UDC 004.032.2

IMAGE RETRIEVAL IN DATABASES

Ye. A. Bashkov, N. S. Kostyukova, O. L. Vovk

Donetsk National Technical University

Abstract

This paper considers the problem of content-based color image retrieval in large databases and describes image database queries on primitive, abstract and logical levels. A detailed description of the approach based on 2D color histograms is provided, image retrieval results are estimated. A multi-level agglomerative method ant its results are considered.

Keywords: image retrieval, images similarity, image features, histogram image features, 2D color histogram, image clustering, multi-level agglomerative method, k- means method

Retrieval of images similar to a given sample is very important and interesting in terms of artificial intelligence systems development. This task is closely related to the domain of visual image analysis. This problem is difficult enough for computers, though people solve it intuitively, taking nearly no time. Its complexity is caused both by the difficulty to adequately describe images on a formula basis, and by the fact that the physiology of human vision, as well as the mechanism of perceiving information with the eyes, have not yet been studied enough to be fully understood.

In general, the problem of retrieving images that look like the sample from a database is stated as follows:

Database S contains information about a number V of uncompressed images:

$$S=\{P, F\},$$

where $P = \{P_k, k=1, 2, \dots V\}$ is a set of images (two-dimensional arrays which contain the image's color, with the dimensions being the x- and y- locations in the image),
$F = \{F_k, k=1, 2,.. V\}$ is a set of image content features (F_k is a scalar or a vector).

Let P_q be the query image (retrieval sample), its contents are described by F_q. For the query image P_{img} and any image from database P_k the similarity of their content features can be calculated:

$$d_k = f(F_q, F_k), k = 1, 2, ..., V .$$

It is necessary to define a set of images Q, which resemble the sample and are arranged in decreasing order of this resemblance:

$$Q = \{P_i, i = 1, 2, \dots \mid \forall\, i_1 > i_2 : d_{i_1} < d_{i_2}\} .$$

Thus, the problem of content-based image retrieval is reduced to the calculation of the values of d_k, $(k=1, 2, \dots n)$, and their following ordering. As this takes place, the calculation of image features F is performed prior to having these images stored in the database.

Contemporary researches in this area are largely aimed at solving two problems: the most reliable formal description of the images content and the efficient comparison of these descriptions from the point of view of time and memory requirements. Some approaches are oriented to specific categories of images, although the most difficult is to solve this problem for unconstrained images. Users' requests to the image collections are traditionally classified according to three levels of abstraction [1]: a **primitive** level (retrieval based on visual variables: color, form, texture – finding images akin to the specimen), a **logical** level (identification of the submitted object; for example, to identify an image of the Eiffel tower), and an **abstract** level (taking into account depictions of scenes – to find images which evoke certain feelings).

At the first, primitive level, researchers used for image content description the color histograms [2], which describe colors of pixels but ignore colors location. One approach for spatial information account is the extraction of image areas features. These areas may be extracted by different means. The least difficult one is a manual selection of the areas, in which the image is handled by humans, and all the necessary information is determined visually. Among other factors, the designers of the QBIC [2] system used this approach. However, the manual selection of areas and objects is rather awkward and requires too much time when large sets of images are considered.

Other decision deals with image decomposition into fixed fragments. This approach proposed by Stricker and Dimai for color image search [3] divides images into five domains. However, in general it is difficult to choose a scale for dividing each particular image. Moreover, this operation is non–invariant for both compressing and re-scaling of the images.

The third approach to domain determination involves image segmentation. Segmentation is a process of image dividing into homogeneous non–overlapping segments. It is necessary that sets of segments should be complete. The samples of the selected domains are not required to be known beforehand during image segmentation. However, segmentation is not a clearly defined process for unconstrained images. The retrieval of homogeneous areas for these images can be considered as an incorrectly formulated problem as the notion "homogeneity" depends on the domain of the image use. Moreover, for unconstrained images there is no formal method of evaluating the images segmentation, and the problem of image segmentation, which has many limitations, does not have only one solution [4]. Many methods of image segmentation have been suggested. Chua, Lim, and Pung [5] suggest segmentation based on color pairs. Hsu, Chua, and Pung [6] enlarged this approach, primarily focusing on the colors of the objects nearest to the viewer. There is one more means for expressing the contents of color images. It is the color correlogram (or "correlogram") – a spatial pair correlation of color changes with distance [7]. In contrast to histograms, correlograms take into account spatial color distribution in the image. Correlogram intersection allows one to see how much one of them differs from another and to measure the difference. It is shown in [7] that for color correlogram comparison, this method is the best.

With the use of histogram features the first stage of retrieval algorithm is color (or lighting levels) quantization. Quantization is the reduction of initial image colors to a basic set of predefined colors or lighting levels. This phase is essential, and without it the comparison of histogram features is senseless. Usually Euclidean distance, histogram intersection, or cosine or quadratic distances are used for the calculation of the images similarity rating [3]. Any of these values does not reflect the similarity rate of two images in itself. It is useful only if compared to other similar values. This is the reason that all the practical implementations of content-based image retrieval must complete computation of all images from the database. It is the main disadvantage of these implementations.

The quantization of lighting levels or color, calculation of histogram image features, and the comparison of images features and their ordering are completed sequentially during the process of content-based image retrieval. In [8] the authors proposed to use 2D-color histograms for color image content representation. 2D-color histogram is analogous to the texture histogram. But it also considers the relation between the pixel pair colors (not only the lighting component). 2D-color histogram is a two-dimensional array, $C_{max}*C_{max}$, where C_{max} is the number of basic colors used in the phase of color quantization. These arrays are treated as matrices, each element of which stores a normalized count of pixel pairs, with each color corresponding to the index of an element in each pixel neighbourhood. For the comparison of 2D-color histograms it is suggested calculating their correlation, because a 2D-color histogram, constructed according to the algorithm described above, is a random vector with dimension C_{max}^2 (in other words, a multidimensional random value). While creating a set of final images, the images should be arranged in decreasing order of the correlation coefficient.

An estimation of retrieval advantages using 2D color histograms was made on the basis of precision–recall plot described in [3] and commonly used to evaluate the quality of image retrieval. These variables, deduced during the retrieval among 2D color histograms, were compared with retrieved variables for traditional color histograms (fig. 1). From the figure it will be obvious that 2D color histogram usage allows higher quality retrieval. Moreover, the analysis of variable changes, made in [8], showed that the average value of Recall variables is 10% higher than normal during 2D- color histograms usage. The average value of a given Recall variable is approximately 2.33 times higher than the average value of this variable using traditional color histograms.

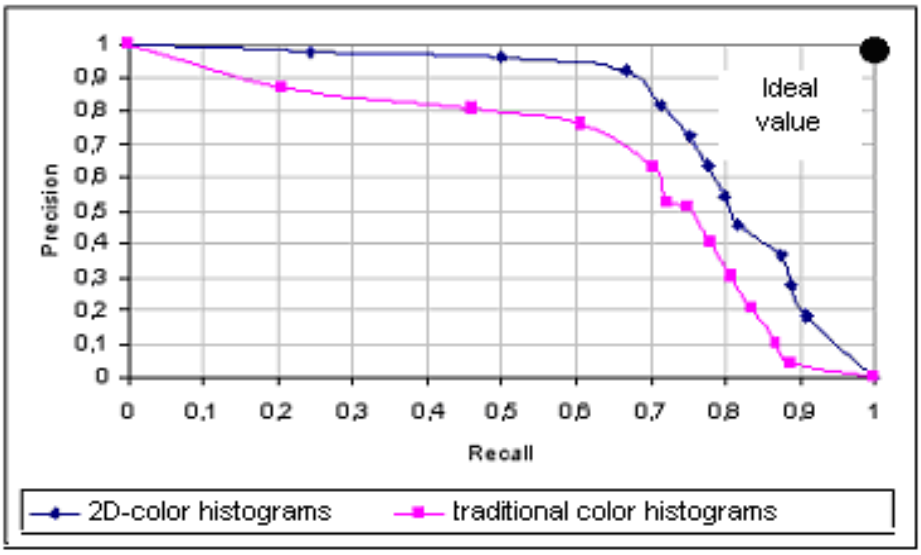

Figure 1. Effectiveness of image retrieval

The above method that uses 2D–color histograms has advantages with regard to quality tracking if compared to traditional color histogram usage.

The problem of how to deal with a particular region of an image has been of importance and concern for many researchers of the second, logic level. The researchers faced the problem of a constantly growing range of applications, which use the files of visual information. This problem leads to the necessity of high-speed methods of image region depiction. One of such methods is a statistical multi-level agglomerative method of image clustering [9].

During image clustering, input objects are pixels, each determined by the vector of color components. In the course of the clustering procedure, the integration of pixels into separate groups (clusters, or regions) based on numeric values of color variables takes place.

Let us consider the method of image region depiction which involves a binary mask of correlations and ranks of the color components of clusters centers [10].

It has been suggested that the mask should be defined as a multidimensional vector for the RGB color space, which includes vectors of ranks $\overline{s}_R, \overline{s}_G, \overline{s}_B$ and vectors of correlation $\overline{s}_{GB}, \overline{s}_{RB}, \overline{s}_{RG}$:

$$\overline{S} = \{\overline{s}_R, \overline{s}_G, \overline{s}_B, \overline{s}_{GB}, \overline{s}_{RB}, \overline{s}_{RG}\},$$

in which the vectors components can accept only two values: 0 and 1.

Let us denote the vectors of ranks $\overline{s}_R, \overline{s}_G, \overline{s}_B$ as a vector $\overline{s}_\alpha$ ($\alpha = R, G, B$), which can be described as:

$$\bar{s}_\alpha = (s_{\alpha 1}, s_{\alpha 2}, s_{\alpha 3}).$$

The vector components $\bar{s}_\alpha$ are computed from the formula:

$$s_{\alpha 1} = \begin{cases} 0, \alpha \in [x_l, GH - eps); \\ 1, \alpha \in [GH - eps, x_h]; \end{cases}$$

$$s_{a2} = \begin{cases} 0, \alpha \in [x_l, GL - eps) \cup (GH + eps, x_h]; \\ 1, \alpha \in [GL - eps, GH + eps]; \end{cases}$$

$$s_{\alpha 3} = \begin{cases} 0, \alpha \in (GL + eps, x_h]; \\ 1, \alpha \in [x_l, GL + eps]. \end{cases}$$

At this point: [x_l, x_h] are a range of fluctuation of the color variables' numeric values (for the range of colors RGB: x_l=0, x_h=255), *GL*, *GH* are the top and bottom values of the numeric values of color variables for predetermined ranks (the author suggests three ranks such as low, middle, and high and three intervals corresponding to the above ranks: [x_l, *GL*], [*GL, GH*], [*GH,* x_h]), *eps* is the method parameters. The latter has been added in order to deal with more than one level for a particular color component, as well as more than one interconnection for a pair of components.

Let us denote the vectors of correlation $\bar{s}_{GB}, \bar{s}_{RB}, \bar{s}_{RG}$ as a vector $\bar{s}_{\alpha\beta}$ ($\alpha = R, G$ и $\beta = B, G$), which can be described as:

$$\bar{s}_{\alpha\beta} = (s_{\alpha\beta 1}, s_{\alpha\beta 2}, s_{\alpha\beta 3}).$$

The vector components $\bar{s}_{\alpha\beta}$ are computed from the formula:

$$s_{\alpha\beta 1} = \begin{cases} 0, \alpha > \beta; \\ 1, \alpha \le \beta; \end{cases}$$

$$s_{a\beta 2} = \begin{cases} 0, |\alpha - \beta| > eps; \\ 1, |\alpha - \beta| \le eps; \end{cases}$$

$$s_{\alpha\beta 3} = \begin{cases} 0, \alpha < \beta; \\ 1, \alpha \ge \beta. \end{cases}$$

The binary mask of correlation and the ranks of the cluster centers color components is aimed at defining the distinctive features of the clustered objects (peculiarities of the color field under study).

The statistical multi-level agglomerative method of image clustering comprises the following:

At the first stage of the method (a stage of complete connection), the processed number of clusters is decreased by means of distributing the pixels with the same binary masks between individual clusters.

It is suggested that initially each pixel of image t is defined as a separate cluster (it is a variable of agglomerativity). First the pixels with equal binary masks are distributed into separate clusters. For this purpose each pixel of the image is inserted in the corresponding binary masks of correlation and rank. Then the pixels with different masks are entered into different clusters (pixels with equal masks are correspondingly united into the same cluster).

Let there be an image of size *t* [*w x h*] whose pixels are defined by:

$$t = \{\ p_{jk} = \{r_{jk}, g_{jk}, b_{jk}\} \mid j \in [1, w],\ k \in [1, h]\ \},\ j \in N, k \in N.$$

Then the integration of pixels with the same binary masks can be nominally recorded in the order:

$$\exists m_1, m_2, ..., m_q : \overline{S}_{j_l k_l} = \overline{S}_{j_e k_e}$$

$$\forall l, e \le m_v, l \ne e, v \in [1, q], m_v \in [1, w \cdot h], m_v \in N, v \in N.$$

In formula: m_v is the number of elements of the group (cluster) with an index *v*; q is the number of clusters; $\overline{S}_{jk}$ is a binary mask of pixel pjk with coordinates (j, k) of image t;*l,e* are pixel indexes inside the cluster.

A set of clusters will be the result of the integration of pixels with the same binary masks:

$$A = \{a_1, a_2, ..., a_q\} : a_v = \cup p_{jk} : \overline{S}_{j_l k_l} = \overline{S}_{j_e k_e}$$

$$\forall l, e \le m_v, l \ne e, v \in [1, q], m_v \in [1, w \cdot h], m_v \in N, v \in N.$$

At the second stage of this method (the stage of unit connection) new clusters are set up by the cluster integration with minimum distance. The stage is repeated while the condition for clusters comparability is being obeyed. This is condition based on the binary mask of correlation and ranks [10].

Figure 2 shows the comparison of the given algorithm with the modification of the k-means method described in [11]. Here image database entropy is on the X-axis, and an average time of image clustering for every entered group is on the Y-axis. The comparison was conducted for the same number of clusters taken for each image processing.

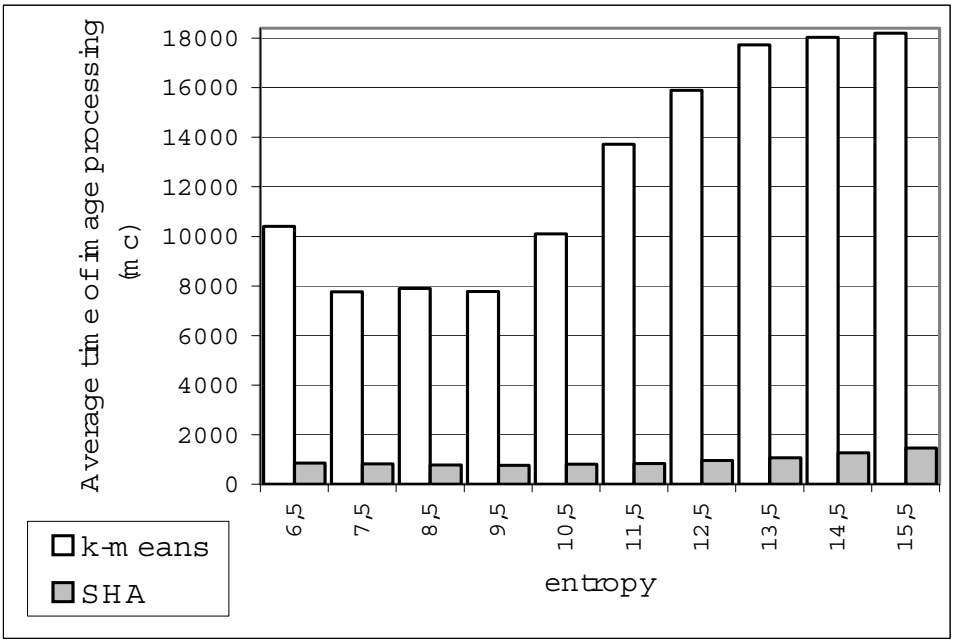

Figure 2. A bar diagram of experimental estimation of the average time used for image clustering by statistical multi-level agglomerate and k-means methods

These experiments of time consumption evaluation of image clustering displayed the superiority of the given method in comparison with the k-means one.

So, the problem of image retrieval is rather interesting. Traditionally, various approaches and methods have been used for solving this task. The method of 2D color histograms and multi-level agglomerative method turned out to have certain advantages if compared to the traditional approaches. Further research in this field should be aimed at reducing time complexity and improving retrieval results by means of advanced query realization on logical level.

References:

1. Chen, C.H.; Pau L.F.; & Wang P.S.P. (1998), „The Handbook of Pattern Recognition and Computer Vision" (2nd ed.). World Scientific Publishing Co.
2. Flickner, M.; Sawhney, H.; Niblack, W.; Ashley, J.; Qian Huang; Dom, B.; Gorkani, M.; Hafner, J.; Lee, D.; Petkovic, D.; Steele, D.; Yanker, P. (1995), "Query by image and video content: the QBIC system". IEEE Computer 28(9).
3. Smith, J.R. (1997), "Integrated Spatial and Feature Image Systems: Retrieval, Analysis and Compression". Graduate School of Arts and Sciences, Columbia University.
4. Stricker, M.; Dimai, A. (1996), "Color indexing with weak spatial constraints". Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases IV.
5. Chua, T.S.; Lim, S.K.; Pung, H.K. (1994), "Content-based retrieval of segmented images". Proceedings of ACM International Conference on Multimedia.
6. Hsu, W.; Chua, T. S.; Pung, H. K. (1995), "An integrated color-spatial approach to content-based image retrieval". Proceedings of ACM International Conference on Multimedia.
7. Huang, J.; Kumar, S.R.; Mitra, M.; Zhu, W.J.; Zabih, R. (1997), "Image indexing using color correlograms". Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition.
8. Bashkov, E.A.; Kostyukova, N.S. (2006), "Effectiveness estimation of image retrieval by 2D color histogram". Jornal of Automation and Information Sciences 6: 84-89
9. Bashkov, E.A.; Vovk, O.L. (2005), "Estimation of new statistic hierarchic aglomerattive clusterization algorithm for image region recognition effectiveness". System Research & Information Technologies 2: 117-130.
10. Vovk, O.L. (2006), "A new approach to visual similar image colors extraction". Jornal of Automation and Information Sciences 6: 100-105.
11. Wang, J. Z.; Li, J.; Wiederhold, G. (2001), „SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries". IEEE Transactions on Pattern Analysis and Machine Intelligence 23**(**9**)**: 947-963.

Received on 21.12.2009