

# Індексування та пошук зображень за структурними коефіцієнтами

Мельник Р.А, Тушиницький Р.Б.

Кафедра програмного забезпечення  
Національний університет «Львівська політехніка», Україна  
ramelnik@polynet.lviv.ua, ruslan.tushnytsky@gmail.com

## Abstract

The 3-stages clustering algorithm for visual patterns to find structural coefficients describing image properties is presented. Structural coefficients and statistical characteristics of visual fragments are proposed for pattern classification and image searching in the databases.

## 1. Вступ

Для великої бази даних із понад десятками тисяч образів ефективна індексація є важливим інструментом в системах знаходження зображень за їх вмістом (Content-based image retrieval – CBIR). Ця проблема є актуальною в сучасних інформаційних системах. Успішна класифікація зображень зменшує час опрацювання зображень фільтруванням зайвих класів образів під час пошуку подібних до них [1]. Інтернет є прикладом розподіленої бази даних, що містить візуальні образи. Пошук в розподіленій базі даних є складною задачею, один з підходів до вирішення якої розглядається в роботі.

CBIR системи працюють у два етапи: індексування та пошук. На етапі індексування, кожний образ у базі даних представляється вектором властивостей. Існуючі універсальні системи CBIR відносять до однієї із трьох категорій залежно від підходу отримання властивостей образу: гістограма, кольорове розташування, і пошук за регіонами. Такими властивостями зокрема, є: колір [2 – 3], форма [4 – 5], структура [6] і розташування [7]. Отримані властивості зберігаються в окремій базі даних візуальних властивостей. На етапі пошуку обчислюються властивості із образу-запиту користувача. Використовуючи критерії подібності, отриманий вектор властивостей порівнюється з векторами у базі даних візуальних властивостей. Користувач у відповідь отримує образи, які максимально відповідають запиту.

Системи пошуку за регіонами використовують локальні властивості регіонів (ідеальних об'єктів) у протилежність глобальним властивостям повного зображення. Прикладом такої системи є SIMPLiCity [8]. Якщо об'єкти в межах зображення сегментовані і кожна властивість об'єкта отримана автоматично, то такі особливості роблять можливою систему пошуку зображень за регіонами [9]. Представлення візуального образу адекватним числом кластерів (об'єкти у зображенні) може краще відобразити його вміст, однак цей підхід є часозалежним.

В роботі [11] запропоновано навчальний компонент для CBIR системи. Він передбачає контролююче тренування системи на різних фрагментах образу. Szummer та Picard [10] розвинули систему класифікації внутрішніх та зовнішніх сцен. Здійснена класифікація із

врахуванням низькорівневих особливостей зображення, таких як кольорова гістограма і коефіцієнти дискретного косинусного перетворення (DCT). 90-процентна норма точності була повідомлена над базою даних у 1300 зображень від Kodak.

Інші приклади семантичної класифікації зображень включають працю порівняння міста та ландшафту [1] і знаходження обличчя [12]. Wang та Fischler [13] показали, що точне семантичне представлення є корисним для задач порівняння зображень.

Робота присвячена дослідженню структурних властивостей зображення, отриманих алгоритмом триступеневої кластеризації. За результатами досліджень запропоновано використовувати структурні властивості та коефіцієнти як критерії для автоматизованого віднесення зображення до певного класу та його пошуку у базах даних.

## 2. Алгоритм триступеневої кластеризації

При розробці алгоритму декомпозиції візуального образу його завданням є вирішення наступних задач: 1) отримати описи складових образу з усіма можливими характеристиками; 2) виділити тільки певну складову з заданими характеристиками; 3) отримати складові частини інтегральної області, що характеризується різною яскравістю.

Для вирішення поставленого завдання розроблено алгоритм триступеневої кластеризації [15] що базується на побудові ієрархічного дерева згортання та засоби керування [16 – 17] для забезпечення доступу до різних його рівнів для одержання фрагментів різної фізичної природи. Значені рівні формування кластерів різної природи позначені на фрагменті ієрархічного дерева, що наведений на рис. 1.

Згортання полягає в поділі візуальних образів на класи – групи об'єктів, схожі між собою за певними ознаками.

На підготовчому кроці вирішуємо оптимізаційну задачу: необхідно отримати множину  $\bar{O}$ , що складається з мікрооб'єктів  $O = \{o_1, o_2, \dots, o_N\}$ :

$$\left. \begin{aligned} \bar{O} &= \{O_1, O_2, \dots, O_n\}, \\ n &\rightarrow \min, \end{aligned} \right\} \quad (1)$$

що задовольняє наступні умови:

$$F_k^- \leq F_k(X_i) \leq F_k^+, \quad (k = \overline{1, m}; i = \overline{1, n}), \quad (2)$$

де  $n$  – число мікрокластерів  $O_1, O_2, \dots, O_n$ ;  $F_k^-, F_k^+$  – граничні значення функцій ознак кластерів, що

формується. Це, наприклад, яскравість, колір, співвідношення заповнених та пустих клітинок тощо.

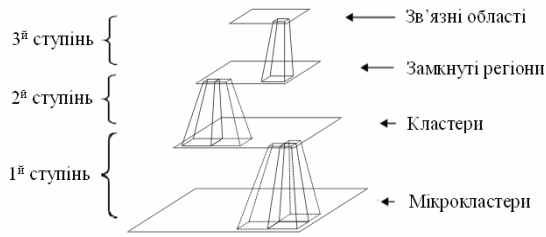


Рис. 1: Фрагмент ієрархічного дерева

### 3. Структурні характеристики образів

В результаті роботи алгоритму триступеневої кластеризації із зображенням отримуємо певні його структурні характеристики, зокрема:  $MC$  – кількість мікрокластерів (об'єктів для формування кластерів),  $C$  – кількість кластерів (об'єктів для формування регіонів) після першого етапу алгоритму,  $CR$  – кількість регіонів (об'єктів для формування областей) після другого етапу алгоритму,  $IA$  – кількість областей після третього етапу алгоритму, а також параметри яскравості всіх об'єктів.

Для оцінки ступеня структуризації зображення або частини введемо ряд параметрів, зокрема:

1) коефіцієнти *ступеня структуризації* одного рівня формування складових об'єктів зображення (індекс вказує на рівень алгоритму): овольняє наступні умови:

$$K_s^1 = C / MC, \quad K_s^2 = CR / C, \quad K_s^3 = IA / CR \quad (3)$$

2) коефіцієнти *ступеня структуризації* двох рівнів формування складових об'єктів зображення:

$$K_s^{12} = CR / MC, \quad K_s^{23} = IA / C \quad (4)$$

3) коефіцієнт *ступеня структуризації* трьох рівнів формування складових об'єктів зображення:

$$K_s^{123}(MC) = IA / MC \quad (5)$$

Результати тестування алгоритму над певними класами зображень показали, що із зменшенням якості зображення, коефіцієнти розмитості збільшуються, а коефіцієнти структуризації зменшуються у порівнянні із початковим зображенням.

Для класифікації повних зображень, в т.ч. що містять білі фрагменти, за їх структурними властивостями коефіцієнт  $K_s^{123}$  структуризації зображення віднесемо до повної кількості пікселів, з яких воно складається. Тобто у формулі (5) замінимо значення  $MC$  на кількість всіх пікселів зображення  $PX$ :

$$K_s^{123}(PX) = IA / PX \quad (6)$$

Кожний виділений  $i$ -ий фрагмент, зокрема, область характеризуємо відносним числом пікселів:

$$vx_i = px_i(IA) / PX, \quad (7)$$

для яких є справедливою нерівність:

$$\sum px_i(IA) \leq PX, \quad (8)$$

де сума береться по всіх  $IA$  сформованих областях.

Якщо визначити середній розмір області в пікселях:

$$M(IA) = (1 / IA) \cdot \sum px_i \quad (9)$$

то для характеристики структуризації введемо характеристику дисперсії розмірів областей, що покривають зображення:

$$D(IA) = \sqrt{(1 / IA) \cdot \sum (px_i - M(IA))^2} \quad (10)$$

### 4. Класифікація зображень за структурними коефіцієнтами

Дослідження алгоритму та структурних коефіцієнтів проведені над базою із 1000 образів, згрупованих у 10 семантичних груп: леви, слони, коні, квіти, їжа, гори, автобуси, дизайн, зображення текстури, медичні образи. Кожна група містить в собі 100 зображень у форматі JPEG із розмірами  $256 \times 384$  пікселів. Приклади образів подано на рис. 2.

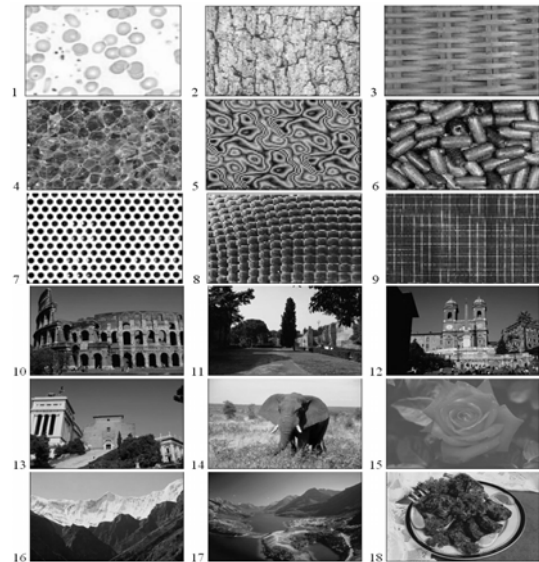


Рис. 2: Пронумеровані зразки досліджуваних образів

Результати декомпозиції образів з рис. 2 та їх сортування за кількістю регіонів  $IA$ , коефіцієнтами структуризації  $K_s^{123}(MC)$  та  $K_s^{123}(PX)$  зведені у таблицю.

Також у таблиці наведені структурні властивості фрагментів образів, а саме: середній розмір фрагмента  $M(IA)$  та дисперсія розмірів фрагментів  $D(IA)$ . Інтегральна класифікація образів у таблиці отримана сортуванням суми місць образів, отриманих класифікацією тільки за одним із параметрів  $IA$ ,  $K_s^{123}(MC)$ ,  $K_s^{123}(PX)$ ,  $D(IA)/M(IA)$ .

Таблиця : Класифікація образів за структурними характеристиками

Місце	IA	№	$K_s^{123}(MC)$	№	$K_s^{123}(PX)$	№	$M(IA)$	$D(IA)$	Інтегр. класифі.
1	193	4	0,01091	7	0,00594	4	168,31	140,24	4
2	190	7	0,00735	1	0,00585	7	91,69	155,23	7
3	183	2	0,00594	4	0,00563	2	170,63	125,99	2
4	175	5	0,00586	2	0,00539	5	185,62	92,81	5
5	170	6	0,00539	5	0,00523	6	189,94	222,73	6
6	169	8	0,00527	8	0,00520	8	189,84	214,51	8
7	162	9	0,00526	6	0,00499	9	200,52	220,76	9
8	151	18	0,00499	9	0,00465	18	215,01	273,49	18
9	140	14	0,00465	18	0,00431	14	231,70	304,53	10
10	123	10	0,00432	14	0,00379	10	264,12	319,70	11
11	121	11	0,00379	10	0,00372	11	266,54	1041,71	1
12	118	13	0,00375	11	0,00363	13	275,31	347,93	14
13	117	12	0,00363	13	0,00360	12	277,46	586,27	13
14	116	17	0,00360	12	0,00357	17	280,06	372,54	12
15	114	3	0,00357	17	0,00351	3	284,97	198,93	17
16	83	16	0,00351	3	0,00255	16	389,60	965,76	3
17	82	1	0,00257	16	0,00252	1	136,11	137,41	16
18	28	15	0,00086	15	0,00086	15	1160,25	3130,79	15

Інтегральну класифікацію образів з рис. 2, наведену в таблиці, подано на рис. 3.

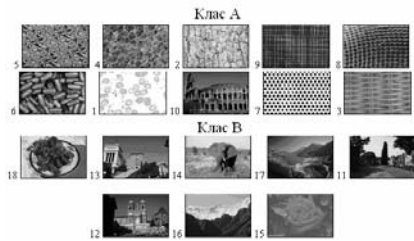


Рис. 3: Інтегральна класифікація візуальних образів

Рис. 3 ілюструє розбиття образи текстур та ландшафтів.

## 5. Процеси індексування та пошуку візуальних образів

### 5.1. Формування ключів образів

Як ключ (вектор властивостей) візуального образу використано його структурні коефіцієнти, зокрема:  $K_s^1$ ,  $K_s^2$ ,  $K_s^3$ ,  $K_s^{123}(MC)$ ,  $K_s^{123}(PX)$ , відношення  $D(IA) / M(IA)$ ,  $D(CR) / M(CR)$ .

На рис. 4 наведено приклади візуальних образів та їх ключів.

$K_s^1$	0.4151	0.5690	0.7617
$K_s^2$	0.9331	0.8523	0.9235
$K_s^3$	0.0033	0.0085	0.0048
$K_s^{123}(MC)$	0.0013	0.0041	0.0033
$K_s^{123}(PX)$	0.0038	0.0123	0.0101
$D(IA) / M(IA)$	4.6019	6.6184	1.7997
$D(CR) / M(CR)$	78.788	70.676	21.118

Рис. 4: Приклади візуальних образів та їх ключів

### 5.2. Процеси індексування та пошуку образів

На рис. 5, 6 зображено схеми процесів індексування та пошуку візуальних образів. Під час індексування до образу застосовується 3-ступеневий алгоритм кластеризації, після роботи якого формується ключ, що є набором структурних властивостей. Далі відбувається класифікація образу та запис його у базу даних класу. Ключ образу записується в іншу базу даних ключів, що відповідає класу, для якого відбувається процес їх згортання.

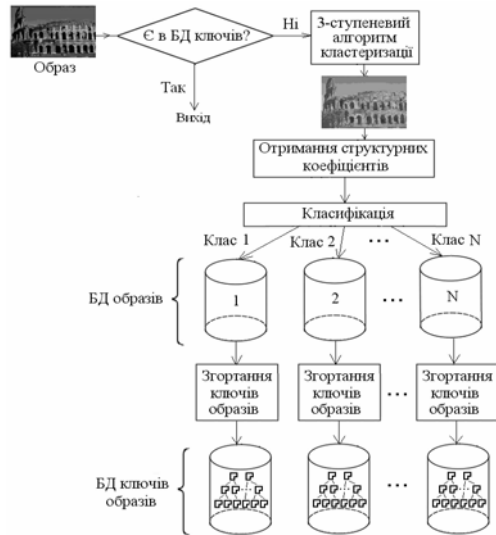


Рис. 5: Схема процесу індексування образів

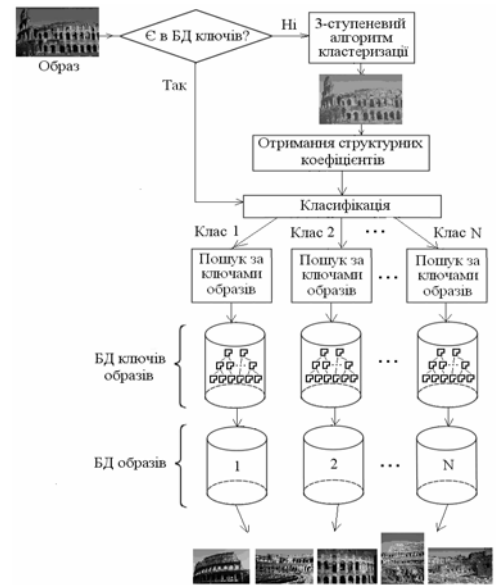


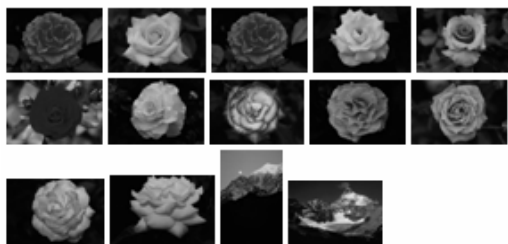
Рис. 6: Схема процесу пошуку образів

Під час пошуку образів перевіряється наявність ключа образу в базі даних ключів. Якщо такого образу немає, застосовується алгоритм кластеризації та отримання структурних коефіцієнтів. За останніми визначається його клас. Далі відбувається пошук образів

за ключами у відповідній базі ключів. За знайденими ключами отримують вихідні візуальні образи які і повертаються як результат запиту користувача. Якщо ж такий образ знайдемо у базі даних, то процес пошуку починається із визначення класу зображення, а отже, і бази даних із відповідними ключами образів та візуальними образами.

## 6. Експериментальні результати

Реалізовано програмний пакет з інтерфейсом користувача, що контролює всі етапи роботи із зображенням: введення образу, керування параметрами, процеси індексування та пошуку, повний звіт про хід алгоритму, числові параметри результатів тощо. На рис. 7 наведено приклади роботи пакету, на якому перше зображення є образом-запитом (без врахування класифікації, база відповідних образу-запиту складає 20 зображень).



11 зображень відповідають образу-запиту, 2 - ні



5 зображень відповідають образу-запиту, 4 - ні



7 зображень відповідають образу-запиту, 4- ні

Рис. 7: Приклади результатів пошуку

## 7. Висновки

Розроблено алгоритм тріступеневої кластеризації візуальних образів, який крім виділення фрагментів різної ієрархії підпорядкованості, дозволяє отримати кількісні та якісні характеристики фрагментів та відношень між ними. Запропоновано коефіцієнти структуризації та розмитості зображення. Експерименти підтвердили доцільність використання структурних коефіцієнтів та характеристики фрагментів для класифікації зображень та їх пошуку. Формування ключів зображення за представленими властивостями образів рекомендуються для застосування у автоматизованих системах пошуку зображень.

## 8. Перелік посилань

- [1] A. Vailaya, A.K. Jain, H.J. Zhang, "On image classification: city vs. landscape", *Pattern Recognition*, vol. 31, p. 1921-1935, 1998.
- [2] M.J. Swain, D.H. Ballard, "Color indexing", *International journal of Computer Vision*, vol. 7, n. 1, p. 11-32, 1991.
- [3] H. Nezamabadi-pour, E. Kabir, "Image retrieval using histograms of unicolor and bicolor blocas and direccional changes in intensity gradient", *Pattern Recognition Letters*, vol. 25, n. 14, p. 1547-1557, 2004.
- [4] F. Mokhtarian, S. Abbasi, "Shape similitaty retrieval under affine transforms", *Pattern Recognition*, vol. 35, p. 31-41, 2002.
- [5] A.K. Jain, A. Vailaya, "Image retrieval using color and shape", *Pattern Recognition*, vol. 29, n. 8, p. 1233-1244, 1996.
- [6] 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.
- [7] J.R. Smith, C.S. Li, "Image classification and quering using composite region templates", *Academic Press, Computer Vision and Understanding*, vol. 75, p. 165-174, 1999.
- [8] 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.
- [9] H.W. Yoo, S.H. Jung, D.H. Jang, Y.K. Na, "Extraction of major object features using VQ clustering for content-based image retrieval", *Pattern Recognition*, vol. 35, p. 1115-1126, 2002.
- [10] M. Szummer, R.W. Picard, "Indoor-outdoor image classification", *IEEE International Workshop on Content-Based Access of Image and Video Database (ICCV'98)*, p. 42-51, 1998.
- [11] T.P. Minka, R.W. Picard, "Interactive learning using a society of models", *Pattern Recognition*, vol. 30, n. 3, p. 565, 1997.
- [12] M.C. Burl, M. Weber, P. Perona, "A probabilistic approach to object recognition using local photometry and global geometry", *Proc. European Conf. Computer Vision*, p. 628-641, 1998.
- [13] J.Z. Wang, M.A. Fishler, "Visual similarity, judgmental certainty and stereo correspondence", *Proc. DARPA Image Understanding Workshop*, 1998.
- [14] I.H. Witten, A. Moffat, T.C. Bell, "Managing gigabytes: Evaluation retrieval effectiveness", Morgan Kaufman Publishers, Inc, p. 188-191, 1999.
- [15] R. Melnyk, R. Tushnytskyu, "Patterns Analysis by Clustering", *Proc. of 5<sup>th</sup> Intern. Conf. ICNNAI-2008*, p. 160-163, 2008.
- [16] Р. Мельник, Р. Тушницький, "Керування яскравістю кластера при декомпозиції зображень спектрів", *Комп'ютерні науки та інформаційні технології*. – 2007. – №598. – С. 3–8.
- [17] Р. Мельник, Р. Тушницький, "Керування параметрами кластера в алгоритмі декомпозиції зображень", *Відбір і обробка інформації*. – 2007. – № 27 (103). – С. 58–62.