

Р.А. Мельник, Р.Б. Тушницький  
Національний університет “Львівська політехніка”,  
кафедра програмного забезпечення

## ПОШУК ЗОБРАЖЕНЬ У БАЗАХ ДАНИХ ЗА СТРУКТУРНИМИ КОЕФІЦІЄНТАМИ НА ОСНОВІ АЛГОРИТМУ ТРИСТУПЕНЕВОЇ КЛАСТЕРИЗАЦІЇ

© Мельник Р.А., Тушницький Р.Б., 2009

Розглянуто методику пошуку зображень у базах даних за структурними коефіцієнтами, що отримані агломеративним ієрархічним триступеневим алгоритмом кластеризації образу. Запропоновано структурні коефіцієнти використовувати для формування ключа зображення.

**Ключові слова** – зображення, база даних, кластеризація, алгоритм

An approach to search images from a databases by structural features got from the three stages clustering algorithm is considered. Structural coefficients as image keys are proposed to be used. Some seaching tests for qualitative estimation of the approach were held.

**Keywords** – image, data base, clastering, algorithm

### Вступ

Традиційно для пошуку зображень використовують їхні текстові характеристики: ім'я файла, заголовок, ключові слова тощо. Однак такий підхід має недоліки. Передусім необхідне втручання людини для опису вмісту зображень відповідно до обраного набору підписів та ключових слів. Здебільшого зображення містить декілька об'єктів, кожен з яких має свій набір атрибутів. Крім цього, потрібно описати просторові відношення між цими об'єктами, щоб зрозуміти його зміст. Оскільки розміри баз даних зображень зростають, використання ключових слів стає не тільки складною, але і недостатньою для представлення зображення. Інша проблема цього підходу полягає у неадекватності єдиного текстового опису зображення. Як результат є необхідність для автоматизованого отримання примітивних властивостей зображень і пошук зображень на основі цих властивостей.

Сучасні системи знаходження зображень за їх вмістом (Content-based image retrieval – CBIR) [1] працюють у два етапи: індексування та пошук. На етапі індексування кожний образ у базі даних представляється вектором властивостей. Існуючі універсальні системи CBIR зараховують 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 можливу систему пошуку зображень за регіонами [9]. Представлення візуального образу адекватним числом кластерів (об'єкти у зображенні) може краще відтворити 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. 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.

У цій роботі запропонована методика пошуку зображень на основі структурних властивостей та коефіцієнтів, отриманих в результаті роботи алгоритму триступеневої кластеризації [18].

### Постановка задачі індексування та пошуку візуальних образів

На рис. 1 зображено архітектуру процесів індексування та пошуку візуальних образів.

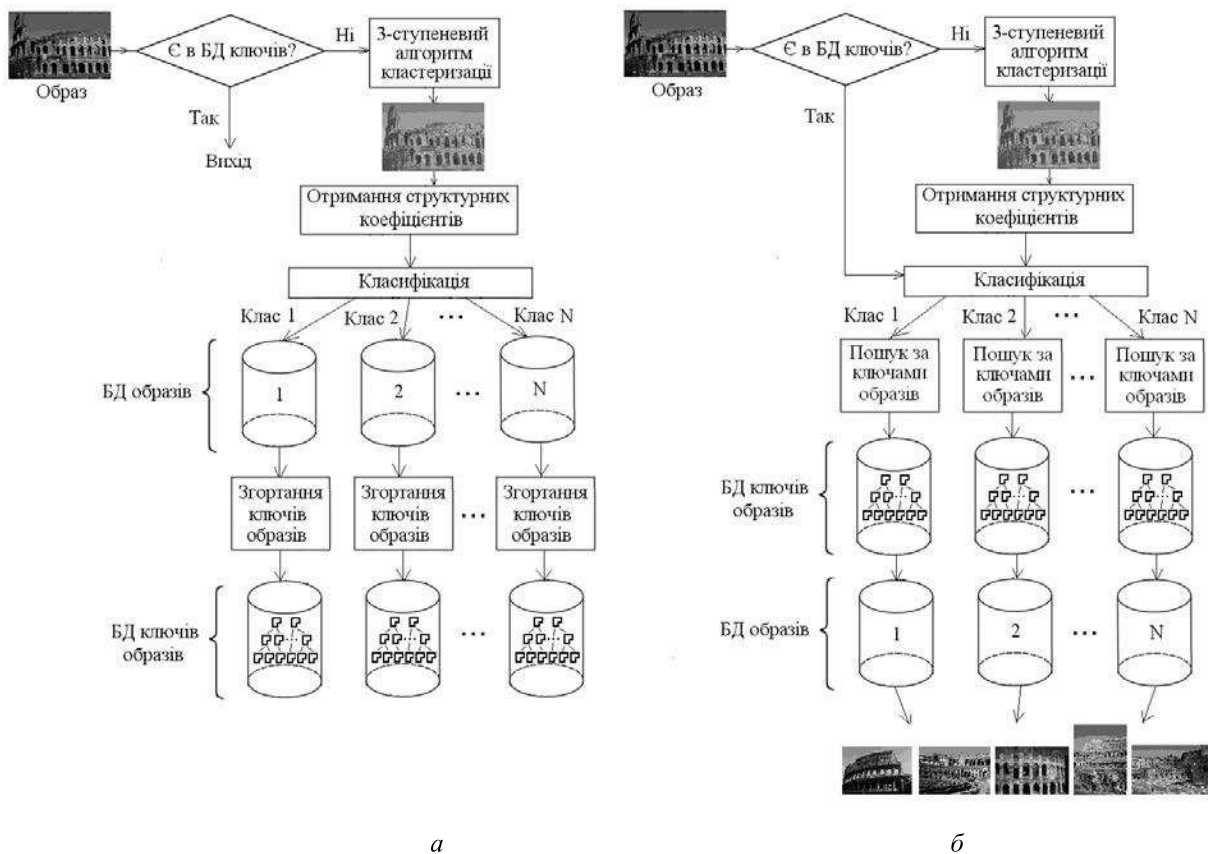


Рис. 1. Архітектура процесів індексування (а) та пошуку візуальних образів (б)

Таким чином запропоновано використати дві бази даних: для зображень та ключів. Кожна з них ділиться в процесі класифікації на відповідні класи зображень, наприклад: структурні зображення, слабкоструктурні медичні образи, лінійчасті спектри, текстурні зображення тощо. Використання дерева згортання ключів дає змогу обійти процес його формування при кожному пошуку образів.

Під час індексування до образу застосовується триступеневий алгоритм кластеризації, після роботи якого формується ключ, що є набором структурних властивостей. Далі відбувається класифікація образу та записування його у базу даних класу. Ключ образу записується в іншу базу даних ключів, що відповідає класу, для якого відбувається процес їх згортання.

### Алгоритм пошуку містить два основних кроки:

Пошук образів здійснюється перевіркою наявності ключа образу в базі даних ключів. Якщо такого образу немає, застосовується алгоритм кластеризації та отримання структурних коефіцієнтів та властивостей. За останніми визначається його клас.

Далі відбувається пошук образів за ключами у відповідній базі ключів. За знайденими ключами отримують вихідні візуальні образи, які 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.

### Формування ключів образів

Для вирішення поставленого завдання розроблено алгоритм триступеневої кластеризації [16], що базується на побудові ієрархічного дерева згортання, та засоби керування [17] для забезпечення доступу до різних його рівнів для одержання фрагментів різної фізичної природи. Зазначені рівні формування кластерів різної природи позначені на фрагменті ієрархічного дерева, що наведений на рис. 2.

В результаті роботи алгоритму триступеневої кластеризації з образу отримуємо різні типи структурних об'єктів: мікрокластери, кластери-прямокутними, замкнуті регіони та зв'язні області. Використовуємо ці кількісні характеристики для отримання структурних властивостей, за якими оцінюємо ступінь структуризації зображення [18].

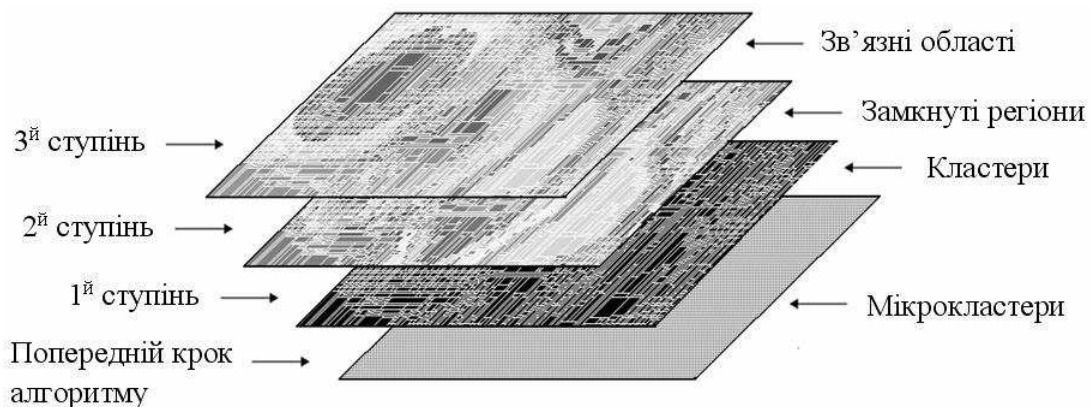


Рис. 2. Триступенева декомпозиція образу

У якості ключа (вектора властивостей) візуального образу запропоновано використати такі два набори із 6 його структурних коефіцієнтів:

**A:**  $IA, K_s^{123} (MC), M(IA), D(IA), R(IA), R_i(IA)$ ;

**B:**  $IA, K_s^1, K_s^2, K_s^3, K_s^{123} (MC), R_i(IA)$ .

Для згортання ключів образів використовуємо абсолютний

$$F_{ij} = \{w_1[a_i - a_j]^2 + w_2[b_i - b_j]^2 + \dots\} / N, \quad (1)$$

та відносний критерії згортання образів:

$$F_{ij} = \left\{ \frac{w_1[a_i - a_j]^2}{\max\{a_i, a_j\}} + \frac{w_2[b_i - b_j]^2}{\max\{b_i, b_j\}} + \dots \right\} / (k + r), \quad (2)$$

де  $a, b, c$  – структурні характеристики образу,  $w$  – зважені коефіцієнти,  $k$  та  $r$  – кількість ключів у відповідному кластері.

Для порівняння двох візуальних образів  $i$  та  $j$  із векторами властивостей  $K_i$  та  $K_j$  (чисел одного порядку) визначаємо відстань:

$$d(i, j) = \sum_{i,j=1}^6 (K_i - K_j)^2 \quad (3)$$

Відповідно, якщо ключ містить два набори властивостей (чисел різного порядку), то визначаємо дві відстані  $d1$  та  $d2$ .

На рис. 3 наведено приклади візуальних образів, їхні структурні властивості зведено у таблицю.



Рис. 3. Приклади візуальних образів

Таблиця 1

### Структурні властивості образів

Коефіцієнт	Зображення		
	а	б	в
$C$	36825	67963	95570
$CR$	35497	61356	94465
$IA$	129	391	231
$K_s^1$	0.37460	0.69597	0.98012
$K_s^2$	0.96393	0.90278	0.98843
$K_s^3$	0.00363	0.00637	0.00244
$K_s^{12}$	0.36109	0.62831	0.96879
$K_s^{23}$	0.00351	0.00575	0.00241
$K_s^{123} (MC)$	0.00131	0.00401	0.00236
$M(CR)$	2.76936	1.59156	1.03221
$D(CR)$	240.51325	54.45091	2.92098
$R(CR)$	0.01151	0.02922	0.35337
$R_A(CR)$	408.72591	1793.39503	33381.90849
$M(IA)$	762.04651	249.74936	422.11255
$D(IA)$	5487.33364	965.82043	511.56099
$R(IA)$	0.13887	0.25858	0.82514
$R_A(IA)$	17.91471	101.10782	190.60874

### Оцінка системи пошуку зображень

Відомі два підходи для оцінки методу індексування зображень. Перший підхід оцінює отримані результати у пошуковій системі образів. Другий ґрунтується на його застосуванні у автоматичній класифікації. Доцільність вибору як ключа образу структурних коефіцієнтів підтверджено експериментально у процесі автоматичної класифікації у роботі [18].

Для якісної оцінки виконання запропонованого підходу, використовуємо визначені у [14] такі критерії:

Точність пошуку (precision)  $P$  характеризує частку зображень, подібних на зразок, серед результатів пошуку:

$$P = \frac{\text{кількість отриманих образів, відповідних запиту}}{\text{кількість всіх отриманих образів}}. \quad (4)$$

Ефективність пошуку (*recall*)  $R$  показує, яка частка зображень, схожих на зразок, була включена у склад результатів пошуку:

$$R = \frac{\text{кількість отриманих образів, відповідних запиту}}{\text{кількість всіх образів у базі, відповідних запиту}}. \quad (5)$$

Для перевірки коректної роботи алгоритмів, а також для отримання кількісних та якісних характеристик пошуку необхідна база зображень, на якій, звичайно, перевіряються інші реалізації пошуку. З цією метою була обрана база зображень Ванга [15], що містить 1000 образів, згрупованих у 10 семантичних груп: людина, пейзаж, архітектура, автобуси, динозаври, слони, квіти, коні, гори, їжа. Кожна група містить в собі 100 24-бітних зображень у форматі JPEG розміром 256×384 пікселів. Однак для того, щоб показати якісні характеристики пошуку, необхідні експертні оцінки, бо саме точка зору людини є останньою метою пошуку зображень. З цією метою база даних була зменшена до 200 зображень, зберігаючи свій кількісний склад у пропорції.

Для роботи алгоритму, кольорові зображення попередньо конвертуються у сірі за допомогою алгоритму BT709 із такими коефіцієнтами  $R$ ,  $G$ ,  $B$ :

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

### Експериментальні результати

Реалізовано програмний пакет з інтерфейсом користувача (рис. 4), що контролює всі етапи роботи із зображенням: введення образу, керування параметрами, процеси індексування та пошуку, повний звіт про хід алгоритму, числові параметри результатів тощо.

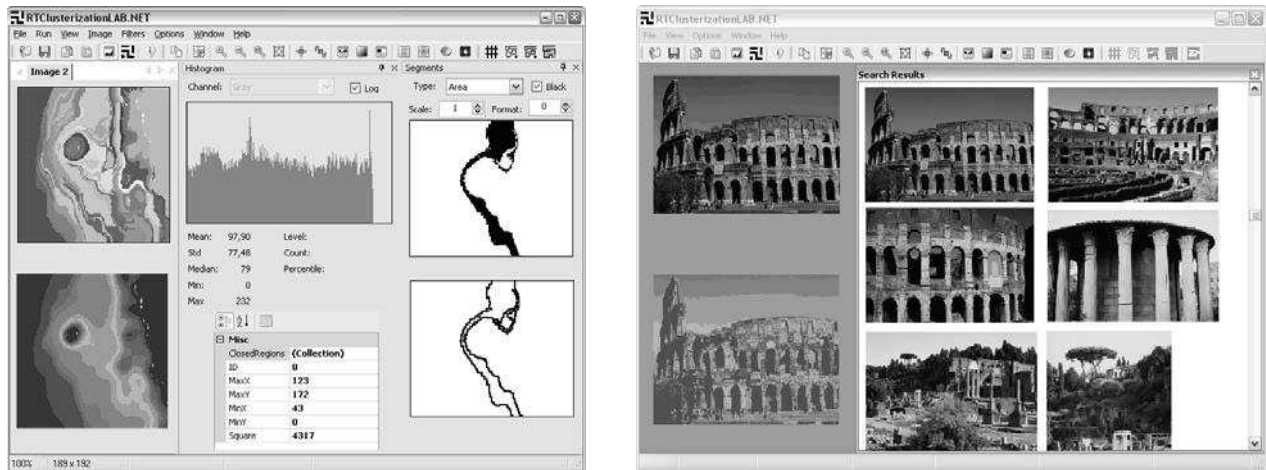


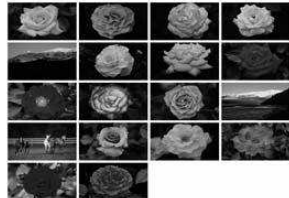
Рис. 4. Інтерфейс програмного пакета

На рис. 5 наведено приклади роботи пакета, на якому перше зображення є образом-запитом.

Попередні дослідження показали доцільність використання структурних властивостей для розділення зображень за формою та контрастом. Для промислового використання методика вимагає виконання подальших досліджень ефективності індексування та пошуку за структурними коефіцієнтами.



Отримано 18 образів:  
15 образів відповідають  
образу-запиту, 3 – ні.  
 $P = 0,83; R = 0,75$



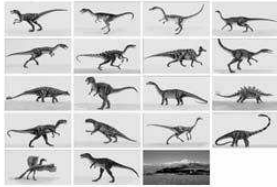
Отримано 18 образів:  
15 образів відповідають  
образу-запиту, 3 – ні.  
 $P = 0,83; R = 0,75$



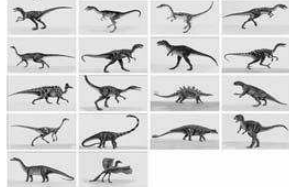
Отримано 17 образів:  
15 образів відповідають  
образу-запиту, 2 – ні.  
 $P = 0,88; R = 0,75$



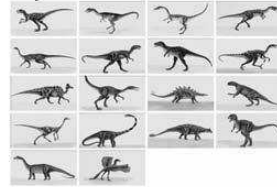
Отримано 18 образів:  
15 образів відповідають  
образу-запиту, 3 – ні.  
 $P = 0,83; R = 0,75$



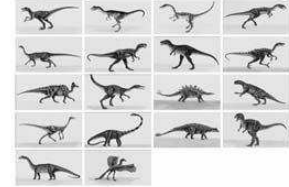
Отримано 19 образів:  
18 образів відповідають  
образу-запиту, 1 – ні.  
 $P = 0,95; R = 0,9$



Отримано 18 образів:  
18 образів відповідають  
образу-запиту, 0 – ні.  
 $P = 1; R = 0,9$



Отримано 18 образів:  
18 образів відповідають  
образу-запиту, 0 – ні.  
 $P = 1; R = 0,9$



Отримано 18 образів:  
18 образів відповідають  
образу-запиту, 0 – ні.  
 $P = 1; R = 0,9$



Отримано 35 образів:  
12 образів відповідають  
образу-запиту, 23 – ні.  
 $P = 0,34; R = 0,6$



Отримано 38 образів:  
13 образів відповідають  
образу-запиту, 25 – ні.  
 $P = 0,34; R = 0,65$



Отримано 25 образів:  
11 образів відповідають  
образу-запиту, 14 – ні.  
 $P = 0,44; R = 0,55$



Отримано 28 образів:  
11 образів відповідають  
образу-запиту, 17 – ні.  
 $P = 0,39; R = 0,55$



Отримано 17 образів:  
7 образів відповідають  
образу-запиту, 10 – ні.  
 $P = 0,41; R = 0,35$



Отримано 12 образів:  
7 образів відповідають  
образу-запиту, 5 – ні.  
 $P = 0,58; R = 0,35$



Отримано 21 образ:  
11 образів відповідають  
образу-запиту, 10 – ні.  
 $P = 0,52; R = 0,55$



Отримано 23 образ:  
13 образів відповідають  
образу-запиту, 10 – ні.  
 $P = 0,57; R = 0,65$

*a*

*б*

*в*

*г*

Рис. 5. Приклади пошуку образів із використанням абсолютного (*a, в*) та відносного (*б, г*) критеріїв згортання ключів для набору структурних властивостей *A* (*a, б*) та *B* (*в, г*)

## Висновки

Запропоновано методику пошуку зображень за їх структурними властивостями. Останні отримуються як результати роботи алгоритму триступеневої кластеризації візуальних образів. Ключі обчислюються за властивостями фрагментів образу різної ієрархії та відношень між ними. Експерименти підтвердили доцільність використання структурних коефіцієнтів та характеристик фрагментів для індексування зображень та їх пошуку. Ключі 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 similitaty 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. Witten I.H., Moffat A., Bell T.C. *Managing gigabytes: Evaluation retrieval effectiveness*. – Morgan Kaufman Publishers, Inc, p. 188-191, 1999. 15. База 1000 тестових зображень (Ванга): <http://wang.ist.psu.edu/~jwang/test1.tar> (2009). 16. Мельник Р., Тушиницький Р. Сегментування зображень алгоритмом ієрархічної кластеризації // *Відбір і обробка інформації*. – 2008. – № 28 (104). – С. 97–102. 17. Мельник Р., Тушиницький Р. Керування параметрами кластера в алгоритмі декомпозиції зображень // *Відбір і обробка інформації*. – 2007. – № 27 (103). – С. 58–62. 18. Мельник Р., Тушиницький Р. Алгоритм триступеневої кластеризації для класифікації зображень за їх структурними властивостями // *Вісник Нац. ун-ту "Львівська політехніка": Комп'ютерні науки та інформаційні технології*. – 2008. – № 629. – С. 46–52.