

ВИЯВЛЕННЯ КЛЮЧОВИХ ТОЧОК НА ЗОБРАЖЕННІ ОБЛИЧЧЯ ЛЮДИНИ

Муригін К.В.

Донецький державний інститут штучного інтелекту
83050, м. Донецьк, проспект Богдана Хмельницького, 84

Тел.: (0622) 926082, факс: (0622) 926082

E-mail: kir@iai.donetsk.ua

У даній роботі описуються кілька схожих підходів до задачі локалізації ключових точок обличчя на зображенні. Усі методи засновані на кореляційному пошуку в просторі векторів ознак, отриманих на основі згорток зображення обличчя людини з набором функцій Габора різних масштабів і кутів повороту. Обговорюються переваги і недоліки описаних методів, приводяться результати їх порівняння.

ВСТУП

Для рішення задачі розпізнавання людини по зображенню обличчя в даний час використовується велика кількість різноманітних методів. Однак більшість з них зіштовхуються з однією з основних труднощів – проблемою відповідності порівнюваних властивостей. Наприклад, невеликі розбіжності в масштабах та кутах зйомки можуть істотно вплинути на результат роботи алгоритмів розпізнавання, що працюють із зображенням обличчя в цілому. Одним з напрямків рішення цієї проблеми є застосування методів, що працюють з окремими деталями обличчя. У рамках таких методів проблема відповідності порівнюваних властивостей трансформується в проблему правильного виявлення і локалізації відповідних деталей зображення обличчя. Подібно тому, як у фізиці не можна порівнювати кілограми з метрами, так і при розпізнаванні не можна порівнювати, наприклад, брова й око людини.

Таким чином, задача визначення відповідності або локалізації точок обличчя, що використовуються при розпізнаванні (ключових точок), представляється дуже важливою, і виступає необхідним етапом на шляху до побудови системи ідентифікації людини по фотографії.

1. ОПИС МЕТОДУ

Серед методів заснованих на локальних властивостях останнім часом велике поширення одержав метод, що використовує як ознаки згортки вихідного зображення з набором функцій Габора [1,2]. Популярність цього підходу багато в чому

порозумівається даними біології про те, що схожу на габоровську форму має рецепторне поле клітини частини головного мозку [3,4], що відповідає за зір.

Як ознаки, що визначають властивості зображення обличчя, метод використовує вектора, що складаються з комплексних елементів, отриманих за допомогою згортки вхідного зображення обличчя з функціями Габора різних масштабів і кутів повороту [1,2]. У даній роботі використовувалися сорок функцій Габора п'яти різних масштабів і восьми орієнтацій. Таким чином, кожна точка зображення перетворювалася в описаний вектор ознак, що у деякій формі містить інформацію про розподіл яскравості навколо даної точки.

Ключові точки розташовувалися на зображенні обличчя у виді сітки (графа обличчя) прямокутної структури (4×7 точок) (див. рис. 1).

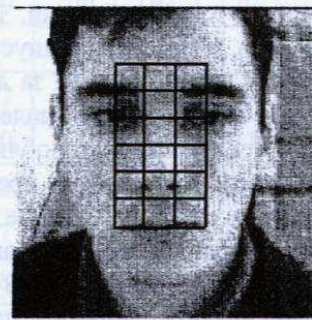


Рис. 1 - Граф обличчя, що використовувався в даній роботі. Вузли графу відповідають обраним ключовим точкам зображення обличчя.

Для зображень облич різних людей вектора ознак у ключових точках будуть відрізнятися, що пов'язано з індивідуальними особливостями обличчя, але, в той же час, вони будуть схожі, тому що описують ту ж саму ключову точку.

Для порівняння векторів ознак використовувалася міра подібності у виді косинуса кута між ними:

$$s(\vec{v}_i, \vec{v}_j) = \frac{\sum_n v_{ni} v_{nj}}{|\vec{v}_i| \cdot |\vec{v}_j|}$$

Така міра дозволяє зробити порівняння інваріантним стосовно лінійних перетворень вхідних зображень,

тому що елементи векторів ознак не залежать від постійного зсуву зображень по шкалі яскравості, а сама міра подібності інваріантна стосовно постійних множників.

2. ПОСТАНОВКА ЗАДАЧІ І РОЗПОДІЛ НА ЕТАПИ

Метою проведених досліджень була побудова алгоритму виявлення ключових точок на зображенні обличчя, незалежно від індивідуальних особливостей.

Як вихідні дані для досліджень використовувалася база даних векторів ознак для кожної ключової точки обличчя, розрахована на основі бази даних, що містить зображення облич 200 різних людей. Усі зображення з бази даних отримані при нормальних умовах освітлення і попередньо приведені до одного масштабу (за масштабний коефіцієнт була обрана відстань між зіницями, еталонна відстань дорівнювала 34 пікселя). На кожному зображенні обличчя були вручну відзначені ключові точки, що відповідають особливим деталям обличчя (зіниці, брови, кінчик носу, краї губ), по яких автоматично будувався граф обличчя, у вузлах якого витягалися вектора ознак. У такий спосіб були враховані розбіжності в пропорціях зображень облич різних людей.

Як відзначалося в роботах [1,2], амплітуда комплексних елементів векторів ознак є повільно мінливою функцією, що дозволяє розділити задачу локалізації ключових точок на два етапи. З огляду на зазначену властивість, можна припустити, що точність розташування графа обличчя за допомогою векторів амплітуд буде слабо залежати від невеликих відмінностей у масштабі обличчя, його пропорціях, невеликого збільшення просторового кроку при кореляційному пошуку. Усе це дозволить швидко знаходити приблизне розташування графа обличчя, побудованого з урахуванням середніх пропорцій облич, що звужить діапазони пошуку для більш точного алгоритму локалізації. Як показують дані, отримані на основі використовуваної бази даних зображень облич, середній граф обличчя має наступні параметри:

$$x_{ij} \approx 11,3 \times i \text{ и } y_{ij} \approx 9 \times j, \text{ де } i = \{0, \dots, 3\}, \text{ а } j = \{0, \dots, 6\}.$$

Таким чином, другий етап локалізації ключових точок на зображенні обличчя зводиться до уточнення отриманих приблизних координат.

3. УТОЧНЕННЯ ПОЛОЖЕННЯ КЛЮЧОВИХ ТОЧОК

Для більш точного визначення координат ключових точок, амплітуд комплексних векторів ознак недостатньо. Виникає необхідність

використання фазової інформації, що має характеристичну частоту, яка відповідає частоті габоровського ядра.

Як перший і найпростіший метод був досліджений метод, що використовує середні по базі даних вектора ознак. На рисунку 2 приведена функція щільності розподілу результатів порівняння (ФЦРРП) середнього вектора ознак і векторів з бази даних для точки ока.

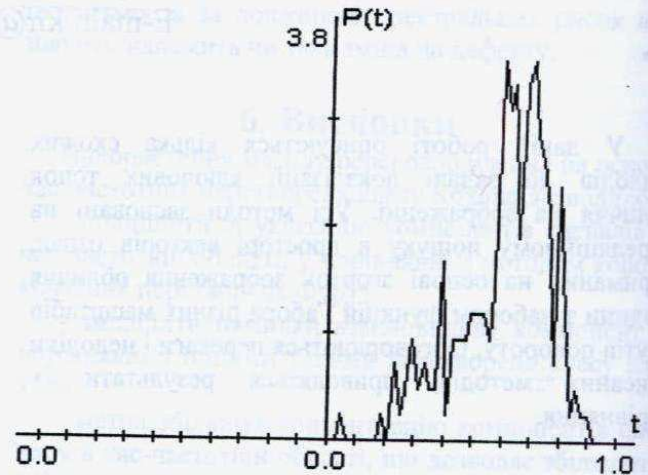


Рис. 2 - ФЦРРП середнього вектора ознак з векторами з БД.

Недоліки цього методу проявилися при першому ж випробуванні. Виявилося, що метод сильно орієнтований на середні пропорції облич. Наприклад, для точки ока пропорції обличчя визначаються взаємним розташуванням центра ока і брови, що проявилось в ефективному визначенні ключової точки для окремого класу облич і майже цілком помилкові результати для інших класів.

Для зменшення впливу на локалізацію локальних пропорцій обличчя були введені вагові коефіцієнти для елементів векторів ознак. Коефіцієнти були знайдені з умови мінімізації середньоквадратичної відстані між зображеннями одного класу, при накладенні обмеження, що виключає тривіальне не змістовне рішення. У монографії [5] відзначається, що така метрика (див. систему (1)) виражає властивість підмножини зображень, у якій вони найбільш подібні.

$$\begin{cases} d(a, b) = \sqrt{\sum_i w_i (a_i - b_i)^2}, \\ \sum_i w_i = 1. \end{cases} \quad (1)$$

Мінімізація дисперсії дає наступні значення для вагових коефіцієнтів:

$$w_i = \frac{1}{\sigma_i^2 \sum_n \frac{1}{\sigma_n^2}}$$

На основі отриманої метрики були зроблені розрахунки ФЦРРП середнього вектора ознак і

векторів з бази даних. Як приклад на рисунку 3 приведена ФЦРРП для точки ока. Отриманий розподіл має менше середнє відхилення від середнього по базі даних вектора ознак, ніж для метрики, що не використовує вагові коефіцієнти. Однак, експерименти також показали сильну залежність точності виявлення від локальних пропорцій обличчя. Крім того, уведення вагових коефіцієнтів знижує кількість аналізованої при кореляційному порівнянні інформації, що підвищує імовірність помилкового виявлення шуму, а також знижує точність локалізації ключових точок.

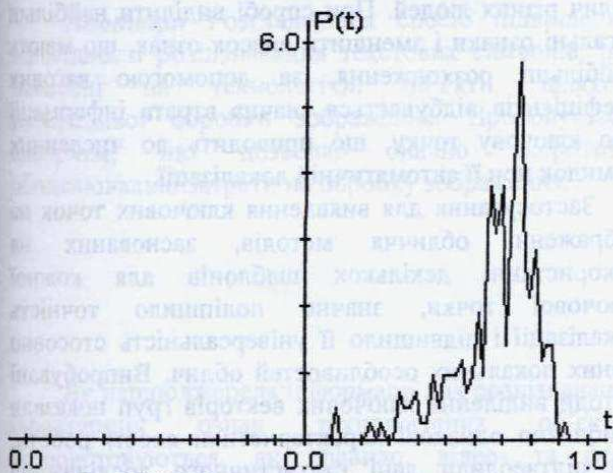


Рис. 3 - ФЦРРП середнього вектора з векторами з БД з урахуванням вагових коефіцієнтів.

Аналіз отриманих результатів показав необхідність уведення декількох шаблонів для кожної ключової точки обличчя. Це привело до задачі розділення всього класу локальних характеристик обличчя людини на групи, що мають найбільш схожі риси. Такий підхід дозволить більш точно знаходити ключові точки на зображенні обличчя, а також за рахунок класифікації локальних рис приведе до спрощення представлення обличчя в базі даних і, у кінцевому рахунку, до скорочення її обсягу. Останнє припускає подальше дослідження можливостей використання розділення на групи для вирішення задачі розпізнавання.

Для розділення векторів ознак на групи використовувалася умова обмеження відмінності між усіма векторами з групи деяким заданим числом ϵ . Після розподілу векторів ознак по групах, усі групи упорядковувалися по кількості членів і вибиралися N груп з найбільшою кількістю членів. Параметром оцінки ефективності розбивки W була кількість векторів властивостей, що потрапили в обрані N найбільш численних груп. У наших експериментах ми задали наступними значеннями параметрів: $N=10$, $W=150$ (з 200). Потім для $N=10$ була обчислена таблиця 1, у якій відбита залежність числа облич, що попадають у N груп, від числа ϵ . Значення ϵ наведені в першому рядку, індекси ключових точок, відповідно до графа рис. 1, – у першому стовпці. Напрямок нумерації вузлів графа: ліворуч - праворуч і зверху - вниз. Після цього для

$W=150$ по таблиці 1 були обчислені значення ϵ для кожної ключової точки.

Таблиця 1 Залежність W від ϵ для кожної ключової точки обличчя.

№	0.1	0.2	0.3	0.4	0.5	0.6
0	200	191	178	148	120	93
1	199	191	167	144	113	79
2	200	191	168	138	111	77
3	200	195	185	160	131	92
4	199	190	169	126	98	55
5	200	191	163	114	83	48
6	200	191	158	129	96	56
7	200	193	182	144	102	61
8	190	165	130	96	61	43
9	200	200	192	158	120	75
10	200	199	191	165	129	78
11	199	171	132	97	68	43
12	190	157	125	98	62	38
13	200	200	197	178	129	96
14	200	200	200	187	152	108
15	199	174	138	107	70	52
16	189	177	134	105	69	48
17	200	200	197	174	140	87
18	200	200	197	179	134	100
19	199	184	152	124	84	50
20	187	159	125	106	69	42
21	200	200	193	177	149	117
22	200	200	198	185	158	120
23	197	170	137	96	72	49
24	148	114	83	61	47	33
25	190	171	149	120	80	52
26	190	177	149	121	82	63
27	139	121	92	62	50	35

Після поділу векторів ознак на групи були проведені два дослідження з визначення ключового вектора групи. У першому випадку ключовий вектор обчислювався як середній по групі. В другому – як середній між максимальним і мінімальним значенням кожної координати векторів групи. Результати обчислення ФЦРРП векторів ознак з бази даних і максимально схожих на них з 10 ключових векторів груп для обох випадків приведені відповідно на рисунках 4 і 5.

Обидва методи виділення ключових векторів груп показали добрі результати при тестуванні алгоритму виявлення ключових точок обличчя на зображеннях різних людей.

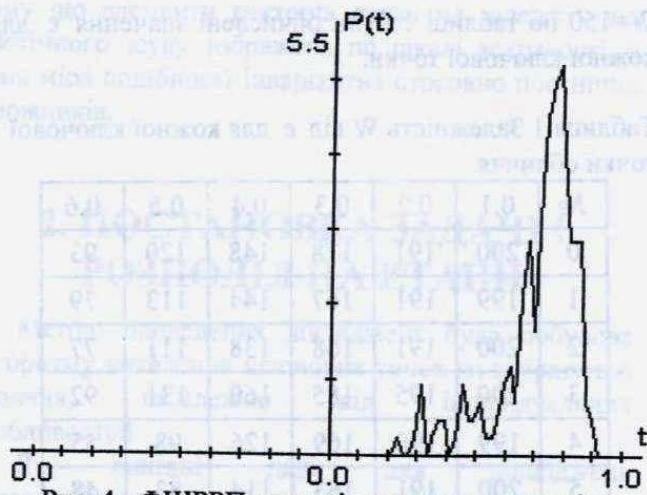


Рис. 4 - ФЩРРП середніх по групах векторів з векторами з БД.

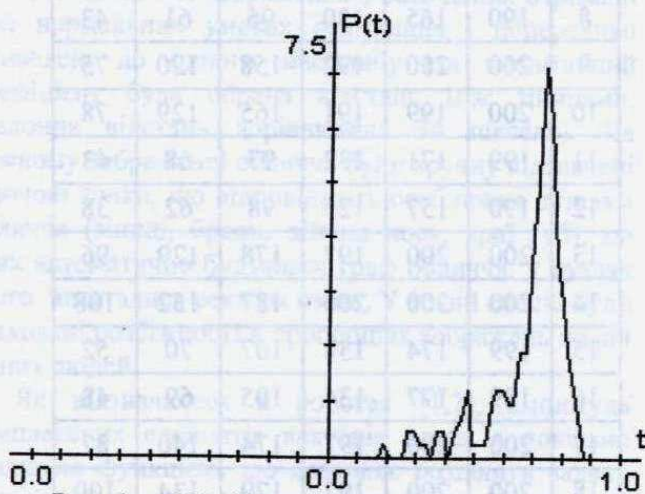


Рис. 5 - ФЩРРП медіан по групах векторів з векторами з БД.

Для порівняння усіх випробуваних методів були обчислені наступні статистичні дані: DM- середня відмінність векторів ознак з бази даних і шаблонних; $P(s > por1)$ - імовірність того, що величина порівняння вектора з бази і шаблону буде більше $por1$; $P(s > por2)$ - імовірність того, що величина порівняння вектора з бази і шаблону буде більше $por2$. Порівняльна характеристика для точки ока приведена в таблиці 2.

Таблиця 2 Порівняльна характеристика алгоритмів локалізації ключових точок на зображенні обличчя. Наведені дані відповідають точці ока.

Характеристика	DM	$P(s > 0.7)$	$P(s > 0.8)$
Метод середнього	0.59	0.35	0.05
Середній з вагами	0.66	0.57	0.15
Метод медіан по групах	0.67	0.69	0.08
Метод середнього по групах	0.70	0.70	0.21

ВИСНОВКИ

Проведені експерименти по тестуванню алгоритмів локалізації ключових точок на зображенні обличчя показали недостатність застосування одного шаблону для кожної ключової точки. Це пов'язано з індивідуальними особливостями зображень обличчя, які необхідно враховувати вже на етапі виявлення деталей обличчя. Такий результат порозумівається значним розходженням між векторами ознак, що відповідають одній ключовій точці, у зображень обличчя різних людей. При спробі виділити найбільш загальні ознаки і зменшити внесок ознак, що мають найбільші розходження, за допомогою вагових коефіцієнтів відбувається значна втрата інформації про ключову точку, що приводить до численних помилок при її автоматичній локалізації.

Застосування для виявлення ключових точок на зображенні обличчя методів, заснованих на використанні декількох шаблонів для кожної ключової точки, значно поліпшило точність локалізації і підвищило її універсальність стосовно різних локальних особливостей обличчя. Випробувані методи виділення ключових векторів груп показали приблизно однакові характеристики якості роботи, що підтвердили дані статистичного дослідження (див. таб. 2, рядки 4 і 5).

Ще одна перевага підходу, заснованого на використанні декількох шаблонів, це можливість використання отриманої класифікації ключових точок обличчя для задачі розпізнавання.

ЛІТЕРАТУРА

1. Wiskott L., Fellous J.M., Krueger N. and von der Malsburg C. Face Recognition and Gender Determination. // Proc. of the Int. Workshop on Automatic Face- and Gesture-Recognition, Zuerich, 1995. -p.92-97.
2. M. Lades et al. Distortion invariant object recognition in the dynamic link architecture. IEEE Trans. Comput., 42(3): 300-311, 1993.
3. J. Jones, L. Palmer An evaluation of the two-dimensional Gabor-filter model of simple receptive fields in cat striate cortex. // J. Neurophysiol., 1987.-p.1233-1258.
4. D. Burr, M. Morrone, D. Spinelli Evidence for edge and bar detectors in human vision. // Vision Res., 1989, p.419-431.
5. Г.С. Себестиан Процессы принятия решений при распознавании образов. // перевод под редакцией В.И. Иваненко, «Техника», Киев, 1965.

Попередня обробка зображень

Котитчук М.Б., Олещук О.В.

Одеський Національний політехнічний університет

65044, пр-т Шевченко 1.

тел. (0482)28-85-37, e-mail: logos104@rambler.ru

Анотація Розглядається спосіб підвищення вірогідності розпізнавання текстових символів, що нанесені на технологічні об'єкти, шляхом попередньої обробки зображення. Пропонується алгоритм, що дозволяє значно скоротити обчислювальні затрати на обробку зображення.

Вступ

Як першоджерела інформації для розпізнавання характерних ознак технологічних об'єктів, використовуються, як правило, відео- та фото зображення [1]. При цьому на вказаних зображеннях можуть бути присутні об'єкти довільного вигляду. Тому перед безпосереднім аналізом зображених предметів необхідно виконати ряд операцій, що дозволяють отримати зображення тільки об'єктів

визначеного типу. Існуючі методи обробки зображень [2]-[3] недостатньо враховують специфіку об'єктів спостереження та небажаних об'єктів, що попадають у поле зору.

В рамках даної статті обмежимося розгляданням зображень текстових символів, хоча розглянуті принципи обробки можуть бути застосовані і до деяких інших типів зображень.

Критерії формування характерних ознак

Для вирішення поставленого завдання потрібно визначити, які критерії класифікації можуть бути використані. Через те, що текст звичайно наноситься чорною чи білою фарбою, то можна очікувати, що найбільш темні чи світлі ділянки зображення є символами. На практиці за рахунок поганого освітлення чи наявності різних елементів конструкції та інших об'єктів, що попадають у поле зору, на зображенні часто присутні більш яскраві чи більш темні ділянки, ніж зображення тексту. З тієї ж причини деякі ділянки з текстом не є найбільш контрастними. З іншого боку, в областях, що безпосередньо містять текст, найбільш контрастними ділянками є переходи від зображення символу до фону зображення.

Також необхідно враховувати, що поверхня, на якій зображений текст, не є абсолютно однорідною, як за рахунок нерівномірного фарбування поверхні, так і за рахунок нерівномірного висвітлення.

Виходячи з вищесказаного, виходить, що одним з перших етапів повинне бути виділення фону зображення і вирівнювання яскравості і контрастності по всій області зображення. Для виділення фону необхідно знати колір символів, що

складають напис (чорний чи білий). Через те, що на цьому етапі колір символів невідомий, необхідно зробити потрібні операції для обох кольорів, а вже на більш пізніх етапах виконати визначення кольору символів.

Виділення фону варто робити на невеликих ділянках (чим менше, тим краще, але так, щоб туди не потрапив символ цілком). Спочатку коротко опишемо основні перетворення над матрицями, що використовуються для виділення фону. Умовимося вважати, що код чорного кольору має більше значення, ніж код білого кольору.

1. Поширення максимуму на сусідні точки:

$$D = DMax(S, k_x, k_y) \quad (1)$$
$$D(x, y) = \max(S(u, v))$$

де

$$u = \frac{x - k_x}{2} + 1, x + \frac{k_x}{2} - 1,$$
$$v = \frac{y - k_y}{2} + 1, y + \frac{k_y}{2} - 1.$$

Зміст перетворення $DMax$ полягає в тому, що значення кожного елемента (x,y) матриці D встановлюється в максимальне значення матриці S , знайдене в околиці точки (x,y) . Величина околиці для пошуку максимуму задається параметрами k_x, k_y .

2. Поширення мінімуму на сусідні точки:

$$D = DMin(S, k_x, k_y)$$

$$D(x, y) = \min(S(u, v)), \quad (2)$$

де

$$u = \overline{x - k_x / 2 + 1, x + k_x / 2 - 1},$$

$$v = \overline{y - k_y / 2 + 1, y + k_y / 2 - 1}.$$

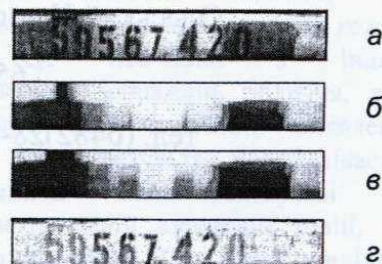
Дане перетворення аналогічне перетворенню $DMax$, за винятком того, що виділяє мінімальне значення в заданій області.

Якщо як розмір області задати значення, що перевищують очікувану товщину ліній символів і застосувати операцію $DMax$ над матрицею M_0 із зображенням цифр білого кольору, то на отриманій матриці M_1 області, що раніше містили символи, будуть містити значення, що відповідають кольору фону. Однак при цьому спостерігається деяке збільшення більш темних ділянок зображення за рахунок зменшення світлих ділянок. Це найбільше помітно при яскраво вираженій неоднорідності фону зображення. Щоб позбутися цього небажаного ефекту, потрібно до матриці M_1 застосувати операцію $DMin$ з тими ж параметрами. На отриманій матриці M_2 темні ділянки зображення повернуться до свого вихідного розміру, але на відміну від M_0 в матриці M_2 будуть відсутні світлі ділянки, розміром менше області пошуку мінімумів і максимумів. Якщо тепер з матриці M_2 відняти матрицю M_0 , то буде отримане зображення темних символів на білому фоні.

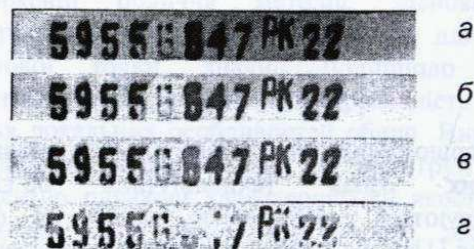
Аналогічно, якщо до зображення з темними символами на світлому фоні послідовно застосувати операції $DMin$ і $DMax$, а потім відняти отриману матрицю з вихідного зображення, то буде отримане зображення темних цифр на білому фоні (мал.1).

Зображення після видалення фону (мал.1г) містить темні ділянки розміром меншим, ніж $k_x * k_y$, де k_x і k_y – параметри в операціях $DMin$ та $DMax$. Вибір параметрів здійснюється виходячи з очікуваної товщини ліній символів на зображенні. Якщо вибрати параметри k_x і k_y занадто великими, то після видалення фону зображення буде містити темні ділянки з фонові області. Чим більше значення параметрів k_x і k_y , тим більше помилкових ділянок фону залишиться на зображенні. При виборі занадто малих значень параметрів k_x і k_y околиці для пошуку

мінімальних і максимальних значень виявляться настільки малими, що лінії символів будуть сприйняті як ділянки фону і будуть вилучені з зображення (мал.2).



Мал. 1. Видалення фону зображення: а – вихідне зображення, б – результат виконання операції $DMin$, в – результат виконання операції $DMax$ над матрицею б, г – різниця зображень а і б.



Мал. 2. Виділення фону з різною величиною околиці пошуку: а – вихідне зображення, б – околиця 64×64 , в – околиця 16×16 , г – околиця 8×8 .

Як правило, вказівка трохи більшого розміру області пошуку не так значно впливає на якість зображення, як область пошуку, менша, ніж товщина лінії символів. Це необхідно враховувати, якщо точно невідома товщина символів чи символи мають різну товщину ліній. У цьому випадку рекомендується вибирати параметри k_x і k_y , виходячи з максимально можливої товщини лінії символів.

Після видалення фону зображення символи, як правило, мають різну яскравість. Це пов'язано як із різною яскравістю символів, так і з різною яскравістю фону на вихідному зображенні (світлі символи на темному фоні будуть яскравіше, ніж світлі символи на світлому фоні). У зв'язку з цим виникає завдання вирівнювання яскравості по всьому зображенню.

Для цього введемо операцію нормалізації області

$$D = NArea(S, k_x, k_y),$$

де

$$D(x, y) = (S(x, y) - M_{Min}(x, y)) / (M_{Max}(x, y) - M_{Min}(x, y)),$$

$$M_{Min} = DMin(S, k_x, k_y),$$

$$M_{Max} = DMax(S, k_x, k_y).$$

(3)

Віднімаючи з матриці S матрицю M_{Min} , тобто мінімальні значення в області, отримаємо, що значення в ній будуть змінюватися від нуля і вище. У знаменнику знаходиться величина, що визначає різницю між максимальним і мінімальним значенням в області обраного розміру. Розділивши на цю величину, одержимо матрицю з діапазоном значень

$$D(x, y) = (S(x, y) - M_{Min}(x, y)) / (M_{Max}(x, y) - M_{Min}(x, y) + \gamma)$$

Недоліком цього способу є залежність результату від співвідношення величини γ і значень матриці S .

$$D(x, y) = \begin{cases} \text{Если } M_{Min}(x, y) \neq M_{Max}(x, y), & \text{то } \frac{S(x, y) - M_{Min}(x, y)}{M_{Max}(x, y) - M_{Min}(x, y)} \\ \text{Если } M_{Min}(x, y) = M_{Max}(x, y), & \text{то } \mu \end{cases}$$

Вибір значення μ варто робити виходячи з особливостей матриці S . Якщо вона є зображенням загального вигляду, то рекомендується встановити μ , рівним 0.5, тобто середині діапазону значень матриці D . Якщо в матриці S попередньо був вилучений фон зображення, то в ній з усіх великих однотонних областей могли залишитися тільки області білого кольору, тобто області із значенням, рівним нулю, тому значення μ варто встановлювати рівним нулю.



Мал. 3. Нормалізація області: а – вихідне зображення, б – нормалізація областей розміром 4x4, в – розміром 8x8, г – розміром 16x16, д – розміром 64x16.

В результаті нормалізації області значення кожної точки (x, y) у матриці D буде залежати тільки від значень точок в області розміром $k_x * k_y$, з центром у точці (x, y) . Параметри k_x і k_y вибираються виходячи з очікуваних розмірів символів на зображенні, відстані між ними, розташування символів стосовно інших об'єктів на зображенні (елементів конструкції, різного роду написів та ін.) і виду цих об'єктів. Вибір

від нуля до одиниці. Необхідно передбачити дії при рівності M_{Min} і M_{Max} у деякій точці (x, y) , щоб запобігти ділення на нуль. Можливі дії в цьому випадку:

Додати деяку постійну величину γ у знаменник формули (3).

Інше рішення - у випадку рівності M_{Min} і M_{Max} у деякій точці (x, y) як значення матриці $D(x, y)$ вибрати деяку постійну величину μ .

малих значень k_x і k_y приводить до того, що між символами можуть з'явитися помилкові темні ділянки і з'явиться велика кількість темних ліній у безпосередній близькості до символів, що зробить важчим подальше виділення символів із зображення. При виборі занадто великих значень k_x і k_y , на якість зображення деяких символів можуть впливати яскраво виражені об'єкти, що знаходяться поблизу від символів, у тому числі сусідні символи.

При необхідності після виконання нормалізації області можна перейти до перетворення зображення в монохромне. Це дозволить спростити подальшу обробку і зменшити витрати пам'яті для збереження зображення. Для перетворення зображення в двоколірне необхідно вибрати граничне значення τ , при перевищенні якого точка зображення вважається чорною і їй привласнюється код, рівний одиниці, інакше точка вважається білою з присвоєнням їй коду, рівного нулю.

Виконана послідовність перетворень дозволяє одержати зображення, що містить яскраво виражені об'єкти з заданими геометричними характеристиками. Результат виконання вищеописаних операцій дозволяє легко робити подальшу обробку на основі зв'язності точок, особливо якщо зображення, що спостерігається, містить послідовність недотичних об'єктів. Більш докладно про аналіз монохромних зображень і розпізнавання зображених на них об'єктів див. у [2], [4].

До недоліків розглянутих операцій можна віднести їх відносно велику обчислювальну вартість. Для усунення цього недоліку необхідно оптимізувати алгоритм їхньої реалізації.

Оптимізація операцій

Час виконання операції $DMax(S, k_x, k_y)$ при її виконанні для кожної точки матриці S окремо пропорційний $k_x k_y n_x n_y$,

де n_x, n_y – кількість точок матриці S по осях x та y .

Обчислювальні витрати при виконанні операції поширення максимуму в такому виді істотно залежать від параметрів k_x і k_y , і є дуже значними при великих значеннях k_x і k_y .

Очевидно, що при обчисленні результату операції для двох близько розташованих точок доводиться порівнювати між собою велику кількість тих самих значень, тому доцільно виконувати цю операцію, використовуючи загальні проміжні результати.

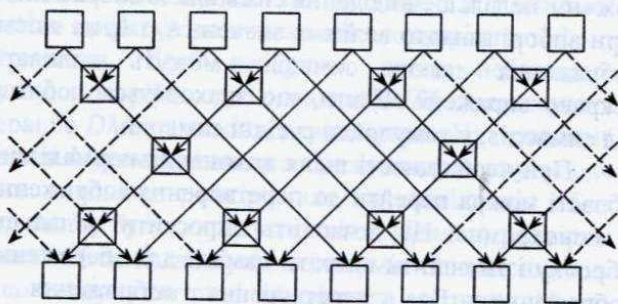
Для простоти розглянемо одномірний вектор V з числом елементів, рівним n ,

де $n = 2^m$, $m = 1, 2, 3, \dots$

і реалізуємо для нього операцію поширення максимуму $DMax(V, k)$.

де $k = 2^p$, $p = 1, 2, 3, \dots$

Запропонована схема реалізації алгоритму показана на мал.4.



Мал.4. Схема реалізації операції поширення максимуму.

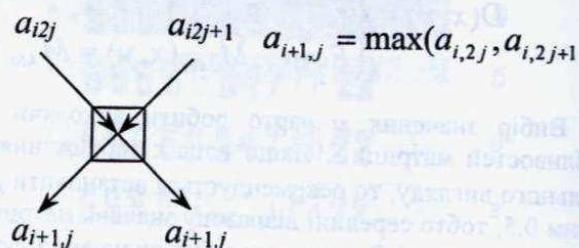
Алгоритм реалізації операції $DMax(V, k)$ умовно можна розділити на два етапи. На першому етапі

здійснюється пошук максимумів серед сусідніх елементів векторів.

На першому кроці попарно порівнюються сусідні елементи вектора V , і максимальне значення з кожної пари записується у вектор V_1 , що має число елементів у два рази менше, чим вектор V . Потім ті ж дії виконуються над вектором V_1 з одержанням вектора V_2 і т.д. доти, поки не буде отриманий вектор V_q ,

де $q = \log_2(k) - 1$.

На другому етапі виконується порівняння елементів векторів різного рівня. Спочатку кожен елемент вектора V_q порівнюється з двома елементами вектора V_{q-1} , що є сусідніми з елементами, з яких отримане значення поточного елемента вектора V_q . Далі отриманий вектор V_{q+1} аналогічно поелементно порівнюється з вектором V_{q-2} і т.д. доти, поки не буде отриманий вектор V_{2q} , що і є шуканим вектором.



Мал.5. Умовне позначення одиничної операції поширення максимуму

Якщо необхідно виконати операцію поширення максимуму, циклічно замикаючи краї матриці, то використовуються зв'язки, зазначені пунктиром. У протилежному випадку пунктирні зв'язки відкидаються.

Швидкість виконання операції поширення максимуму при використанні оптимізованого алгоритму пропорційна $\log(k_x) \log(k_y) n_x n_y$.

Висновки

Запропонований спосіб попередньої обробки зображень символів, нанесених на технологічний об'єкт, дозволяє поліпшити якість зображення, за рахунок чого вірогідність розпізнавання символів зростає до 85%.

Алгоритм оптимізації запропонованих операцій при обробці областей зображень розміром 16×16 дозволяє скоротити час обробки зображень у 16 разів, а при розмірі областей 32×32 – у 40 разів.

Література

1. Копитчук М.Б., Олещук О.В. "Использование метода развертки изображения для распознавания символов" – Праці УНДІРТ, 2001, №3(27). – С. 63 – 67.
2. Дуда Р., Харт П. "Распознавание образов и анализ сцен" – М.: Мир, 1976.
3. Василенко Г.И. Тараторин А.М. "Восстановление изображений" – М.: Радио и связь, 1986.
4. Вапник В.Н., Червонекис А.Я. "Теория распознавания образов" – М.: Наука, 1974.

ГІСТОГРАМНИЙ МЕТОД РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ І СИМВОЛІВ

Олександр Ільченко

Українська Академія Друкарства

79039 м. Львів вул. Підголоско 19 тел.59-94-20 e.mail:alex_i2002@ukr.net

Абстракт. У роботі описаний розроблений гістограмний метод і сформульовані основні принципи побудови системи розпізнавання рукописного тексту і символів.

1. ВСТУП

Розпізнавання образів і, зокрема, розпізнавання рукописного тексту і символів на даний час дуже актуальна науково-технічна проблема. Розв'язання її значно полегшить взаємодію людини із сучасною обчислювальною технікою, комп'ютерами, телекомунікаційним обладнанням, тощо. Відомо, що, як правило, введення інформації у комп'ютерну техніку здебільшого виконується вручну. Це займає багато часу і потребує великої кількості ресурсів.

2. РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ І СИМВОЛІВ

Рукописним текстом вважаємо сукупність слів, окремих літер, чисел, цифр, розділових знаків написаних від руки, кириличним або латинським алфавітом. До рукописного тексту відноситься не тільки рукописний текст, як ми його уявляємо, а й текст написаний від руки друкованими літерами, а також текст, написаний від руки стилізованими літерами або цифрами, наприклад, на поштових конвертах.

Рукописні символи це більш широке поняття, у яке входять усе вищеперелічене, а також спеціальні символи, наприклад, математичні, тощо.

3. ОГЛЯД МЕТОДІВ РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ І СИМВОЛІВ

Відомо декілька методів розпізнавання рукописного тексту і символів. Докладніше зупинимося на трьох із них, а власне на методі

вертикального сканування, хвильовому методі знаходження скелета растрового зображення, і на методі, так званого "фонтанного перетворення" (від англійського font – шрифт).

Принцип дії цих методів розпізнавання рукописного тексту і символів полягає у зведенні рукописного тексту і рукописних символів до спрощеного, стандартизованого вигляду.

Хвильовий метод знаходження скелета растрового зображення базується на принципі відслідковування за проходженням сферичної хвилі по контуру растрового зображення. Докладно цей метод описаний у [1].

Метод вертикального сканування викладений у [2] та полягає у побудові стандартизованого скелету растрового зображення окремих рукописних літер, шляхом пошуку екстремальних значень контура зображення.

Метод "фонтанного перетворення" заснований на принципі структурно-плямових еталонів, і використовується у широко відомій системі розпізнавання рукописного тексту "Fine Reader Рукопис", компанії АBBYY, яка широко використовується для обробки анкет Пенсійного фонду та Податкової інспекції [3,4].

Основним недоліком вищеперелічених методів є те, що вони мало придатні для розпізнавання рукописного тексту, складні в практичній реалізації, потребують значних системних ресурсів.

Розроблений автором гістограмний метод розпізнавання рукописного тексту і символів, який описаний нижче, не має вищеперелічених недоліків.

4. ГІСТОГРАМНИЙ МЕТОД РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ І СИМВОЛІВ

Гістограмний метод розпізнавання рукописного тексту і символів заснований на побудові графіка (гістограми) залежності кількості інформативних пікселів матриці бінарного растрового зображення, відкладеної по осі ординат, від координати рядка або стовпця, відкладеної по осі абсцис (рис.1) та порівнянні її з набором еталонних гістограм навчальної вибірки.

Під бінарним растровим зображенням розуміється двовимірна матриця з чорних та білих точок, в якій об'єкт задається чорними пікселями растру, а фон – білими. Під інформативними пікселями розуміються чорні пікселі растру.

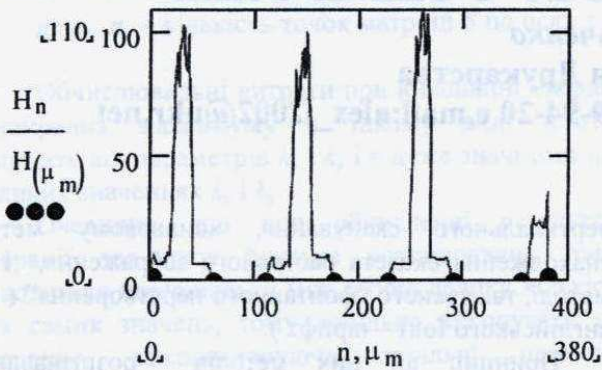


Рис.1. Гістограма рукописного тексту, наведеного на рис.2.

На рис.1 наведена гістограма тексту "В останні роки розпізнавання образів знаходить усе більше застосування в повсякденному житті", зісканованого із роздільною здатністю 300 dpi (рис.2).

В останні роки розпізнавання образів знаходить усе більше застосування в повсякденному житті

Рис.2. Рукописний текст, гістограма якого наведена на рис.1.

Гістограмний метод розпізнавання рукописного тексту і символів дозволяє із 100%-ю вірогідністю розділяти рядки у суцільному тексті і окремі слова, літери, цифри і символи у рядку, шляхом знаходження локальних мінімумів на гістограмі (рис.3,4).

В останні роки

Рис.3. Зразок рукописного тексту для ілюстрації розділення слів і окремих літер у рядку.

На рис.4 наведена ілюстрація розділення слів у рядку рукописного тексту, знаходженням локальних

мінімумів на гістограмі. Локальні мінімуми на гістограмі позначені крапками.

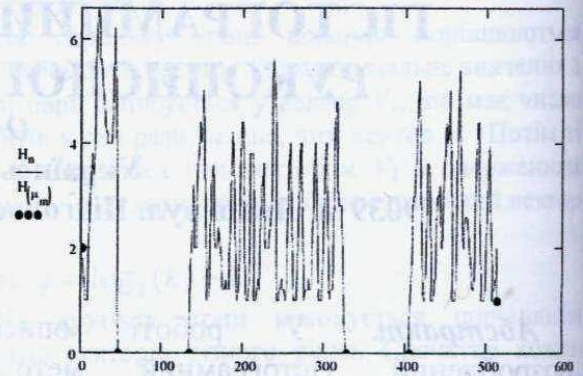
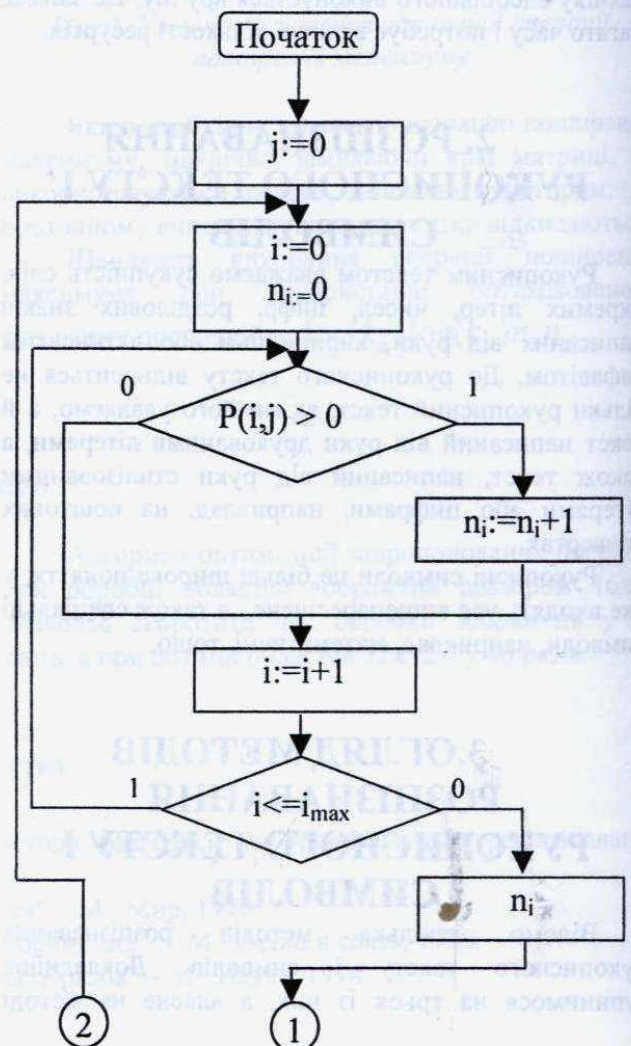


Рис.4. Гістограма окремих слів тексту, наведеного на рис.3.

5. АЛГОРИТМ ПОБУДОВИ ГІСТОГРАМИ БІНАРНОГО РАСТРОВОГО ЗОБРАЖЕННЯ

На рис.5 наведена блок-схема алгоритму побудови гістограми бінарного растрового зображення.



ВИКОРИСТАННЯМ ГІСТОГРАМНОГО МЕТОДУ

Система розпізнавання рукописного тексту і символів (рис.7) складається із сукупності програмних і апаратних засобів. Апаратна частина системи стандартна, і складається із комп'ютера, сканера і принтера і в межах даної статті розглядатися не буде.

Загальний алгоритм роботи системи розпізнавання рукописного тексту і символів приведений на рис.6.

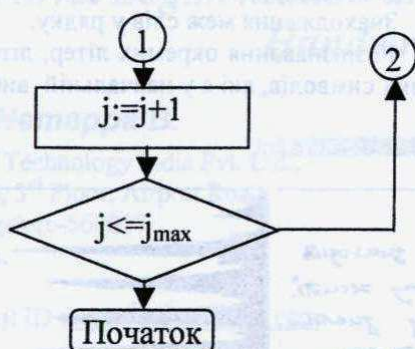


Рис. 5. Блок-схема алгоритму побудови гістограми бінарного растрового зображення.

Де

n_i - кількість інформативних пікселів у рядку матриці бінарного растрового зображення;

i - номер (координата) рядка матриці;

j - номер (координата) стовпця матриці;

$P(i,j)$ - координата кожного пікселя матриці;

У блоках 1 та 2 блок-схеми алгоритму побудови гістограми бінарного растрового зображення відбувається початкова ініціалізація змінних.

У блоці 3 перевіряється чи є кожен піксель матриці інформативним, тобто чорним.

Якщо піксель є інформативним, тоді до кількості інформативних пікселів у i - му рядку матриці - n_i додається одиниця (блок 4).

У блоці 5 відбувається перехід на наступний рядок матриці.

У 6-му блоці перевіряється чи даний рядок не є останнім у матриці.

Виводиться значення кількості інформативних пікселів у i - му рядку матриці бінарного растрового зображення (блок 7).

У блоці 8 відбувається перехід на наступний стовпець матриці.

У 9-му блоці перевіряється, чи даний стовпець не є останнім у матриці. Гістограмний метод розпізнавання рукописного тексту і символів заснований на побудові графіка (гістограми) залежності кількості інформативних пікселів матриці бінарного растрового зображення, відкладеної по осі ординат, від координати рядка або стовпця, відкладеної по осі абсцис (рис.1) та порівнянні її з набором еталонних гістограм навчальної вибірки.

6. СИСТЕМА РОЗПІЗНАВАННЯ РУКОПИСНОГО ТЕКСТУ І СИМВОЛІВ ІЗ



Рис.6 Блок-схема алгоритму роботи системи.

1. Введення навчальної вибірки шляхом сканування необхідного набору літер, цифр,

розділових знаків, символів тощо, введення їх значення з клавіатури.

2. Побудова гістограми кожного символу навчальної вибірки і присвоєння їй введеного з клавіатури значення.

3. Введення зі сканера тексту, який буде розпізнаватися.

4. Побудова гістограми тексту, який буде розпізнаватися.

5. Знаходження меж рядків у тексті.

6. Знаходження меж слів у рядку.

7. Розпізнавання окремих літер, літер у словах та інших символів, які є у навчальній вибірці.

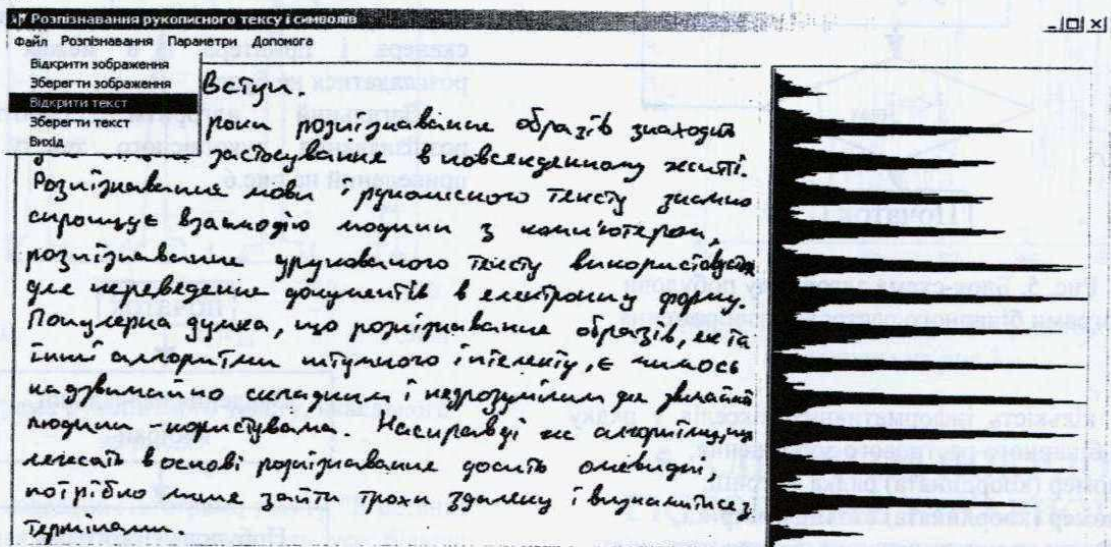


Рис. 7. Інтерфейс системи розпізнавання рукописного тексту і символів

8. Перевірка правильності розпізнавання слів з допомогою вбудованого словника.

Запис розпізаного тексту у файл, або вивід на друк.

Розроблена математична модель системи у середовищі MathCad 2001Pro.

Фрагмент програми на внутрішній мові програмування середовища MathCad 2001Pro.

Знаходження меж рядків у тексті і меж слів у рядку відбувається шляхом знаходження локальних мінімумів на гістограмах. Ці локальні мінімуми і є межами рядків або слів (рис.7).

Розпізнавання літер у словах відбувається так. Будується гістограма слова. Від цієї гістограми відтинається відрізок, ширина якого дорівнює мінімальній ширині літери. Гістограма цього відрізка порівнюється зі всіма гістограмами символів навчальної вибірки. При збігу гістограм, гістограмі відрізка, яка відповідає першій літері слова присвоюється значення символу гістограми навчальної вибірки із якою вона збіглася. Якщо відрізок мінімальної ширини не збігся ні з одною гістограмою навчальної вибірки, тоді до відрізка мінімальної ширини додається певний крок, будується нова гістограма відрізка і вся процедура порівняння нової гістограми із набором гістограм символів навчальної вибірки повторюється. Так відбувається доти, поки не відбувається збігання гістограми відрізка із набором гістограм символів навчальної вибірки. Після цього береться наступний відрізок мінімальної ширини і всі дії повторюються знову. Так відбувається доти, поки не буде розпізнано останню літеру слова.

Після цього відбувається перевірка правильності розпізнавання слів шляхом порівняння набору літер розпізаного слова із вбудованим словником системи розпізнавання.

7. ВИСНОВКИ

1. Розроблений метод розпізнавання рукописного тексту і символів, який ґрунтується на побудові гістограм тексту, щорозпізнається, і порівнянні її з набором еталонних гістограм навчальної вибірки.

2. Сформульовані основні принципи побудови системи розпізнавання рукописного тексту і символів із використанням розробленого методу.

7. ЛІТЕРАТУРА

1. Применение волнового алгоритма для нахождения скелета изображения. <http://www.ocrai.narod.ru/vectoriy.html>

2. Макуцький Я., Овсяк В. Алгоритми пошуку екстремальних точок і побудови стандартизованого скелету символів рукописного тексту. // Комп'ютерні технології друкарства. Збірник наукових праць №3. Львів, 1999р. с. 90 – 97.

3. Абраменко А. Компьютер читает. <http://www.ocrai.narod.ru/fr.html>.

Novel Segmentation and Feature extraction Techniques for Handwritten Text Recognition.

Hemappa B.

Intel Technology India Pvt. Ltd.,
#136, 3rd Floor, Airport Road,
Bangalore-560017
India.

Email ID : hems_bs@hotmail.com

Abstract: *In this paper, we proposed novel slant detection and correction techniques, segmentation and feature extraction methods for handwritten text recognition. The segmentation and feature extraction module works on data, which is to be preprocessed. This is necessary because of the output of the scanned document is noisy and usually contains too many unwanted impression happened due to humidity and uneven ink distribution on paper. The novel techniques are used to detection and correction of any rotation that may occur during the writing or scanning process. The slant correction algorithm then generates a non_slant image by rotating the blocks, rather than the individual pixels. The proposed segmentation algorithm will count the block pixels in horizontal lines of the binarized image and estimate the width and base line of each text line in the document. This ensures easy to segment the document into text of sentences. Each sentence again segmented into words. Each word processed and extract the prominent feature for further classification. The proposed system based on neural network feature extraction gives promising results in terms of classification and recognition.*

Keywords: *pattern recognition, Preprocessing, Trimming, Segmentation, Feature extraction, Slant detections, Slant correction, Neural network.*

1. Introduction

The area of handwriting recognition can be divided into online and offline [2,3]. In online recognition the writer is connected via an electronic pen or a mouse to the computer and the handwriting is recorded as a function of time. By contrast, in offline recognition the handwriting is captured by means of a scanner and becomes available for processing and recognition in the form of an image. Because of the availability of temporal information, online recognition is often considered the easier problem [1]. Automatic processing of handwritten forms requires form analysis [4, 5],

Dr.N.V.Subba Reddy

Professor and Head,
Department of Computer Science
Manipal Institute of Technology
MAHE (Deemed University)
Manipal-576 119, India.

Email Id: dr_nvsreddy@rediffmail.com

field extraction [6,7], handwritten recognition, and so on. Texts/Lines separation has two non-trivial problems: one is to detect and remove the lines, and the other is to restore some character strokes distorted by the deleted lines [8]. A number of methods have been reported for sentences/lines separation [9, 10], but most of them can handle lines with in $\pm 5^\circ$ because of computational burden of their algorithm and fails in some exceptional situations where a straight line is curved a little or has a non-uniform thickness.

2. Developed system

The structure of the developed text recognition system shown in fig 1. Notice that both the segmentation and feature extraction modules works on data, which is preprocessed. This is necessary because of the output of the scanned document is noisy and usually contains too many unwanted impression happened due to humidity and uneven ink distribution on paper. The novel techniques are used to detection and correction of any rotation that may occurred during the writing or scanning process. The volume of the data to processed can be reduced and prominent features of patterns (Texts) can be extracted by incorporating trimming and removal of underlines, by which efficiency of preprocessor improves and it will have influence on the over all performance of the pattern recognition system(11). The proposed segmentation algorithm will count the block pixels in horizontal lines of the binarized image and estimate the width and base line of each text line in the document. So that is easy to segment the document into text of sentences. And each sentence again segmented into words. Each words processed and extract the prominent feature for further classification and recognition system and to improve the accuracy of the classification, we will propose neural network based feature extraction system. The proposed

methods work well for multi lingual language documents and pattern recognition system.

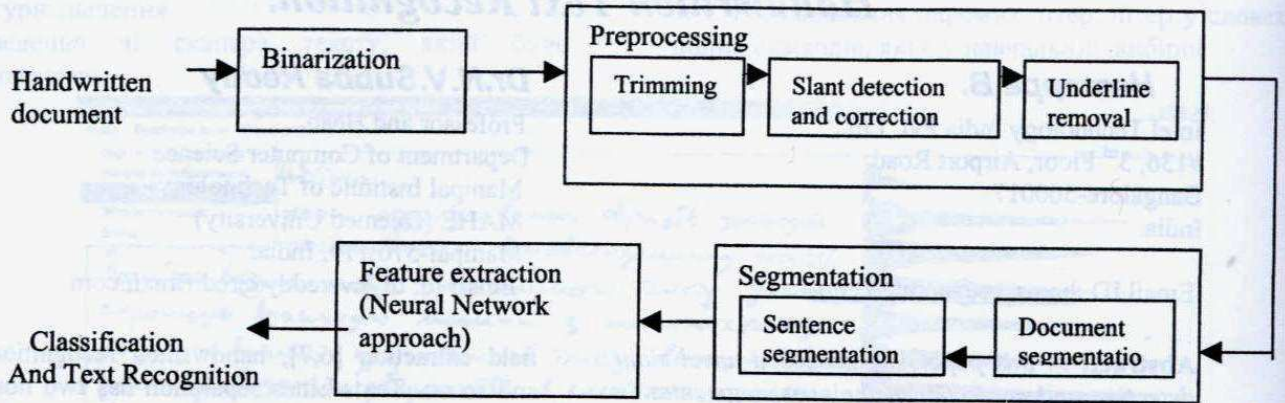


Fig 1. Overview of proposed model for printed and handwritten text recognition system.

3. Preprocessing and Normalization

The preprocessing model takes as input data and gives as output pattern containing the word to be recognized without any other element irrelevant to the recognition process. The tasks performed at this level depend on the data: when the document from which the word is extracted presents a background pattern, the latter should be removed. In some other cases, the words are extracted from forms showing boxes and lines that should be eliminated. For our data, binarization, trimming[11] the pattern to remove the unwanted impression happened due to humidity and uneven ink distribution on paper.

Once the image is preprocessed, it must be normalized. The normalization is deals with removing slant and slope of the document, line of texts and words. The slant detection and correction of the document images is particularly crucial among the document processing operations. Following algorithms explains the slant detection and corrections technique of document , sentence and words.

It is a violation of the copy right

A sequential input presentation definition for a learning machine dealing with a non-linear problem can be, at its best, a loose or a masterpiece of an expert. In real world applications of intelligent machines, a good solution approach is depend on a good representation of knowledge. Many algorithms have been proposed, but some of them defines robust and precise steps for an input presentation definition.

Fig 2. Sample inputs

we are

Fig 3. Slanted input

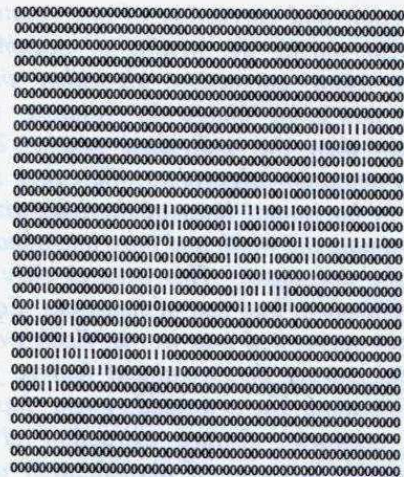


Fig.4 After Binarization

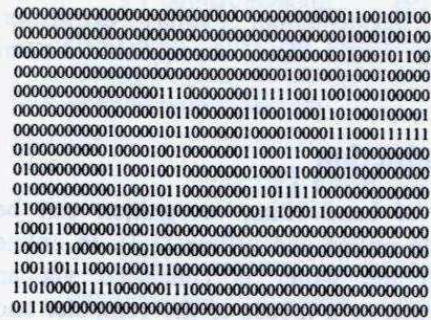


Fig 5 After Trimming

3. 1. Slant Detection Algorithm

The Slant detection algorithm takes preprocessed data as input and gives the degree of the slant towards right or left of the image.

Algorithm:

1. Initialize the TopLeft, TopRight, BottomLeft, BottomRight flags to zero.
2. Scan the 1/8th of binary pattern row wise from top to bottom.
3. if rows contains continues 0's
 $A[I][j] = 1;$
 Else
 $A[I][j] = -1;$
 Where $I=0 \dots \text{int}(M \times 1/8)$
 $J=0 \dots N$
 $M = \text{height of the pattern}$
 $N = \text{width of the pattern}.$
4. if each row of array A contains 1,s in left side ,
 Set TopLeft=1 else TopRight=1;
5. Scan the 1/8th of binary pattern row wise from bottom to top.
6. Initialize the Array A to 0, Repeat the Step 3.
7. if each row of array A contains 1,s in left side ,
 Set BottomLeft=1 else BottomRight=1;
8. if TopLeft = 1 and BottomRight=1 then Return LeftDownward Slant.
 Else if TopRight = 1 and BottomLeft =1 then return RightDownward Slant
 Else if((TopLeft=1 & bottomRight =0)or(TopRight=1 & BottomLeft =0)) then return Partial slant.
 Else return no slant.

3.2 Slant correction

Algorithm :

1. Assign the pattern into M x N matrix
2. if (LeftDownward slant)
 for($I=0; I < M * 0.5; I++$) {
 for($j=0; j < N; j++$) {
 if($\text{patten}[I][j] \neq 1$)
 $\text{pattern}[I][j] = \text{pattern}[I+1][j];$
 else
 break;
 }
 }
3. else if(RightDownward slant)
 for($I=0; I < M * 0.5; I++$) {
 for($j=N; j > 0; j--$) {
 if($\text{patten}[I][j] \neq 1$)
 $\text{pattern}[I][j] = \text{pattern}[I+1][j];$
 else
 break;
 }
 }

```
00000000000000000000111000000011110010001100111100
0000000000100000111000000011110010001100100100
0100000000100000111000000011110010001000100100
01000000001000001110000000110110010001000101100
11000000001000001110000000100010010001000100000
110000000010000010110000001000010000111009111111
11000100000010001010000000001110001100000000000
100011000001000100000000000000000000000000000000
100011000001000100000000000000000000000000000000
100110111000100011100000000000000000000000000000
110100001111000001110000000000000000000000000000
011100000000000000000000000000000000000000000000
```

Fig 6. After Slant correction.

4. Segmentation

Text regions possess a unique texture because they typically follow a specific arrangement rule: each region consists of text line of the same orientation with approximately the same spacing between them and each text line consists of characters of approximately the same size. The above observations suggest that the primary component of the page layout – Text lines can be discriminated using the following novel segmentation algorithm.

4.1 Page segmentation.

Algorithm:

1. Scan the binarized data row wise.
2. Count the number of on cells in each row.
 $D[i] = \text{OC};$
 Where $I = 0 \dots R$
 $\text{OC} = \text{Total number of on Cells in a row}.$
 $R = \text{Total number of rows in page}.$
3. Mark the minimum or no on cells rows.
 If ($D[I] = 0 \parallel D[I] < \text{Minimum}$)
 $M[j] = I;$
 $J = 0 \dots R/4;$
4. Find the width between the lines and separate the line of text.

4.2 Segmenting the text lines into words

To find the word boundaries within a text, We first set the distance between each words in a text line. Scan the binarized text line pattern vertically to identify the space between words and divide the text line into individual words to use for feature extraction. Preprocessed text line gives as input to the text line segmentation, which trimmed and underline removed pattern.

1. Set the space between the words
2. Scan the text line pattern column wise
3. if the space between words is equal to assumed space. Then divide the word.

```
000000000000000000001110 0111110010001100111100
000000000001000001110 0111110010001100100100
010000000001000001110 0111110010001000100100
010000000001000001110 0110110010001000101100
110000000001000001110 0100010010001000100000
110000000001000001011 1000010000111000111111
110001000000100010100 0011000011000000000000
100011000000100010000 0000000000000000000000
100110111000100011100 0000000000000000000000
110100001111000001110 000000000000000000000000
011100000000000000000 00000000000000000000000000
```

Fig 7. segmented and trimmed words

5. Feature extraction

In most systems, the word image is segmented into small parts supposed to be the basic information units. In our feature extraction, preprocessed whole word is segmented into 16 numbers of regions, extraction of the feature from each region and generates the feature

vectors. This feature vectors are especially useful in classification of words using neural network system.

Preprocessed pattern divided into 16 numbers of regions.

Feature value = on cells /total number of on and off cels in a regions.

Feature vector = { G1, G2,G3,G4,G5,G6, D1, D2,D3,D4,D5,D6, M1,M2,M3}

1. Divide the preprocessed pattern into number of regions.
- 2.find the total number of On cells in each regions
- 3.Find the total number of ON and OFF cells in a region,
- 4.Calculate the feature value

Feature value = Total number of ON cell in a region/Total number of ON and OFF cells.

Repeat the step 4 for all the regions and generate the feature vector.

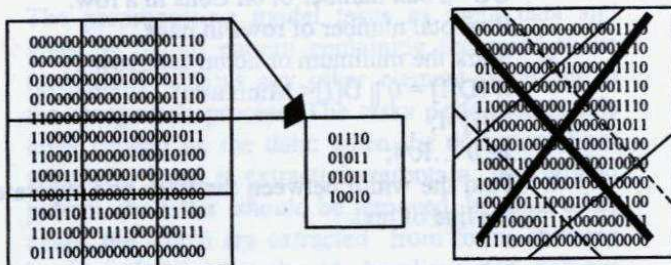


Fig 8. Segment the word into Grids(G1,G2..G6).

Doted lines segments the word as D1,D2,D3 regions, Solid lines divide as D3,D4,D5 regions
Solid Thick line segments M1,M2,M3,M4. regions

Conclusion:

The segmentation and feature extraction mechanisms developed are very useful in the recognition of text. The segmentation and feature extraction are two main components of the preprocessor. The overall performance of the text recognition system will have influence on better use of these two components. In this work proper care has been taken to improve the overall development of the text recognition system.

Reference:

1. U.V.Marti and H. Bunke, "Using A Statistical Language Model to Improve the Performance of an HMM-Based Cursive Handwriting Recognition System," Inter Journal of pattern recognition and Artificial Intelligence, Vol.15, No. 1(2001) pp.65-90.
2. I Guyon, M. Schenkel and J.Denker, "Overview and synthesis of on-line cursive handwriting recognition techniques," Handbook of character Recognition and

document Image Analysis, eds H.Bunke and P.S.P Wang, World Scientific, 1997, Chap 7, pp. 227-258.

3. J.C.Simon, "Off-line cursive word Recognition", Proc. IEEE80, 7(1992) 1150-1161.
4. A. Dengel and G.Barth, High level document analysis guided by geometric aspects, Int'l J. Pattern Recognition and Artificial Intelligence, Vol. 2, No. 4, pp 641-655, 1988.
5. T. Watanabe, Q. Luo, and N.Sugie, "Layout recognition of multi-kinds of table-form documents", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 17, no. , pp. 432-445, 1995.
6. F. Cesarini, M. Gori, S. Marinai and G.Soda, Data extraction from form images, Proc. Conference DEXA 95, pp. 438-448, London(UK), September 1995.
7. L.Y. Tseng and R.C. Chen, "Recognition and data extraction of form documents based on three types of line segments, Pattern Recognition", Vol.31,no. 10, pp. 1525-1540, 1998.
8. S.H Kim, S.H. JEONG, H.K. Kwag , "Line Removal and Character Restoration using Bag Representation of Form Images", Proceedings of seventh international Workshop on frontiers in handwritten recognition, Sept-2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 43-52.
9. B. Yu and A.K. Jain, A Generic system for form dropout, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18,no 11, pp. 1127-1131, 1996.
10. M.D. Garris, "Method and evaluation of character stroke preservation on handprint recognition", Technical Report NISTIR 5687, July 1995.
11. Hemappa B, Dr.N.V.Subba Reddy, "Preprocessing Technique for Handwritten Pattern Recognition," National Conference on Intelligent and Efficient Electrical system, January 19-20,2001. No. II-5.

ІНВАРІАНТНА МОДЕЛЬ ОБРОБЛЕННЯ ТА РОЗПІЗНАВАННЯ ІЗОПЛАНАТИЧНИХ ЗОБРАЖЕНЬ

В. В. Гнатушенко

Дніпропетровський національний університет
49050, Дніпропетровськ, пров. Науковий, корпус 12
E-mail: vvgnat@ukr.net
Тел. (0562) 43-36-30. Факс (0562) 46-55-23

АНОТАЦІЯ

Робота присвячена побудові інформаційно-геометричної моделі форм ізопланатичних зображень.

1. ВСТУП

Як відомо [1], дифракційні явища відіграють важливу роль при роботі оптичних приладів, призначених для одержання зображень об'єктів. Зображення, отримані за допомогою фокусувальних пристроїв (лінз), ніколи не відтворюють об'єкт з ідеальною точністю. Вони бувають спотворені унаслідок усякого роду недосконалостей оптичних систем (аберацій). Але навіть ідеальна лінза, вільна від аберацій, не може дати ідеального зображення через хвильову природу світла. Унаслідок дифракційних явищ порушуються закони геометричної оптики та відображення точок об'єкта в точки зображення перестає бути проєктивним. В термінах теорії хвильових процесів це означає, що зображення перестає бути стигматичним. Таким чином, виникає важлива задача відновлення й ідентифікації нестигматичних зображень, які не охоплюються розвинутою в [2, 3] інформаційно-геометричною моделлю.

2. ІНВАРІАНТНА ГЕОМЕТРИЧНА МОДЕЛЬ ПОДАННЯ ІЗОПЛАНАТИЧНИХ ЗОБРАЖЕНЬ

2.1. ОСНОВНІ СПІВВІДНОШЕННЯ ФОРМУВАННЯ ІЗОПЛАНАТИЧНИХ ЗОБРАЖЕНЬ

Розглядається некогерентне джерело з інтенсивністю випромінювання $I_0(\mathbf{x})$, яке

розташоване у предметній площині (некогерентність означає відсутність статистичного зв'язку між випромінюваннями різних точок джерела). Випромінювання від елемента поверхні об'єкта, розташованого у точці \mathbf{x} , поширюється через середовище і падає на лінзу діаметра D , центр якої розташовано у точці \mathbf{q} фотограмметричної системи координат XYZ . Зображення з інтенсивністю $I(\mathbf{r})$ спостерігається в площині зображення, паралельній площині лінзи на відстані d за нею (рис. 1), причому \mathbf{r} відлічується від головної точки зображення A — основи перпендикуляра з точки \mathbf{q} на площину зображення. Через ρ позначено радіус-вектор точки в площині лінзи.

Як відомо з теорії хвильових процесів, у загальному випадку інтегральний зв'язок між інтенсивністю зображення $I(\mathbf{r})$ та інтенсивністю випромінювання точок об'єкта $I_0(\mathbf{x})$ має вигляд [3]:

$$I(\mathbf{r}) = \int_{(\Omega)} I_0(\mathbf{x}) P(\mathbf{r}, \mathbf{x}) d\mathbf{x}, \quad (1)$$

де $P(\mathbf{r}, \mathbf{x})$ — функція розмиття точки (ФРТ). Інтегрування ведеться по області розташування джерела (поверхні об'єкта).

У загальному випадку $P(\mathbf{r}, \mathbf{x})$ є функцією двох змінних (\mathbf{r}, \mathbf{x}) , але, якщо довжина хвилі випромінювання λ суттєво менша за характерні розміри іконічної системи та відстані, з якої формується зображення, ФРТ з високою точністю може бути подана у вигляді $P(\mathbf{r}, \mathbf{x}) = P(\mathbf{r} - \mathbf{x}')$

(ізопланатичне наближення [4]), де $\mathbf{x}' = -\frac{d}{q}\mathbf{x}$. З

урахуванням цього функція яскравості ізопланатичного зображення має вигляд:

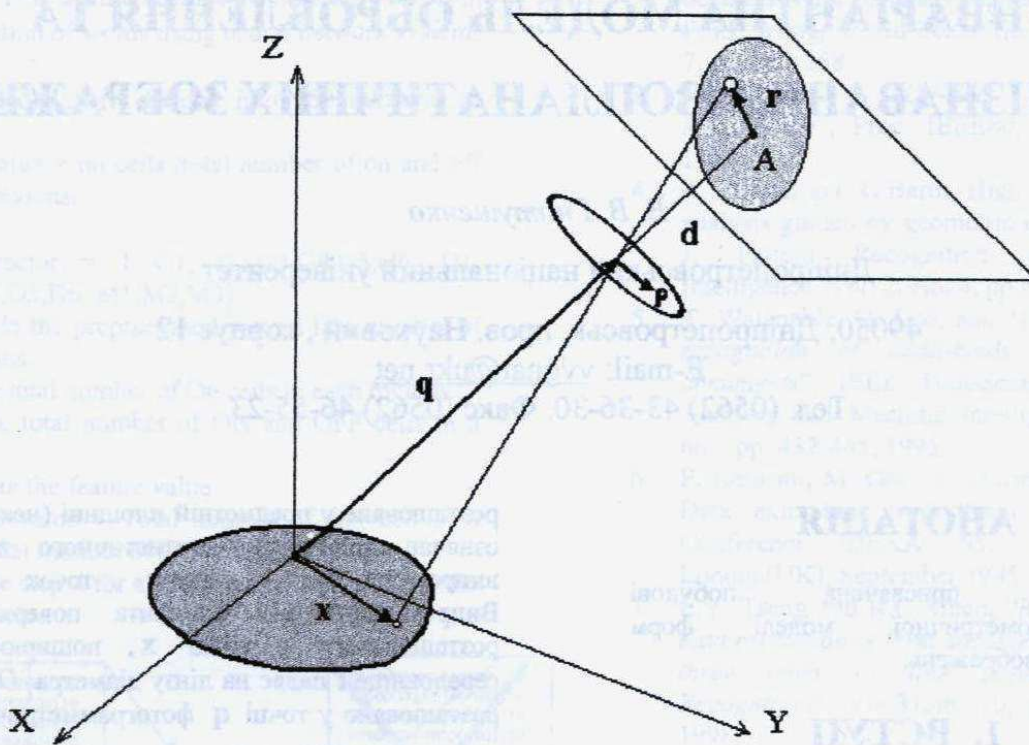


Рис. 1 Формування зображення в дальній фотограмметрії

$$I(\mathbf{r}) = \frac{1}{\lambda^4 q d^3} \int_{(\Omega)} I_0\left(-\frac{q}{d} \mathbf{x}'\right) P(\mathbf{r} - \mathbf{x}') d\mathbf{x}', \quad (2)$$

де ФРТ для круглої лінзи з апертурою діаметра D

$$P(\mathbf{r} - \mathbf{x}') = \frac{\pi D^2}{4} \int \exp\left[\frac{jk\rho d}{d}(\mathbf{r} - \mathbf{x}')\right] \times \exp\left[\frac{-jkq\rho d}{q}\right] K\left(\frac{\rho d}{D}\right) d\rho d \quad (3)$$

Через $K\left(\frac{\rho d}{D}\right)$ у виразі (3) позначено функцію

$$K\left(\frac{\rho d}{D}\right) = \begin{cases} \frac{2}{\pi} \left[\arccos\left(\frac{\rho d}{D}\right) - \left(\frac{\rho d}{D}\right) \times \sqrt{1 - \left(\frac{\rho d}{D}\right)^2} \right], & \rho d \leq D \\ 0, & \rho d > D, \end{cases} \quad (4)$$

яка дорівнює площі перетину двох кругів діаметра D з відстанню між центрами ρd .

Зазначимо, що у випадку $D \rightarrow 0$, $K\left(\frac{\rho d}{D}\right) \rightarrow \delta(\rho d)$ і відповідність точок об'єкта й зображення описується виразом $\mathbf{r} = -\frac{d}{H} \mathbf{x}$, (H — висота центра проєціювання над предметною площиною), що збігається із загальним виразом для стигматичних зображень [2, 4].

2.2. ІНВАРІАНТНІ БАГАТОВИДИ ПОДАННЯ ГЕОМЕТРИЧНИХ ФОРМ ІЗОПЛАНАТИЧНИХ ЗОБРАЖЕНЬ

Інформаційно-геометрична модель подання зображень проєкційної природи, розвинена в [5], будується у просторі інформаційних характеристик — безрозмірних семіінваріантів, які визначаються на основі інтегрального перетворення Фур'є відповідних функцій яскравості. З метою узагальнення моделі на випадок ізопланатичних зображень, виконаємо перетворення Фур'є від виразу (2), який пов'язує функцію яскравості зображення та власну яскравість точок об'єкта. Врахуємо при цьому, що цей вираз має вигляд згортки власної яскравості точок об'єкта $I_0(\mathbf{x})$ й функції розмиття точки $P(\mathbf{r})$. Зі спектрального аналізу відома теорема, у відповідності з якою перетворення Фур'є від згортки двох функцій

дорівнює добутку перетворень Фур'є від них [6]. З урахуванням цього одержуємо:

$$F(\mathbf{u}) = \frac{\pi D^2}{q^2 d^2} F_0\left(-\frac{d}{q} \mathbf{u}\right) \mathcal{P}(\mathbf{u}), \quad (5)$$

де $F(\mathbf{u}), F_0(\mathbf{u})$ — відповідно перетворення Фур'є від функції яскравості зображення та власної яскравості об'єкта. Через $\mathcal{P}(\mathbf{u})$ позначено перетворення Фур'є від ФРТ (3):

$$\mathcal{P}(\mathbf{u}) = \frac{\pi D^2}{4} \left(\frac{\lambda d}{2\pi}\right)^2 K\left[\frac{\lambda d}{2\pi D} \sqrt{u_1^2 + u_2^2}\right]. \quad (6)$$

На підставі цього зв'язок між семіінваріантами зображення S_{nm} , скалярної яскравості об'єкта $S_{nm}^{(0)}$ та функції розмиття точки \mathbf{S}_{nm} має вигляд:

$$S_{nm} = (-1)^{n+m} \left(\frac{f}{q-f}\right)^{n+m} S_{nm}^{(0)} + \mathbf{S}_{nm}. \quad (7)$$

де f — фокусна відстань лінзи.

У відповідності з принципом формування інформаційних характеристик зображень, які описуються семіінваріантами парного порядку $N = n + m$ [2, 5], маємо:

$$Z_i = \frac{S_{i-1, N+1-i}}{F^{N/2}(\mathbf{0})}, i = \overline{1, N+1}. \quad (8)$$

З урахуванням співвідношення (7) зв'язок між векторами подання двох зображень $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}$, отриманих при різних позиційних умовах їхньої фіксації, набуває вигляду:

$$Z_i^{(2)} = \left(\frac{q_1 - f}{q_2 - f}\right)^{n+m} \left(\frac{M_{00}^{(1)}}{M_{00}^{(2)}}\right)^{\frac{n+m}{2}} Z_i^{(1)} + \frac{\mathbf{S}_{i-1, N+1-i}^{(2)} - \left(\frac{q_1 - f}{q_2 - f}\right)^{n+m} \mathbf{S}_{i-1, N+1-i}^{(1)}}{\left(M_{00}^{(2)}\right)^{\frac{n+m}{2}}}, \quad (9)$$

де $M_{00}^{(1)}, M_{00}^{(2)}$ — декартові моменти нульового порядку від функцій яскравості відповідних зображень:

$$M_{00}^{(1)} = \frac{I_0(\mathbf{0})}{(\lambda^2 q_1 d_1)^2} \mathcal{P}_1(\mathbf{0}) = \frac{D_1^2 I_0(\mathbf{0})}{16\pi\lambda^2 q_1^2},$$

$$M_{00}^{(2)} = \frac{I_0(\mathbf{0})}{(\lambda^2 q_2 d_2)^2} \mathcal{P}_1(\mathbf{0}) = \frac{D_2^2 I_0(\mathbf{0})}{16\pi\lambda^2 q_2^2}. \quad (10)$$

Зазначимо, що вектори $\mathbf{Z}^{(1)}, \mathbf{Z}^{(2)}$ залежать від позиційних умов фіксації відповідних зображень і геометричних параметрів іконічної системи: в розглядуваному випадку — q, D . Далі розглядаємо випадок, коли $D_1 = D_2 = D$.

Приймаючи $\mathbf{Z}^{(1)}$ за подання еталонного зображення (далі \mathbf{Z}), для якого зазначені параметри формування вважаються відомими, покладемо співвідношення (9) в основу узагальнення інформаційно-геометричної моделі на випадок ізопланатичних зображень. Далі ризкою позначаємо величини, що пов'язані з аналізованим зображенням.

В загальному випадку довільної вимірності простору моделі зв'язок між векторами $\mathbf{Z}^{(1)}$ та $\mathbf{Z}^{(2)}$ є досить громіздким і тут не наводиться. В окремому випадку тривимірного простору моделі ($N = 2$), маємо:

$$\left. \begin{aligned} \bar{Z}_1 &= \left(\frac{\bar{q}(q-f)}{q(q-f)}\right)^2 Z_1 + \frac{16\pi\lambda^2 \bar{q}^{-2}}{D^2 I_0(\mathbf{0})} \times \\ &\quad \times \left[\bar{\mathbf{S}}_{02} - \left(\frac{q-f}{q-f}\right)^2 \mathbf{S}_{02} \right], \\ \bar{Z}_2 &= \left(\frac{\bar{q}(q-f)}{q(q-f)}\right)^2 Z_2, \\ \bar{Z}_3 &= \left(\frac{\bar{q}(q-f)}{q(q-f)}\right)^2 Z_3 + \frac{16\pi\lambda^2 \bar{q}^{-2}}{D^2 I_0(\mathbf{0})} \times \\ &\quad \times \left[\bar{\mathbf{S}}_{20} - \left(\frac{q-f}{q-f}\right)^2 \mathbf{S}_{20} \right], \end{aligned} \right\} \quad (11)$$

Виходячи з вигляду функції розмиття точки (3) можна показати, що її семіінваріанти симетричні за першим та другим індексами, $\bar{\mathbf{S}}_{02} = \bar{\mathbf{S}}_{20}$. Здійснимо лінійну заміну координат у просторі моделі

$$Y_1 = \frac{1}{2}(Z_1 + Z_3); \quad Y_2 = Z_2;$$

$Y_3 = \frac{1}{2}(Z_1 - Z_3)$. Врахуємо також поворот на кут γ картинної системи координат, в якій задане

зображення, що аналізується, стосовно координатної системи еталонного зображення. Виходячи з означення семіінваріантів [7] та виразу для перетворень координат картинної площини при повороті на кут γ :

$$\bar{x} = x \cos \gamma - y \sin \gamma, \quad \bar{y} = x \sin \gamma + y \cos \gamma,$$

можна показати, що зв'язок між інформаційними характеристиками зображення, підданого такому перетворенню, має вигляд:

$$\left. \begin{aligned} \bar{y}_1 &= y_1; \\ \bar{y}_2 &= y_2 \cos(2\gamma) + y_3 \sin(2\gamma); \\ \bar{y}_3 &= -y_2 \sin(2\gamma) + y_3 \cos(2\gamma) \end{aligned} \right\}. \quad (12)$$

На підставі виразів (11), (12), одержимо:

$$\left. \begin{aligned} \bar{Y}_1 &= kY_1 + V(\bar{q}, q), \\ \bar{Y}_2 &= kY_2 \cos(2\gamma) + kY_3 \sin(2\gamma), \\ \bar{Y}_3 &= -kY_2 \sin(2\gamma) + kY_3 \cos(2\gamma), \end{aligned} \right\}, \quad (13)$$

$$\text{де } k = \left(\frac{\bar{q}(q-f)}{q(q-f)} \right)^2,$$

$$V(\bar{q}, q) = \frac{32 \pi \lambda^2 \bar{q}^2}{D^2 I_0(\mathbf{0})} \left[\mathbf{s}_{02} - \left(\frac{q-f}{q-f} \right)^2 \mathbf{s}_{02} \right].$$

Застосовуючи методику побудови афінно-інваріантних багатомірних другого степеня, розвинену в роботі [5], одержимо рівняння квадрики, інваріантної до перетворень (13) компонент вектора подання ізопланатичних зображень

$$F_2^2 = \frac{1}{k^2} \left\{ [Y_1 - V(\bar{q}, q)]^2 - Y_2^2 - Y_3^2 \right\}. \quad (14)$$

Порівнюючи вираз (14) з рівнянням афінно-інваріантної квадрики для стигматичних зображень (точкової відповідності об'єкта та зображення) [5, 8], легко бачити, що урахування ізопланатичності веде до зміщення центра квадрики вздовж осі OY_1 на величину $V(\bar{q}, q)$ та перетворення інваріантної інформаційної ознаки F_2^2 .

Зазначимо, що, як випливає з (13), (14), при збільшенні відстані, з якої одержано зображення, квадрика (14) стягується в точку $[V(\bar{q}, q); 0; 0]$.

3. ВИСНОВКИ

Встановлені закономірності геометричної моделі подання ізопланатичних зображень дозволяють здійснювати їх морфологічний (позиційно-інваріантний) аналіз на основі належності точки, яка подає розподіл яскравості зображення, інваріантному багатомірному другому степеня (14), що визначається еталонним зображенням, та розташування цього багатомірну у просторі інформаційних характеристик. Взаємне розташування точок подання еталонного зображення та зображення, що аналізується, на інваріантному багатомірному дає підстави для обчислення відмінностей цих зображень у випадку їх належності одному класу еквівалентності геометричних форм.

ЛІТЕРАТУРА

1. Борн М., Вольф Э. "Основы оптики": Пер. с англ. — М.: Наука, 1973. — 720 с.
2. Михайленко В.С., Корчинський В.М. "Багатомірні геометричні моделі в інваріантному розпізнаванні проєкційних зображень" // Proc. of the Third All-Ukrainian International Conf. "Signal/Image Processing and Pattern Recognition"(UkrOBRAZ'96). — Kiev (Ukraine). — 1996. — P.94-96.
3. Korchinsk'kyj V. "Informative features of multiplied plane images for their invariant recognition" // Proc. of the Third All-Ukrainian International Conf. "Signal/Image Processing and Pattern Recognition"(UkrOBRAZ'96). — Kiev (Ukraine). — 1996. — P.87-88.
4. Исимару А. "Распространение и рассеяние волн в случайно-неоднородной среде". Том 2: Пер с англ. — М.: Мир, 1981. — 318 с.
5. Михайленко В. С., Корчинський В. М. "Концепція геометричного об'єкту в морфологічному аналізі проєкційних зображень" // Прикладна геометрія та інженерна графіка: К.: КДТУБА, 1997. — Вип.61. — С. 59 – 63.
6. Васильев В. Н., Гуров И. П. "Компьютерная обработка сигналов в приложении к интерферометрическим сигналам." — СПб.: БХВ — Санкт-Петербург, 1998. — 240 с.
7. Гнатушенко В. В. "Багатомірні геометрична модель напівтонових зображень у вигляді центральних проєкцій". // Збірка праць міжнар. наук.-практ. конф. "Сучасні проблеми геометричного моделювання". — Харків, 2001. — С. 202-204.
8. Гнатушенко В. В. "Графічне подання перетворень проєкційних зображень у багатомірній інформаційно-геометричній моделі." // Прикладна геометрія та інженерна графіка. К.: КНУБА — 2001. — Вип. 69. — С. 217 – 219.

ВИДІЛЕННЯ ІНФОРМАТИВНИХ ОЗНАК ТЕКСТУРНИХ ЗОБРАЖЕНЬ НА ОСНОВІ МЕТОДА РЕКУРЕНТНОГО КІЛЬКІСНОГО АНАЛІЗУ

О.М.Ахметшин, М.Г.Комарова

Дніпропетровський національний університет
49050 м. Дніпропетровськ, пер. Науковий 13
Email: akhm@mail.dsu.dp.ua

АНОТАЦІЯ

У роботі розглянуто новий метод кількісного аналізу текстурних зображень, який ґрунтується на ідеї аналізу нелінійних динамічних процесів у багатовимірному фазовому просторі, що дозволяє підвищити чутливість аналізу. Подані результати досліджень на прикладах текстурних зображень та даних рентгеновської комп'ютерної томографії для виявлення ранніх стадій розвитку остеопорозу.

1. ВСТУП

Проблема аналізу текстурних зображень традиційно викликає велике зацікавлення та знаходиться на межі методів комп'ютерного бачення та штучного інтелекту. Особливо актуальною ця проблема виявляється при застосуванні до задач ранньої медичної діагностики та контролювання динаміки процесу лікування онкологічних захворювань. У роботі [1] було розглянуто підхід, заснований на ідеї відображення зрізів текстурних зображень у двовимірному просторі ядер Вольтерра другого порядку, а в роботі [2] для відображення одновимірних зрізів використана двовимірна площина кумулянтів третього порядку. Обидва розглянутих підходи є достатньо чутливими, але за своєю суттю подають апарат якісного аналізу, оскільки базуються на тому, що синтезовані зображення візуально відрізняються від еталонних.

Метою поданої роботи є подальший розвиток цього спрямування, але з напірним на підвищення загальної чутливості аналізу на основі кількісних характеристик, що обчислюються за допомогою метода рекурентного кількісного аналізу (РКА).

2. МЕТОД

РКА, запропонований Дж.П.Збілутом та Ч.Л.Веббером [3, 4], являє собою технологію для аналізу нелінійних динамічних сигналів. Головною перевагою метода є можливість його застосування до часових та просторових рядів незалежно від їх статистичної природи, наявності шуму від будь-якого джерела (самої системи або спостереження) та

довжини ряду. Метод РКА дає локальний огляд поведінки сигналу, надає можливість досліджувати короткочасову поведінку системи.

При дослідженні нелінійної динаміки процесів вважається, що змінні, які описують систему, змінюються за часом залежно одна від одної. Якщо систему описують за допомогою d змінних x_1, x_2, \dots, x_d , то така взаємозалежність може бути подана у формі вектора (x_1, \dots, x_d) . Але на практиці для багатьох систем можливо замірювати лише поведінку за однією змінною. Для подальшого дослідження динаміка такої системи повинна бути змодельована у відбудованому фазовому просторі. Процедура побудови такої моделі спирається на концепцію, що скалярний сигнал, що спостережується, є одновимірною проекцією динамічної поведінки системи, що подається d вимірами. Динаміка може бути реконструйована як багатовимірна за допомогою даних лише одного скалярного часового сигналу, який було спостережено. Для застосування процедури відбудови записані дані повинні являти собою квантований сигнал $s(n)$. Вектор $Y(n)$ формується з затриманих версій сигналу $s(n)$ [5]:

$$Y(i) = (s(i), s(i+1), \dots, s(i+(d-1)L)), \quad (1)$$

де d – обрана кількість вимірів, L – затримка. Теорема Такенса [6] встановлює математичне співвідношення між розмірністю побудованого вектора та реальною розмірністю атратора відповідної динамічної системи (D): $d=2D+1$, де d є мінімальною розмірністю, що забезпечує можливість реконструювання атратора. Під атратором розуміємо множину точок в фазовому просторі, на яку виходить система при нескінченному збільшенні часу спостереження. Наведене вище співвідношення є справедливим для детермінованих систем при відсутності шуму. При наявності шуму від будь-якого джерела (самої системи чи вимірювання) необхідний більш високий порядок.

Технологія рекурентного кількісного аналізу не передбачає побудовання атратора у реконструйованому фазовому просторі, але на основі даних, достатніх для цього, дозволяє визначити деякі

показники хаотичності системи (ентропію процесу та значення максимальної експоненти Ляпунова).

Рекурентний кількісний аналіз ґрунтується на обчисленні матриці Євклідових відстанів D_{ij} між всіма можливими комбінаціями пар рядків відбудованої матриці $Y(n)$.

$$D(i, j) = \sqrt{\sum_{k=1}^n (i_k - j_k)^2}, \quad (2)$$

де i, j – вектори, що відповідають рядкам матриці $Y(n)$, n – розмірність цих векторів.

Отримана матриця використовує функцію Хевісайда для значень відстаней між скалярними рядами, що є меншими за деякий напередвизначений поріг. За значеннями, що повертає така функція, будується графік, що носить назву рекурентного, а затемнені точки на ньому, які відповідають відстаням, меншим за значення порогу, – рекурентних точок. Графічне подання являє собою квадратну матрицю з затемненою головною діагоналлю ($D_{ij}=D_{ji}$) та може бути розглянуте як глобальна автокореляційна структура системи. Найбільш важливими для розглядання є рекурентні точки, що утворюють діагональні структури, які є паралельними відносно головної діагоналі. Наявність ділянок с такими структурами на рекурентному графіку вказує на наявність детермінізму у динамічному процесі, що розглядається.

На основі побудованого рекурентного графіка можуть бути обчислені дескриптори динамічного процесу.

Оскільки рекурентний графік є симетричним відносно головної діагоналі, до розглядання залучається тільки верхня трикутна область за виключенням нульових євклідових відстанів, що утворюють головну діагональ.

1. Рекурентність (R) являє собою відношення кількості затемнених (рекурентних) точок до загальної кількості точок у верхній трикутній області графіка. Рекурентні точки відповідають "рівнинам" фазового простору, в таких точках значення величини, що спостерігається, визначається початковими умовами (детерміновано процедурою введення), відповідно, величина, що дорівнює $100-R$, визначає пропорцію точок, значення в яких зумовлено стохастичною поведінкою системи.

$$R = \frac{N_r}{N}, \quad (3)$$

де N_r – кількість рекурентних точок, N – загальна кількість точок.

2. Детермінізм (D) визначає відсоток рекурентних точок, які утворюють на графіку лінії, паралельні головної діагоналі. Наявність діагональних ліній є сигнатурою детермінованої структури системи. Значення цього параметру визначається у значній мірі напередвизначеним параметром LINES, що вказує, яку кількість підряд розташованих точок слід вважати лінією.

$$D = \frac{N_d}{N_r}, \quad (4)$$

де N_d – кількість рекурентних точок, що формують діагональні структури, N_r – кількість рекурентних точок.

3. Ентропія (E) розподілення довжини діагоналей, паралельних головної, є мірою невизначеності (хаотичності) ймовірносної системи. По мірі наближення системи до стану рівноваги значення ентропії зростає. Визначається за формулою Шеннона для інформаційної ентропії.

$$E = \sum_{i=1}^n P_i \log_2 \frac{1}{P_i} \quad (5)$$

4. Максимальна довжина діагональної лінії (M) зворотно пропорційна максимальній експоненті Ляпунова (L) для процесу, що розглядається,

$$L = \frac{1}{M} \quad (6)$$

являє собою міру швидкості розходження сусідніх траєкторій у фазовому просторі, тобто швидкість "забування" початкових умов системи (ступень "самостійності" поведінки системи). Аналіз цього показника дозволяє оцінити можливість передбачення подальших значень ряду.

5. Тренд локальної рекурсії (T) є мірою стаціонарності процесу, визначає міру зміни густини рекурентних точок в напрямку від головної діагоналі графіка. Стаціонарним сигналам відповідає близьке до нуля значення цього показника.

3. Застосування

запропонованого методу

Метод РКА був застосований для оцінки стохастичних процесів, поданих текстурними зображеннями та томографічними знімками кісток людини, хворої на остеопороз.

Текстурні зображення, обрані для порівняння, подані на Рис.1.

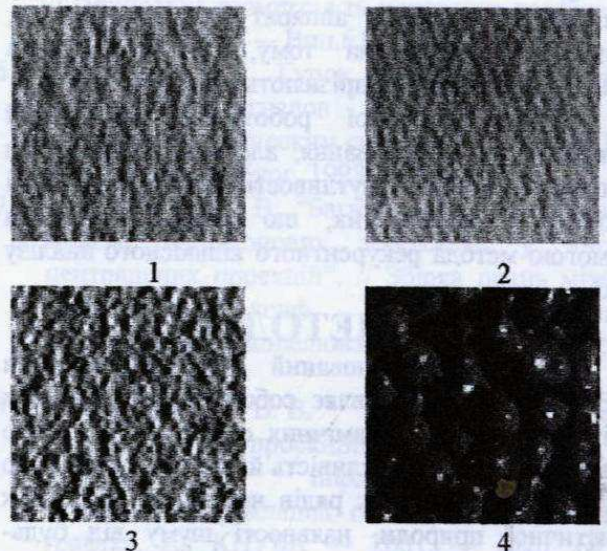


Рис.1. Текстурні зображення

Оскільки метод РКА призначений для оцінювання рядів даних, для визначення параметрів текстурних зображень потрібно працювати з їх зрізами. Зріз зберігає інформацію про спектральний склад текстури, але втрачаються дані про об'єкти, що присутні на зображенні.

Поданим текстурним зображенням відповідають рекурентні графіки, приведені на Рис.2.

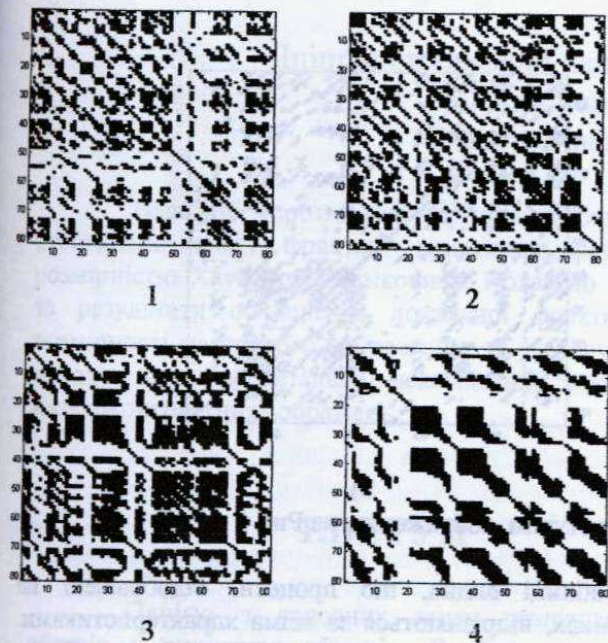


Рис.2. рекурентні графіки, побудовані для зрізів текстур, зображених на Рис.1.

Графіки, що мають структуру майже однорідної густини, відповідають рядам з невеликою дисперсією значень. Пробіли у рекурентному графіку свідчать про наявність "скачків" даних у розглянутій послідовності. для значення

відбудованої розмірності більше одиниці така залежність переноситься на поведінку траєкторій у фазовому просторі. Для побудовання рекурентних графіків (Рис.3) була обрана розмірність $d=2$, оскільки при виборі одиничної розмірності кількісні показники зображень 1 та 2 відрізняються на незначні величини. При обиранні занадто малої розмірності відбудованого фазового простору деяка інформація про структуру даних може бути загублена. Значення порогу визначення рекурентних точок дорівнює 20 відсотків від максимальної з отриманих евклідових відстанів. У Табл.1 наведені п'ять основних дескрипторів РКА та величина, зворотна довжині максимальної діагональної лінії на графіку, яка відповідає максимальній експоненті Ляпунова для процесу, що розглядається.

На основі кількісних характеристик текстурних зображень можна зробити висновки про характер процесів, які вони відображають. По-перше, як за візуальною оцінкою, так і за кількісним аналізом рекурентних графіків схожі початкові зображення виявляються досить різними. Найбільше значення параметру R відповідає текстурі 3, найменше – текстурі 1, тобто вона виявляє найбільш стохастичну поведінку. За коефіцієнтом детермінізму D більш детермінований процес серед усіх розглянутих являє текстура 4, текстура 3 має найменшу швидкість "забування" початкових умов. Ентропія текстури 4 найбільша, що свідчить про те, що процес, який вона відображає, знаходиться ближче до стану рівноваги, ніж решта розглянутих. З точки зору стаціонарності найближче до нуля значення тренду локальної рекурсії відповідає текстурі 1.

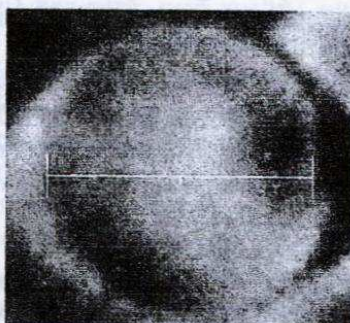
Кількісні характеристики аналізованих текстур

Таблиця 1

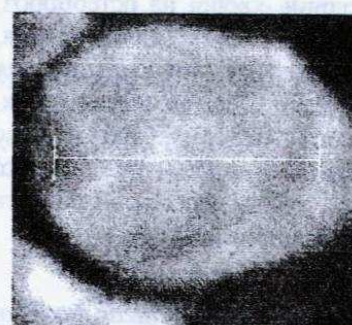
	R	D	M	L	E	T
текстура 1	27.500	21.473	16	0.0625	1.890	9.066
текстура 2	29.475	26.492	12	0.0833	2.392	78.586
текстура 3	42.469	42.078	19	0.0526	2.527	141.510
текстура 4	29.557	64.240	16	0.0625	2.970	34.566

Метод РКА був застосований також для аналізу томографічних знімків двох кісток людини (Рис.3), які знаходяться на різних ступенях

захворювання на остеопороз. Білі лінії на зображеннях вказують розташування на знімках зрізів, які були взяті для аналізу.



a)

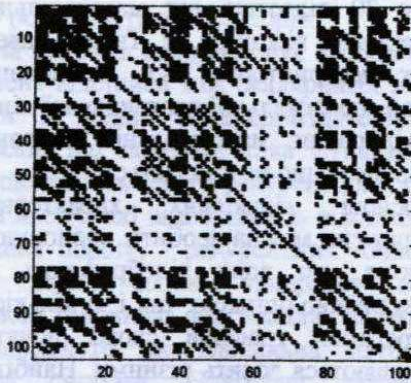


b)

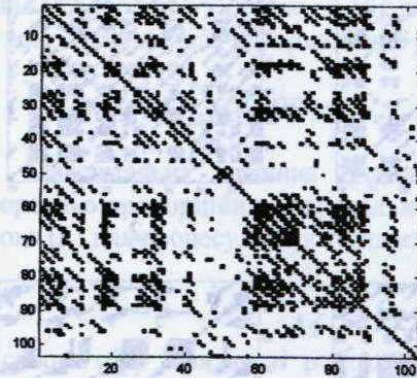
Рис.3. Томографічні знімки кісток людини
a) – нормальна кістка, b) – кістка, хвора на остеопороз.

Предметом розглядання є знаходження та зіставлення кількісних характеристик цих знімків з метою виділення характерних ознак, які б відрізняли їх один від одного. Як і при аналізі текстурних зображень, кількісні характеристики обчислені для зрізів. Перетворення двовимірного зображення на одновимірний сигнал хоча і відповідає умовам застосування методу, але при такому поданні повністю загублюється

достовірність інформації – обробці підлягає вже деякий періодичний ряд. Оскільки кістка має пористу структуру, деякі ознаки повинні зберігатися на обох зображеннях, і відхилення значень кількісних показників буде незначним. Тому для дослідження були взяті градієнтні перетворення зрізів зображень (тільки високочастотні компоненти).



a)



b)

Рис.4. Рекурентні графіки, що відповідають текстурам, зображеним на Рис.4

Для побудовання рекурентних графіків, поданих на Рис.4, були обрані такі вхідні параметри: відбудована розмірність дорівнює трьом, поріг рекурентних відстаней – 30 відсотків від максимальної евклідової відстані для кожного зображення, лінією вважаються розташовані поряд 10 рекурентних точок. Кількісні характеристики, обчислені для графіків (Рис.4), наведені у Таблиці 2.

З таблиці видно, що процеси, відображені на знімках, відрізняються за всіма характеристиками. Кістка на зображенні b, яка знаходиться на більш пізньому ступені захворювання, являє собою процес, який є менш стаціонарним та детермінованим, ніж процес, що відображений на знімку a, але водночас більш залежить від "початкових умов".

Таблиця 2

Кількісні характеристики аналізованих зображень

	R	D	M	L	E	T
Зображення a	55.797	38.315	23	0.0435	3.264	108.598
Зображення b	34.875	12.828	17	0.0588	2.247	123.867

4. ВИСНОВКИ

В ході роботи проілюстровано застосування методу РКА для підвищення чутливості та отримання кількісних характеристик зображень. Метод був успішно застосований для аналізу текстурних зображень. Результати аналізу томографічних знімків двох кісток людини, хворої на остеопороз, які знаходяться на різних ступенях захворювання, підтвердили можливість застосування методу для медичної діагностики навіть на ранніх стадіях. Зручність підходу полягає в тому, що аналізу підлягають дані, не описані за допомогою будь-якої математичної моделі. Це робить область застосування методу дуже широкою.

ЛІТЕРАТУРА

[1]. Akhmetshin A.M., Shutin D.V., *Nonlinear model of qualitative classification on base of wave-packet decomposition in space of Volterra kernels*, UkrOBRAZ'2000, Kyjiv, pp. 153-156, 2000.

[2]. Akhmetshin A.M., Mikhalev D.A., *Classification of texture images with the usage of artificial neural nets and higher-order statistics*, UkrOBRAZ'2000, Kyjiv, pp. 173-176, 2000.

[3]. Manetti, C., M.-A. Ceruso, A. Giuliani, C.L. Webber, Jr., and J.P. Zbilut (1999). "Recurrence quantification analysis as a tool for the characterization of molecular dynamics simulations", *Physical Rev. E*. 59: pp. 992-998.

[4]. Manetti, C., A. Giuliani, M.-A. Ceruso, S. Cannistraro, C. L. Webber, Jr., and J.P. Zbilut (2001). "Recurrence analysis of hydration effects on nonlinear protein dynamics: multiplicative scaling and additive processes". *Phys. Lett. A* 281: 317-323

[5]. F. Takens, *Dynamical Systems and Turbulence* (D.A. Rand and L.-S. Young, eds.) pp. 366-381. (Springer-Verlag, New York, Heidelberg, Berlin, 1980).

[6] П. Берше, И. Поло, К. Видаль, *Порядок в хаосе*, - М.: Мир, 1991. - с.366

СЕГМЕНТАЦІЯ НИЗЬКОКОНТРАСТНИХ ЗОБРАЖЕНЬ НА ОСНОВІ МЕТОДУ МУЛЬТИФРАКТАЛЬНОЇ ДЕКОМПОЗИЦІЇ

Ахметшин О.М. Киргизов І.О.

Дніпропетровський національний університет, Україна, 49050,
м. Дніпропетровськ, пр. Науковий 13.
Електронна пошта: akhm@mail.dsu.dp.ua

Наведені короткі теоретичні відомості, що стосуються поняття фрактал – множини з дробовою розмірністю Хаусдорфа-Безіковича. Показано метод та результати обчислення локальної фрактальної розмірності растрового зображення та застосування методу мультифрактальної декомпозиції текстур і низькоконтрастних зображень.

1. ВСТУП

Однією з головних задач розпізнавання образів є виділення об'єктів. Особливо гостро і актуально ця проблема існує при сегментації низькоконтрастних зображень.

Кожний об'єкт на зображенні, який потрібно виділити, має властиві йому і тільки йому деякі характеристики. Точки зображення, що мають спільні характеристики відносяться до одного об'єкту, а точки з іншими характеристиками – до інших об'єктів.

Одним із простих методів сегментації зображення є його квантування. Тоді точки з однаковими рівнями інтенсивності об'єднують в один клас або сегмент даних. Зрозуміло, що цей метод занадто тривіальний, оскільки, багато різних об'єктів можуть мати одну і ту ж саму інтенсивність і навпаки, один об'єкт може мати різні рівні інтенсивності. Але об'єкти можуть відрізнятися не тільки рівнями інтенсивності, а й своєю формою.

Деякі методи використовують знаходження краю (грані) для того, щоб знайти контури об'єкта. Одним із недоліків таких методів є те, що вони дають небажані грані при дослідженні складних об'єктів або текстур.

Застосування фрактального аналізу для сегментації та класифікації об'єктів дозволяє уникнути вищезгаданих проблем.

Фракталом називають структуру, складену з частин, які в деякому сенсі подібні цілому. Фрактал має однаковий вигляд, незалежно від масштабу спостереження [1]. Прикладом фракталів є гілки або кора дерева, морозний візерунок на склі. Структури з такою геометрією є самоподібними і їх можна

класифікувати таким критерієм, як фрактальна розмірність D .

Відомо, що просторово ізотропні поверхні дають зображення, інтенсивність яких є фрактальною функцією тієї ж дробної розмірності, що і початкова поверхня. Це дозволяє, вимірюючи характеристики зображення, проводити аналіз самої поверхні.

2. ФРАКТАЛЬНА РОЗМІРНІСТЬ

Фрактали можна розглядати як множину точок, вкладених у простір. Множина точок, що утворюють площину в евклідовому просторі має топологічну розмірність $D_1=2$. Евклідова розмірність простору $E=3$. Для визначення розмірності Хаусдорфа-Безіковича (фрактальної розмірності D) потрібно визначити поняття відстань між точками в просторі. Потрібно визначити величину множини точок \mathcal{Z} у просторі. Один із засобів виміряти площу поверхні полягає в тому, щоб розділити простір на невеликі куби із стороною δ . Якщо помістити центр куба в деякій точці множини, то всі точки в кубі будуть покриті цим кубом. Підраховуючи число кубів, необхідних для покриття множини точок, ми отримуємо міру величини множини. Площу кривої можна визначити як число кубів $N(\delta)$. Продовжуючи підрахунок і, знайшовши $N(\delta)$ при менших значеннях δ . Для фрактальних поверхонь буде виконуватись співвідношення:

$$N(\delta) \equiv \frac{1}{\delta^D} \quad (1)$$

Звідси, фрактальна розмірність D визначається:

$$D = \lim_{\delta \rightarrow 0} \frac{\ln(N(\delta))}{\ln\left(\frac{1}{\delta}\right)} \quad (2)$$

Тобто, D – кутовий коефіцієнт графіка $\ln N(\delta)$ як функції від $\ln(\delta)$.

Розмірність D , яка визначається по формулі (2)

через підрахунок кількості кубів, необхідних для покриття множини \mathfrak{Z} в залежності від розміру куба, називають клітковою розмірністю.

Кожен об'єкт (підмножина \mathfrak{Z}_i) зображення \mathfrak{Z} характеризується своєю фрактальною розмірністю D_i і множина точок \mathfrak{Z} є мультифрактал та являє собою об'єднання підмножин \mathfrak{Z}_i [1]:

$$\mathfrak{Z} = \bigcup_i \mathfrak{Z}_i \quad (3)$$

Суть розробленого методу мультифрактальної декомпозиції полягає в обчисленні D_i та в об'єднанні (групуванні) точок в підмножини \mathfrak{Z}_i .

2. РЕАЛІЗАЦІЯ МЕТОДУ

2.1. ОПИС ПОЧАТКОВИХ ДАНИХ

Спочатку дамо опис даних, над якими буде проводитись дослідження.

- 1) Зображення – це двовимірна матриця значень інтенсивності;
- 2) Зображення представляється в градаціях сірого: інтенсивність має значення цілого числа в інтервалі $[0, 255]$, де 0 – чорний, а 255 – білий.
- 3) Координати пікселя (ряд, стовбець). Початок координат точка (0, 0) – верхній лівий кут;
- 4) Точка матриці I з координатами (i, j) – $I[i, j]$.

2.2. АЛГОРИТМ

Структура методу полягає в наступному. В зображенні для кожного пікселя (i, j) обчислюється локальна фрактальна розмірність $D_{i,j}$ для деякого невеликого вікна P . Псевдокод алгоритму методу має вид:

1. I – це матриця розміром $[m, n]$ інтенсивностей.
2. ID – матриця фрактальних розмірностей для кожного пікселя матриці I .
3. Для кожного пікселя $I[i, j]$:
 - 3.1. P – невелике вікно розміром $[k, k]$, що проходить по зображенню $I[i, j]$.
 - 3.2. Обчислюється фрактальна розмірність D для вікна P .
 - 3.3. Назначаємо $D \rightarrow ID[i, j]$.
4. Отримуємо розподілення розмірностей ID і кожній групі назначаємо свій колір. Проводимо

сегментацію зображення $I[i, j]$, оскільки, кожній точці зображення $I[i, j]$ відповідає точка $ID[i, j]$.

2.3. АЛГОРИТМ ОБЧИСЛЕННЯ ФРАКТАЛЬНОЇ РОЗМІРНОСТІ

Особливу увагу потрібно звернути на крок 3.2. Оскільки, фрактальна розмірність є тільки апроксимацією на наших дискретних даних.

Розмір квадратного вікна P вибираємо $[k, k]$, де k -число степені 2 або $k=2^v$, для деяких цілих чисел v . Слід звернути увагу на те, що розмір вікна є парним, отже немає центрального пікселя якому можна поставити у відповідність обчислене значення фрактальної розмірності. Оскільки, вікно пробігає всі піксели по черзі, то обчислене значення розмірності присвоюється лівому верхньому пікселю вікна P .

Ідея обчислення фрактальної розмірності для вікна P базується на покритті множини точок зображення цим вікном та визначенні діапазону інтенсивності, що займають піксели у вікні. Далі, розбиваючи вікно на дрібніші підвікна, проводимо ті ж обчислення. Потім будуюмо залежність $\ln(N(\delta))$ -кількість зайнятого діапазону від $\ln(1/\delta)$ -масштабу та, використовуючи метод найменшого квадрата (МНК), визначаємо кут нахилу даної прямої, яка являє собою фрактальну розмірність вікна.

Вибір розміру вікна рівний числу степені 2 пояснюється зручністю розбиття вікна P на менші підвікна однакового розміру. Наприклад, вікно розміром $P[16,16]$ можна розбити на 4 підвікна розміром $[8, 8]$, на 16 підвікон розміром $[4,4]$ і т.д.

Алгоритм:

1. Для різних значень дробового числа $h \in (0, 1]$
 - 1.1. Ділимо вікно P на $(1/h)^2$ квадратів.
 - 1.2. Розділяємо рівні інтенсивності зображення $[0, 255]$ на $1/h$ рівнів. Нумеруємо ці рівні від 1 до $1/h$.
 - 1.3. Для кожного підвікна $(a,b) \in P$ виконуємо:
 - 1.3.1. Позначаємо u та w мінімальний та максимальний рівні інтенсивності $(1..1/h)$ відповідно в даному підвікні.
 - 1.3.2. Записуємо $w-u+1 \rightarrow n_{a,b}(h)$.
2. Будуюмо залежність $\ln[N(h)]$ та $\ln(h)$ і використовуючи (МНК) знаходимо фрактальну розмірність пікселя для покриваючого множини вікна P .

Вікно $P[k \times k]$ розбивається на підвікна наступним чином $[(k \cdot h) \times (k \cdot h)]$. Для розбиття вікна P h приймає значення $h=1, 1/2, 1/4, \dots, (1/h)^v$. Де параметр v потрібен для визначення розміру вікна в пікселях 2^v .

Для отримання фрактальної розмірності виконуємо підбір параметрів лінійної функції $y=ax+b$ методом найменших квадратів [5], де $x=\ln[1/h]$, $y=\ln[N(h)]$, а параметр $a=-D$. Отже,

$$a = \frac{K_{xy}}{D_x}; \quad (4)$$

де K_{xy} – кореляційний момент, D_x – дисперсія

$$K_{xy} = \frac{\sum_{i=1}^n (x_i - m_x)(y_i - m_y)}{n}; \quad (5)$$

$$D_x = \frac{\sum_{i=1}^n (x_i - m_x)^2}{n}. \quad (6)$$

m_x та m_y – математичне очікування

$$m_x = \frac{\sum_{i=1}^n x_i}{n} \quad (7)$$

$$m_y = \frac{\sum_{i=1}^n y_i}{n}. \quad (8)$$

2.4. РЕЗУЛЬТАТИ

Описані вище алгоритми дозволяють отримати оцінку фрактальної розмірності зображення з різними її значеннями та на цій основі проводити сегментацію і класифікацію зображень.

Але не визначеним питанням є те, яким слід вибирати розмір вікна для покриття множини, по якому буде знаходитись фрактальна розмірність. Оскільки до розміру ставляться зовсім протилежні вимоги:

- для більш точного визначення границь між різними текстурами необхідно зменшувати розмір вікна;
- для підвищення точності обчислення фрактальної розмірності необхідно його збільшувати.

Вибір розміру вікна робиться емпірично при задовільних результатах виділення контурів та сегментації текстур.

Для сегментації були вибрані чотири різних за своєю структурою текстур (рис.1).

На рис.2. представлена карта фрактальних розмірностей – фрактальні розмірності кожної точки текстури.

Заключною стадією є сегментація. З рис.3. видно, що після цієї операції кожен сегмент має декілька точок іншого сегмента. Цей результат можна пояснити тим, що в текстарах зустрічаються подібні по своїй структурі частини зображення.

Однією з переваг такого метода є те, що сегментація текстур не залежить від інтенсивності (кольору) текстури, а залежить тільки від її форми.

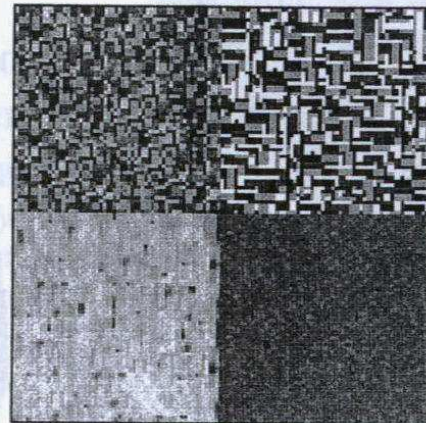


Рис. 1. Об'єднання текстур.
Розмір 120 x 120.

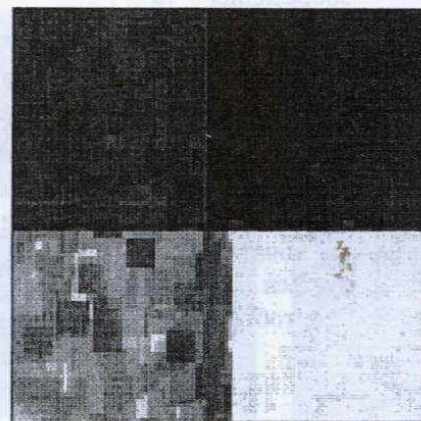


Рис. 2. Карта фрактальних розмірностей текстури

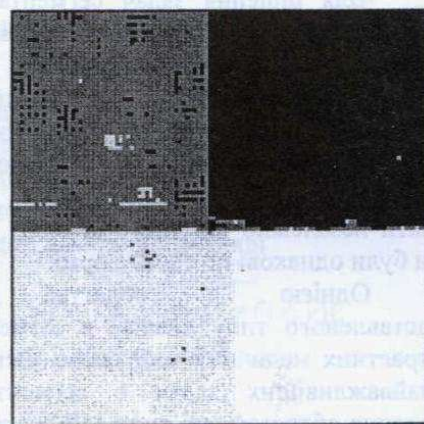


Рис. 3. Результат сегментації текстур.



Рис. 4. Низькоконтрастне зображення



Рис. 5. Карта фрактальних розмірностей

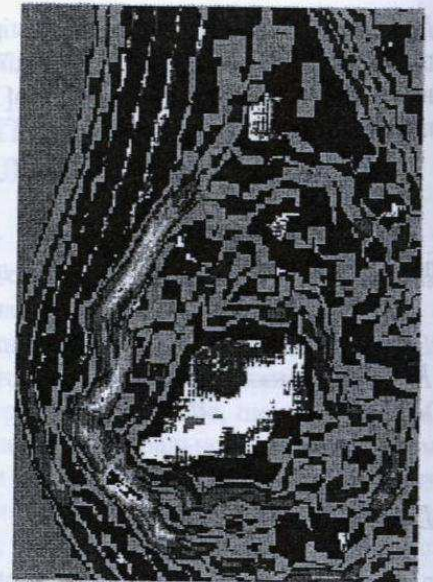


Рис. 6. Результат сегментації

По низькоконтрастному зображенню (рис. 4) було обчислено фрактальні розмірності (рис. 5) на основі яких зроблено сегментацію зображення (рис. 6). На рис. 7 чітко виділені групи розмірностей. Бачимо, що початкове зображення є множина, яка характеризується набором фрактальних розмірностей, тобто є мультифракталом.

ВИСНОВОК

В цій статті був розроблений метод мультифрактальної декомпозиції з допомогою якого була проведена сегментація текстур та низькоконтрастних зображень.

Для рішення задач сегментації зображень застосовується багато різних методів, але тільки з допомогою фрактального аналізу, застосовуючи метод мультифрактальної декомпозиції, при сегментації зображення та класифікації об'єктів, буде враховуватись їх структура (геометрія).

Даний метод відносить до одного класу об'єкти незалежно від його кольору, головне, щоб вони були однакові по своїй формі.

Однією з областей застосування представленого типу аналізу є сегментація низькоконтрастних медичних зображень. Оскільки одними з найважливіших задач є сегментація пухлин, виявлення областей, які вони займають та виділення країв цих областей.

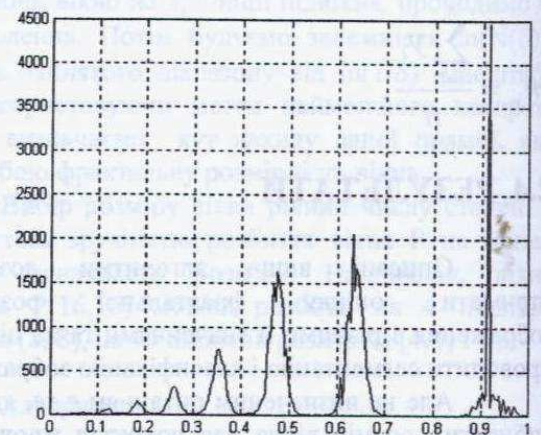


Рис. 7. Розподіл фрактальних розмірностей, розбитих в діапазоні [0..1]

ЛІТЕРАТУРА

1. Федер Е. "Фракталы": Пер. с англ.-М.; Мир, 1991.-254 с.
2. Горобець Ю.І., Кучко А.М. "Вступ до фізики фрактальних структур".-Київ; 2000.
3. Потапов А.А. Герман А.А. // "Радиотехника и электроника".-М.; 2000. Т.-45 N8, с 946-953.
4. Kenkel N.C. Walker D.J."Fractal in the biological sciences"
5. Вентцель Е.С. "Теория вероятностей".-М.;1969.-576 с.
6. Chen, C. C., J. S. Daponte, and M. D. Fox, "Fractal Feature Analysis and Classification in Medical Imaging" IEEE Transactions on Medical Imaging, 8, pp. 133-142, 1989.
7. Коваленкова Н.В.// "Оптика механическая промышленность".-1991.-N4 с.42-46.

Класифікація текстурних зображень в інваріантному просторі моментів Цернакі аналітичних автокореляційних функцій.

Довженко Оксана Володимирівна, Олександр Мубаркович Ахметшин

Дніпропетровський національний університет, Радіофізичний факультет, вул. Наукова 13,
корпус 12, м. Дніпропетровськ 49050. тел. 056-7769092. e-mail: dovzhenko@mail.dsu.dp.ua,
akhm@mail.dsu.dp.ua.

В роботі розглянуто новий метод класифікації текстурних зображень, представлених аналітичними автокореляційними функціями (АКФ) в інваріантному просторі ознак моментів Цернакі на основі апарату векторної підтримки. В рамках експериментів було досліджено інваріантну стійкість ознак АКФ еталонних текстурних зображень до перетворень зсуву, масштабу і повороту.

1. Вступ

Виділення стійких інформаційних ознак текстурних зображень представляє собою одну з найбільш важливих і актуальних проблем на перетині областей штучного інтелекту і "комп'ютерного бачення", що обумовлено її великою практичною значимістю для задач ранньої медичної діагностики, дистанційного зондування і контролю якості матеріалів.

З необхідності автоматизації процедур класифікації текстур постає вимога переконливих теоретичних гарантій інваріантності наборів описових ознак до зсуву, масштабу і повороту, у зв'язку з чим у цій області було досліджено цілий ряд методів, до найбільш популярних з котрих можна віднести статистичний, спектрально-кореляційний підходи [1, 2], а також поширені останнім часом методи виділення інформаційних ознак багаторівневої вейвлет-декомпозиції [3], однак проблема залишається далекою від розв'язку.

Остання обставина стимулює до відшукування нових підходів, в тому числі оснований на стандартних методах спектрально-кореляційної класифікації.

Візуальний аналіз АКФ вихідного зображення або енергетичного Фур'є-спектра (рис. 1) показує, що головним недоліком даних підходів є недостатня чутливість до виділення характерних інформативних ознак кожної з текстур, крім того, ці характеристики не мають властивості інваріантності.

Успішне використання математичного апарату кореляційного аналізу, в першу чергу, пов'язане з відшукуванням методу підвищення чутливості АКФ до виділення характерних ознак зображень текстур, подруге, вирішенням проблеми інваріантності описових характеристик, для чого в рамках даної роботи було використано метод моментів Цернакі.

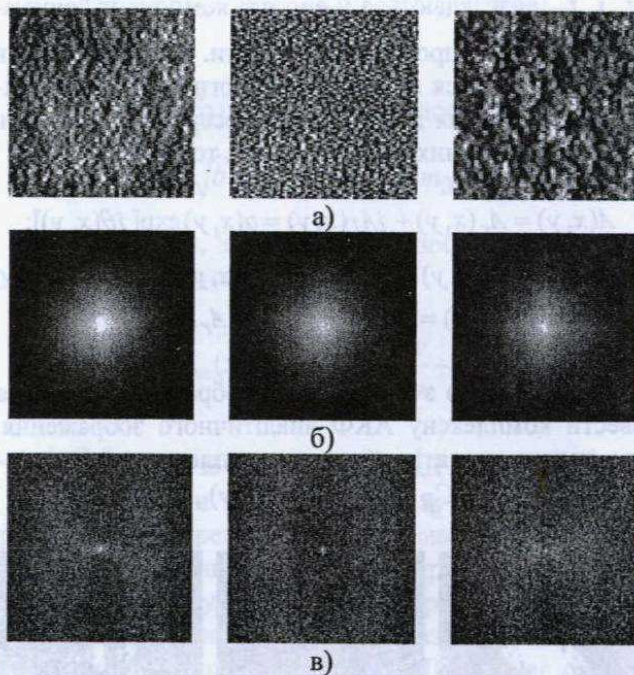


Рисунок 1. Кореляційні і спектральні характеристики, обчислені на основі стандартних підходів: а – вихідні зображення текстур; б – АКФ; в – енергетичні Фур'є спектри.

2. Аналітичні АКФ як апарат виділення інформаційних ознак текстур.

Задачу підвищення чутливості АКФ було вирішено на основі використання концепції "аналітичного зображення", що відома як метод амплітудно-частотної демодуляції [4], в рамках котрої зображення $S(x, y)$ представляється у вигляді

$$S(x, y) = a(x, y) \cos[2\pi f(x, y)];$$
$$f(x, y) = f_{cp} + \Delta\theta(x, y), \quad (1)$$

де припускається, що функція $a(x, y)$ повільно змінюється у зрівнянні з $\cos[2\pi f(x, y)]$. Функція $f(x, y)$ розглядається у вигляді аналогу "віртуальної" несучої частоти, а $\Delta\theta(x, y)$ – компонента, що відповідає фазовій модуляції. В цьому розумінні $a(x, y)$ представляє собою компоненту амплітудної модуляції, а $f(x, y)$ –

частотної. Оскільки f_{cp} в (1) не являє інтересу, виділення інформативних характеристик $a(x,y)$ та $\theta(x,y)$ базується на формуванні нового комплексного зображення $z(x,y)$

$$z(x,y) = S(x,y) + jq(x,y) = A(x,y) \exp[-j(xf'_x + yf'_y)], \quad (2)$$

де $q(x,y)$ розглядається як зображення, сполучене за Гільбертом з вихідним $S(x,y)$, або

$$q(x,y) = \frac{1}{\pi^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{S(u,v)}{(u-x)(v-y)} du dv,$$

f_x і f_y розглядаються у вигляді компонент "віртуальної" несучої просторової частоти. Функція $A(x,y)$ в (2) розглядається як комплексна огинаюча аналітичного зображення $z(x,y)$, що безпосередньо залежить від інформативних характеристик, тобто

$$A(x,y) = A_r(x,y) + jA_I(x,y) = a(x,y) \exp[j\theta(x,y)];$$

$$a(x,y) = (A_r(x,y) + A_I(x,y))^2;$$

$$\theta(x,y) = \arctg[A_I(x,y)/A_r(x,y)].$$

За аналогією з аналітичним зображенням, можна ввести комплексну АКФ аналітичного зображення, що також характеризується комплексною огинаючою $C(x,y) = C_R(x,y) + jC_I(x,y)$.

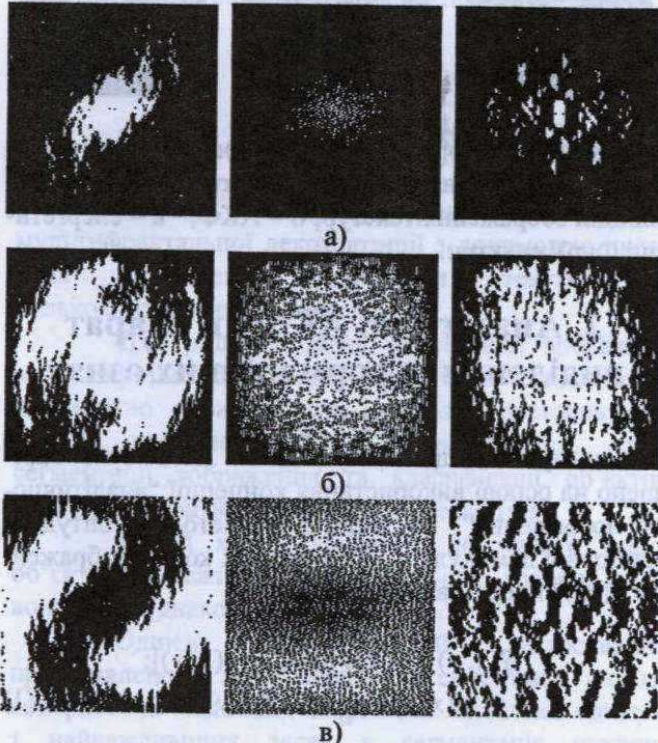


Рисунок 2. Характеристики комплексної огинаючої АКФ текстурних зображень: а - бінаризовані $C_R(x,y)$; б - амплітудно-просторові характеристики $C_R(x,y)$; в - фазо-просторові характеристики $C_R(x,y)$.

На рис. 2 представлені бінаризовані характеристики АКФ, також її амплітудно-просторові і фазо-просторові характеристики, що відповідають вихідним текстурним зображенням, показаним на рис. 1, а, візуальний аналіз котрих дозволяє зробити висновок про високу чутливість нового підходу до виділення значних інформативних характеристик.

3. Представлення аналітичних АКФ в інваріантному просторі моментів Церніке.

Моменти і функції моментів в якості інформаційних характеристик широко використовуються в цілому ряді практичних задач [4]. Один з найбільш поширених підходів базується на відшуванні так званих регулярних моментів, що в дискретній формі можуть бути представлені

$$m_{pq} = \sum_x \sum_y x^p y^q f(x,y), \quad (3)$$

де $(p+q)$ - порядок моменту дискретної функції зображення $f(x,y)$. На основі виразу (3) ґрунтуються сім інваріантних до зсуву, масштабу і повороту функцій, відомих як інваріантні моменти, однак, базис $x^p y^q$ не має властивості ортогональності, у зв'язку з чим реконструкція зображення на основі цих моментів вимагає значного обсягу обчислень. Для вирішення цієї проблеми було запропоновано ортогональний базис поліномів Церніке [5], коефіцієнти котрого можуть бути використані як інваріантні ознаки зображень.

Набір поліномів Цернікі $\{V_{nm}(x,y)\}$, визначених на одиничному колі $x^2 + y^2 = 1$ в загальному вигляді може бути представлений

$$V_{nm}(x,y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta),$$

де n - порядок поліному, m - параметр повтору, ρ - відстань від центру одиничного кола до пікселя (x,y) , θ - кут між вектором ρ та віссю x в напрямку проти годинникової стрілки. Радіальні поліноми $R_{nm}(\rho)$ визначаються

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! s! \left(\frac{n-|m|}{2} - s\right)!} \rho^{n-2s},$$

причому $R_{nm}(\rho) = R_{nm}(-\rho)$.

Ці поліноми ортогональні і задовільняють

$$\iint_{x^2+y^2 \leq 1} [V_{nm}(x,y)]^* V_{pq}(x,y) dx dy = \frac{\pi}{n+1} \delta_{np} \delta_{mq}$$

$$\text{при } \delta_{ab} = \begin{cases} 1 & a = b \\ 0 & \text{в іншому випадку.} \end{cases}$$

Моменти Церніке – це проекція функції зображення на базис ортогональних поліномів; для дискретного зображення вираз для комплексних коефіцієнтів розкладення записується

$$A_{nm} = \frac{n+m}{\pi} \sum_x \sum_y f(x,y) V_{nm}^*(\rho, \theta), \quad x^2 + y^2 \leq 1$$

В якості інформативних ознак (характеристик Церніке) використовуються амплітудні характеристики $|A_{nm}|$, обчислені для різних порядків поліномів.

Припустимо, вихідне текстурне зображення повертається на довільний кут α , в такому разі зв'язок між вихідним і перетвореним зображенням може бути представлений у вигляді виразу:

$$f^*(\rho, \theta) = f(\rho, \theta - \alpha).$$

Моменти Церніке повернутого зображення, представлені в полярній системі координат, записуються

$$A_{nm}^* = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^1 f(\rho, \theta - \alpha) R_{nm}(\rho) \exp(-jm\theta) \rho d\rho d\theta. \quad (4)$$

Якщо позначити

$$\theta_1 = \theta - \alpha, \quad p = \frac{n+1}{\pi},$$

то (4) набуде вигляду:

$$\begin{aligned} A_{nm}^* &= p \int_0^{2\pi} \int_0^1 f(\rho, \theta_1) R_{nm}(\rho) \exp(-jm(\theta_1 + \alpha)) \rho d\rho d\theta_1 = \\ &= \left[p \int_0^{2\pi} \int_0^1 f(\rho, \theta_1) R_{nm}(\rho) \exp(-jm\theta_1) \rho d\rho d\theta_1 \right] \exp(-jm\alpha) = \\ &= A_{nm} \exp(-jm\alpha). \end{aligned} \quad (5)$$

Рівняння (5) показує, що перетворення повороту зображення може бути виражене через зсув фазової характеристики в той час, як амплітудна характеристика $|A_{nm}|$ лишається незмінною.

Інваріантність до зсуву та масштабу досягається попередньою нормалізацією зображення на основі регулярних інваріантних моментів m_{pq} [6]. Спочатку вихідне зображення $f(x,y)$ трансформується за допомогою моментів першого порядку до вигляду $f(x+\bar{x}, y+\bar{y})$, де \bar{x} і \bar{y} представляють собою центроїди $f(x,y)$

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}.$$

Іншими словами, початок координат переміщується в точку з координатами центроїдів.

Властивість масштабної інваріантності отримуємо збільшенням/зменшенням зображення так, щоб регулярний момент нульового порядку дорівнював змінній масштаба β .

Нехай $f(x/a, y/a)$ представляє собою масштабовану версію зображення. Тоді регулярні моменти порядку pq вихідного і перетвореного зображення будуть пов'язані на основі виразу:

$$\begin{aligned} m_{pq}^* &= \iint_{xy} x^p y^q f\left(\frac{x}{a}, \frac{y}{a}\right) dx dy = \\ &= a^{(p+q+2)} \iint_{xy} x^p y^q f(x,y) dx dy = a^{(p+q+2)} m_{pq}. \end{aligned}$$

Оскільки необхідно досягти рівності $m_{00} = \beta$, то

$$a = \sqrt{\beta / m_{00}}, \quad m_{00}^* = a^2 m_{00} = \beta.$$

Таким чином, нормалізоване зображення в результаті має наступний вигляд

$$f_{norm}(x,y) = f\left(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y}\right).$$

4. Метод векторної підтримки.

Класифікатори на основі методу векторної підтримки (МВП) представляють собою чутливий і потужний в розумінні спроможності працювати в просторі параметрів дуже високої розмірності апарат для вирішення задач розпізнавання образів [7].

В рамках спрощеного опису, МВП являє собою задачу побудови гіперплощини, що розділяє вхідні дані $\{x_i, y_i\}$, $i=1, \dots, m$; $y_i \in \{-1, +1\}$, $x_i \in R^n$ таким чином, що забезпечується максимальна відстань між найближчими точками.

В загальному (лінійно нероздільному випадку) задача побудови гіперплощини базується на вирішенні оптимізаційної проблеми виду:

$$\min\left(\frac{|w|^2}{2} + C(\sum_i \xi_i)\right),$$

за обмежень $y_i((w \cdot x_i) + b) \geq 1$, для $\forall i$.

де w – нормаль до гіперплощини, b – зміщення, ξ_i – «ковзаюча» змінна, що разом з константою C визначає припустиму похибку класифікації. Для вирішення оптимізаційної проблеми використовується математичний апарат Лагранжа [8].

Узагальнення методу до нелінійного випадку досягається відображенням вхідних ознак x_i в деякий багатовимірний простір, розмірність якого вище, ніж розмірність простору вхідних значень на основі виразу $\Phi: X \rightarrow F$; після чого застосовується стандартний алгоритм МВП.

Для здійснення відображення використовуються так звані функції ядра, для котрих виконується умова Мерсера. Аналіз публікацій показує, що найбільш поширеними практично є наступні функції ядра: ядра на основі радіальних базисних функцій, поліноміальні ядра, гіперболічні тангенціальні ядра.

4. Експериментальна частина.

Експерименти проводилися на прикладах 61-го зображення різних фізичних текстур з міжнародної бази даних.

Кожна бінаризована характеристика аналітичної АКФ текстур представлялася в просторі інваріантних моментів Церніке, обчислених для різних порядків поліномів і представлених амплітудними характеристиками.

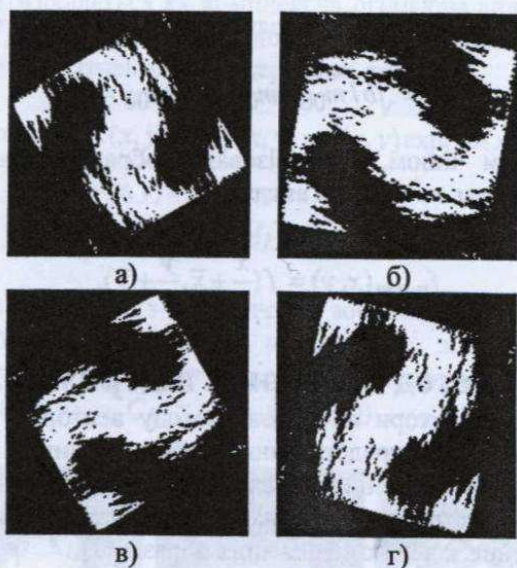


Рисунок 2. Поворот бінаризованої аналітичної кореляційної функції на а) 30°; б) 80°; в) 300°; г) 345°.

В табл. 1 ілюструється інваріантність амплітудних характеристик Церніке порядків $n=2, 3, 4$ до перетворення повороту бінаризованої АКФ, окремі приклади котрого представлені на рис. 3. В табл. 2 наведені значення стандартного відхилення кожного з наборів характеристик і його відношення до маточкування, аналіз котрих показує дуже низький ступінь розсіювання характеристик для різних кутів повороту АКФ відносно середнього значення.

Таблиця 1. Амплітудні характеристики моментів аналітичної кореляційної функції для перетворень повороту і масштабу.

Перетворення	Амплітудні характеристики моментів			
	$ A_{20} $	$ A_{22} $	$ A_{31} $	$ A_{42} $
Поворот				
0°	1605.4	2192.2	7506.9	2005.0
30°	1593.2	2225.9	7644.8	1997.0
45°	1614.0	2229.8	7657.5	1974.1
60°	1611.4	2230.0	7660.2	1989.6

Таблиця 1. Продовження.

Перетворення	Амплітудні характеристики моментів			
	$ A_{20} $	$ A_{22} $	$ A_{31} $	$ A_{42} $
Поворот				
70°	1594.9	2221.2	7629.9	1991.8
80°	1615.5	2220.0	7620.8	1999.4
300°	1593.2	2225.9	7644.8	1997.0
345°	1604.5	2226.2	7649.6	1970.8

Для класифікації, котра проводилася на основі характеристик Церніке до тридцятого порядку (256 ознак для кожної текстури; як було показано в [5], саме такого порядку поліномів достатньо для відновлення зображення, візуально подібного вихідному) було вибрано метод векторної підтримки, що обумовлено його високою точністю, на основі котрого була досягнуто 99% класифікація текстур.

Таблиця 2. Аналіз амплітудних характеристик моментів Церніке.

Параметри	$ A_{20} $	$ A_{22} $	$ A_{31} $	$ A_{42} $
σ	9.3013	12.308	50.215	12.16
σ/μ	0.0058	0.0055	0.0065	0.0061

4. Висновки

На основі проведених експериментів можна заключити, що метод класифікації на основі інваріантних характеристик Церніке аналітичних АКФ представляє чутливий і стійкий спосіб виділення інформативних ознак текстурних зображень.

Література

1. Haralick R. H. *Statistical and structural approaches to texture*. // Proc. IEEE, 1979, v. 67, pp. 786-804.
2. Tomita F., Tsuji S. *Computer analysis of visual textures*. – Kluwer Academic Publishers, Boston, 1990.
3. Simoncelli E. P. and Portilla J. *Texture Characterization via Joint Statistics of Wavelet Coefficient Magnitudes*. // 5th IEEE Int'l Conf on Image Processing. - Chicago, IL. Oct 4-7, 1998.
4. S. A. Dudani, K. J. Breeding, and R. B. McGhee, *Aircraft identification by moment invariants*, IEEE Trans. Comput., vol. C-26, no. 1, pp. 39-45, Jan. 1983.
5. A. Khotanzad and Y. H. Hong, *Rotation invariant pattern recognition using Zernike moments*, in Proc. 9th ICPR, Rome, Italy, Nov. 1998, pp. 326-328.
6. M. K. Hu, *Visual pattern recognition by moments invariants*, IRE Trans. Inform. Theory, vol. IT-8, pp. 179-187, Feb. 1962.
7. Ахметшин О. М., Довженко О. В., *Підвищення ефективності класифікації в багатовимірному просторі ознак великої розмірності на основі методу векторної підтримки*, Збірка праць МКІМ-2002, т. 4, сс. 36-41, 2002.
8. V. Vapnik, *The nature of statistical learning theory*, Springer – Verlag, N. Y., 1995.

Перетворення палітри кольорових зображень

Роман Воробель, Ігор Журавель, Костянтин Сущик

Фізико-механічний інститут НАН України, 79601, Львів, Наукова5,
vorobel@ah.ipm.lviv.ua, тел. (0322) 654803

Transformation of the color image palette. Roman Vorobel, Igor Zhuravel, Kostantyn Suschyk. The article is devoted to methods, connected with transformation palette of image. The method of optimum transformation palette of image by using a theory of clasterization. The results of test of proposed image processing are shown.

1. Вступ

Одним з важливих чинників, які суттєво впливають на швидкість обробки та передачі інформації, є її об'єм. В даній роботі об'єктом досліджень є цифрові зображення, як один з найбільш зручних способів представлення інформації при неруйнівному контролі, медичній діагностиці та дистанційному зондуванні Землі.

Графічні файли таких зображень, в більшості випадків, потребують значних об'ємів пам'яті, що зумовлює певні незручності при обробці та передачі інформації. Тому актуальною є задача зменшення об'єму графічних файлів зображень при мінімальному погіршенні їх візуальної якості. Вирішення цього питання є також важливим у задачах сегментації кольорових зображень.

Усі зображення можна умовно розділити на чотири типи - бінарні, напівтонові, палітрові та повноколірні. Бінарні та напівтонові зображення представляють собою двохвимірні масиви чисел, які є еквівалентами яскравостей. Повноколірні зображення зберігаються у вигляді тривимірних масивів. Для доступу до значень інтенсивностей складових кольору піксела зображення потрібно вказати рядок - k , стрічку l та номер складової: 1 - для R , 2 - для G та 3 - для B . Палітрові зображення зберігаються у вигляді двохвимірних масивів індексів. Для кожного палітрового зображення існує двовимірний масив палітри, в трьох стовпцях якого розміщені значення інтенсивностей R , G , B .

Кожний графічний файл складається із двох основних частин - заголовка та власне даних. Заголовок містить інформацію про структуру графічного файлу. Здебільшого після заголовку у файлі палітрового зображення міститься палітра кольорів. В даній роботі ми розглядатимемо палітрові повноколірні зображення з метою зменшення кількості кольорів при мінімізації втрати їх якості. Тому з аналізу структури графічного файлу випливає, що одним із шляхів зменшення його

об'єму, є оптимізація палітри зображень. Розглянемо відомі підходи до вирішення цього питання [1].

Для більшості зображень, кількість кольорів, яка необхідна для їх відтворення, менша за 256. Забираючи невикористані кольори із палітри сприяємо зменшенню об'єму файлу зображення. Однак даний підхід неефективний для зображень, які містять велику кількість кольорів та відтінків.

Нами запропоновано інший підхід, який базується на тому, що палітра зображення вибирається фіксованою або оптимізується. У випадку фіксованої палітри кожному елементу зображення вибирається у відповідність найближчий за значенням колір. Основною вимогою до такої палітри є те, що вона повинна містити основні кольори, які присутні на зображенні. Виконання цієї вимоги буде в значній мірі впливати на якість відтворення зображення. В більшості випадків використовують так звану "безпечну палітру", яка містить набір часто використовуваних кольорів та їх відтінків. Оскільки розмір палітри є фіксований, а колірне наповнення зображень є найрізноманітнішим, то не завжди вдається провести перетворення палітри зображення із допустимою мінімальною візуальною похибкою. Тому доцільно оптимізувати палітру.

У випадку оптимізованої палітри, формується список усіх кольорів, які містяться на зображенні. Далі на підставі аналізу частоти появи цих кольорів будується конкретна палітра.

Розглянемо процес аналізу та вибору кольорів оптимізованої палітри детальніше.

2. Оптимізація палітри зображення

Однією з найбільш зручних видів представлення яскравісних характеристик зображення є його гістограма. У випадку кольорових (RGB) зображень гістограма представляється у вигляді куба (рис. 1), всередині якого розміщена інформація про частоту появи кольорів на зображенні.

Побудова тривимірної гістограми зображення є важливим елементом при формуванні оптимальної палітри. Далі на підставі аналізу будують нову палітру зображення. Розглянемо ряд методів визначення нової палітри зображення.

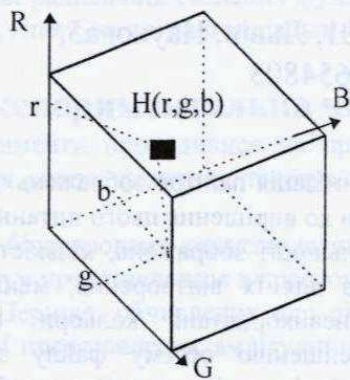


Рис. 1. Структура даних тривимірної гістограми $H(r, g, b)$ зображення.

Один із підходів базується на тому, що формується нова палітра розмірності $n \times 3$, яка складається з найбільш вживаних на зображенні кольорів. Такий алгоритм ефективний для побудови нової палітри зображень, які не містять відтінків основних кольорів. Але більшість зображень містить велику кількість відтінків, тому використання даного методу призводить до того, що палітра перетвореного зображення складатиметься здебільшого з відтінків одного кольору, який найчастіше зустрічається на вхідному зображенні. В результаті такого перетворення, візуальна різниця вхідного та перетвореного зображень є суттєвою, що зумовлює його низьку якість.

Наступним підходом до створення нової палітри, використання якої призвело б до зменшення об'єму графічного файлу зображення, є створення "умовно фіксованої палітри". Суть цього підходу



а)



б)



в)



z)

Рис. 2. Результат комп'ютерного моделювання оптимізації палітри зображення: а) вхідне зображення (розмір графічного файлу зображення 131126 байт); б) зображення а), оброблене не методом бінаризації з використанням алгоритму Флойда-Стейнберга (8254 байт); в) зображення а), палітра якого зменшена до 8 кольорів (32886 байт); з) зображення а), палітра якого зменшена до 8 кольорів та використано алгоритм Флойда-Стейнберга (32886 байт).

полягає у тому, що нова палітра потрібного (довільного) розміру формується на основі використання білого, чорного, *RGB*-кольорів та їх відтінків. Цей метод доцільно застосовувати щодо перетворення палітри зображень, які містять всі сім основних кольорів. Недоліком цього методу є те, що у ньому не проводиться аналіз кольорів зображення. Це призводить до того, що побудована палітра в більшості випадків не є оптимальною.

Як було зазначено вище, більшість зображень містить невелику кількість чистих кольорів та велику кількість їх відтінків. В результаті цього елементи тривимірної гистограми розподілені нерівномірно, а певними скупченнями (кластерами). Ця властивість і лягла в основу запропонованого нами методу. У цьому методі для аналізу гистограми зображення використовують один з методів кластеризації. Він полягає у тому, що всі елементи тривимірної гистограми розбиваються на *k*-кластерів таким чином, що мінімізується евклідова відстань між об'єктами всередині кластерів та максимізується евклідова відстань між ними. Вибираючи координати центра мас заданої кількості *k* найбільших кластерів, проводиться побудова оптимальної палітри із заданою розмірністю *k*.

Зменшення об'єму графічного файлу досягається через оптимізацію палітри призводить до погіршення візуальної якості сприйняття зображення. Для зменшення помітності похибки та створення враження, що на зображенні присутня більша кількість різноманітних кольорів, ніж є насправді, використовується дифузійне псевдозмішування кольорів.

3. Дифузійне псевдозмішування кольорів

Потрібний ефект зменшення помітності похибки досягається за рахунок групування в певній локальній області зображення елементів з кольорами, представленими в палітрі, змішування яких дає близький до необхідного колір. Для цього використовується відомий алгоритм Флойда-Стейнберга [1]. В цьому алгоритмі елемент зображення замінює свій колір на найближчий в евклідовій метриці, який є у палітрі. Також обчислюється похибка колірності *Error_RGB* для кожної складової як різниця між значеннями величин складових кольорів палітри *MAP_new(k,:)* та вхідного зображення *L(i,j,:)*. У випадку *RGB*-складових, алгоритм представляється наступним чином:

$$Error_RGB(i,j,:) = L(i,j,:) - MAP_new(k,:). \quad (1)$$

Далі похибка розповсюджується на сусідні елементи зображення шляхом додавання деякої долі

похибки до значень відповідних складових і відбувається так звана "дифузія похибки".

$$L(i,j+1,:) = L(i,j+1,:) + \frac{3 \cdot Error_RGB(i,j,:)}{8}, \quad (2)$$

$$L(i+1,j,:) = L(i+1,j,:) + \frac{3 \cdot Error_RGB(i,j,:)}{8}, \quad (3)$$

$$L(i+1,j+1,:) = L(i+1,j+1,:) + \frac{Error_RGB(i,j,:)}{4}. \quad (4)$$

Алгоритм застосовується рекурсивно до всіх пікселів зображення. Використання псевдозмішування кольорів дозволяє покращати колірне сприйняття за рахунок погіршення роздільної здатності та підвищення рівня високочастотного шуму.

4. Висновки

Запропонований у роботі метод побудови оптимальної палітри на основі кластерного аналізу гистограми зображення та використання дифузійного псевдозмішування кольорів дозволяє представляти будь-яке кольорове чи напівтонове зображення заданою кількістю кольорів з мінімальною візуальною похибкою. Похибка перетворення зображення обернено пропорційна кількості елементів оптимальної палітри.

Області застосування методу різноманітні. Важливе місце серед них займає передача відеоданих по каналах зв'язку, де об'єми графічних файлів повинні бути невеликі. Приклад ефективності запропонованого методу ілюструє рис. 2. Тут на рис. 2-а показано вхідне зображення з розмірами файлу у 128 КВ. На рис. 2-б наведено застосування класичної технології дифузійного псевдозмішування кольорів через бінаризацію і застосування алгоритму Флойда-Стейнберга [1-3]. На рис. 2-в показано вхідне зображення 2-а зі зменшеною до восьми кольорів палітрою. Результуюче зображення з) отримане із вхідного шляхом оптимізації палітри та застосуванням алгоритму псевдозмішування кольорів Флойда-Стейнберга. За візуальною якістю воно практично не поступається вхідному зображенню, а об'єм його графічного файлу становить лише 32 КВ, що майже у чотири рази менше, ніж у вхідного зображення. Вибором розміру оптимальної палітри шукається компроміс між об'ємом файлу та якістю перетвореного зображення. При кількості кольорів оптимальної палітри більше ніж 16, втрати якості стають візуально непомітними, а виграш у об'ємі значним.

1. Роджерс Д. Алгоритмические основы машинной графики. М.: Мир, 1989.
2. Шлихт Г.Ю. Цифровая обработка изображений. М.: Эком, 1997.
3. Красильников Н.И. Цифровая обработка изображений. М.: "Вузовская книга", 2001.

