Content Based Image Retrieval

Media Integration and Communication Center
University of Firenze
Firenze, Italy

The rapid advancements in multimedia technology have increased the relevance that repositories of digital images are assuming in a wide range of information systems. Effective access to such archives requires that conventional searching techniques based on textual keywords be complemented by content-based queries addressing visual features of searched data [1]. To this end, a number of models have been experimented which permit to represent and compare images in terms of quantitative indexes of visual features [2]. In particular, different techniques have been identified and experimented to represent content of single images according to low-level features, such as color, texture, shape and structure, intermediate-level features of saliency and spatial relationships, or high-level traits modeling the semantics of image content.
In so doing, extracted features may either refer to the overall image (e.g., a color histogram), or to any subset of pixels constituting a *spatial entity* with some apparent visual cohesion in the user perception (e.g., an object).

In this demonstrator, an image retrieval system is presented which supports content based queries according to features of color and shape.
The system provides to the user a graphical interface which enables queries by using the "query by example" paradigm. In particular, for color based searches, the user can draw regions, colorize and locate them in a drawing panel, in order to find images with a predefind arrangement of color patches (see the example of Fig.1). Database images are automatically segmented to determine relevant color patches. These regions are then separately represented according to features of color, size and position [3].
On the right of Fig.1, results for the red region query are reported. Images in the result set are ranked in increasing order of the distance scored comparing the query region and the regions extracted from database images.

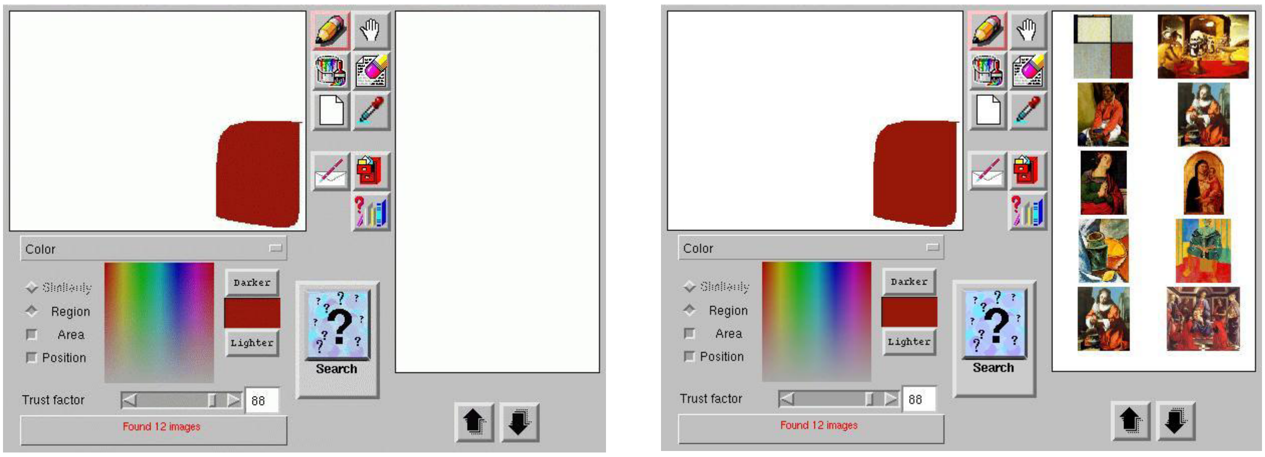

Figure 1. An image retrieval example based on region attributes of color, area and position is shown.

In addition to the region-based image representation, image chromatic content is globally represented by a histogram which enables image retrieval by global color similarity [4]. In this latter case, the user can select an example image and search for images similar to the query on the base of color distribution.

Shape based retrieval is supported by letting the user sketch the contour of a shape he/she is looking for. The shape is represented by breaking its boundary into *tokens* according to points of minima of the curvature function computed on the shape boundary. Shape tokens are then individually represented using their curvature and orientation (see Fig.2).
Shape comparison is supported by combining distances computed on individual tokens of two matching shapes.
In addition to this, shape tokens are combined with an indexing structure (*metric tree*) capable to speed up the search by reducing the number of distance computation between the query shape to the database shapes from a linear scan of the database shapes to a sub-linear scan.

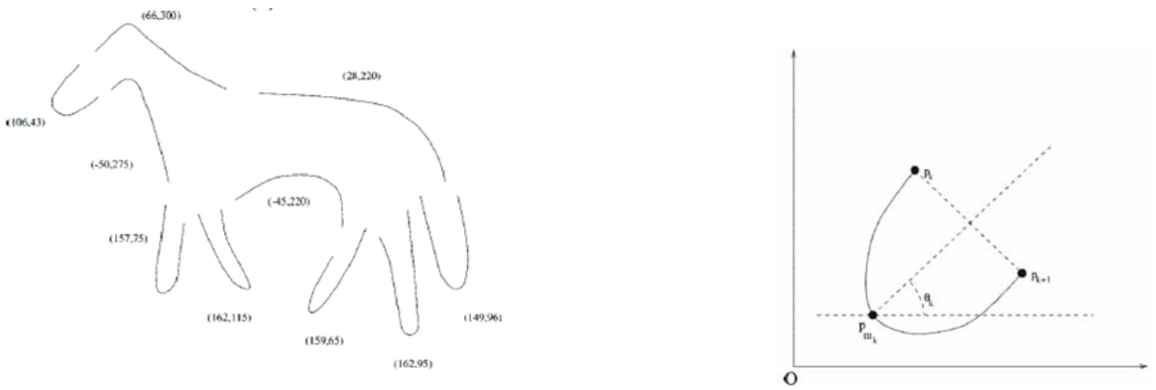

Figure 2. On the left, a shape is broken into tokens according to the minima of the curvature function. On the right, each token is represented using its curvature and orientation.

Fig.3 shows the retrieval results for a query shape. In this approach, the user can draw the query shape in the white panel on the left, while results are shown on the right in decreasing order of similarity. The shapes identified in each result image are delimited by their minimum embedding rectangles.

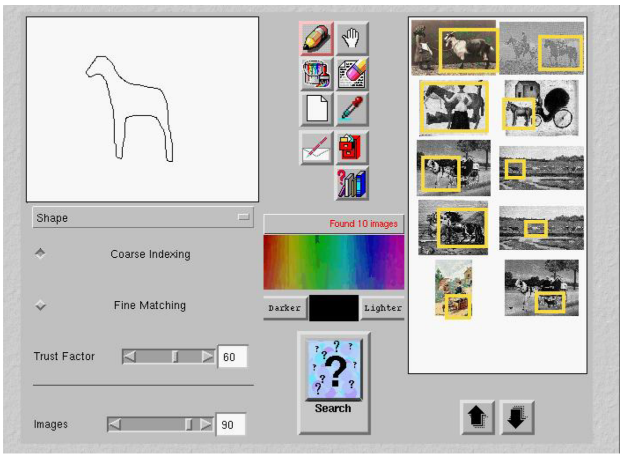

Figure 3. A shape retrieval example.

This system has been developed at the Dipartimento di Sistemi e Informatica of the University of Firenze, Italy, before the start of the DELOS project (reference person: *Prof. Alberto Del Bimbo*). However, this research line is still ongoing under the activities of DELOS Task 4.5a, with specific reference to the issue of relevance feedback.

References

[1] A. Del Bimbo, "Visual Information Retrieval," *Academic Press*, San Francisco 1999.

[2] A.W.M. Smeulders, M.Worring, S.Santini, A.Gupta and R.Jain, "Content Based Image Retrieval at the End of the Early Years," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.22, n.12, pp.1349-1380, December 2000.

[3] S. Berretti, A. Del Bimbo and E. Vicario, "Efficient Matching and Indexing of Graph Models in Content Based Retrieval," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.23, n.10, pp.1089-1105, October 2001.

[4] A. Del Bimbo, M. Mugnaini, P. Pala and F. Turco, "Visual Querying by Color Perceptive Regions", *Pattern Recognition*, vol.31, n.9, pp.1241-1253, September 1998.

[5] S.Berretti, A.Del Bimbo and P.Pala. "Retrieval by Shape Similarity with Perceptual Distance and Effective Indexing," *IEEE Transactions on Multimedia*, vol.2, no.4, pp.225-239, December 2000.

Content Based Retrieval of 3D Objects

Media Integration and Communication Center
University of Firenze
Firenze, Italy

In recent years, the use of 3D models has been progressively spreading throughout many application domains: in manufacturing industries, 3D models are used in the design of many different objects and components, and are archived to enable reuse and rapid prototyping of new products; in medicine, many tests that are fundamental for disease study and diagnosis provide output data in the form of 3D models; in videogames design, reuse of available models may speed up the development of new games; in video production, where high quality 3D models are used for special effects, considerable resources, in terms of time and effort, can be saved by reusing or adapting existing models; in cultural heritage, regular acquisition of the 3D structure of statues, bas-reliefs and art-pieces supports monitoring potential deformations of object structure due to unappropriate preservation conditions. In addition to this, it should be noticed that the availability of 3D models is increasing thanks to a growing number of techniques enabling their generation. These solutions differ in terms of costs, resolution and type of acquired information (the un-textured external surface of the object, the textured external surface or even the interior composition of the object) and include CAD, tomography, magnetic resonance, 3D laser scanners, structured light systems and photogrammetry.
Holding these assumptions, retrieval by content of 3D models will emerge as a key issue in the next future and some solutions have already been proposed to address this problem [3]. Many solutions proposed so far, rely on the extraction of content descriptors capturing global properties of 3D object surface: moments, distributions of vertex distances, surface curvature or angles between faces. Alternatively, surface properties can be described in transformed domains, like the wavelets and the spherical harmonics.

In this work, we define a new solution combining the advantages of view-based and structure-based approaches to description and matching of 3D objects. The new solution relies on Spin Image signatures and clustering to achieve an effective, yet efficient representation of 3D object content.
For each mesh vertex V, a *spin image* is built mapping any other mesh vertex x onto a two-dimensional space (see the left of Fig.1). The grey-level spin images are derived by considering the density of mesh vertices that map on the same point of the spin image, and evaluating the influence of each vertex over the neighboring pixels of its projection, according to a bilinear interpolation scheme. On the right of Fig.1 a sample model of a teapot is shown with spin images computed from four different vertices of the model.

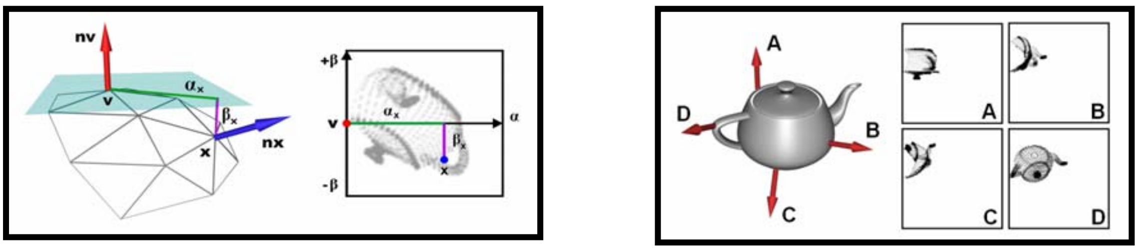

Figure 1. A spin image computed on the vertex of a 3D model is shown on the left. On the right, spin images for the same object computed along the four normals A, B, C, D to the object surface, are shown.

Spin images are then partitioned into: sectors of circular crowns for the upper ($\beta>0$) and lower ($\beta>0$) half-planes and circular sectors centered in the origin (see Fig.2). For each of them, we have considered the number of vertex projections that fall in the region: ($Cp=<cp1, \ldots , cpn>$, $Cn=<cn1, \ldots , cnn>$, and $S=<s1, \ldots , sns>$, respectively). Experimentally, $np=nn=ns=6$ has been found to be a satisfactory trade-off between representation compactness and selectivity. This leads to compress the spin image informative content into a 18-dimensional vector $D= <Cp, Cn, S>$.
Description vectors have been clustered using fuzzy clustering so as to take the centers of the clusters as signatures of the spin image representation. The optimal number of clusters is derived considering two functions that express a measure of under- and over-partitioning, respectively. The optimal number of clusters is the number that minimizes the sum of the two functions representing the trade-off between under- and over-partitioning. Finally, similarity between spin image signatures of 3D objects is obtained considering the permutation that minimizes the sum of distances between the corresponding cluster centers.

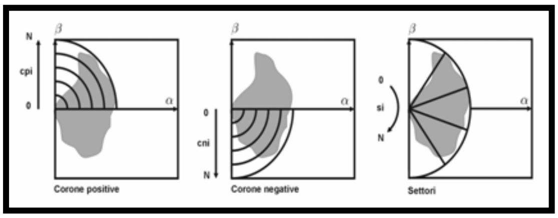

Figure 2. Circular crowns and sectors used to derive a vector representation of the spin image.

In order to validate the retrieval performance of the proposed approach, a compound 3D repository of VRML models has been constructed by gathering 3D models featuring different characteristics mainly in terms of resolution and complexity. In particular, the 3D repository includes 3D models collected from the Web, from the 3D Princeton Shape Benchmark Archive [2] and from the De Espona Encyclopedia.
All models of the compound 3D repository have been processed in order to extract the Spin Image signature and signatures of four other approaches for 3D objects representation and retrieval that we used for comparison purposes. A variety of dissimilarity measures to compare content descriptors have been also considered. In addition, models of the compound 3D repository have been manually annotated so as to represent semantic information about their content. Annotations are used to automatically extract precision and recall curves. These have been used to provide a representative comparison of the performance of prototype retrieval engines.
Retrieval performance based on Spin Image signatures has been compared against the performance of four other prototype retrieval engines, using Light Field (LI), Curvature Histogram (CH), Shape Functions (SF) and Curvature Moments (CM). Comparison has been carried out in terms of figures of precision and recall. Although comparison is still in progress, preliminary results suggest that retrieval based on Spin Image signatures outperforms the other prototype retrieval engines for medium- and high-resolution 3D models. In particular, on the right of Fig.3 the average precision/recall curves are reported for the five methods under investigation. This has been obtained by using objects from every model category of the princeton shape benchmark database as query model.

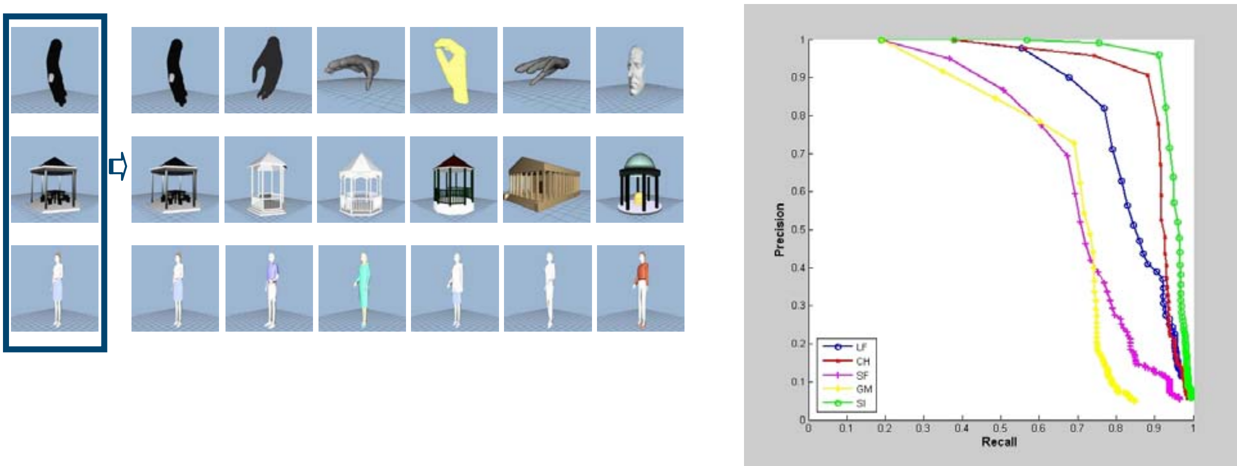

Figure 3. On the left, retrieval results are reported for three query models. Average precision/recall curves are reported on the right, comparing Spin Images (SI) against four other approaches for 3D object retrieval, namely, Light Field (LI), Curvature Histogram (CH), Shape Functions (SF) and Curvature Moments (CM).

This prototype has been developed at the Media Integration and Communication Center of the University of Firenze, Italy, in the activities carried out during year 2005 for the DELOS Task 3.8 (Reference person: *Prof. Alberto Del Bimbo*).

References

[1] A. Del Bimbo, P.Pala. "Content based Retrieval of 3D Models," accepted for publication on *ACM Transactions on Multimedia Computing, Communications, and Applications, 2006*.

[2] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser. "The Princeton Shape Benchmark," *Shape Modeling International*, Genova, Italy, June 2004.

[3] J.W.H. Tangelder, R.C. Veltkamp. "A Survey of Content Based 3D Shape Retrieval Methods," In *Proceedings Shape Modeling International*, pp.145-156, 2004.