CHAMELEON: A Region-Based Image Retrieval System *

J.Freixenet, X.Muñoz, X.Cufí, J.Martí and X.Lladó
University of Girona.
Institute of Informatics and Applications.
Campus de Montilivi s/n. 17071. Girona, Spain
{jordif,xmunoz,xcuf,joanm,llado}@eia.udg.es

Abstract

A major research subject in image databases is to support efficient and effective access to images based on their visual content. In this paper, a new Content-Based Image Retrieval (CBIR) system is presented. In contrast to the classical approach of most of the CBIR systems which relies on global attributes (i.e. color histograms), CHAMELEON is region based and takes into account the spatial composition of the image. A segmentation technique integrating region and boundary information is used to extract homogeneous regions of images. Then, images are represented on indexes by the features of principal regions and their spatial relationships. Experimental results show that the proposed approach captures correctly the content of images and achieves promising retrieval results.

Keywords: Image databases, Content-based image retrieval, segmentation, color and texture features.

1 Introduction

Recent years have seen a rapid increase of the size of digital image collections, and their use in a wide range of applications areas is currently more and more habitual. Stating that Internet is a major source for such databases, the development of searching engines with the capability to retrieve relevant images from large databases is becoming increasingly important.

The existing search engines can be classified into two categories: text-based or image-content based systems. Text-based systems can be traced back to the late 1970's. In this approach, a set of keywords are first assigned to every image. Then, it can be retrieved by its corresponding annotations. However, this approach entails

Work partially funded by *Departament d'Universitats, Recerca i Societat de la Informació de la Generalitat de Catalunya*

two main difficulties: the first one, the large amount of manual effort required in image annotations; the other drawback is related with the need to express the content of the images in few words and the inherent subjectivity of the human perception. So, conventional attempts to cast visual features into textual keywords have been far recognized to be inadequate in indexing pictures. [1]. To overcome these difficulties content-based systems were proposed. A content-based image retrieval (CBIR) system uses information from the content of images such as color, texture, shape, etc. as the image index. We focus on image retrieval by image example, where an example query image is given by the user as the input. For a given query image, its content is directly compared with that of the images in the database and a desired number of images close to the query image are retrieved.

The current content-based image retrieval systems often rely on global image characteristics such as overall colour and texture histograms. Although simplicity is the major gain obtained from these approaches, they fail to capture the spatial layout and the structure of the image. To address this challenge, recent works have proposed an strategy consisting on using objects or regions to accurately represent the visual content of the image. This approach requires that images are first segmented into homogeneous regions, and afterwards, chromatic, geometric and spatial features of these regions are matched against corresponding regions in the query image. Despite this strategy holds the image layout, a reliable segmentation often involves such procedural drawbacks (e.g. associated computational time) that its use can be unsuitable. In this sense, the CHAMELEON project constitutes a region-based CBIR system intended to deal with reliable and fast segmentation.

The rest of the paper is organized as follows: section 2 gives a description of the CBIR engine, detailing the segmentation, the region description, and the retrieval algorithms. Results of queries in a collection of TV images are presented in section 3. Finally, further improvements and conclusions are discussed.

2 The CHAMELEON Project

Basically, images are first segmented into regions by a gross-segmentation technique integrating region and boundary information. Following, color, texture and spatial features are extracted from the regions of the images of the database, and these features are summarized in a reduced set of indexes. Then, given an image example, the retrieval process consists on segmenting and extracting features from this image, projecting them onto the indexes space and looking for the nearest neighbors.

2.1 Segmentation

To fully represent the content of an image with regions, the images are first segmented into several regions suitable for further analysis. Cooperative segmentation techniques have demonstrated, in different areas of application, its superiority regarding the traditional techniques. So, in this project a method based on the integration of region and boundary information is used (see [2] for details).

The method begins by detecting the main contours of the scene, which are characterized to have an important length and to separate two regions with appreciable differences in chromatic and textural features. The contours obtained represent the boundaries between regions, so every contour theoretically separates two adjacent regions in the image. Then a set of growing centers are chosen at each side and along the contour. Here, we suppose the seeds belonging to a determined side of the contour are associated to the same region. Then, these seeds are used as samples of the corresponding regions and analyzed in the chromatic space. The idea consists on to know a priori some characteristics of regions in order to establish appropriated criteria for the posterior growing processes. This allows to adjust the homogeneity criterion to the region's characteristics using specific criteria on each region. Finally, the seeds simultaneously start a concurrent growth using the criterion established for each region which is based on clustering analysis. From obtained regions, only those which have a superior size at 10% of the entirely image are considered sufficiently relevant to represent the content embedded in the image.

2.2 Region description

Color and texture are the features considered in our system for automated content extraction and database searching. These features are really important because colors and textures are fundamental characteristics of the content of many images, giving this work general application towards databases of images from a variety of domains. Color property of regions is extracted using the average of their pixels in the $l_1l_2l_3$ color space. Gevers and Smeulders [3] argued, based on theoretical and experimental results, the convenience of using this color space on image retrieval applications. Robustness was demonstrated against changes in imaging conditions: shadows, shading, highlights and intensity of illumination.

Textural information of regions is captured using the gray level co-occurrences matrices proposed by Haralick et al. In order to use the information contained in the gray level co-occurrence matrices Haralick et al. defined 14 statistical measures. Since many distances and orientations result in a very large number of values, computation of co-occurrence matrices and extraction of many features from them

become infeasible for an image retrieval application which requires fast computation. This aspect was noted by Aksoy and own Haralick in [4], where they argued the convenience of using only the variance as statistic measure of texture. We compute this feature for two distances, and for $0°$,$45°$,$90°$, and $135°$ orientations to constitute a 8-dimensional feature vector.

2.3 Spatial Relations

Spatial information is a critical component of image description and subsequent matching which has only recently been addressed by researchers in the field. Obviously, if the spatial structures of query regions are maintained, the accuracy of retrieval result be further improved.

The whole structure of an image is represented by the spatial relational graph (SRG). Psychological studies affirm that the vertical spatial relation is more important than the horizontal one. So, for simplicity we only consider the vertical relation. Let $V = R_1, R_2 \ldots R_N$ be the set of nodes (regions) and $E = (R_i, R_j), 1 \leq i, j \leq N$ be the set of edges between nodes which have a vertical spatial relation between them. Then, $G = (V, E)$ is a graph representing the SRG of an image.

2.4 Retrieval

In order to speed up the efficiency of retrieval, the method suggested in this paper is to first rank the desired images based on the similarity measure and then return the final target results based on the spatial measure. In real implementation, the entire database is first pruned to obtain a subset of compatible images using the visual similarity score. Then, based on the spatial measure, the subset of images are further ranked for improving the final retrieval results.

In the on-line search the indices permit to find similar images without looking the whole database. In our approach each one of the regions of interest of the query image becomes a descriptor of the image. In this sense, indices retrieve a subset of images of the database which contain descriptors of the query image. The indices space is set up by a classical hierarchical clustering algorithm and is organized in a R-tree structure. The visual similarity score S^t, calculated based on the approach of [5], is computed as follows: given the image query Q, consisting of n regions $(R_1, R_2, \ldots, R_n)$, the similarity between Q and a target image I^t of the database is

$$S^t(Q, I^t) = exp(-\sum_{j=1}^{n} dist(R_j, I^t)) \tag{1}$$

where $dist(R_j, I^t)$ is the distance between R_j and image I^t, and is calculated as

$$dist(R_j, I^t) = \min dist_E(R_j, R_k^t), \tag{2}$$

where R_k^t is one of the regions of I^t, and $dist_E(R_j, R_k^t)$ is the Euclidean distance between the feature vectors of R_j and R_k^t.

To improve the retrieval results, a subset of scored images can be re-ranked by using the spatial constraint imposed by the query image. Here, computing similarity between two images is equivalent to graph matching. As Jolion discussed in [6], many different algorithms have been proposed in the past for this kind of problem. Although their high complexity is generally a strong limitation due the number of nodes, here we have simplified the problem: 1) the graphs have few nodes corresponding to representative regions of images, and 2) previous pruning limits the matching to a reduced set of images.

3 Results

In order to analyze the performance of the retrieval system, CHAMELEON has been used to search images coming from television. The database contains about 2000 images belonging to different themes. Fig.1.a shows a set of example images randomly extracted from the database, while Fig.1.b shows the result of the retrieval using the football image as the image query. The smaller left window on the browser gives the user the possibility to graphically define and edit some parameters. For instance, the user can either retrieve images which respect the spatial configuration of the query image, or alternatively, disable spatial ranking to simply retrieve images containing similar types of regions.

4 Conclusions and further work

In this paper, we have presented the CHAMELEON project, a region-based CBIR system, which uses multiple regions as the key to retrieve images. Color and textural features of regions are combined with spatial relationships to capture the structural semantics of the content of images. Experimental results show than the results are much better than other methods, which utilize global features of the image.

Further work in CHAMELEON targets to improve the interface user-system in order to offer an easier way to query at the database. In this sense, we are working in a kind of queries (query by object) where the goal is to retrieve images which contain some specific objects, eg. retrieve images with a red car. The possibility

(a) (b)

Figure 1: CHAMELEON image retrieval system: (a) Set of random images for the query; (b) Retrieval results of football images.

of user interaction with the retrieval system is another aspect to improve. It is generally accepted that one cannot expect that the results obtained in response to a query, using only one example image, are fully satisfactory. The interaction, called *relevance feedback*, allows to refine the original query on the basis of the retrieved images.

References

[1] A. Del Bimbo, M. Mugnaini, P. Pala, and F. Turco. Visual querying by color perceptive regions. *Pattern Recognition*, 31(9):1241–1253, 1998.

[2] X. Cufí, X. Muñoz, J. Freixenet, and J. Martí. A concurrent region growing algorithm guided by circumscribed contours. *Int. Conf. on Pattern Recognition*, I:432–435, September 2000.

[3] T. Gevers and W.M. Smeulders. Content-based image retrieval by viewpoint-invariant color indexing. *Image and Vision Computing*, 17:475–488, 1999.

[4] S. Aksoy and R.M. Haralick. Textural feature for image database retrieval. In *IEEE Workshop on Content-Based Access of Image and Video Libraries, in CVPR'98*, pages 45–49, June 1998.

[5] J. Hsieh, W.E.L. Grimson, C. Chiang, and Y. Huang. Region-based image retrieval. In *Int. Conf. on Image Processing*, volume I, pages 77–80, 2000.

[6] J.M. Jolion. Feature similarity. *Principles of Visual Information Retrieval*, Springer, 2000.