

Р. Мельник, Ю. Каличак
Національний університет "Львівська політехніка",
кафедра програмного забезпечення

ПОШУК ОБРАЗІВ ЗА ЇХ ІНВАРІАНТНИМИ ТА ПАРАМЕТРИЧНИМИ ОЗНАКАМИ

© Мельник Р., Каличак Ю., 2012

Наведена методика отримання розподілених інваріантних та параметричних властивостей образів під час фрагментації об'єму інтенсивності зображення. Показано застосування отриманих результатів у разі пошуку зображень за вмістом.

Ключові слова: зображення, образ, простір інтенсивності, об'єм інтенсивності, фрагментація, розподілені інваріантні ознаки, розподілені параметричні ознаки.

An extraction method of distributed invariant and parametric features of patterns based on the image intensity volume fragmentation is presented.

Key words: image, pattern, intensity space, intensity volume, fragmentation, distributed invariant features, distributed parametric features.

Вступ

Системи знаходження зображень за їх вмістом (Content-based image retrieval – CBIR) мають два етапи: індексування та пошук. Центральним елементом систем є вектор властивостей зображення. Тип вектора та алгоритм його визначення ґрунтуються на категоріях та властивостях образу та апарата їх опрацювання: гістограм інтенсивності чи кольору, взаємного розташування кольорових регіонів об'єктів, їх форми, коефіцієнтів вейвлет- чи Фур'є-функцій тощо. Робіт, присвячених алгоритмам екстракції ознак зображень, сьогодні дуже багато. Зокрема в роботах [1–3] здійснюється побудова регіонів, гістограм блоків кольору та їх інваріантних коефіцієнтів. У роботах [4–6] наведені алгоритми побудови форми основних об'єктів уваги та їх кількісних характеристик. Різноманітні характеристики текстури і алгоритми їх обчислення запропоновані в роботах [7–9]. Просторове розташування об'єктів уваги також використовується для класифікації зображень. Зокрема в роботах [10,11] введено поняття „верхній”, „нижній” тощо.

Прикладами систем CBIR є SIMPLIcity [12] VisualSEEK [13]. В обох системах зображення сегментується на регіони, колір та розмірність яких входять у вектори властивостей. На етапі пошуку обчислюються властивості із образу-запиту користувача. Отриманий вектор властивостей порівнюється з векторами у базі даних візуальних властивостей. Користувач у відповідь отримує образи, які максимально відповідають запиту.

Зазначимо, що більшість перерахованих підходів є доволі складними для реалізації та потребують значних затрат часу. У роботі запропонована значно простіша методика отримання декількох векторів властивостей зображення, яку можна застосувати на етапі порівняння образу-запиту з наявними в базі даних. Методика базується на фрагментації тривимірного простору інтенсивності зображення, формуванні сегментів з фрагментів та обчисленні для кожного з них статистичних характеристик, які в сукупності утворюють розподілені властивості зображення.

1. Визначення об'єму інтенсивності зображення

Для отримання тривимірної поверхні зображення спочатку кольорове зображення перетворюється у відтінки сірого. Кожна елементарна клітинка, піксель, приймає значення від чорного до білого кольору, яке позначимо як b -інтенсивність. Діапазон всіх можливих значень інтенсивності перебуває в межах 0–255.

Для перетворення до кожного пікселя застосуємо алгоритм BT709 з коефіцієнтами $R = 0,2125$; $G = 0,7154$; $B = 0,0721$:

$$c_i = R \cdot c_R + G \cdot c_G + B \cdot c_B \quad (1)$$

де c_i ($i = 1, 2, 3$) – значення червоної, зеленої та синьої компонент сірого кольору пікселя, c_R, c_G, c_B – початкові значення компонент кольору пікселя (червоної, зеленої та синьої відповідно).

Значення інтенсивності пікселя обчислюється з рівняння відносного заповнення клітинки у відсотках від білого кольору:

$$b = c \times 100 / 256 \quad (2)$$

де c – значення однієї з компонент сірого кольору пікселя.

Поділимо простір інтенсивності зображення на n фрагментів горизонтальними площинами XOY з інтервалом (кроком фрагментації інтенсивності) $d=255/n$ (рис. 1). На рис. 1, а показано тестове зображення і його поверхню інтенсивності у тривимірному просторі, розділену на фрагменти (рис. 1, б).

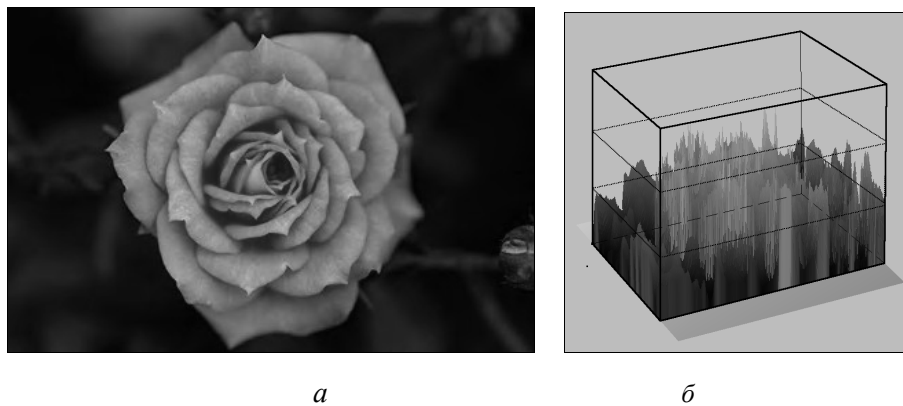


Рис.1. Тестове зображення (а), тривимірне зображення його інтенсивності (б)

Об'єм фрагмента поверхні інтенсивності визначасмо, наприклад, таким способом: за перерізу фрагмента поверхні інтенсивності зображення горизонтальними площинами XOY (кількість площин $d-1$) об'єм є сумою відповідних площ поперечного перерізу поверхні зображення у фрагменті (всіх пікселів, інтенсивності яких попадають у відповідний фрагмент, разом з контуром перерізу).

$$V_F(s) = \sum_{i=s*d}^{s*d+d-1} S_L(b_f(i)), \quad (3)$$

де $V_F(s)$ – об'єм фрагмента, $S_L(b_f(i))$ – площа поперечного перерізу для рівня інтенсивності $b_f(i)$.

Можливий інший спосіб: об'єм визначається підсумовуванням частин значень інтенсивності пікселів, що належать цьому фрагмента:

$$V_F(s) = \sum_{b(x,y) \geq b_f(s)+d} d + \sum_{b_f(s) \leq b(x,y) < b_f(s)+d} (b(x,y) - b_f(s)), \quad (4)$$

де $b(x,y)$ – інтенсивність пікселя, $b_f(s)$ – початковий рівень інтенсивності s -го фрагмента.

Формування сегмента інтенсивності відбувається шляхом його покрокового збільшення на один фрагмент. На рис. 2 нарощування сегментів відбувається на моделі фрагментації за руху згори донизу, тобто від білого до чорного кольорів:

$$V_s(s) = \sum_{i \in 1, \dots, s} V_f(i), \quad (5)$$

де $V_f(i)$ – об'єм одного фрагмента.

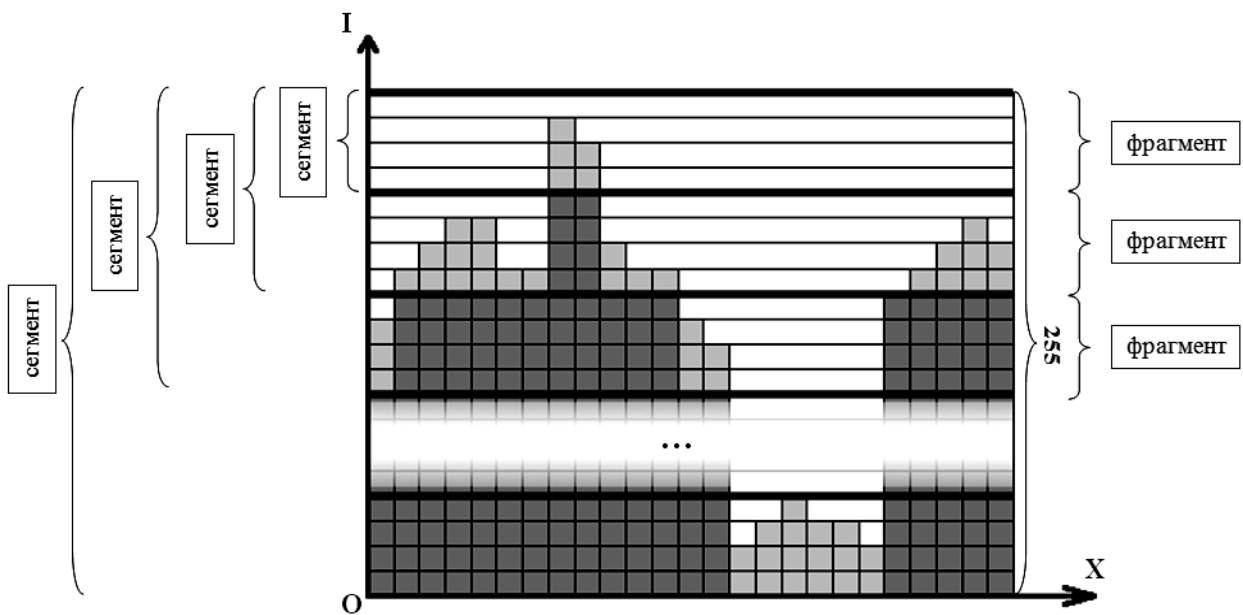


Рис. 2. Фрагменти та сегменти в перетині простору інтенсивності

Об'єми фрагментів та сегментів нормалізуємо максимальним можливим значенням об'єму відповідного фрагмента та сегмента:

$$V_F(s) = V_F(s) / V_{F_{MAX}}, s = 1, n \quad (6a)$$

$$V_S(s) = V_S(s) / V_{S_{MAX}}, s = 1, n \quad (6a)$$

2. Інваріантні ознаки образу

Сегменти містять значні кількості пікселів. Для них визначаємо розподілені ознаки координат пікселів: математичне сподівання координат пікселів сегмента:

$$\bar{x}(s) = 1/k_s(s) \sum_{i \in I(s)} x_i(s), \quad \bar{y}(s) = 1/k_s(s) \sum_{i \in I(s)} y_i(s) \quad (7)$$

та дисперсію координат пікселів у сегменті:

$$E^2(s) = (1/k_s(s)) \cdot \sum_{i \in I(s)} (x_i(s) - \bar{x}(s))^2 + (y_i(s) - \bar{y}(s))^2 \quad (8)$$

де $k_s(s)$ – кількість пікселів у сегменті, $\bar{x}(s)$, $\bar{y}(s)$ – математичне сподівання координат пікселів сегмента, $x_i(s)$, $y_i(s)$ – координати пікселя сегмента, E^2 – дисперсія координат пікселів у сегменті. Дисперсія виступає як інваріантна ознака.

Визначивши математичне сподівання координат пікселів сегмента, додаткову інваріантну ознаку як відстань між двома сусідніми сегментами:

$$D(s-1, s) = \sqrt{(\bar{x}(s-1) - \bar{x}(s))^2 + (\bar{y}(s-1) - \bar{y}(s))^2}, s=2, \dots, n \quad (9)$$

де $D(s-1, s)$ – відстань між точками математичного сподівання (7), двох сусідніх сегментів.

На рис.3 подано тестові зображення облич. Для них на рис. 4–5 показано залежності інваріантних ознак – стандартного відхилення та відстані між двома сусідніми сегментами.

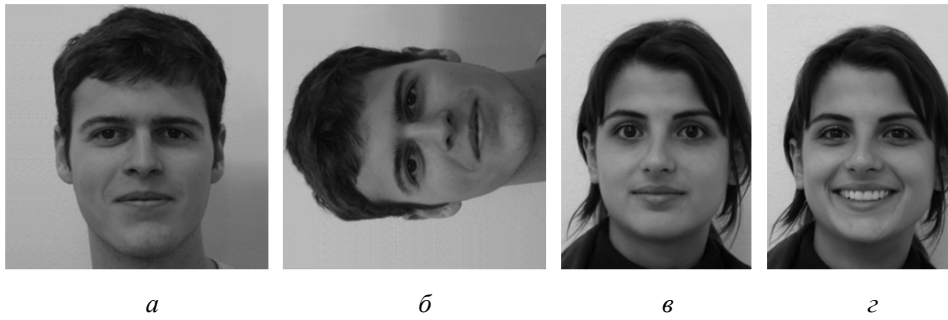


Рис. 3. Тестові зображення облич

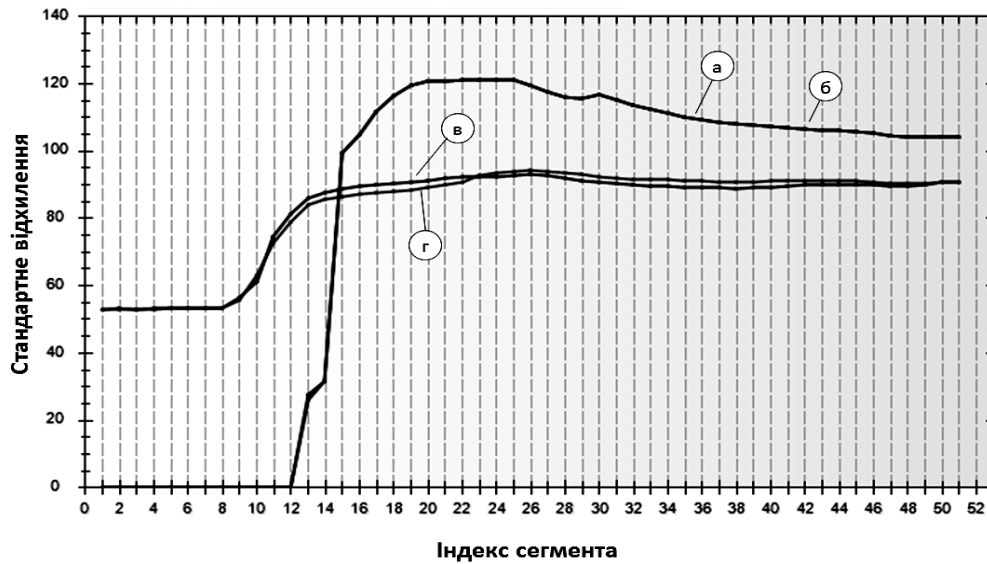


Рис. 4. Стандартне відхилення координат пікселів у сегменті

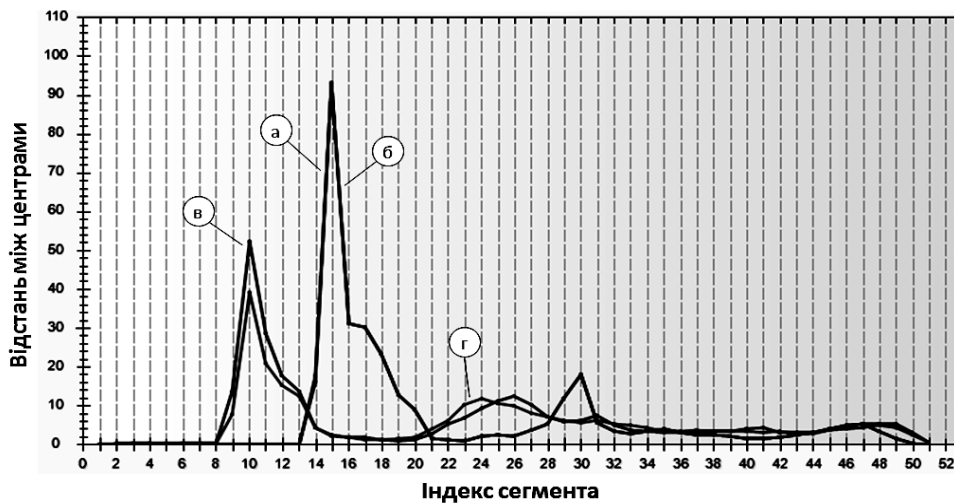


Рис. 5. Відстані між сусідніми сегментами

3. Параметричні ознаки образу

Для отримання ознак, залежних від координат пікселів, поділимо поверхню інтенсивності сіткою з різним кроком розбиття (рис.6). Кожна клітинка має свій порядковий номер. Отже, точка математичного сподівання координат пікселів сегмента (фрагмента) відповідає певній клітинці. Номер клітинки є ознакою сегмента (фрагмента). Точність позиції залежить від розмірів клітинки.

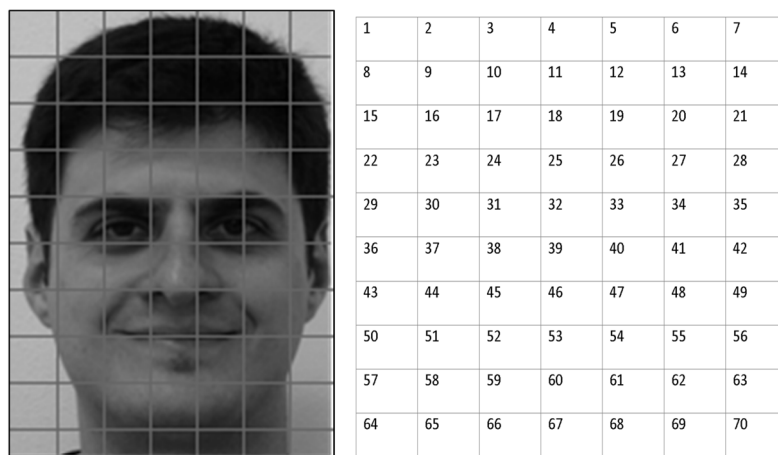


Рис. 6. Тестове зображення обличчя (а) та розбиття його сіткою (б)

На рис. 7 наведено залежності параметричної ознаки для образів із рис. 3. Під час побудови залежності для виокремлення кожного з образів (рис. 3) використовувалося розбиття сіткою 12x12 клітинок з рівномірним кроком.

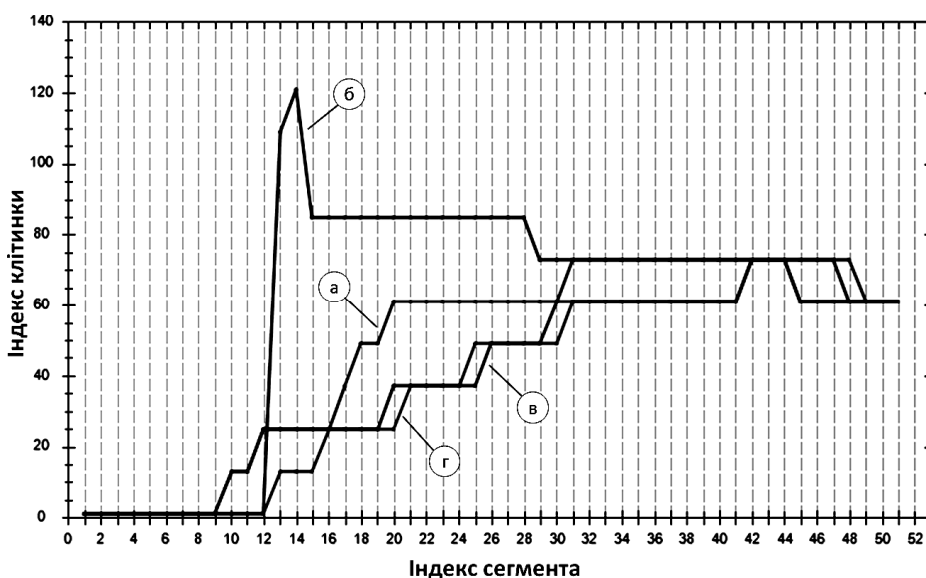


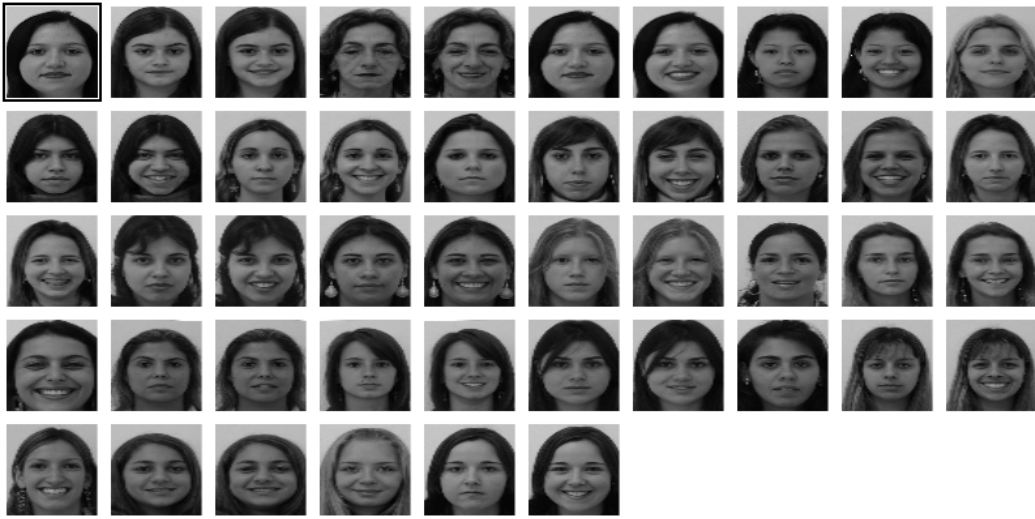
Рис. 7. Індекс клітинки за індексом сегмента

4. Пошук зображень за індексами кластерів

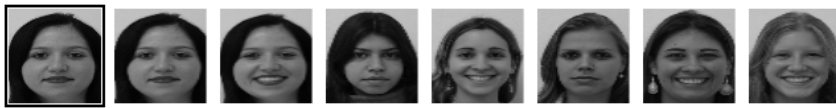
Для тестування прийнято базу з 200 знімків пар облич з різними виразами [14] та базу Ванга [15] з 1000 зображень, розбитих на 10 категорій. Розмірність кривих становить 51 сегмент (5 значень інтенсивності на один фрагмент). Критерієм пошуку використано середньоквадратичне відхилення шуканого зразка від наявних в базі даних екземплярів. Для кривих різних характеристик вирази для похибки тотожні. Використано алгоритм пошуку, наведений у роботі [16]. На ілюстраціях прикладів результатів пошуку перше зображення є образом запитом, решта – результат.

Для оцінювання отриманих результатів як критерії використовуються мінімальна похибка для результатів з усіма релевантними до запиту образами і точність отриманих результатів – відношення кількості релевантних образів до загальної кількості отриманих образів.

На рис. 8 наведені приклади результатів пошуку жіночого обличчя за ознакою середньоквадратичного відхилення координат пікселів у сегменті. Рис. 8, а показує результат пошуку зі 45 зображеннями (26 різних жіночих облич) при похибці 0,25. Рис. 8, б містить 7 результуючих зображень (6 різних облич) при похибці 0,1. Зменшення похибки пошуку до 0,06 звужує ряди знайдених знімків до двох присутніх (рис. 8, в).



а



б



в

Рис. 8. Приклади результатів пошуку облич за стандартним відхиленням

Для цього набору зображень похибка пошуку результату, що містить всі релевантні образи, коливається в межах 0,03–0,17.

Рис. 9 показує приклад результату пошуку обличчя з використанням ознаки відстаней між сегментами. Тут похибка для отримання результату з двома виразами становить 0,86.

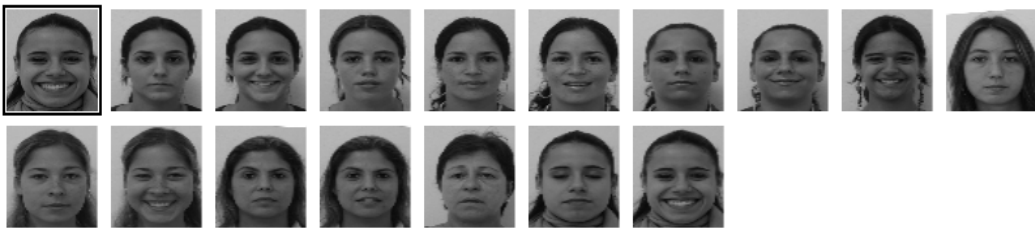


Рис. 9. Приклад результатів пошуку облич за відстанями між центрами

Загалом для відстаней похибка результату з релевантними образами становить 0,18–1,87.

Рис. 10 показує приклад результатів пошуку з використанням індексів клітинок. Пошук виконувався із розбиттям сітки на 100 клітинок (10x10). Для цієї ознаки похибка пошуку обох виразів обличчя коливається в межах 0,58–3,15.

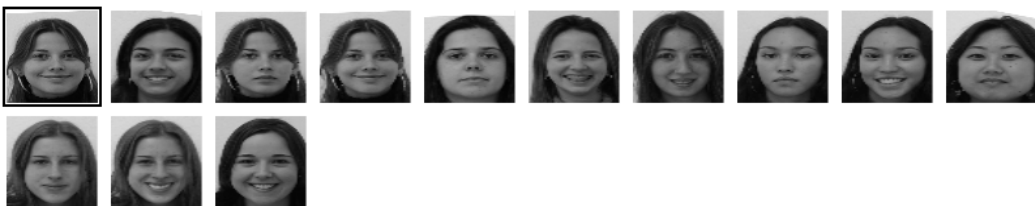


Рис. 10. Приклад результатів пошуку облич за параметричною ознакою

Для бази зображень Ванга результати пошуку з використанням стандартного відхилення коливались від 20 до 100 відсотків залежно від типу образу. Зокрема, найкращі результати отримані для динозаврів (80–100 відсотків) троянд (80–90) (рис. 11, а), коней (60–80). Середніми вважаємо результати пошуку автобусів (50 відсотків) (рис. 11, б).



а



б

Рис. 11. Приклади результатів пошуку зображень з бази Ванга за стандартним відхиленням

Пошук із використанням відстаней між центрами сегментів дає такі результати: для динозаврів точність становить 70–100 відсотків, троянд – 10–30, коней – 40–70. Рис. 12 ілюструє приклад результатів пошуку з використанням відстаней між центрами.



Рис. 12. Приклад результатів пошуку зображень з бази Ванга за відстанями між центрами

Схожі результати отримуємо, застосувавши параметричну ознаку для пошуку. Так для категорії динозаврів результати коливаються в межах 80–100 відсотків, троянд – 60–80 відсотків, коней – 40–60 відсотків.

Для зображень зі значним перемішуванням інтенсивностей у площині (прикладом можна вважати африканські краєвиди, харчі, міські пейзажі тощо) точність пошуку коливалась в межах 10–30 відсотків для усіх зазначених ознак.



Рис. 12. Приклад результатів пошуку зображень з бази Ванга за відстанями між центрами

Результати пошуку залежать від виду зображення: для виразного поділу інтенсивностей, характерного для групи образів, ця група формується алгоритмом як відповідь; якщо зображення не має виразних регіонів, відповідь містить образи різних категорій. У цих випадках однієї статистичної ознаки замало і пошук необхідно здійснювати за двома і більше ознаками. Зменшення похибки призводить до знаходження запиту.

Висновки

Запропоновано підхід до формування ключів графічних образів на основі інваріантних та параметричних ознак, отриманих під час фрагментації простору інтенсивності зображень. Запропоновано використання статистичних залежностей координат пікселів як коефіцієнтів для індексації та пошуку зображень в СВІР-системах. Наведено експериментальні результати проведення такого пошуку.

1. Smith J. R. and Chang S.-F., *Tools and techniques for colour image retrieval* // *Proceedings of Symposium on Electronic Imaging: Science and Technology, Storage & Retrieval for Image and Video Databases IV*, vol. 2670, p.1–12, 1996.
2. H. Nezamabadi-pour, E. Kabir, “Image retrieval using histograms of unicolor and bicolor blocs and directional changes in intensity gradient” // *Pattern Recognition Letters*, vol. 25, n. 14, p. 1547–1557, 2004.
3. T. Gevers, A. Smeulder, *Content-based image retrieval by viewpoint invariant color indexing* // *Image Vision Comput*, n.17, p. 475–488, 1999.
4. F. Mokhtarian, S. Abbasi, “Shape similarity retrieval under affine transforms” // *Pattern Recognition*, vol. 35, p. 31–41, 2002.
5. A.K. Jain, A. Vailaya, “Image retrieval using colour and shape” // *Pattern Recognition*, vol. 29, n. 8, p. 1233–1244, 1996.
6. V. Mezaris, I. Kompatsiaris, M.G. Strintzis, *An ontology approach to object-based image retrieval* // *Proceedings of the ICIP*, vol. II, p. 511–514, 2003.
7. B.S. Manjunath, W.Y. Ma, “Texture feature for browsing and retrieval of image data” // *IEEE PAMI*, vol. 8, n. 18, p. 837–842, 1996.
8. J. Liu, X Zhou, W.Y. Ma, *Extraction of texture features from arbitrary-shaped regions for image retrieval*, *International Conference on Multimedia and Expo* p. 1891–1894, 2004.
9. P.W. Huang, S.K. Dai, *Image retrieval by texture similarity* // *Pattern Recognition* 36 (2003) 665–679.
10. Y. Song, W. Wang, A. Zhang, *Automatic annotation and retrieval of images* // *J. World Wide Web*, n. 6 (2), p. 209–231, 2003.
11. A. Mojsilovic, B. Rogowitz, *ISee: perceptual features for image library navigation*, *Proceedings of the SPIE, Human Vision and Electronic Imaging*, vol. 4662, p. 266–277, 2002.
12. J.Z. Wang, J. Li, G. Wiederhold, “SIMPLiCity: semantic sensitive integrated matching for picture libraries” // *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, n. 9, p. 947–963, 2001.
13. John R. Smith, Shih-Fu Chang, *VisualSEEk: a fully automated content-based image query system* // *Proceedings of the fourth ACM international conference on Multimedia for Image Technology for New Media*, 1996.
14. FEI data base, <http://fei.edu.br>.
15. Wang database, <http://wang.ist.psu.edu/~jwang/test1.tar>, 2009.
16. Мельник Р., Каличак Ю., *Пошук образів за силуетами яскравості зображень* // *Відбір і обробка інформації* // *Вип.32 (108)*. – 2010. – С.86–92.