

Р. Мельник, Р. Тушницький

Національний університет "Львівська політехніка",  
кафедра програмного забезпечення

## АЛГОРИТМ ТРИСТУПЕНЕВОЇ КЛАСТЕРИЗАЦІЇ ДЛЯ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ ЗА ЇХ СТРУКТУРНИМИ ВЛАСТИВОСТЯМИ

© Мельник Р., Тушницький Р., 2008

Наведено алгоритм триступеневої кластеризації візуальних образів для знаходження структурних коефіцієнтів, що описують властивості зображень. Запропоновано використати структурні коефіцієнти для класифікації та пошуку зображень.

The 3-stages clustering algorithm for visual patterns to find structural coefficients describing properties of images is presented. It is suggested to use structural coefficients for classification and searching of images.

### 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 складним завданням, один з підходів до вирішення якого 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 належать 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) на противагу глобальним властивостям повного зображення. Прикладом такої системи є SIMPLcity [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 адекватною кількістю кластерів (об'єкти у зображенні) може краще відобразити його вміст, однак цей підхід є часозалежним.

У 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. Здійснено класифікацію 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 містять процес 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.

### Алгоритм триступеневої кластеризації

Під час розроблення алгоритму декомпозиції візуального образу його завданням є вирішення таких завдань: 1) отримати описи складових образу з усіма можливими характеристиками; 2) виділити тільки певну складову із заданими характеристиками; 3) отримати складові частини інтегральної області, що характеризуються різною яскравістю.

Для вирішення поставленого завдання розроблено алгоритм триступеневої кластеризації [14], який ґрунтується на побудові ієрархічного дерева згортання та засоби керування [15] для забезпечення доступу до різних його рівнів для одержання фрагментів різної фізичної природи. Зазначені рівні формування кластерів різної природи позначені на фрагменті ієрархічного дерева (рис. 1).

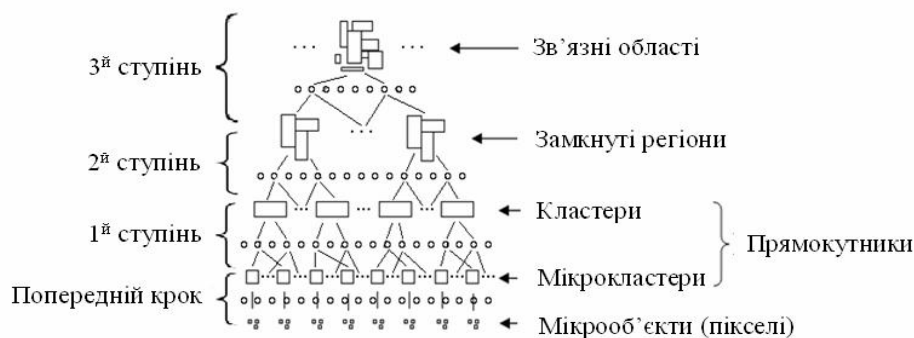


Рис. 1. Фрагмент ієрархічного дерева

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 такі 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 формуються. Це, наприклад, яскравість, колір, співвідношення заповнених та пустих клітинок тощо. Вершини, що не підлягають подальшому згортанню, відповідають кластерам результату.

Мікрооб'єкти образу утворюються у разі його поділу сіткою. Кроки сітки вибираються з ряду величин  $1 \times 1, 1 \times 2, 2 \times 1, 2 \times 2$  тощо (керує кроком користувач). Кожний елементарний квадрат –

піксель набуває значення у діапазоні від чорного до білого, яке позначимо як  $b$  – яскравість мікрооб’єкта. Діапазон яскравості зображено на рис. 2.

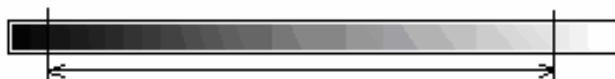


Рис. 2. Діапазон значень яскравості мікрооб’єктів та поріг чутливості алгоритму

Керуючи діапазоном чутливості алгоритму, можна аналізувати не все зображення цілком, а окремі сегменти за яскравостями. Збільшення кроку сітки призводить до виявлення найхарактерніших мікрокластерів, що використовується для формування скелетону зображення.

Кольорові зображення попередньо конвертуються у сірі за допомогою алгоритму *BT709* з наступними коефіцієнтами  $R, G, B$ :

$$R = 0,2125; G = 0,7154; B = 0,0721.$$

Значення яскравості мікрооб’єкта (процент заповнення) обчислюється за формулами відносного заповнення клітинки у відсотках від чорного:

$$b = (256 - c_i) \times 100 / 256,$$

або у відсотках від білого:

$$b = c_i \times 100 / 256,$$

де  $c_i$  ( $i=1,2,3$ ) – значення однієї із компонент  $R$  ( $G$  або  $B$ ) кольору сірого пікселя.

### Структурні характеристики образів

У результаті роботи алгоритму тріступеневої 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, що підлягає декомпозиції, введемо ряд параметрів, якими характеризуємо зображення з погляду його структурних властивостей, зокрема:

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

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

$$K_s^2 = CR / C,$$

$$K_s^3 = IA / CR.$$

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

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

$$K_s^{23} = IA / C.$$

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

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

Обернені до коефіцієнтів структуризації величини назовемо коефіцієнтами *ступеня розмитості* зображення на різних рівнях ієрархії розбиття зображення, наприклад:

$$K_k^{123}(MC) = MC / IA.$$

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

Для класифікації повних зображень, зокрема 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  $PX$ :

$$K_s^{123}(PX) = IA / PX.$$

Кожний виділений  $i$ -й фрагмент, зокрема область характеризуємо відносною кількістю пікселів:

$$vx_i = px_i(IA) / PX,$$

для яких є справедливою нерівність:

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

де сума береться за всіма  $IA$  сформованими областями.

Якщо визначити середній розмір області в пікселях

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

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

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

Отже, характеризувати зображення можемо за величинами  $C$ ,  $CR$ ,  $IA$ ,  $K_s^1$ ,  $K_s^2$ ,  $K_s^3$ ,  $K_s^{12}$ ,  $K_s^{23}$ ,  $K_s^{123}(MC)$ ,  $K_s^{123}(PX)$ ,  $M(IA)$ ,  $D(IA)$ .

### Класифікація зображень за структурними коефіцієнтами

Алгоритм та структурні коефіцієнти досліджено за базою із 1000 образів, згрупованих у 10 семантичних груп: леви, слони, коні, квіти, їжа, гори, автобуси, дизайн, зображення текстури, медичні образи. Кожна група містить 100 зображень у форматі JPEG із розмірами  $256 \times 384$  пікселів. Приклади образів подано на рис. 3.

Результати декомпозиції образів з рис. 3 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  $IA$ ,  $K_s^{123}(MC)$ ,  $K_s^{123}(PX)$ .

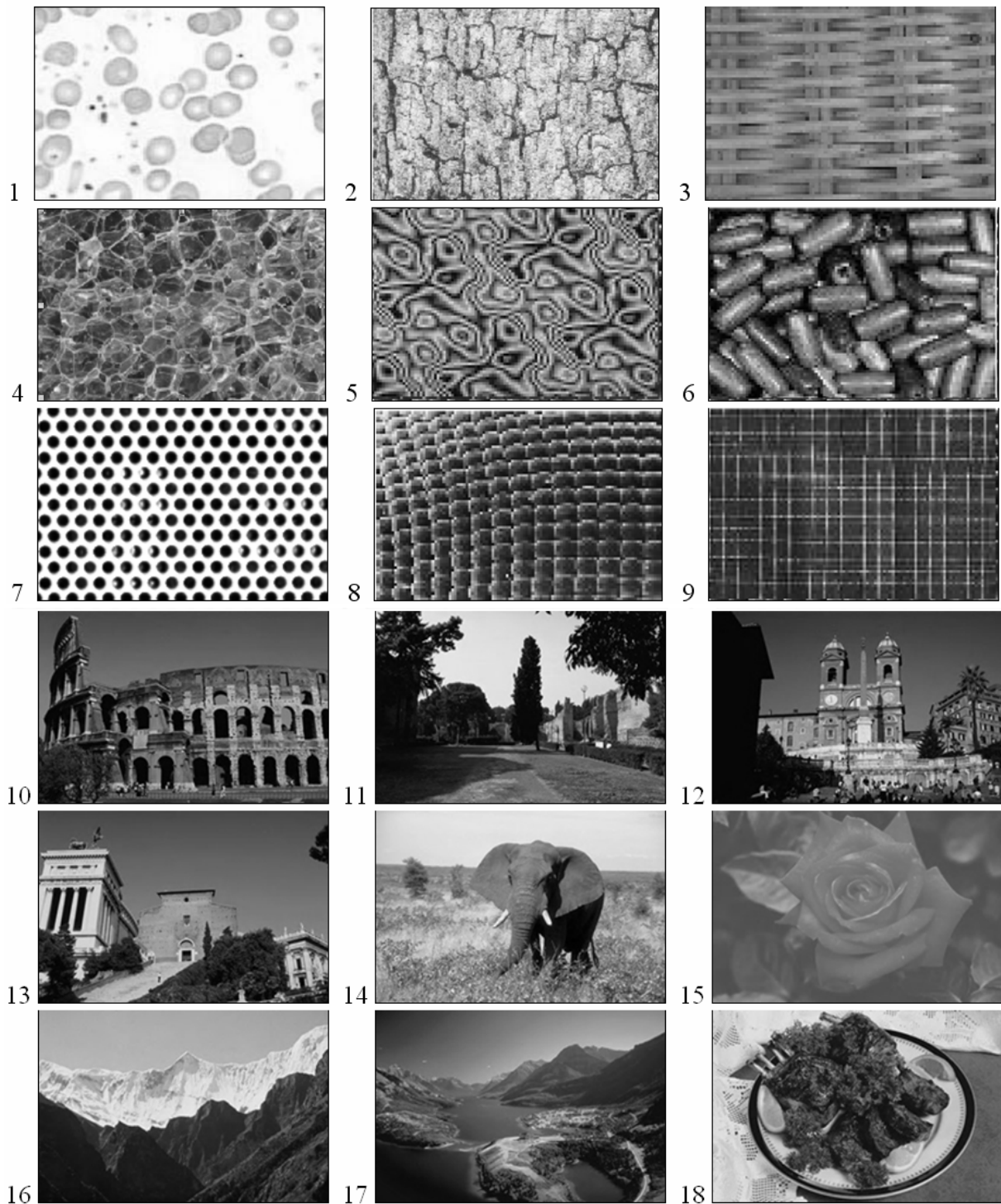


Рис. 3. Пронумеровані зразки досліджуваних образів

З таблиці отримуємо збіг класифікації деяких образів. Це означає, що ми однозначно можемо класифікувати образи текстур, медичні зразки, ландшафти або архітектурні комплекси. Комбінуючи структурні коефіцієнти та параметри об'єктів – фрагментів образів можна створити правила класифікації медичних зображень, структур матеріалів тощо. Зокрема, беручи до уваги параметри образу  $M(IA)$  та  $D(IA)$  на рис. 4, клас С можна розбити на окремий клас зображення клітин крові (зображення № 1) та зображення текстур (зображення № 3, 9).

### Класифікація образів за структурними характеристиками

Місце	$IA$	№ образу	$K_s^{123}(MC)$	№ образу	$K_s^{123}(PX)$	№ образу	$M(LA)$	$D(LA)$	Інтегральна класифікація, № образу
1	178	2	0,00838	7	0,00548	2	175,43	124,91	2
2	170	4	0,00735	1	0,00523	4	191,08	166,48	4
3	148	5	0,00570	2	0,00456	5	219,49	177,61	7
4	146	7	0,00523	4	0,00449	7	119,32	244,53	5
5	143	6	0,00456	5	0,00440	6	225,80	571,13	6
6	131	14	0,00443	6	0,00403	14	247,61	376,77	14
7	130	8	0,00405	8	0,00400	8	246,79	1440,20	8
8	126	18	0,00404	14	0,00388	18	257,67	1013,51	18
9	108	9	0,00388	18	0,00332	3	300,81	194,52	3
10	108	3	0,00332	3	0,00332	9	300,78	1066,08	9
11	97	10	0,00332	9	0,00299	10	334,92	1098,05	1
12	94	13	0,00299	10	0,00289	13	345,61	717,35	10
13	91	12	0,00289	13	0,00280	12	356,74	1067,25	13
14	84	17	0,00280	12	0,00259	17	386,75	1093,51	12
15	83	11	0,00259	17	0,00255	11	388,57	1339,15	17
16	82	1	0,00257	11	0,00252	1	136,11	137,41	11
17	63	16	0,00195	16	0,00194	16	513,29	2004,53	16
17	4	15	0,00012	15	0,00012	15	8121,75	14019,37	15

Інтегральну класифікацію образів з рис. 3, наведену в таблиці, подано на рис. 4.

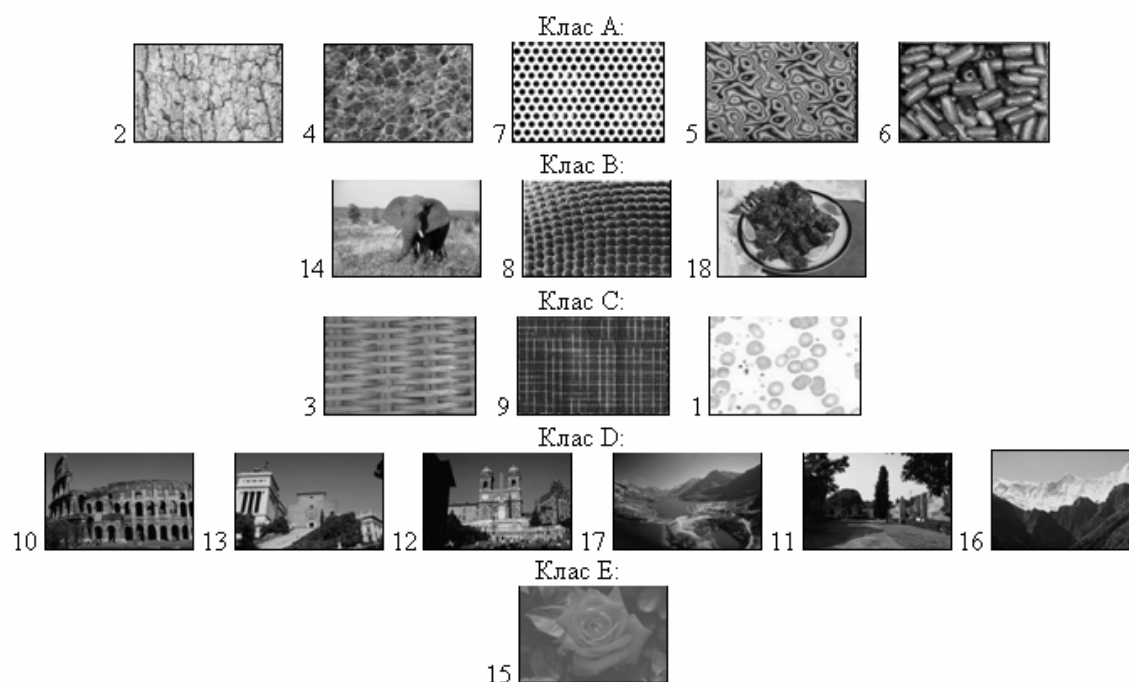


Рис. 4. Інтегральна класифікація візуальних образів

### Висновки

Розроблено алгоритм триступеневої кластеризації візуальних образів, який, крім виділення фрагментів різної ієрархії підпорядкованості, дає змогу 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. Формування ключів зображення за представленими властивостями образів рекомендуються для застосування у автоматизованих системах пошуку зображень.

1. Vailaya A., Jain A.K., Zhang H.J. On image classification: city vs. landscape // *Pattern Recognition*, vol. 31, p. 1921-1935, 1998. 2. Swain M.J., Ballard D.H. Color indexing // *International journal of Computer Vision*, vol. 7, n. 1, p. 11-32, 1991. 3. Nezamabadi-pour H., Kabir E. 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. Mokhtarian F., Abbasi S. Shape similitarity retrieval under affine transforms // *Pattern Recognition*, vol. 35, p. 31-41, 2002. 5. Jain A.K., Vailaya A. Image retrieval using color and shape // *pattern Recognition*, vol. 29, n. 8, p. 1233-1244, 1996. 6. Manjunath B.S., Ma W.Y. Texture feature for browsing and retrieval of image data // *IEEE PAMI*, vol. 8, n. 18, p. 837-842, 1996. 7. Smith J.R., Li C.S. Image classification and quering using composite region templates // *Academic Press, Computer Vision and Understanding*, vol. 75, p. 165-174, 1999. 8. Wang J.Z., Li J., Wiederhold G. 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. Yoo H.W., Jung S.H., Jang D.H., Na Y.K. Extraction of major object features using VQ clustering for content-based image retrieval // *Pattern Recognition*, vol. 35, p. 1115-1126, 2002. 10. Szummer M., Picard R.W. Indoor-outdoor image classification // *IEEE International Workshop on Content-Based Access of Image and Video Database (ICCV'98)*, p. 42-51, 1998. 11. Minka T.P., Picard R.W. Interactive learning using a society of models // *Pattern Recognition*, vol. 30, n. 3, p. 565, 1997. 12. Burl M.C., Weber M., Perona P. A probabilistic approach to object recognition using local photometry and global geometry // *Proc. European Conf. Computer Vision*, p. 628-641, 1998. 13. Wang J.Z., Fishler M.A. Visual similarity, judgmental certainty and stereo correspondence // *Proc. DARPA Image Understanding Workshop*, 1998. 14. Мельник Р., Тушиницький Р. Керування яскравістю кластера при декомпозиції зображень спектрів // *Комп'ютерні науки та інформаційні технології*. – 2007. – №598. – С. 3–8. 15. Мельник Р., Тушиницький Р. Керування параметрами кластера в алгоритмі декомпозиції зображень // *Відбір і обробка інформації*. – 2007. – № 27 (103). – С. 58–62.