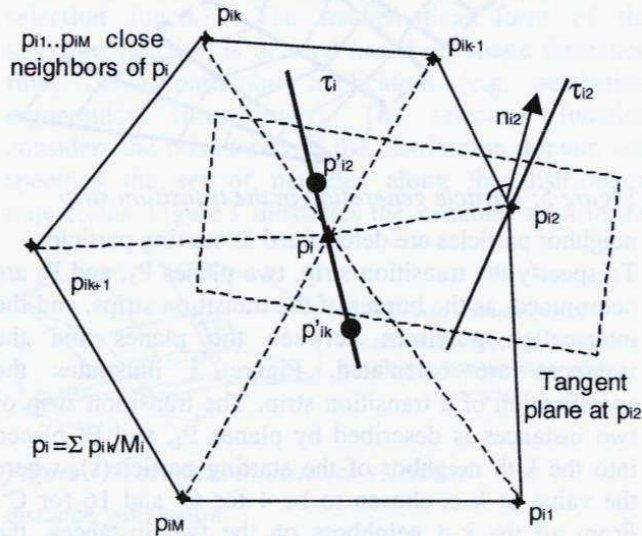


Figure 6: Principle of the correction function

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

$$\mathbf{p}'_{ik} = \frac{(\mathbf{p}_k - \mathbf{v}_i) \cdot \mathbf{n}_k}{\mathbf{n}_k \cdot \left( \mathbf{n}_k \cdot \frac{\tau_i}{|\tau_i|} \right)} \cdot \mathbf{n}_k = \frac{(\mathbf{p}_k - \mathbf{v}_i) \cdot \mathbf{n}_k}{A_{ik}} \cdot \mathbf{n}_k \quad (4)$$

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:

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

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.

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

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

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

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

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

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).

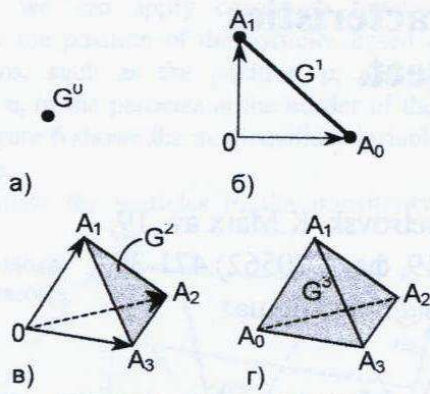


Fig. 1. Examples of rectilinear simplexes: a) – zero-dimensional, b) – one-dimensional, c) – two-dimensional, d) – three-dimensional

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

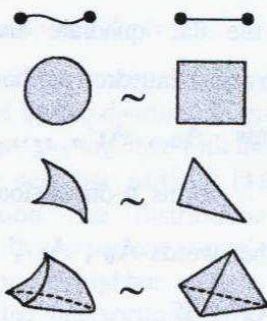


Fig.2. Examples of homeomorphic curvilinear and rectilinear simplexes.

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

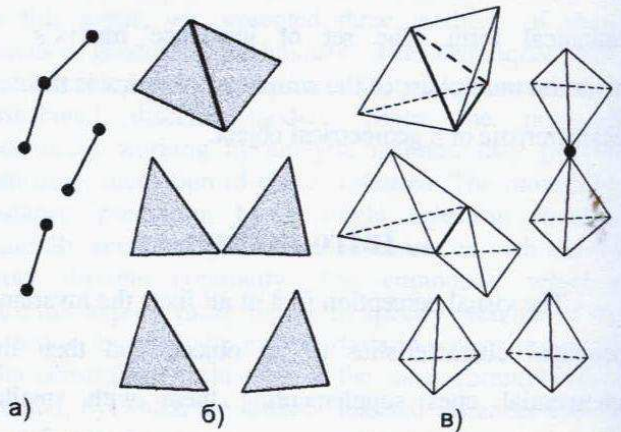


Fig.3. Variants of positional relationships of one-dimensional (a), two-dimensional (b) and three-dimensional (c) simplexes.

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.

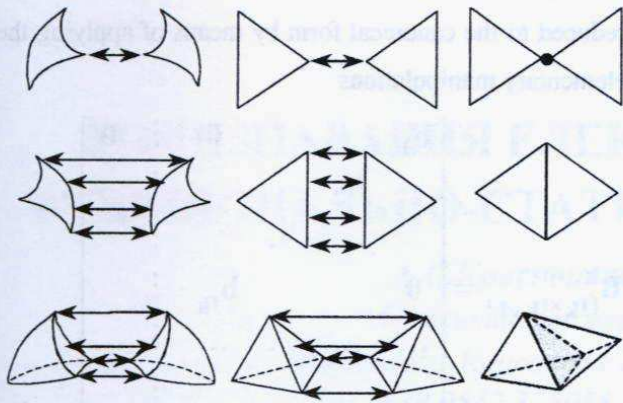


Fig. 4. Splice by zero-dimensional (a), one-dimensional (b), two-dimensional (c) facet.

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,  $1 \leq i \leq t_k$ ,  $t_k < \infty$ ,  $t_k$  is the number of  $k$ -dimensional simplexes in  $X$ ,  $1 \leq k \leq n$ ,  $n = \dim X$ .

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

The margin  $\partial G^k$  of oriented simplex  $G^k$  is the union of all its  $(k-1)$ -dimensional facets taking into consideration their orientation:

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

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

$$\partial G^2 = A_0 A_1 + A_1 A_2 - A_0 A_2 .$$

In this case

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

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

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  $a_{ij}^k$ ,  $1 \leq j \leq t_{k+1}$ ,  $1 \leq i \leq t_k$ , 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:  $B^0, B^1, \dots, B^{n-1}$ . For example for the simplex  $G^2$  (fig.5)

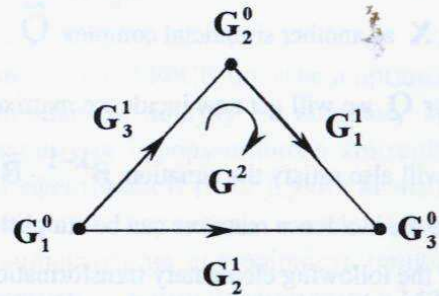


Fig. 5. Two-dimensional oriented simplex  $G^2$

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  $B^0, B^1, \dots, B^{n-1}$  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

$$\begin{aligned} \partial (\partial G_j^{k+1}) &= \sum_{i=1}^{t_k} a_{ij}^k \partial G_i^k = \\ &= \sum_{m=1}^{t_{k-1}} \sum_{i=1}^{t_k} a_{ij}^k a_{mi}^{k-1} \partial G_m^{k-1} = 0. \end{aligned}$$

Hence,

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

At the settled  $k$

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

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

For  $\tilde{Q}$  we will get new incidence matrixes  $\tilde{B}^k$ ,

but they will also satisfy the equation  $\tilde{B}^{k-1} \cdot \tilde{B}^k = 0$ .

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  $\tilde{B}^{k-1}$  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  $\tilde{B}^k$  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  $\tilde{B}^{k-1}$  is obtained from  $B^{k-1}$  by multiplying the  $i$ -th column of  $B^{k-1}$  by  $(-1)$ , and the matrix  $\tilde{B}^k$  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

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

where  $r_k = \text{rang } B^k$ ;  $b_1, b_2, \dots, b_{r_k}$  -

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

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

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

#### 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 розпізнавання 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 загально-го визначення нечіткого розбиття  $\tilde{R}^{lM1}$ , яке допус-

кає перетин класів розпізнавання, необхідне виконання наступних умов:

$$(\forall \tilde{X}_m^0 \in \tilde{R}^M) \{ \tilde{X}_m^0 \neq 0 \}$$

$$\{ (\exists \tilde{X}_k^0 \in \tilde{R}^M) (\exists \tilde{X}_l^0 \in \tilde{R}^M) \{ \tilde{X}_k^0 \neq \tilde{X}_l^0 \rightarrow \tilde{X}_k^0 \cap \tilde{X}_l^0 \neq 0 \} \}$$

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

$$\bigcup_{\tilde{X}_m^0 \in \tilde{R}^{lM1}} \tilde{X}_m^0 < X .$$

Ідея навчання за МФСВ полягає в організації ітераційної процедури пошуку на кожному кроці 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

$$d_m = \sum_{i=1}^N (x_{m,i} \oplus \lambda_{m,i}), \text{ де } x_{m,i} - i\text{-та координата вектора}$$

$x_m \in \tilde{X}_m^0$ ;  $\lambda_{m,i}$  -  $i$ -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  $d_m = d(x_m \oplus \lambda_m)$ ;

- 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:

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

де  $H_0 = -\sum_{l=1}^M p(\gamma_l) \log_2 p(\gamma_l)$  - апіорна (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  $\gamma_l$ ;  $p(\mu_m / \gamma_l)$  - апостеріорна ймовірність прийняття рішення  $\mu_m$  за умови, що прийнята гіпотеза  $\gamma_l$ .

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 апіорні за формулою Байєса отримаємо [1]:

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

де  $\alpha$ ,  $\beta$ ,  $D_1$ ,  $D_2$  - 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):

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

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

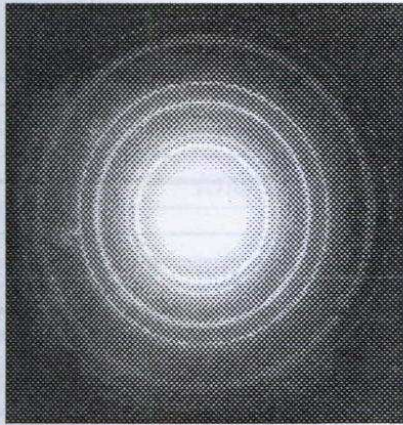
$$E_m = 1 - \frac{1}{2} \left( \frac{K_1}{K_1 + K_4} \log_2 \frac{K_1}{K_1 + K_4} + \frac{K_2}{K_2 + K_3} \log_2 \frac{K_2}{K_2 + K_3} + \frac{K_3}{K_2 + K_3} \log_2 \frac{K_3}{K_2 + K_3} + \frac{K_4}{K_1 + K_4} \log_2 \frac{K_4}{K_1 + K_4} \right). \quad (4)$$

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].

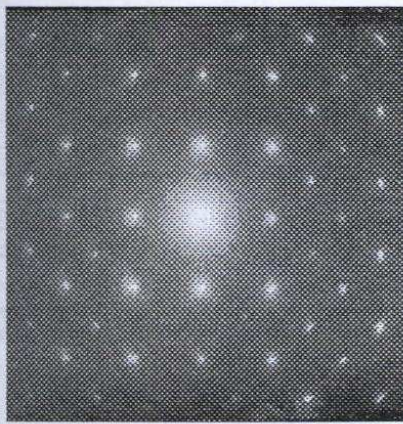
## 2 Обробка зображення

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 зчитування 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.

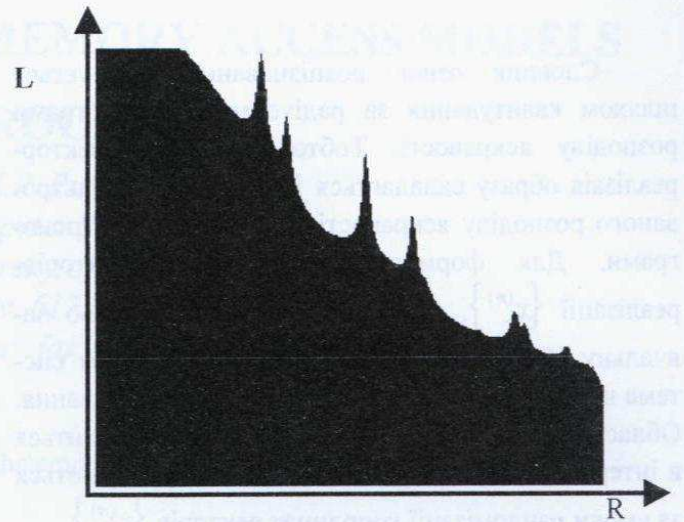
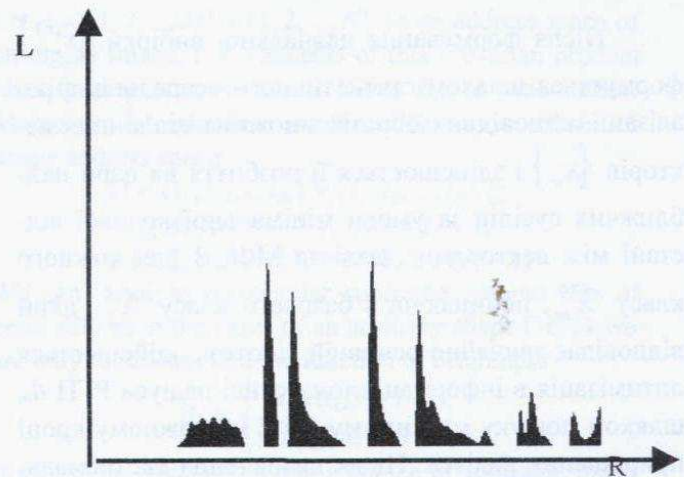
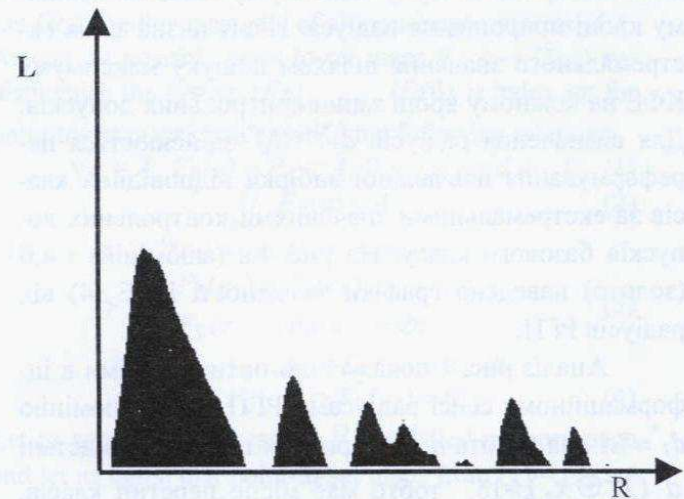


Рисунок 2. Вихідні дані другого етапу обробки зображення (алюміній).



а)



б)

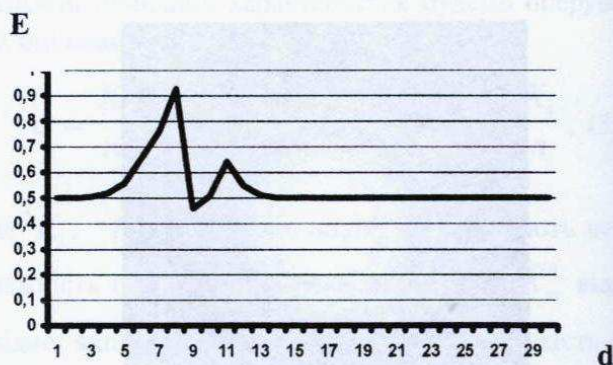
Рисунок 3 - Вихідні дані третього етапу обробки зображення: а) алюміній; б) золото.

Словник ознак розпізнавання формується шляхом квантування за радіусом електронограми розподілу яскравості. Тобто первинний вектор-реалізація образу складається із дискрет фідфільтрованого розподілу яскравості відповідної електронограми. Для формування вторинних векторів-реалізацій  $\{x_m^{(n)}\}$ , 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  $\{x_m^{(n)}\}$ .

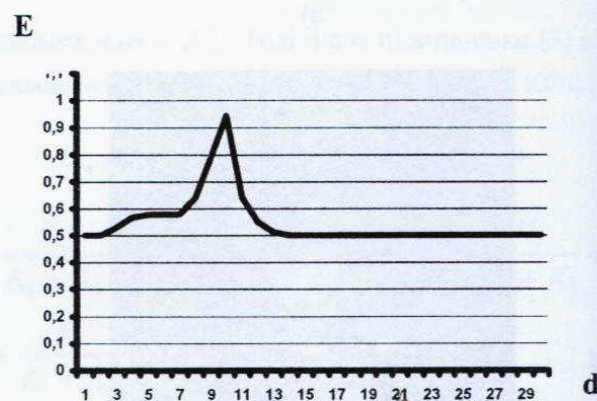
### 3 Реалізація алгоритму навчання.

Після формування навчальної вибірки  $\{x_m^{(n)}\}$  формується шляхом статистичного усереднення реалізацій 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  $X_m^0$ , 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  $d_1^*, \dots, d_M^*$  здійснюється перетворення 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 складало 10 градацій яскравості.



а)



б)

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. Краснопопсовський 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. Краснопопсовський 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  $A = \{1, 2, \dots, M\} \times \{1, 2, \dots, N\}$  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  $\Omega$  as a subset of image address space

$$\Omega = \{i_1, i_2, \dots, i_m\} \times \{j_1, j_2, \dots, j_n\} \subseteq A, \quad i_{s+1} = i_s + 1, \quad j_{r+1} = j_r + 1, \\ s \in \{1, 2, \dots, m-1\}, \quad r \in \{1, 2, \dots, n-1\}.$$

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

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

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  $F: A \rightarrow \Pi(A)$  we distinguish the system  $\{F_\beta\}_{\beta \in D(\beta)}$  ( $D(\beta)$  is index set for point-to-set maps choice) satisfying following relations

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

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

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

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

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

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

$$= \text{rem} \{ (i + (j - 1)m - 1) / \text{card } \Omega \},$$

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

$$= \text{rem} \{ (j + (i - 1)n - 1) / \text{card } \Omega \},$$

where «rem( $\circ$ )» 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

$$\forall a_{ij}, a_{pq} \in \Omega \subseteq A \Rightarrow \begin{cases} F_{\beta'}(a_{ij}) = P_{i'} \\ F_{\beta''}(a_{ij}) = P_{i''} \end{cases} \Rightarrow i' \neq i''$$

where

$$i' = \text{rem} \{ (i + (j - 1)m - 1) \},$$

$$i'' = \text{rem} \{ (p + (q - 1)m - 1) \}.$$

Let us assume the contrary, namely,  $i' = i''$ . Then there is a divisibility of the value  $|i + (j - 1)m + (p - (q - 1)m)|$  by  $\text{card } \Omega$  or, that is the same, the following equality holds

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

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

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

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

$$|(i - p) + m(j - q)| \leq |i - p| + m|j - q| \leq$$

$$\leq m - 1 + m(n - 1) = mn - 1.$$

At the same time  $mn - 1 < \text{card } \Omega$ , consequently,  $r = 0$  and therefore  $|i - p| = m|j - q|$ . From inequalities (6) it follows  $|m(j - q)| \leq m - 1$  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)  $\text{card} \{F_{\beta}(a)\}_{a \in A} \geq \text{card } \Omega$ , hence we shall say: point-to-set map  $F : A \rightarrow \Pi(A)$  induces an optimal page scan if  $\text{card} \{F_{\beta}(a)\}_{a \in A} = \text{card } \Omega = mn$ , 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

$$F_{\beta} \Delta \mathfrak{Z} : A \rightarrow \Pi(A) \times \mathfrak{R}^{\tau}$$

where  $F_{\beta}$  is page scan,  $\mathfrak{Z}$  is some map which restriction

$$\mathfrak{Z}|_{F_{\beta}^{-1}(a) \in A} : \{F_{\beta}(a)\}_{a \in A} \rightarrow \mathfrak{R}^{\tau}$$

is injection, i.e.

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

Introducing of  $\tau$ -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  $P_{n+1}, P_{n+2}, \dots, P_{2n}, \dots, P_{(m-1)n+1}, P_{(m-1)n+2}, \dots, P_{mn}$  are already fixed and it is necessary to fulfil pixel readdressing to pages  $P_1, P_2, \dots, P_n$ . 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

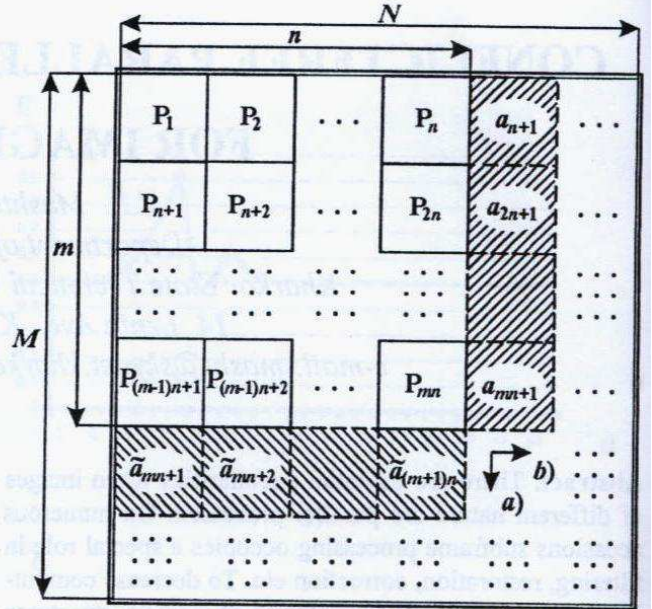


Fig. 1. Page scans number evaluation

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

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

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

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

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  $\beta$ , selecting page scans  $\{F_{\beta}(a)\}_{a \in A}$ , as vector-function  $f(i, j) \in \mathfrak{R}^{\tau}, (i, j) \in A$ . 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

$$f|_{\{1, 2, \dots, m\} \times \{1, 2, \dots, n\}} : \{1, 2, \dots, M\} \times \{1, 2, \dots, N\} \rightarrow \mathfrak{R}^{\tau}$$

is injection, i.e.

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

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

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

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  $1 \leq i' \leq |p - i|, 1 \leq j' \leq |q - j|$ . Taking into account that  $a_{ij}, a_{pq} \in \Omega$  we have

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

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)

$$f(i, j) = n \operatorname{rem} \left\{ \frac{i}{m} \right\} + \operatorname{rem} \left\{ \frac{j-1}{n} \right\}. \quad (7)$$

We first note that condition i) is obvious, indeed

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

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

$$\begin{aligned} n \left| \operatorname{rem} \left( \frac{i_1}{m} \right) - \operatorname{rem} \left( \frac{i_2}{m} \right) \right| &\geq n, \\ \left| \operatorname{rem} \left( \frac{j_1-1}{n} \right) - \operatorname{rem} \left( \frac{j_2-1}{n} \right) \right| &\leq n-1 \end{aligned}$$

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

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

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  $f(A) \leq mn - 1$ . Further according to property i) the equality

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

holds and since

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

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  $F_\beta$  is optimal if

$$\begin{aligned} \forall \Omega \subset A, \forall F_\beta(a_{ij}) \operatorname{card} \{ \Omega \cap F_\beta(a_{ij}) \} \leq \\ \leq \operatorname{Int} \{ d/\delta + (d-1)/\delta \}, \end{aligned} \quad (8)$$

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

Similarly, a map  $\varphi : A \rightarrow \mathfrak{R}^r$ , satisfying i), is an address quasiindicator if

$$\begin{aligned} \forall \xi \in \varphi(\{1, 2, \dots, m\} \times \{1, 2, \dots, n\}) \\ \operatorname{card} \varphi^{-1} \{ \xi \} \leq \operatorname{Int} \{ d/\delta + (d-1)/\delta \}. \end{aligned}$$

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

$$F_\beta \Delta \Phi_\beta : A \rightarrow \mathfrak{R}^r \times \{1, 2, \dots, \operatorname{Int} \{ d/\delta + (d-1)/\delta \} \}$$

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,  $N = 2^{s'}$ ,  $M = 2^{s''}$ ,  $n = 2^{s'}$ ,  $m = 2^{s''}$  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  $\{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$  and connect with it the set of permutation of order  $m$ :

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

where meaning of  $s_h$ ,  $h \in \{1, 2, \dots, m\}$  lies in a number of memory pages. Moreover, let us consider a cyclic permutation  $\pi_c = \begin{pmatrix} 1, & 2, & \dots, & m-1, & m \\ 2, & 3, & \dots, & m, & 1 \end{pmatrix}$ . It is clear that

permutation  $\pi = \pi_c^\alpha \pi_j$  allows to get new columns from all other subframes. Here the degree  $\alpha$  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,  $K \leq n + m - 1$ .

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

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

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

$$\text{where } \overline{\text{rem}}\{p/r\} = \begin{cases} \text{rem}\{p/r\}, & \text{if } \exists \omega \in \mathbb{N} : p \neq \omega r, \\ r & \text{if } \exists \omega \in \mathbb{N} : p = \omega r, \end{cases}$$

« $\overline{\text{rem}}$ » denotes rounding up integer values;  $\mathbb{N}$  is set of positive integer numbers. It should be noted that  $\alpha_0([i/n] - 1) = \alpha$  is value from the basic permutation  $\pi$  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  $\lambda \in \mathbb{R}^2$  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  $\times_{u \in U} F_\beta\{\Omega_{\lambda_u}\}$ , where « $\times$ » denotes Cartesian product of page scans

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

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  $\times_{u \in U} \times_{v \in V} F_\beta(\Psi(\Omega_{\lambda_u}^v))$ , where map

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

is optimal page scan for subframe chosen viz.

$$\Omega_{\lambda_u} = \bigcup_{v \in V} \Omega_{\lambda_u}^v, \quad V = \{1, 2, \dots, \text{Int}(d/\delta) + (d-1)/\delta\},$$

$$\forall v_1, v_2 \in V \quad \Omega_{\lambda_u}^{v_1} \cap \Omega_{\lambda_u}^{v_2} = \emptyset,$$

$$\forall v \in V \quad \text{card } F_\beta\{\Omega_{\lambda_u}^v\} = \text{card } \Omega_{\lambda_u}^v.$$

Parallel subframe processing we can indicate as a set

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

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

$$\forall \gamma \in \{1, 2, \dots, \Gamma\} \quad \Phi_\gamma : \bigcup_{\lambda_u} (\Omega_{\lambda_u}) \rightarrow \Pi(A),$$

$$\Phi_\gamma\left(\bigcup_{\lambda_u} \Omega_{\lambda_u}\right) = \bigcup_{\lambda_u} \Phi_\gamma(\Omega_{\lambda_u}), \quad (9)$$

$$\forall \gamma \in \{1, 2, \dots, \Gamma\} \quad \text{card } \Phi_\gamma\{\Omega_{\lambda_u}\}_{u \in U} =$$

$$= \text{card } \{\Omega_{\lambda_u}\}_{u \in U}, \quad \forall \gamma, \forall \lambda \in \Lambda \quad \Phi_\gamma(\Omega_\lambda) = P_q \in \Omega_q, \quad (10)$$

$$\forall \gamma \in \{1, 2, \dots, \Gamma\}, \forall u_1, u_2 \in U : u_1 \neq u_2 \Rightarrow$$

$$\Rightarrow F_\beta(\Phi_\gamma(\Omega_{\lambda_{u_1}})) \neq F_\beta(\Phi_\gamma(\Omega_{\lambda_{u_2}})), \quad (11)$$

$$\forall \lambda \in \Lambda, \forall \gamma_1, \gamma_2 \in \{1, 2, \dots, \Gamma\} : \gamma_1 \neq \gamma_2 \\ \Rightarrow \Phi_{\gamma_1}(\Omega_\lambda) \neq \Phi_{\gamma_2}(\Omega_\lambda). \quad (12)$$

Now explain properties (9) – (12). Essence of relationship (9) consists in additiveness which, as a consequence, provides quasioptimal parallel-subframe processing. Formulae (10) describe structure of maps  $\Phi_\gamma$ , 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  $\Phi_\gamma$  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 structures of data with purpose of subset parallel processing.

## REFERENCES

1. Kuck D. J. Iliac IV software and application programming // IEEE Transactions on Computers. – 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., Gösler M. On conflictfree parallel memory access // Computer and Artificial Intelligence. – vol. 2, 1983. – P. 359-401.
5. Gösler 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 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 quantum 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 quantum 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 quantum 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  $\alpha$ -undefinition, that stipulates 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 outputting 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 linguistic information in the form of variety of levels on complication **probable algorithm quantum 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 inbuilt procedures for realization of exact countings of intermediate and concluding decisions' probables on the base of equivalent transformations of logical functions statistically depended v-quantums chains to independent by ortogonizing of corresponding logical functions.

## 2. ORGANIZATION OF THE TASK

It is necessary to decide two basic tasks of formally presenting different-levels vk-knowledges ( $A_v$ -task) and algorithm synthesis of the rule of

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  $[0, 1]$ ;

$\Pi_v$  - 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 algorithmization.

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 reliability's degree from  $\tilde{V}$  on the base of using rules  $\Pi_v$  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

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

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

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 outputting finding decisions with some appraisal of their probable from other sources, and also **quantity and contents 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), outputting 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

$$K_v = \{vk_0 a, vk_1 y, vk_1 b\}. \quad (3)$$

Here **terminal choosing v-quantum of knowledges of the nil level  $vk_0 a$  with semantic:** «from  $\ell$ -measured totality observed probable meanings of signs  $\alpha_1, \alpha_2, \dots, \alpha_\ell$  to choose meaning  $\alpha_k$  with reliable's index  $p(\alpha_k) \in [0, 1]$ »

has such form:

$$vk_0 a = [V_k^{(\ell)}(\alpha_1 | p(\alpha_1), \dots, \alpha_k | p(\alpha_k), \dots, \alpha_\ell | p(\alpha_\ell)) = \alpha_k | p(\alpha_k)], \quad (4)$$

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

**Terminal vector v-quantum of knowledges of the first level  $vk_1 y$  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

$$vk_1 y = [d_1^{x_1} d_2^{x_2} \dots d_n^{x_n}] = [\alpha_1^1 | p(\alpha_1^1), \dots, \alpha_1^{r_1} | p(\alpha_1^{r_1}); \alpha_2^1 | p(\alpha_2^1), \dots, \alpha_2^{r_2} | p(\alpha_2^{r_2}); \dots; \alpha_n^1 | p(\alpha_n^1), \dots, \alpha_n^{r_n} | p(\alpha_n^{r_n})], \quad (5)$$

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

**Terminal characteristic v-quantum of the first level  $vk_1 b$  with the name  $b$  has a form**

$$vk_1 b = [\chi_{\tilde{Y}_j}(\alpha_k^j | p(\alpha_k^j))] = \begin{cases} 1, & \text{if } \alpha_k^j | p(\alpha_k^j) \in \tilde{Y}_j, \\ 0, & \text{if } \alpha_k^j | p(\alpha_k^j) \notin \tilde{Y}_j, \end{cases} \quad (6)$$

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  $\alpha_k^j$  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,  $CON(\bullet)$ -operator and  $CON[\bullet]$ -operator in frameworks  $A_v$ -task (1), are called different-levels vk-knowledges. Class vk-knowledges  $M_v$ , made on definition 1, let's call **standard class PAQK-models** of presenting of knowledges about object under recognition (OR).

For example, by expression

$$\begin{aligned}
 vk_1c1 = & \overbrace{[0,1|p(\alpha_2^1),1|p(\alpha_3^1),\emptyset:1|p(\alpha_1^2),0,\emptyset:\emptyset|\emptyset:\alpha_1^{3u}|p(\alpha_1^{3u}),\emptyset:}^{x_1} \overbrace{\quad\quad\quad}^{x_2} \overbrace{\quad\quad\quad}^{x_3=x_u} \\
 = & \underbrace{p(\rightarrow c1);A(c1)p(c1)}_{\text{output domain}} \quad (7)
 \end{aligned}$$

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  $\alpha_2^1$  sign  $x_1$  with IR  $p(\alpha_2^1)$  **OR** third knowledges  $\alpha_3^1$  with OR  $p(\alpha_3^1)$  **AND** first meaning  $\alpha_1^2$  of sign  $x_2$  **THEN** OR has one component purposed sign  $x_u$  with IR  $p(\alpha_u^1)$ ; with it output IR  $p(c1)$  v-quantum is counted by algorithm  $A(c1)$ , using set probability  $p(\rightarrow c1)$  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  $A(c1)$ .

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  $N_1, N_2, \dots, N_k \subset E$  of summits, and also next **properties**:

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

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

3) all knots  $X_i \in N_k \subset E$  with  $\Gamma N_k = \emptyset$  are **purposing** (output) knots-conclusions  $C_s, (1 \leq s \leq S)$  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  $X_i \in N_1 \subset E_{vk}, \Gamma^{-1}N_1 = \emptyset$  correspond to **input vk-knowledges-sendings** with names  $e_i, (1 \leq i \leq n'_n)$  relatively vk-knowledges-consequences with names  $C_j, (1 \leq j \leq m'_c)$  with given indexis of reliability;

3) all  $X_i \in N_k \subset E_{vk}, \Gamma N_k = \emptyset$  are **purpose vk-knowledges-conclusions** with names  $C_s, (1 \leq s \leq S)$  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 outputting 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 *инпредставляется 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  $e := \langle \text{object } \omega \in \Omega \text{ has sign } x \rangle$  is made with probability  $p(e)$ , and event  $c := \langle \text{if } \omega \text{ ohas sign } x, \text{ that } \omega \text{ corresponds to category } A \rangle$  is set with conditional probability  $p(e \rightarrow c)$ , 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

$$p(z) = P(e \wedge c) = p(e) \cdot p(e \rightarrow c). \quad (8)$$

If now for showed events  $e$  and  $c$  to correspond logical variables  $\alpha$  and  $\beta$  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  $\alpha$  and  $\beta$  with giving probabilities of appearing their true meanings.

For more complicated event  $z1 := \langle \text{if object } \omega \text{ has sign } x_1 \text{ or } x_2, \text{ or } \dots \text{ or sign } x_k, \text{ then } \omega \text{ belongs to class } A \rangle$  probability  $p(z1)$  can be counted on logical function  $\alpha_1 \cdot \beta_1 \vee \alpha_2 \cdot \beta_2 \vee \dots \vee \alpha_k \cdot \beta_k$ , transformed to equivalent to it ODNF

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

At the result of deducting outputting **probable identified decisions** relatively OR  $\omega$  with given sending vk-knowledges  $vk_1 Y_{B\omega}$  in  $B_v$ -task with help of algorithm AJL( $B_v$ ) at the output v-QNDO purpose vk-knowledges are formed  $vk_s R_{B\omega}$ , ( $s = 0, 1$ ) as one or some v-quantums with counted resulted IR  $p(R_{B\omega})$ .

Semantic  $vk_s R_{B\omega}$  shows at the **belonging identified OR**  $\omega$  to definition category or class  $K$  from  $\Omega$  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  $\alpha$ -undefinition using IND-operator [3-4] as

$$BPQK(B_v) = \\ = SETK(B_v) \xrightarrow[\text{ALOBUCHALAKVA, ALOPT}]{\text{IND}} BPQK \quad (9)$$

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

$$vk_s R_{B\omega} = \\ = DED(BPQK(B_v), vk_1 Y_{C\omega}; AL(B_v), ALUPR, vk_s R_{B\omega}) \\ = BPQK(B_v) \xrightarrow[\text{vk}_1 Y_{B\omega}; AL(B_v), ALUPR]{\text{DED}} vk_s R_{B\omega}. \quad (10)$$

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.



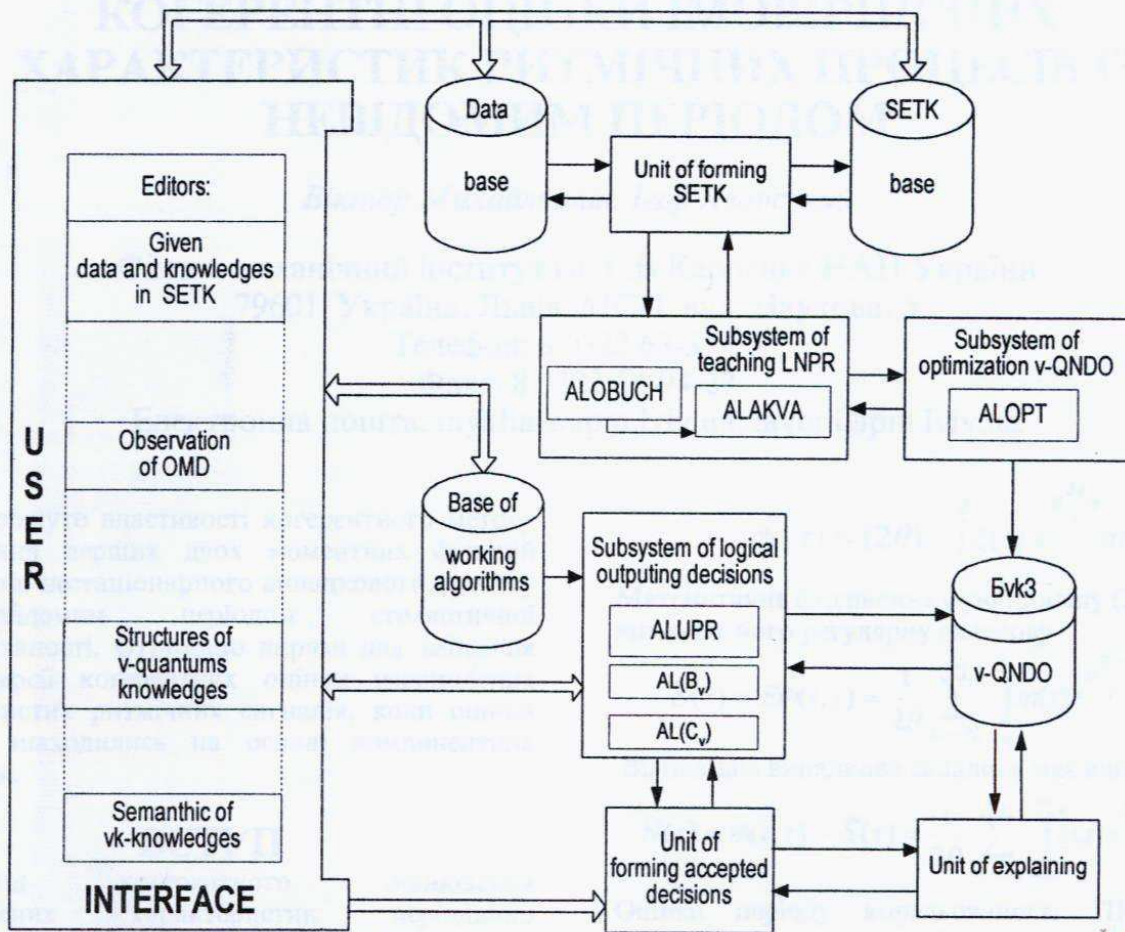


Fig. 1. Total architecture of IPK "V-KVANT"

⇨ - management; → - functional connections

## 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 advantageges 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 алгоритмічних квантів знань // Праці Четвертої Всеукраїнської Міжнародної конференції "Оброблення сигналів і зображень та розпізнавання образів" (УкрОбраз'98). – Київ.: Кіб. центр НАН України.–1998.– С. 17-19.

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

