Pattern-Based Retrieval of Cultural Heritage Images

Dimitris K. Iakovidis [1], Evangelos E. Kotsifakos[2], Nikos Pelekis[2], Haralampos Karanikas[1], Ioannis Kopanakis [3], Thomas Mavroudakis[1], Yannis Theodoridis[2]

[1] National & Kapodistrian Univ. of Athens, Knowledge Management Lab., Athens, Greece
dimitris.iakovidis@ieee.org, bkaranikas@phs.uoa.gr
[2] Univ. of Piraeus, Piraeus, Greece
{ek, npelekis, ytheod}@unipi.gr
[3]Technological Educational Institute of Crete, Heraklion Crete, Greece
i.kopanakis@emark.teicrete.gr

Abstract

Conventional Content-Based Image Retrieval (CBIR) systems make use of similarity measures estimated directly from low-level image features, involving multidimensional and exhaustive, nearest neighbor searching. In this paper we present an image retrieval methodology suited for efficient search in cultural heritage images that utilizes similarity measures defined over higher-level patterns associated with clusters of low-level image features. The similarity between two patterns is estimated as a function of the similarity between both the structure and the measure components of the patterns. We evaluate our system using cultural heritage images derived from the repository of the Foundation of Hellenic World (FHW). For the experiments we use Local Binary Pattern (LBP) and contrast distributions as image features and the results show that the proposed pattern-based approach efficiently retrieves images compared to commonly employed methods.

Keywords: Content-based image retrieval, Cultural images, Cohen's distance, Clustering.

1. Introduction

Content-Based Image Retrieval (CBIR) of cultural heritage images is an emerging field of research bridging society, culture and information technology [Ching et. al. (2004)]. Querying by example databases of paintings, sculptures, photographs, and documents of historical value from different civilizations, would facilitate both educational and research and enable the exploration of unknown inter and intra cultural associations.

In the last decade the advances in information technology allowed the development of Content-Based Image Retrieval (CBIR) systems, capable of retrieving images based on the similarity their features have with the features of one or more query images. Some of these systems are QBIC [Faloutsos et. al. (1994)], VisualSEEK [Smith et. al. (1996)], Virage [Hampapur et. al. (1997)], Netra [Ma et. al. (1999)], PicSOM [Laaksonen et. al. (2000)], SIMPLicity [Wang et. al. (2001)], CIRES [Iqbal et. al. (2002)], and FIRE [Deselaers et. al. (2004)] . More than fifty CBIR systems are surveyed in [Veltcamp et. al. (2000)].

Recently, studies targeting especially to the retrieval of cultural heritage images have appeared. Most of these studies propose methods based on color image features [Ardizzone et. al. (2004)][Valle et. al. (2006)]. More sophisticated approaches include the utilization of wavelet domain feature descriptors in conjuction with mixtures of stochastic models for the retrieval of Chinese paintings [Jia et. al. (2004)].

Common ground for most of the systems cited above is that image retrieval is based on similarity measures estimated directly from low-level image features, whereas it involves multidimensional, usually exhaustive, nearest neighbor searching over the whole set of the available feature vectors. However, such an approach can be time consuming with large image databases.

Research on improving the efficiency of the image retrieval process has mainly focused on image indexing techniques by utilizing data structures, such as R-trees [Faloutsos et. al. (1994)] [Petrakis et. al. (1997)], feature index trees [Grosky et. al. (1990)], iconic index trees [Wu et. al. (1994)], and meshes of trees [Jeng et. al. (2005)]. Other approaches to improving efficiency, include clustering of the image feature spaces [Stehling et. al. (2001)] [Zhang et. al. (2002)], and utilization of alternative similarity measures, usually dependent on feature sets [Berman et. al. (1997)] [Stehling et. al. (2002)].

In this paper, we propose a novel approach to efficient content-based retrieval of cultural heritage images that utilizes similarity measures defined over higher-level *patterns* associated with clusters of low-level image feature spaces. The term *pattern* is defined in the context of a state of the art framework named PANDA (PAtterns for Next generation Database systems) [Bartolini et. al. (2004)].

The rest of this paper comprises four sections. Section 2 outlines the PANDA framework. The proposed approach to CBIR is described in Section 3, and the results obtained from its application for the retrieval of cultural heritage images provided by the Foundation of Hellenic World (FHW) are apposed in Section 4. Finally, Section 5 summarizes the conclusions and the future perspectives of this study.

2. The PANDA Framework

Efficient management of patterns extracted from image databases is of vital importance due to the large storage requirements. Taking advantage of the PANDA framework [Bartolini et. al. (2004)] we adopt the idea of a Pattern Based Management System (PBMS) as the infrastructure for managing patterns extracted from our CBIR scheme, in contrast to DataBase Management Systems (DBMS). The key concept of PANDA is that any type of pattern can be represented in a compact and unified way. This can be achieved through a Pattern-Base (PB) keeping information about extracted patterns. Such a PB introduced in [Bartolini et. al. (2004)] consists of three basic layers: the *pattern*, the *pattern type* and the *class*. A *pattern type* is a description of the pattern structure. A *pattern* is an instance of the corresponding pattern type and class is a collection of semantically related patterns of the same pattern type.

Formally, a pattern type pt is a quintuple $pt = <n, ss, ds, ms, f>$, where n is the name of the pattern type. The structure schema ss defines the pattern space by describing the structure of the pattern type, while the source schema ds defines the data source space by describing the dataset wherefrom patterns of this pattern type are derived. The measure schema ms, quantifies the quality of the source data representation achieved by patterns of this pattern type and the formula f describes the relationship between the source space and the pattern space. In this notation, if pt is a pattern type then $p = <pid, s, d, m, e>$ is an instance of pt, where pid is a unique pattern identifier, s, d, m are the corresponding structure, source and measure of the pattern, while e is an expression indicating the section of the source space related to pattern p.

Aside from the physical storage of the patterns we appropriately utilize the PANDA framework for the comparison of patterns [Bartolini et. al. (2004)] as the base of our image retrieval methodology in our CBIR scheme. In PANDA the similarity expressed as a distance dis, where the minimum distance indicates the maximum best maching, between two patterns p_1, p_2 of the same type can be computed by combining, by means of an aggregation function f_{aggr}, the distance between both the structure s and the measure m components:

$$dis(p_1, p_2) = f_{\text{aggr}}(dis_{\text{struct}}(p_1.s, p_2.s), dis_{\text{meas}}(p_1.m, p_2.m)) \quad \textbf{(1)}$$

Efficient definition of the structure and measure of patterns extracted from cultural heritage images, as well as appropriate selection of aggregation logic and distance functions to assess the respective distances, are some challenges faced in this paper.

3. Retrieval of Cultural Images using Patterns

The proposed CBIR approach is outlined in Fig. 1. It involves four steps: a) feature extraction from each of the stored and the query images, b) clustering of the extracted feature vectors per image, c) pattern instantiation of the clusters, and d) computation of pattern similarities. The registration of a new image in the database involves the first three of the four steps described for image retrieval (a, b, and c).

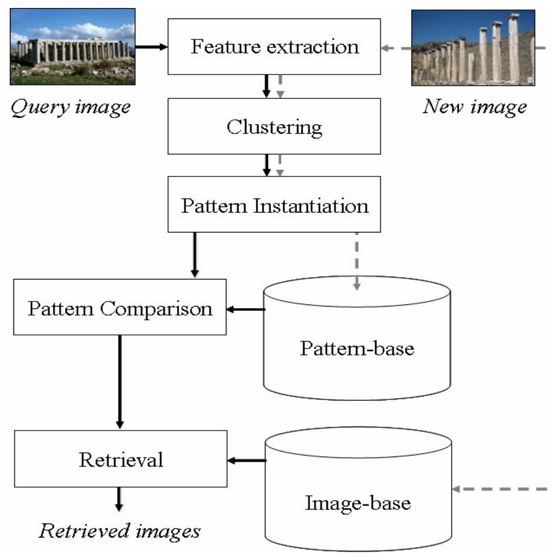

Figure 1. *Outline of the proposed pattern-based CBIR approach. The solid arrows indicate the data flow for image retrieval, whereas the dashed arrows indicate the data flow for the registration of a new image.*

3.1 Feature Extraction

Each of the images stored in the database, as well as the query image is raster scanned with a sliding window of user-defined size and sliding step. The sliding step may allow windows to overlap between each other. For each window N features f_i, i=1,2, N are calculated to form a single feature vector F. The number of feature vectors produced for each image depends on the size, the dimensions and the step of the sliding window.

Aiming to illumination invariant representation of the images we have considered Local Binary Pattern (LBP) distributions as features. LBP features are calculated from the weighted mean of pixel values over a small neighborhood, in which each pixel is considered separately. The LBP features were supplemented by an orthogonal measure of local contrast according to which the average of the gray levels below the center pixel is subtracted from that of the gray levels above (or equal to) the center pixel [Ojala et. al. (1996)]. Comparative studies have demonstrated that the use of LBP along with contrast distributions may result in higher classification accuracy than the Gabor and wavelet features, with a smaller computational overhead [Maenpaa et. al. (2004)][Iakovidis et. al. (2005)].

3.2 Clustering

Feature extraction is followed by clustering using the Expectation Minimization (EM) algorithm [Dempster et. al. (1997)]. The EM algorithm is a widely-used statistical clustering method. It performs clustering by estimating the mean and standard deviation of each cluster, so as to maximize the likelihood of the observed data.

3.3 Pattern Instantiation

Given a clustering of an image comprising M clusters C_i, i=1,2,..,M, according to the PANDA formalization and with respect to the output of the EM algorithm a pattern $object_i$ is instantiated for each cluster C_i representing an object depicted in a cultural heritage image:

$$object_i = \begin{pmatrix} SS : (D : [[mean : [\text{Real}], s\text{tdDev} : [\text{Real}]]_1^N), \\ MS : (pp : \text{Real}) \end{pmatrix} \quad \textbf{(2)}$$

where *mean* and *stdDev* are the mean and the standard deviation of the distribution D_{j} for every one of the N features (j=1,2,..,N) in cluster C_i, respectively, and *pp* is the prior probability of C_i. Here prior probability is defined as the fraction of the feature vectors of the image that belong to cluster C_i. Intuitively, prior probability *pp* is equivalent with the *support* measure widely used in data mining models. In our case, in addition to the qualitative aspect of the prior probability, it also provides an indication of the size of the object.

In this connection, a medical image is considered as a complex pattern defined by (3), consisting of a set of simple clusters each one of them represented by the mean and standard deviation values of a distribution.

$$image = \begin{pmatrix} SS : \{object\}, \\ MS : \perp \end{pmatrix} \quad \textbf{(3)}$$

3.4 Computation of Pattern Similarities

Aiming at the definition of the similarity of two images (i.e., complex patterns), we have first to define the similarity between the measures and the structures among two clusters C_1 and C_2 (i.e., simple patterns). This similarity is expressed as the distance between two images and the components of the distance computation are analyzed below. The distance between the measures of two clusters is computed using the Euclidean distance as in Eq.(4)

$$dis_{meas}(C_1, C_2) = \left| C_1.pp - C_2.pp \right| \quad \textbf{(4)}$$

Rephrasing the problem of defining the structural distance between C_1 and C_2 we need to find a measure for evaluating the closeness of two sets of distributions, as C_1 and C_2 are. Further decomposing the problem, we should first define a method of computing the distance between just two distributions D_1 and D_2. To achieve this, we use the standardized difference *Cohen's d* between two distributions as it has been defined by Cohen [Cohen (1988)]. *Cohen's d* is defined as the absolute difference between the means of the distributions, divided by the root mean square of the two standard deviations. When the two standard deviations are similar the root mean square will not differ much from the simple average of the two variances (Eq. 5).

$$d = \frac{\left|D_1.mean - D_2.mean\right|}{\sqrt{\frac{D_1.stdDev^2 + D_2.stdDev^2}{2}}} \quad \textbf{(5)}$$

Cohen's d is a non-negative real number interpreting the overlap between two distributions. If d is zero, the distributions are identical. Small d's indicate more similar distributions whereas larger d's indicate less similar distributions. If the two standard deviations are both zero, we use the absolute difference of the means as the distance between the distributions.

Given the above, we define the distance between distributions D_1 and D_2, by Eq.(6).

$$dis(D_1, D_2) = \begin{cases} d = \frac{\left|D_1.mean - D_2.mean\right|}{\sqrt{\frac{D_1.stdDev^2 + D_2.stdDev^2}{2}}}, \\ if\ D_1.stdDev \neq 0\ and\ D_2.stdDev \neq 0 \\ \\ \left|D_1.mean - D_2.mean\right|, otherwise \end{cases} \quad \textbf{(6)}$$

This is a means to automate and materialize the intuitive overlap between two distributions. Having this, we let the structural distance between two sets of distributions (i.e. two clusters C_1 and C_2) be the average among the distances computed for each pair of the N features:

$$dis_{struct}(C_1, C_2) = \sum_{j=1}^{N} dis(D_j^1, D_j^2) \Big/ N \quad \textbf{(7)}$$

We aggregate the distances between the qualitative dis_{meas} and the structural dis_{struct} distances between the clusters by using the following aggregation function f_{aggr}, which gives the same weight to either of the above distances, while further weights the overall distance by the mean of prior probabilities of the clusters, as a bias towards similar and concurrently big clusters.

$$dis(C_1, C_2) = \frac{dis_{struct}(C_1, C_2) + dis_{meas}(C_1, C_2)}{2} \cdot (C_1.pp + C_2.pp)/2 \quad \textbf{(8)}$$

Having defined the similarity between clusters (i.e., simple patterns), to compare two images I_1 and I_2 (i.e., complex patterns) we need to determine the coupling methodology between the different clusters of each image. Though various coupling types can be applied in the context of the PANDA framework [Bartolini et. al.(2004)], we adopt the matching of Eq.(9), allowing each cluster of the first image to match more than one cluster of the second, and vice versa.

$$dis(I_1, I_2) = \frac{1}{M^2}\left(\sum_{i=1}^{M}\sum_{j=1}^{M} dis\left(C_i^{I_1}, C_j^{I_2}\right)\right) \quad \textbf{(9)}$$

4. Results

The experiments aim to demonstrate the efficiency of the proposed pattern-based approach to CBIR over the approach used by conventional CBIR systems. Cultural heritage images originating from the database of the Foundation of Hellenic World (FHW) comprise the dataset used in the experiments. The images span five classes, namely ancient monuments, coins, photo portraits, drawings, and marbles. These include both color and grey level images of different sizes, inconsistently acquired from different sources. All images have been converted to 8-bit grey level format and have been downscaled to fit into a 256×256 bounding box.

Figure 2. *Sample images from the cultural image database used in the experiments.*

A total of 5,000 regions were sampled from the available images using 128×128-pixel windows with a 96-pixel overlap. The feature vectors extracted from each image were clustered by means of the EM algorithm implemented in the WEKA data mining tool [Witten et. al. (2005)]. A binary clustering approach was followed, considering that the images contain one or more objects of interest of the same kind (e.g. one or two coins), and background information.

For each cluster a pattern $object_i$, i=1,2 (Eq.2) was assigned, and each image was represented by a complex pattern *image* (Eq.3). The collection of patterns originating from the images registered in the database was used to build the pattern-base. Subsequent queries were executed to evaluate the performance of the proposed approach.

The performance of the proposed pattern-based approach to CBIR in comparison with the conventional, exhaustive approach is illustrated in Fig. 4, in terms of the number of comparisons between the query and the registered data. It can be observed that the proposed approach achieves approx. 156 times less comparisons.

The retrieval performance of the proposed CBIR approach with the LBP and contrast distributions was estimated 80.4%. The respective performance obtained using standard 3-level Discrete Wavelet Transform (DWT) energy features was 62.2%.

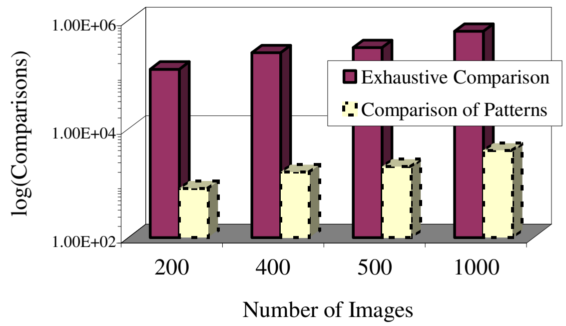

Figure 3. *Number of comparisons between the query and the registered data for the conventional and the proposed approaches.*

5. Conclusions

We presented a novel approach to efficient content-based retrieval of cultural images. This scheme utilizes *patterns*, defined in the context of the state of the art framework PANDA. Image retrieval is based on similarities defined over complex patterns. Searching for similar images involves comparison of patterns instead of exhaustive, nearest neighbor searching over the whole set of the feature vectors stored in the image database. The results advocate to the efficiency of the proposed scheme for image retrieval from large databases over the commonly employed methods.

Storing clustering patterns along with the low-level features set in a unified format facilitates further processing and analysis. To date, most data mining algorithms have concentrated on the extraction of interesting rules directly from low-level data [Spiliopoulou et. al. (2000)]. Our approach provides the means for deriving rules from the results of other data mining algorithms that is, mining from rules set. In the current work our initial low-level feature set is further processed and represented via clustering by higher-level patterns which are in a machine processible format. A significant advantage of this approach is that the nature of the rules to be extracted by this process contains different higher order semantics.

Future perspectives of this work include research on ontology-based translation of the high-level patterns into semantics in order to bridge the gap between low-level features and the higher-level real world concepts.

Acknowledgment

This research work is supported by the project MetaOn, founded by the Operational

Programme “Information Society” of the Greek Ministry of Development, General Secretariat for Research and Technology, co-funded by the European Union. Research also supported by the General Secretariat for Research and Technology of the Greek Ministry of Development under a PENED’2003 grant. The collection of images used in this study provided by Foundation of the Hellenic Worlds (FHW).

References

Ardizzone E., Chella A., Pirrone R., Gambino O., (2004), *An image retrieval system for artistic database on cultural heritage*, Atti Conferenza Italiana sui Sistemi Intelligenti, (CISI 2004), Perugia, Italia, 2004.

Bartolini I., Ciaccia P., Ntoutsi I., Patella M., and Theodoridis Y., (2004), *A Unified and Flexible Framework for Comparing Simple and Complex Patterns*, in Proc. 8th Eur. Conf. on Principles and Practice of Knowledge Discovery in Database, pp. 496-499, 2004.

Berman A., and Shapiro L. G., (1997), *Efficient Image Retrieval with Multiple Distance Measures*, in Proc. SPIE Conf. on Storage and Retrieval for Image and Video Databases, pp. 12-21, 1997.

Chen C.C, Wactlar H., Wang J., and Kiernan K., (2004), *Digital Imagery for Significant Cultural and Historical Materials - An Emerging Research Field Bridging People, Culture, and Technologies*,' Int. J. Digital Libraries, 2004.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Dempster A., Laird N., and Rubin D., (1997), *Maximum likelihood from incomplete data via the EM algorithm*, Jour. Royal Stat. Soc., Series B, vol. 39, no.1, pp.1–38, 1977.

Deselaers T., Keysers D., and Ney H., (2004), *FIRE - Flexible Image Retrieval Engine: ImageCLEF 2004 Evaluation*, LNCS 3491, pp. 688-698, 2004.

Faloutsos C., Barber R., Flickner M., Hafner J., Niblack W., Petkovic D., and Equitz W., (1994), *Efficient and Effective Querying by Image Content*, J. Intell. Inf. Systems, vol. 3, pp. 231-262, 1994.

Grosky W.I. and Mehrotra R., (1990), *Indexed-Based Object Recognition in Pictorial Data Management*, Comp. Vision, Graphics, Image Proc., vol. 52, pp. 416-436, 1990.

Hampapur A., Gupta A., Horowitz B., Shu C. F., Fuller C., Bach J., Gorkani M., and Jain R., (1997), *Virage video engine*, in Proc. SPIE: Storage and Retrieval for Image and Video Databases V, pp. 188-197, San Jose, February 1997.

Iakovidis D.K., Maroulis D.E., Karkanis S.A., (2005), *A Comparative Study of Color-Texture Image Features*, in Proc. IEEE Int. Workshop on Systems, Signal and Image Processing (IWSSIP), Halkida, Greece, 2005, pp. 205-209.

Iqbal Q., and Aggarwal J. K., (2002), *CIRES: A System for Content-based Retrieval in Digital Image Libraries*, in Proc. Int. Conf. Control, Automation, Robotics and Vision (ICARCV), pp. 205-210, December 2-5, 2002.

Linacre, J.M. (1996), “Overlapping Normal Distributions,” *Rasch Measurement*

Transactions, vol. 10, no. 1 pp. 487-488.
Jeng W.-M., and Hsiao J.-H., (2005), *An Efficient Content Based Image Retrieval System Using the Mesh-of-Trees Architecture*, J. Inf. Science Eng., vol. 21, pp. 797-808, 2005.
Jia L. and Wang J. Z., (2004), *Studying Digital Imagery of Ancient Paintings by Mixtures of Stochastic Models*,' IEEE Trans. on Im. Proc., vol.12, no.2, pp., 2004.
Laaksonen J., Koskela M., Laakso S., and Oja E., (2000) *PicSOM - Content-Based Image Retrieval with Self-Organizing Maps*, Pattern Recognition Letters, vol. 21, pp. 1199-1207, 2000.
Ma W.Y., and Manjunath B.S., (1999), *Netra: A Toolbox for Navigating Large Image Databases*, Multimedia System, vol. 7, pp. 184-198, 1999.
Mäenpää, T., and Pietikäinen, M. (2004), *Classification with color and texture: Jointly or separately?*, Pattern Recognition, Vol. 37, No. 8, pp. 1629–1640.
Ojala T., Pietikainen M., Harwood D., (1996), *A Comparative Study of Texture Measures with Classification based on Feature Distributions*, Pattern Recognition, vol. 29, 1996, pp. 51-59.
Petrakis E. G.M., Faloutsos C., (1997), *Similarity Searching in Medical Image Databases*, IEEE Trans. Knowl. Data Eng., vol. 9, no. 3, pp. 435-447, 1997.
Smith J.R., and Chang S.F., (1996), *Visualseek: A Fully Automated Content-Based Image Query System*, in Proc. ACM Int'l Multimedia Conf., pp. 87-98, 1996.
Spiliopoulou M., and Roddick J. F., (2000), *Higher Order Mining: Modellingand Mining the Results of Knowledge Discovery*, in Proc. Int. Conf. Data Mining, vol. 2, pp. 309-320, 2000.
Stehling R.O., Falcao A.X., and Nascimento M.A., (2001), *An Adaptive and Efficient Clustering-Based Approach for Content-Based Image Retrieval in Image Databases*, in Proc. Int. Symp. Database Eng. & App. (IDEAS 01), pp. 56-365, 2001.
Stehling R.O., Nascimento M.A., and Falcão A.X., (2002), *A Compact and Efficient Image Retrieval Approach based on Border/Interior Pixel Classification*, in Proc. 11th Int. Conf. Inf. Knowledge Man. (CIKM'02), ACM Press, NY, pp. 102-109.
Valle E., Cord M., and Philipp-Foliguet S., (2006), *Content-Based Retrieval Of Images For Cultural Institutions Using Local Descriptors*, Int. Conf. on Geometric Modeling and Imaging, London, 2006.
Veltcamp R., and Tanase M., (2000) "Content--Based Image Retrieval Systems: A Survey" Tech. Report UU-CS-2000-34, Dept. of Comp. Sci., Utrecht Univ., 2000.
Wang J.Z., Li J., and Wiederhold G., (2001) *SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries*, IEEE Trans. Patt. Anal. Machine Intell., vol. 23, no. 9, pp. 947-963, Sept. 2001.
Witten I. H., and Frank E., (2005), *Data Mining: Practical machine learning tools and techniques*, 2nd Edition, Morgan Kaufmann, San Francisco, 2005.
Wu J.K., and Narasimhalu A.D., (1994), *Identifying Faces Using Multiple Retrievals*, IEEE Multimedia, vol. 1, no. 2, pp. 20-38, 1994.
Zhang R., Zhang Z.(M), (2002) *A Clustering Based Approach to Efficient Image Retrieval*, Proc. 14th IEEE Int. Conf. Tools Artif. Intel. (ICTAI'02), p. 339, 2002.