CONTENT-BASED DESCRIPTION OF IMAGES FOR RETRIEVAL IN LARGE DATABASES : MUVIS

Mejdi Trimeche, Faouzi Alaya Cheich, Moncef Gabbouj, Bogdan Cramariuc
Dept. of Information Technology, Tampere University of Technology,
P.O.Box 553, Tampere, FINLAND
Tel: +358 3 3652111; fax: +358 3 3652112
e-mail: t161699@cs.tut.fi

ABSTRACT

The rapid increase in the size of digital image and video collections is urging for the development of efficient browsing and search tools that skip the subjective task of keyword indexing, paving the way for the ambitious and challenging idea of content based description of imagery. With this goal in mind the **MUVIS**[1] system attempts to provide an integrated solution for the indexing and retrieval of images based on their content within large databases. In this paper we describe briefly the overall structure of the system and expose the promising results obtained so far, demonstrating the similarity retrieval capabilities of the system based on separate image features.

Keywords: Content, Indexing, Query, Retrieval, features, shape, color, texture

1 INTRODUCTION

Nowadays, with the advances in digital imagery and communication technologies, the information conveyed through images is gaining in importance, also the size of image and video collections is growing at an incredible rate. However, we can not access or make use of this information unless it is efficiently indexed in order to ease the task of searching and browsing through these collections. Traditional methods of indexing images in databases rely on a number of descriptive keywords, associated with each image. This manual annotation approach is subjective (relative to the person indexing the images) and recently, due the emergence of large scale image collections, it is becoming outdated.

To overcome these difficulties, content-based Image retrieval was proposed. That is, instead of being manually annotated by text-based keywords, images are indexed by their visual content, such as color, texture, edges, regions, object shape, etc. These image attributes form feature vectors, which are stored with the associated image or on a remote location to be used to organize and search through the image database. This approach is adapted in the MUVIS system, with an emphasis on the performance of the individual features describing shape, color and texture together with the associated similarity measures in retrieving images, as a step towards the successful integration of these features in the system.

The rest of this paper is organized as follows: in section 2 we present the MUVIS block diagram and explain the functionalities of each block, in section 3, we review the features used and tested. in section 4, we describe the experimental results demonstrating the shape similarity retrieval of the system and in section 5 conclusions are drawn.

2 The MUVIS SYSTEM

MUVIS[1] is a modular system intended to be developed on different platforms, the block diagram of the system is shown on figure 1. The image processing toolbox is composed of a collection of image analysis, filtering and editing functions, as well as functions for image segmentation, edge detection, object contour extraction etc., feature extraction is carried out subsequently to build up the feature vectors used in indexing and retrieval. The internal database at its final version will contain a unique indexing structure, most likely based on the pyramid technique. Speedup indices provide for an extremely fast response to the query if the user opts to restrict his/her search out of a given set of images. The graphical user interface allows non-expert users to use the system for browsing and searching for images without extra knowledge in image processing, it allows also more advanced users to tune their search by using a recursive search and to refine the search results. It is possible through the user interface to perform image retrieval based on a single or a combination of image features (shape, color, texture), this allows for a maximum of flexibility for the users to define their queries based on the expected visual relevance of the system output.

3 FEATURES USED

MUVIS is developed with the purpose of indexing images and objects inside images based on their content.

[1]MUVIS stands for Multimedia Video Indexing and Retrieval System

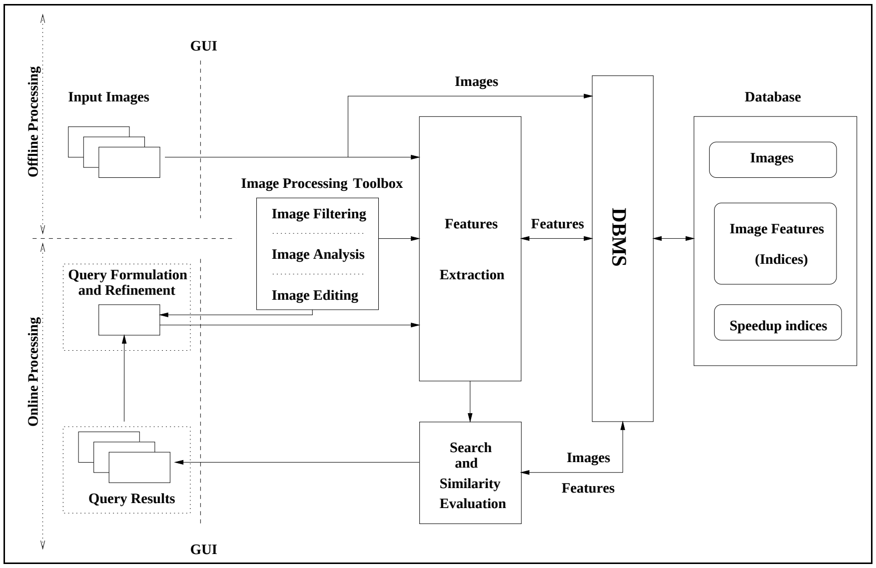

Figure 1: Block diagram of the MUVIS system

Therefore, features at both levels, image level and object level, are extracted from the given images and used for the indexing of images. The features extracted from the whole image are called global features, and they include: color texture or image layout. The global features convey information about the image as a hole while the local features extracted from an object, or a segment of an image convey information just about that object. the local features describe the location of the object in the image, its intersection with its neighbors, its shape .... When querying large collections of images, the global features can be very useful to narrow down the search space, on the other hand, they can not provide enough accurate information to estimate the similarity between images. In the next subsections we review the color and texture features, which can be considered to be global and then we review an object dependent attribute: the shape feature.

3.1 Color

If, for instance, we are interested to search in a database containing natural scene imagery or art works, then, among all features, retrieval based on color is expected to give the best results. In fact, for this class of images, color is probably the most useful feature to be used for measuring similarity within images. In the MuVi-project, an analysis of different Color histogram representations was carried out, we focused on the efficiency of the retrieval and on complexity of the matching algorithms to be used. It was found that some adaptive histogram methods[3] allow to compare histograms with different number of bins, but this increases computational complexity which is not a desirable property when similarity analysis will be performed on-line on a large number of images.

Simulations[4] were carried out to assess the visual similarity through histograms based on RGB, L*a*b*, XYZ and HSV color spaces. From these results it can be seen that HSV and RGB (64 and 125 bins) produced best results. With fewer bins HSV is a better space to analyze data. This is an expected result, because of the nature of the color spaces used, RGB color space is for instance hardware oriented, but still reflects color perception adequately, on the other hand, the HSV color space is user-oriented and represents colors closest to the way we humans analyze the color stimulus.

Simulations where the color feature was summarized in the three dominant colors of the histogram have also shown that most efficient color spaces to use (for representing the dominant colors) are the HSV (166 bins) and the RGB (125 bins).

3.2 Texture

Texture is one of the basic characteristics for the analysis of many types of images. It provides important information for segmentation of scenes in distinct objects or regions as well as for classification or recognition of surface materials. A large range of descriptors have been used for texture analysis, such as: Fourier power spectrum, spatial texture energy, autocorrelation function, structural elements, Markov random field model, fractal dimension, spatial gray-level co-occurrence probability and multichannel representation. In the MUVIS system we have investigated different representations that might show relevant in texture-based retrieval: Gabor filter-banks and different moments for texture features extraction, including the Orthogonal Fourier-Mellon Moments, Zernike Moments and a combination of the latter two through a nonlinear transformation function. Results have shown that Zernike and Orthogonal Fourier-Mellon moments have performed efficiently. For the MUVIS system we will opt for the

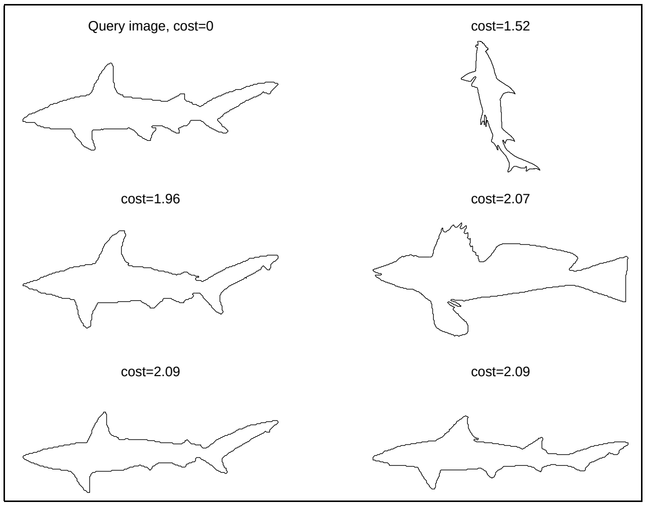

Figure 2: Example of querying based on shape: similarity retrieval results using CSS technique.

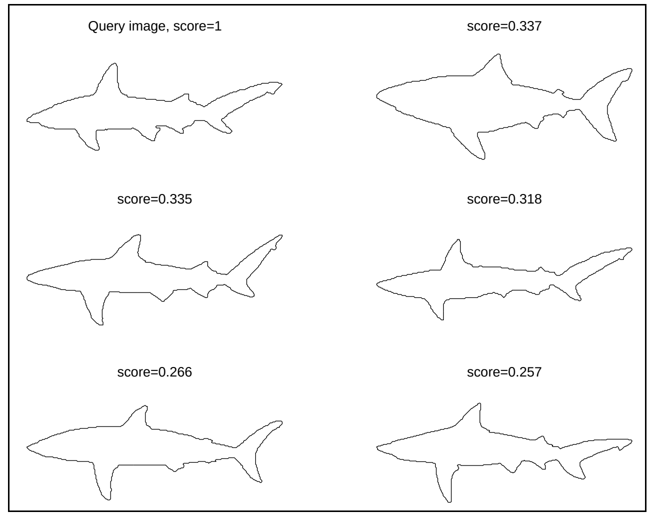

Figure 3: Example of querying based on shape: similarity retrieval results using WTMM corners method.

integration of the computationally least expensive representation among the above to be used as the texture feature vector.

3.3 Shape

The human perception and understanding of objects and visual forms relies heavily on their shape properties, hence the robustness of the shape representation within an integrated content-based system have been studied carefully. To be useful, an efficient shape representation has to achieve the following properties:

- robustness to noise,
- invariance with respect to geometric transformations, such as translations, rotations and scaling,
- similarity relevance: the shape representation should capture the salient perceptual properties of a shape and allow for the design of fast and efficient similarity algorithms,
- robustness and relevance for arbitrary and different shape structures.
- reasonable performance on occluded shapes,
- accessibility: the representation should provide for a significantly short feature vector that allows the over-all system to obey real-time constraints.

With these evaluation criteria in mind, several shape abstraction techniques are being investigated within the MUVIS. Region based representations such as simple geometric attributes (area, symmetrical axis, roundness ...etc.), moments invariants and Zernike moments have been the most popular in content-based systems so far, however the limitations of this approach are clear. Boundary-based representations of shape have to be included if we want perform efficient retrieval of content based on shape, especially if we want to retrieve occluded shapes. In the MUVIS project, three approaches are being investigated: Curvature Scale Space[2], wavelet-based and the stochastic approach. The first two methods are Multi-resolution techniques that decompose the planar curve contour into components of different scales, the CSS method summarizes the contour of a shape into inflection (zeros of curvature) points associated with the scales at which they first appear, in the second method, the WTMM[5] (Wavelet Transform Modulus Maxima) of the contour orientation profile are used to extract high curvature points at multiple scales, these points are then used in a sophisticated similarity matching algorithm to determine the similarity between shapes. The stochastic method is based on the Hidden Markov Modeling of the WTMM points. Using these techniques we have developed several shape representations that compress significantly the data and when associated with their respective matching algorithms, reflect successfully the human similarity perception. In the next section a sample result of the query based on shape is shown.

4 Experimental Results

In this section, The Experimental results of the query results based on shape are presented. We used a database of more than 1100 shapes [2] containing fish contours to test the efficiency of the shape representations and the associated similarity algorithms. In figure 2, the shape similarity retrieval results using CSS technique are shown, figure 3 shows the retrieval results on the same query image using WTMM corner technique. The similarity algorithm associated with the latter representation has proved very efficient on the retrieval of occluded shapes, results are shown on figure 5. On figure 4,

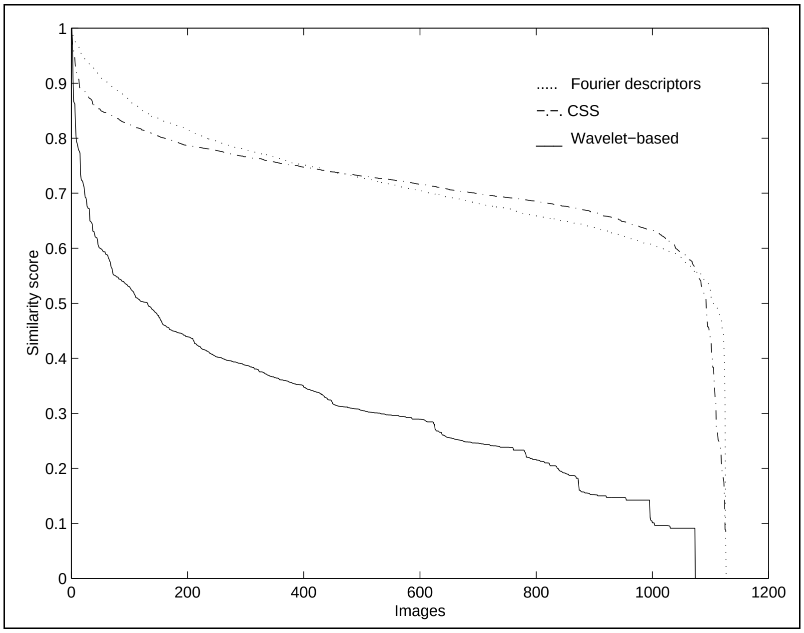

Figure 4: Comparison of the Fourier Descriptors method, CSS technique and the WTMM corners method. Our technique has the strongest discrimination power in a large database.

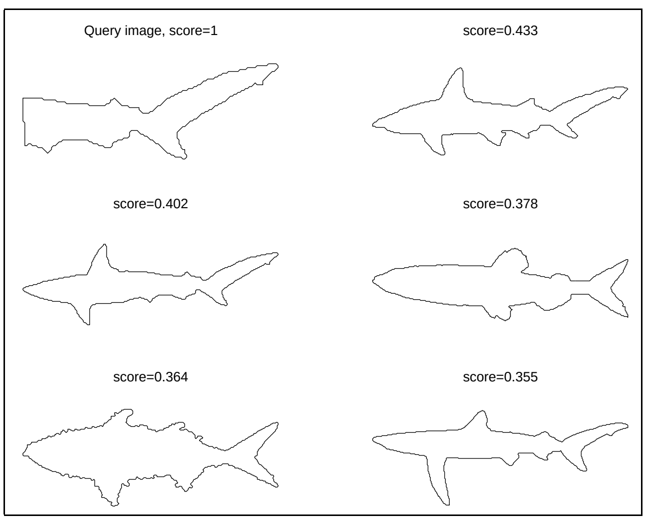

Figure 5: Similarity retrieval results on an occluded shape query, using WTMM corners method.

we have plotted the normalized similarity scores for the two methods discussed above, together with the Fourier Descriptors method, the similarity scores performed by our algorithm had the steepest descent, thus isolating most efficiently the similar shapes from the rest of the database. This is a desirable property for content based retrieval in very large databases.

5 Conclusions

An integrated system, **MUVIS**, for content-based image indexing and retrieval is being at progress. In this paper, the shape, color and texture features have been briefly presented and the retrieval capabilities of the system based on the shape feature have been demonstrated, promising for an integrated system with very sharp retrieval capability based on individual or combined visual features. In its final version, the MUVIS system will allow for indexing of images based on key words, color, texture, shape and their layout in the image space.

References

[1] F. Alaya Cheikh, B. Cramariuc and M. Gabbouj, "MUVIS: A System for Content-Based Indexing and Retrieval in Large Image Databases", *Proc. VLBV'98 workshop*, pp. 41-44, Urbana, USA, October 1998.

[2] F. Mokhtarian, S. Abbasi and J. Kittler, "Efficient and Robust Retrieval by Shape Content through Curvature Scale Space, *Proc. First Int. Workshop on Image Databases and Multimedia search*, pp. 35-42, Amsterdam, Netherlands, August 1996.

[3] V. Ng and et. al., "Adaptive Histogram Indexing", *SPIE*, vol. 2606: 202-211.

[4] P. Kerminen and M. Gabbouj, "Image Retrieval Based on Color Matching", *Proc. FINSIG'99*, Oulun yliopistopaino, pp. 89-93, 1999.

[5] A. Quddus, F. Alaya Cheikh and M. Gabbouj, "Wavelet-Based Multi-level Object Retrieval In Contour Images", *Proc. VLBV'99 Workshop*, Kyoto, Japan, October, 1999.