Block-Diagonal Form of Distance Matrix for Region-Based Image Retrieval

D. Kinoshenko, V. Mashtalir, E. Yegorova
Kharkov National University of Radio Electronics, Computer Science faculty,
Lenin Av., 14, Kharkov, 61166, Ukraine
{Kinoshenko, Mashtalir, Yegorova}@kture.kharkov.ua

Abstract

There are two substantial open issues in the field of the image retrieval: semantic gap between computationally extracted low-level features and human operated high-level concepts, and high retrieval speed independent from the volume of the database. Search of the images on the level of objects or regions (segmentation) is a step towards semantic-based retrieval. In this paper we propose a new cluster-like indexing algorithm in metric space which preliminary transforms distance matrix into block-diagonal form and ensures the minimum number of matches at the retrieval stage. This form can be used separately or embedded into existent indexing methods.

1. Introduction

In the recent decade we could often see an elucidation of content-based image retrieval as region-based procedures. Latter implemented systems like Blobworld, SIMPLIcity [1,2, etc.] outperformed long-standing ones (e.g. QBIC or MARS). Still the problem remained is how to preprocess database content in order to speedup a search. In some cases the only information we have is a distance between pairs of images and multidimensional indexing methods like VA-file or X-trees [3] can not be applied. If at that the distance used is also a metric one can apply metric indexing methods. The most promising metrics for segmentations matches are EMD [4], Meila's variation of information [5] and metric on arbitrary measurable quotient sets [6]. Existent metric indexing methods could be roughly classified into two groups [7] viz pivot-based and compact partitioning. In the first case a number of reference objects called pivots is chosen, distances from all of non-pivot objects to them are calculated and stored. The most popular variants are full-matrix index AESA, sparse matrix LAESA and VP-tree [8,9]. Compact partitioning indexing algorithms like M-tree, D-index [10] form a sort of clusters around predefined reference objects. From many performance characteristics of indexing methods we are most interested in those which measure the number of distance evaluations at the search stage since operation of distance calculation between two sets of image regions is often time-consuming. For example, for EMD metric one needs to solve transportation problem. Algorithms like AESA which utilize the full matrix of distances between database objects have shown absolutely best results for this type of performance criteria but unfortunately they are impractical for large-scale databases due to quadratic space requirements. Other indexing methods keep a part of the distance matrix and create auxiliary data structures based on intrinsic correlation of distances between objects. Their performance relies heavily on a set of chosen pivot points. We propose a new indexing method in metric space which transforms distance matrix into block-diagonal form by certain compactness criteria and predicts the maximum number of distance evaluations at the search stage what hasn't yet been done by any of known indexing methods. We emphasize that our method does not rely on pivot selection strategy: its construction policy is aimed at creation of compact blocks minimum number with distance between any elements of the block less then predefined threshold value. Later the capacity of those blocks can be refined by solving the optimization task.

2. Block-Diagonal Form of Distance Matrix

The problem statement may be formulated as follows. Let $X=\{x_1, x_2, ..., x_N\}$ be an image collection in which content-based retrieval is carried

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

out. Suppose that the set U ($X \subseteq U$) defines problem-oriented field of image understanding. Result of ε- search with the query $y \in U$ is any element (all elements) $x_i \in X$ if $\rho(y, x_i) \le \varepsilon$ for given so-called search radius $\varepsilon \ge 0$ (here $\rho(\circ,\circ)$ is a metric). Let us introduce a notation $\rho_{i,j} = \rho(x_i, x_j)$ for elements of distance matrix $d(X)$ and consider some arbitrary subset of database elements for which $\rho_{i,j} \le \delta$. These objects can be used as a δ- search result, yet under that a formalization of all such groupings search procedures for some given criterion is necessary. As a basic criterion we shall use the number of matches between the query and database elements. In other words, the problem lies in the preliminary clustering of the database with the search aiming to find the closest cluster and if necessary to continue searching inside the chosen cluster. Here the basic feature of clustering is matches number minimization. Let us consider the clustering procedure.

We shall call symmetrical l-range matrix as a Δ_l^k-block of distance matrix $d(X)$

$$\Delta_l^k[d(X)] = \begin{pmatrix} 0 & \rho_{k,k+1} & \cdots & \rho_{k,k+l-1} \\ \rho_{k+1,k} & 0 & \cdots & \rho_{k+1,k+l-1} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{k+l-1,k} & \rho_{k+l-1,k+1} & \cdots & 0 \end{pmatrix}$$

which is the result of rows and columns transposition with indices $\{i_1, i_2,..., i_l\}$ such that

$$\forall i', i'' \in \{i_1, i_2,..., i_l\} \Rightarrow \rho_{i',i''} \le \delta .$$

In addition we shall consider Δ_l^k-block of matrix $d(X)$ as maximal if

$$\nexists r \in \{1, 2,..., n\} \setminus \{i_1, i_2,..., i_l\} : \rho_{r,i'} \le \delta \forall i' \in \{i_1, i_2,..., i_l\}.$$

There may exist some elements which can belong to two or more different Δ_l^k-blocks of matrix $d(X)$. It is obvious that two variants are possible: these elements are included into all possible blocks or they are included into those blocks for which sum of elements is minimal what meets compactness criterion (2).

As Δ-representation we shall call a block-diagonal type of matrix $d(X)$

$$\Delta[d(X)] = \begin{pmatrix} \Delta_{l_1}^{k_1} & 0 & & \\ 0 & \Delta_{l_2}^{k_2} & 0 & \\ & 0 & \cdots & 0 \\ & & 0 & \Delta_{l_m}^{k_m} \end{pmatrix}$$

where $k_1 = 1,\ k_i = \sum_{j=1}^{i-1} l_j + 1,\ \sum_{j=1}^{m} l_j \ge n$.

It is clear that under δ- search the Δ-representation with the minimal number of blocks will be the best one in respect to the number of matches. In other words forming Δ-representation of matrix under given δ should provide

$$\min_{\Delta_l^k \in d(X)} m . \qquad (1)$$

In case when we have several maximal Δ_l^k-blocks one can considered a criterion

$$\min_{\Delta_l^k \in d(X)} \sum_{i,j \in \{i_1, i_2,..., i_l\}} \rho_{i,j} \qquad (2)$$

which does not change the goal function (1) value, but allows to get more sufficient database clustering.

Introduce the procedure of forming the maximal Δ_j^i-block of matrix $d(X)$ on set $\{p_1, p_2,..., p_r\} \subseteq \subseteq \{1, 2,..., N\}$. We find a row α^* such that for all $q \in \{p_1, p_2,..., p_r\}$

$$\alpha^* = arg \max_{\alpha \in \{p_1, p_2,..., p_r\}} \{card\{\rho_{\alpha,q}\} : \rho_{\alpha,q} \le \delta\} . \qquad (3)$$

Denote by $\{\alpha_1,...,\alpha^*,..., \alpha_\beta\}$ indices found in (3). If there exists more than one of such indices kit we shall randomly chose one of them. Two cases are possible:

$$\forall \alpha', \alpha'' \in \{\alpha_1,...,\alpha^*,..., \alpha_\beta\} \Rightarrow \rho_{\alpha',\alpha''} \le \delta , \qquad (4)$$

$$\exists \alpha', \alpha'' \in \{\alpha_1,...,\alpha^*,..., \alpha_\beta\} \text{ such that } \rho_{\alpha',\alpha''} > \delta . \qquad (5)$$

Implication (4) denotes that choice of α^* provides forming of maximal $\Delta_{\beta+1}^i$-block of matrix $d(X)$ on set $\{p_1, p_2,..., p_r\}$. Thus having redefined the search domain

$$\{p_1, p_2,..., p_r\} \leftarrow \{p_1, p_2,..., p_r\} \setminus \{\alpha_1, \alpha_2 ,..., \alpha_\beta\} \qquad (6)$$

we can move on to forming the next maximal Δ_j^i-block starting from (3) if $\{p_1, p_2,..., p_r\} \ne \varnothing$.

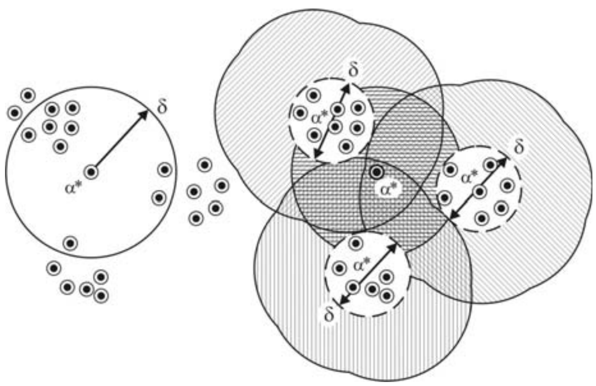
Note that situation described in (5) is much more complicated, but it can be brought to (4) by iterative elimination of block outliers. Here again two situations are possible: equality of the eliminated elements amount on this step or their disparity.

Under disparity we sequentially eliminate α'_γ, ($\gamma = 1,...,\Gamma$, $\{\alpha'_0\}=\varnothing$, Γ: $\nexists \alpha', \alpha'' \in \{\alpha_1,...,\alpha_\beta\} \Rightarrow \rho_{\alpha',\alpha''} > \delta$, $\{\alpha_1,..., \alpha_\beta\} \leftarrow \{\alpha_1,..., \alpha_\beta\} \setminus \{\alpha'_{\gamma-1}\}$) such that

$$\alpha'_\gamma = arg \max_{s \in \{\alpha_1,..., \alpha_\beta\}} \{card\{\rho_{q,s}\} : \rho_{q,s} > \delta, q \in \{\alpha_1,..., \alpha_\beta\}\}$$

until (4) is fulfilled.

If cardinality of indexing set reduced in this way still exceeds the number of compact elements of next distance matrix row according to (3) criterion, then the next block is obtained. Elsewise having α'_γ temporary eliminated from $\{p_1, p_2,..., p_r\}$, we repeat considered steps till the next Δ^i_j -block is obtained. After that all deleted rows are brought back for further analysis. On figure 2, a) a geometrical interpretation of this case is shown. Fig. 2, b) illustrates the result of Δ^i_j -blocks forming.



a) b)

Figure 1. To the explanation of Δ^i_j -blocks forming

Consider choice of α^* when having multiple rows in (3). Emphasize that for (4) we have exactly β rows and α^* choice is not crucial as either all these elements will be simultaneously assigned to one Δ^i_j -block (see fig. 1) or there are several blocks with the same cardinality and they all will be obtained sequentially. If (5) holds then choice α^* is also arbitrary since up to numeration all maximal blocks will be sequentially formed by reduction till cases (4).

Further eliminating from consideration set of indices $\{1, 2,..., k_1\}$ we get matrix of $(n-k_1)\times(n-k_1)$ dimension. Repeating the procedure for every new distance matrix firstly we get required representation in result, secondly by virtue of blocks maximality on each step (with accuracy up to enumeration) it provides fulfillment of (1).

We shall order obtained Δ^i_j -blocks under their dimensions ascending. Let us have s_0 elements which were not included into any blocks, s_1 blocks of dimension l_1, s_2 blocks of dimension $l_2,\dots,s_t$ s_2 – of dimension l_t, i.e. $1{=}l_0 \le l_1 < l_2 < l_t \le \mathrm{N}$. Then supposing that blocks with dimension starting from l_i are to be partite, and denoting maximal dimension of resulting blocks by $\mathrm{M} \in [l_{i-1}, l_i] \cap \mathbb{N}$, we get two possible search strategies. One lies in searching the best Δ-representation block, and then choosing a block of nested Δ-representation and matching in the closest Δ^i_j -block. Second strategy implies choosing a block among the union of all blocks of two-level distance matrix Δ-representation. We shall concentrate on the second one i.e. maximal number of matches is equal to the sum of blocks and their maximal dimension

$$f(\mathrm{M}) = \mathrm{M} + \sum_{j=1}^{i-1} s_j + \sum_{j=i}^{t} \left\lceil l_j/\mathrm{M} \right\rceil s_j\,,\ i = \overline{1, t}\,. \quad (7)$$

Thus it is necessary to find the M value providing minimum of (7). We shall fulfill the search of $f(\mathrm{M})$ minimum on each $[l_0, l_1],\ [l_1, l_2],\dots,[l_{t-1}, l_t]$. Among the best obtained results the minimal one which provides nested partition of given Δ-representation blocks is chosen eventually. Let us analyze the $f(\mathrm{M})$ local minimum search procedure.

Let us chose the next partial interval $[l_{i-1}, l_i] \cap \mathbb{N}$ and replace goal function $f(\mathrm{M})$ with continuous one namely $\varphi(x) = x + u/x + v$ where $u = \sum_{j=i}^{t} l_j s_j > 0$, $v = \sum_{j=1}^{i-1} s_j > 0$. It is easy to see that

$$\forall x \in [l_{i-1}, l_i],\ \forall \mathrm{M} \in [l_{i-1}, l_i] \cap \mathbb{N} \Rightarrow \varphi(x) \le f(\mathrm{M})\,.$$

We have proved that the minimum $\varphi(x)$ is

$$\varphi^* = \min_{[l_{i-1}, l_i] \cap \mathbb{N}} \varphi(x) = \min\{\varphi(l_{i-1}), \varphi(l_i),$$
$$\varphi(\max\{\min\{\mathrm{M}_1^*, l_i\}, l_{i-1}\}), \varphi(\max\{\min\{\mathrm{M}_2^*, l_i\}, l_{i-1}\})\}$$

where $\mathrm{M}_1^* = \lfloor x^* \rfloor$, $\mathrm{M}_2^* = \lceil x^* \rceil$, $x^* = \sqrt{\sum_{j=i}^{t} l_j s_j}$.

If the error $f^* - \varphi^*$ (f^* is calculated at the same points) is not small we restrict the search domain

$$[l', l''] = \begin{cases} [\lfloor x' \rfloor, \lceil x'' \rceil], & m^* \in \{\mathrm{M}_1^*, \mathrm{M}_2^*\}; \\ [l_{i-1}, \lfloor x' \rfloor], & m^* = l_{i-1}; \\ [\lceil x'' \rceil, l_i], & m^* = l_i \end{cases}$$

and repeat the offered procedure to find the optimum.

3. Results of Experiments and Discussion

We have carried out a number of experiments to test distance matrix block-diagonal form construction and its application to images collection indexing. In our image retrieval system we provided a region-based capability of the search, i.e. proximity of images was determined by distance between their partitions produced by some segmentation. Metric [6] was chosen to calculate the similarity between partitions (segmentations) $\alpha = \{\mathrm{A}_i\}_{i=1}^{n}$ and $\beta = \{\mathrm{B}_j\}_{j=1}^{m}$

$$\rho(\alpha, \beta) = [\sum_{i=1}^{n} \sum_{j=1}^{m} \mu(\mathrm{A}_i \cap \mathrm{B}_j)\mu(\mathrm{A}_i \Delta \mathrm{B}_j)] / (\mu(\alpha))^2$$

where μ is any measure (here $\mu(\mathrm{A}) = card(\mathrm{A})$), Δ denotes symmetrical difference of sets. We have used

Berkley Segmentation Dataset which consisted of 12000 ground-truth segmentations of 1000 Corel images as image database. The question was how much a shape of block-diagonal matrix depended on the segmentation algorithm? For this purpose we compared result matrix formed on the base of ground-truth segmentation with that formed on the base of well-known color-texture JSEG segmentation algorithm [11]. Experiments showed that capacity and content of each block of matrix was approximately the same for both types of segmentation. For example, on Fig. 2 you can see different segmentation type results included into the first maximum Δ-blocks.

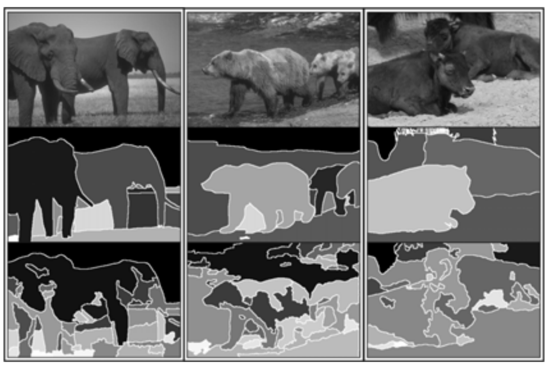


Figure 2. Example of images from Δ-block. First row – original images, second row – ground-truth segmentation, third row – JSEG segmentation

Simple strategy of discarding blocks on a search stage for range queries with radius ε was implemented. We selected median element x_i^* in every block and calculated distance $\rho(x_i^*, y)$ between query object y and these median elements. Then by the triangular inequality we could discard block i if $\rho(x_i^*, y) > \varepsilon + \delta$ or accept all its elements if inequality $\rho(x_i^*, y) < \varepsilon - \delta$ holds. Multiple experiments confirmed considerable speedup of the search.

We have presented a new indexing method in the metric space. It takes a distance matrix between database elements as input and transforms it to block-diagonal shape with compactness property of each block (distance between any of its elements is not less then specified threshold). Following the requirement of minimum number of distance evaluations on the search stage, we constructed a target matrix in such a way that number of blocks is minimal and then one could refine capacity of each block by solving formulated optimization task. As a result we obtain compact representation of distance matrix and can preserve the number of distance evaluations in worse-case conditions. This has a really high significance for large-scale image databases retrieval applications, where more and more complicated image abstractions and metrics for their comparison are used in order to bridge a semantic gap what leads to increase of search complexity. We successfully applied developed indexing method to speedup region-based search of images with no degradation of search quality (e.g. recall/precision measure). It should be noted, however, that there are still a lot of open issues here: selection of the δ value in order to construct more optimal in sense of matches number block-diagonal shapes of distance matrix, comparison with novel indexing methods like D-index. Finally, we think that our method could be effectively combined with existent indexing methods, for example, elements inside every Δ-block can be preprocessed via some other indexing method.

References

[1] C. Carson, and al. *Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying*. IEEE Trans. on PAMI, 24 (8): 1026–1038, 2002.

[2] J. Wang, J. Li, G. Wiederhold. *SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries*. IEEE Trans. on PAMI, 23 (9): 947–963, 2001.

[3] C. Yu, and al. *Indexing the Distance: An Efficient Method to KNN Processing*. Int. Conf. on Very Large Data Bases: 421–430, 2001.

[4] Y. Rubner, C. Tomasi, L. Guibas. *The Earth Mover's Dstance as a Metric for Image Retrieval*. Int. Journal of Computer Vision, 40 (2): 99–121, 2000.

[5] M. Meila. *Comparing Clusterings by the Variation of Information*. COLT/Kernel '03: 173–187, 2003.

[6] D. Kinoshenko, V. Mashtalir, V. Shlyakhov. *A Partition Metric for Clustering Features Analysis*. Int. Journal 'Information Theories and Applications', 14 (3): 230–236, 2007.

[7] G. Hjaltason, H. Samet. *Index-driven similarity search in metric spaces*. ACM Trans. on Database Systems (TODS), 28 (4): 517–580, 2003.

[8] M.L. Mico, J. Oncina, E. Vidal. *A New Version of The Nearest-Neighbour Approximating and Eliminating Search Algorithm (AESA) with Linear Preprocessing Time and Memory Requirements*. Pattern Recognition Letters, 15 (1): 9–17, 1994.

[9] P. Yianilos. *Data structures and algorithms for nearest neighbor search in general metric spaces*. Proc. of the fourth annual ACM-SIAM Symposium on Discrete Algorithms: 311–321, 1993.

[10] V. Dohnal. *D-Index: distance searching index for metric data sets*. Multimedia Tools and Applications, 21 (1): 9–33, 2003.

[11] Y. Deng, B.S. Manjunath. *Unsupervised segmentation of color-texture regions in images and video*. IEEE Trans. on PAMI, 23 (8): 800–810, 2001.