Proceedings of IEEE International Conference on Multimedia and Exposition ICME 2000, U.S.A.

Image Retrieval By Shape: A Comparative Study *

Maytham Safar Cyrus Shahabi and Xiaoming Sun
Integrated Media Systems Center - Department of Computer Science
University of Southern California
Los Angeles, California 90089–0781
$[safar, shahabi, xiaomins]@usc.edu$

Abstract

Besides traditional applications (e.g., CAD/CAM and Trademark registry), new multimedia applications such as structured video, animation, and MPEG-7 standard require the storage and management of well-defined objects. We focus on shape-based object retrieval and conduct a comparison study on four of such techniques: FD, GB, DT, and $MBC-TPVAS$. Our results show that the similarity retrieval accuracy of our method ($TVPAS$) is as good as other methods, while it has the lowest computation cost to generate the shape signatures of the objects. Moreover, it has low storage requirement, and a comparable computation cost to compute the similarity between two shape signatures. In addition, $TPVAS$ requires no normalization of the objects, and is the only method that has direct support for RST query types. In this paper, we also introduce a new shape description taxonomy.

1 Introduction

Shape of an object is an important feature for image and multimedia similarity retrievals. Therefore, this study focuses on shape-based object retrieval techniques. There is a variety of techniques that has been proposed in the literature for shape representation [4, 3]. In [3], shape representation techniques are divided into two categories: boundary-based and region-based. Boundary based methods use only the contour or border of the object shape and completely ignore its interior. On the other hand, the region based techniques take into account internal details (e.g., holes) besides the boundary details. The purpose of this paper is to evaluate and compare the performance of four boundary based methods for shape representation and retrieval: Fourier descriptors methods FD [6] (based on objects' shape radii), grid-based method GB [5] (based on chain codes), Delaunay triangulation method DT [4] (based on corner points), and Touch-point-vertex-angle-sequence $TPVAS$ (based on minimum bounding circles MBC and angle sequences AS). For further details on the different shape representation techniques, please refer to Tech. [2]. $TPVAS$ approach to shape representation and similarity measure was proposed by the authors in [1]. We compare its retrieval performance with that of the other 2 more established methods (i.e., FD, and GB methods) that were used in some commercial systems [3] and as a basis for different comparison studies in [5, 3]. In addition, it was compared to a new method based on a new indexing technique (i.e., DT) [4]. Although $TPVAS$ method uses simple attributes such as minimum bounding circles and angle sequences (simple to extract), these attributes were shown to be translation, rotation, and scale invariant in [1]. Furthermore, $TPVAS$ has low storage requirement and has the lowest computation cost to generate the shape signatures of the objects. Moreover, it has a comparable computation cost to compute the similarity between two shape signatures. In addition, the similarity retrieval accuracy of $TVPAS$ is comparable to the other methods. In some application domains, the support of RST query types in which the matching objects could be within a specified rotation angle, scaling factor, translation vector, or any combination of the three is important (see [1] for further details). An example is, searching for similar tumor shapes in a medical image database [9]. Tumors are represented by a set of 2D images each of which represents a slice of the 3-D tumor. A method for retrieving similar tumor shapes would help to discover correlations between tumor shape and certain diseases. Besides the shape, the size and the location of the tumor would help in the identification of patients with similar health history

*This research has been funded in part by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152, by NASA/JPL contract nr. 961518, and unrestricted cash/equipment gifts from Intel and NCR.

(*i.e.*, RST query types). With $TVPAS$ method, a subset of the MBC features (*e.g., radius r, center C, and start point SP*) could be utilized to support RST query types. On the other hand, the other three methods lack the support of RST queries. This is due to either the use of normalization, or lack of features that could support such queries. In addition, this study introduces a new shape description taxonomy capturing techniques that were not classified in previous taxonomies [3, 4] (e.g., turning angle, collinearity, ...etc.).

2 Shape Description Techniques

The main objective of shape description in object recognition is to measure geometric attributes of an object, that can be used for classifying, matching, and recognizing objects. Shape description techniques tend to perform better in some application domains than others. For example, region-based techniques (take into account internal details like holes) are more suitable than the boundary-based techniques when internal details of the objects are as important as their contour. Another example is that transformation domain techniques are more suitable than direct representation (spatial) techniques when the size the of shape signatures is important. Transformation domain techniques tend to reduce the size of the shape signatures and fix their dimension to be suitable to index using different SAMs. Finally, some techniques support partial similarity of objects, while others only support complete similarity. Therefore, it is very useful to have a taxonomy of the different shape descriptions to help the users in choosing the proper techniques depending on their application domains and their restrictions.

The study in [3] categorizes the techniques into boundary based and region based methods. The region-based methods are further broken into spatial and transform domain sub-categories depending on whether direct measurements of the shape are used or a transformation is applied. A drawback of this categorization is that it does not further sub-categorizes boundary based methods into spatial domain and transform domain methods. For example, Fourier descriptors method can be considered as a transform domain technique, while chain codes can be considered as spatial domain technique. Another drawback of this categorization is that it assumes that structural techniques (e.g., 2D-strings) are a sub-category of region-based, spatial domain techniques. However, we believe that structural techniques as 2D-Strings are spatial similarity based techniques, in which the retrieval of objects is performed based on the spatial relationships among objects and not shape retrieval techniques (see Fig.1). In another study [4], shape description techniques were broken into two different categorizes: transformation based and measurement based categories. They further breakdown the transformation based category into two sub-categories: functional and structural categories, however, it is not clear what criteria is used to this end. Some drawbacks of this categorization are that it does not distinguish between boundary or region based techniques, and sometimes it miss categorizes some techniques. For example, chain code technique is categorized as a transformation-based technique, while it is a measurement-based technique. Another example is that *silhouette moments* is considered as a region-based technique but not as a boundary-based technique. Therefore, we introduce a new shape description taxonomy in Fig.1. We also included additional techniques that were not considered in [3, 4] (e.g., turning angle, collinearity, ...etc.).

3 Discussion

Normalization For the GB and FD methods, the shape representations obtained for the same object with different orientation in space or with a different scale are different. Therefore, the normalization of the object boundaries prior to indexing is crucial to meet the uniqueness criteria of the shape descriptor. The DT and $TPVAS$ methods use simple shape attributes such as the angles of the resulting Delaunay triangles of a set of points, MBC and AS. However, these attributes were shown to be translation, rotation, and scale invariant. Therefore, no normalization is required.
Support for RST queries With RST query type we are interested in finding all objects in a set of objects that are exactly identical to a query object regardless of its size and orientation, or with a specified size and/or orientation. With $TPVAS$, a subset of the MBC features (i.e., r, C, and SP) could be utilized to support RST queries (see [1] for further details). On the other hand, the GB, FD, and DT methods are unable to support such query type. With the GB and FD methods, normalization is used so that the objects fit into a prespecified mesh, or force the boundaries to have a standard size, and orientation, respectively. The DT method representation also provides no information about the size, and orientation of the original objects.

4 Experiments

To measure the effectiveness of the similarity retrieval of the four different shape representation methods (*i.e.*, FD, GB, DT, and $TPVAS$), an experiment was designed to measure the performance in terms of recall and precision. We implemented a prototype shape retrieval system in Java, on a SUN UltraSparcII workstation. The shape database of the system consists of 296 fish and tool shapes provided to us by [7, 8]. In order to reduce the number of vertices used to represent the shapes while maintaining its general characteristics, we employed a three pass algorithm. In each pass, we try to identify the straight lines on the boundary

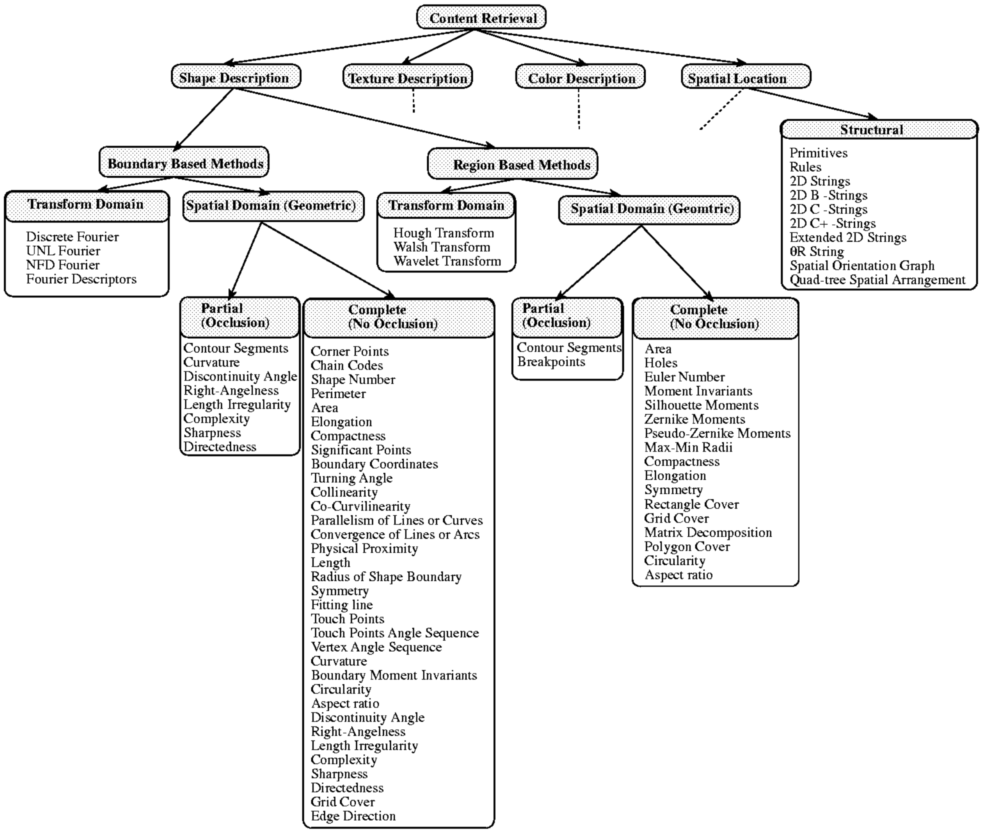


Figure 1. Taxonomy of Shape Description Techniques

of the shape and eliminate the extra points that the edge detection algorithm usually introduce to represent a straight line (jagged line like saw edge).

Fifteen shape queries were selected randomly from the database, where each query object has 10 similar (relevant) objects in the database. The ten similar objects' variants were constructed as: three rotation variants, three rotation scaled up variants, three rotation scaled down variants, and one translated variant. For our experiments, we simply assume that each shape is relevant only to itself and to its ten variants. As advised in [5, 3, 4, 1], for GB we employed a grid cell size of 12X12 pixels and the length of the standardized major axis was fixed at 192 pixels. With FD, we used radius-based signature with 64 uniformly sampled boundary points. For both FD and $TPVAS$, we kept 8 low frequency coefficients of the Fourier transform. Finally, for DT, we used two largest angle histogram with 18 bins of size 10 degrees. For further details on the experimental setup, please refer to Tech. [2].

The recall and precision results for 15 query shapes are averaged and shown in Fig.2. From Fig.2(a), it is observed that FD was outperformed by all the other three methods. This poor performance is explained by the fact that the shape signature is constructed using shape radii of points uniformly sampled along the object boundary. These points are not the exact vertices of the object, therefore, using different sampled points leads to different shape signatures. In other words, the rotation normalization is not truly achieved by taking only the magnitude information from the Fourier coefficients ($F(u)$s) and ignoring its phase information as explained in [5]. Fig.2(b), shows the recall-precision curves of the methods excluding FD. From Fig.2(b), we make the following observations. First, that the accuracy of our proposed method ($TVPAS$) is as good as other methods. Second, unlike what was expected, the results obtained for DT were better than what was reported in [4]. This is due to the fact that in their approach, they reduce the number of vertices used to represent an object by identifying high- -curvature points called *corner points*. On the other hand, in our experiments, we used the three pass algorithm. In their experiments, the number of vertices used to represent an object was approximately 20-26 points, however, in our approach it was from 50-90 points. Although our approach used more vertices to represent the objects, however, the

Table 1. Costs

a) Storage cost (n objects)

GB	DT	FD	TPVAS
$2*n*SS+8*n$ and $temp\ px/8$	$4*n*nb$ and $temp\ 3*N$	$8*n*fn$	$8*n*fn$

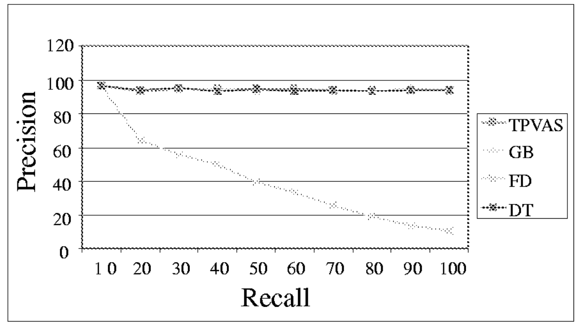
b) Shape signature generation cost

GB	DT	FD	TPVAS
MJ: $O(N^2)$, MN: $O(N)$, NR: $O(N)$, IN: $O(N)$, and SG: $O(N^3)$	DT: $O(NlogN)$, and SG: $O(N)$	DT: $O(NlogN)$, FC: $O(r^2)$, and SG: $O(r)$	MBC: $O(N)$, and SG (DFT): $O(N)$

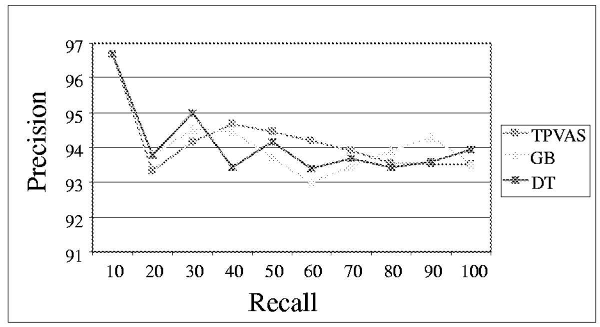
c) Similarity computation cost

GB	DT	FD	TPVAS
$8*SS\ bit\ xor$, and $8*SS\ bit\ add$	$nb\ int\ sub$, $nb\ int\ mult$, and $nb-1\ int\ add$	$fn\ float\ sub$, $fn\ float\ mult$, and $fn-1\ float\ add$	$fn\ float\ sub$, $fn\ float\ mult$, and $fn-1\ float\ add$

MJ: major axis, MN: minor axis, CN: centroid, NR: normalization, IN: inside function, FC: FD coefficients, SG: shape signature



(a): Four Methods



(b): Three methods

Figure 2. Recall-Precision Curves

representation is more discriminative in shape signatures. We can conclude that DT performance highly depends on the technique used to find the representative vertices of the objects.

From Table 1(a) we can conclude that both $TPVAS$ and FD have the lowest storage requirement, while DT requires the largest storage requirement. GB requires larger storage than DT if the objects have multiple major axes. Both GB and DT, require some additional temporary storage to perform their algorithms. Table 1(b) shows the computation cost to generate the shape signatures of the objects. We can conclude that $TPVAS$ has the lowest computation cost and GB the highest. Both DT and FD have comparable computation costs due to the use of Delaunay triangulation algorithm. However, Table 1(c) shows that GB requires the lowest computation cost to compute the similarity between two shape signatures, while the other three methods have comparable computation costs. A complete cost computation analysis for the different methods that compares their efficiency in terms of computation and storage requirements as well as more performance experiments will be included in the full paper.

5 Conclusion and Future Work

This study compared four shape-based object retrieval techniques (FD, GB, DT, and $TPVAS$). The similarity retrieval accuracy of our method ($TVPAS$) was comparable to the other methods, while it had the lowest computation cost to generate the shape signatures of the objects. Moreover, it has the lowest storage requirement, and a comparable computation cost to compute the similarity between two shape signatures. In addition, $TPVAS$ requires no normalization of the objects, and is the only method that has direct support for RST query types. We also introduced a new shape description taxonomy.

We intend to extend this work in several directions. First, we plan to investigate the issues of robustness and stability of the alternative shape representation techniques, and study the behavior of such techniques under "uncertainty" issues and situations (e.g., the presence of noise, parameters used to identify a shape and its vertices). Second, we plan to investigate how to handle objects with curved parts, holes, and open curves or lines. Finally, intend to incorporate human perceptions in the evaluation of the effectiveness of the shape retrieval methods.

References

[1] C. Shahabi, M. Safar Efficient Retrieval and Spatial Querying of 2D Objects *Proc. of IEEE ICMCS*, June 7-11, 1999, Florence, Italy. pp.611-617.

[2] M. Safar, C. Shahabi and X. Sun Image Retrieval By Shape: A Comparative Study *University of Southern California Technical Report 99-719*

[3] B.M. Mehtre, M.S. Kankanhalli, W.F. Lee Shape Measures for Content Based Image Retrieval: A Comparison *Information Processing and Management*, Vol.33, No.3, pp. 319-337, 1997

[4] Y. Tao, and W.I. Grosky Delaunay triangulation for image object indexing: a novel method for shape representation *Proc. of the Seventh SPIE Symposium on Storage and Retrieval for Image and Video Databases*, San Jose, California, January 1999, pp.631-642.

[5] A. Sajjanhar, and G. Lu A Grid Based Shape Indexing and Retrieval Method *Special issue of Australian Computer Journal on Multimedia Storage and Archiving Systems*, Vol. 29, No.4, November 1997, pp. 131-140.

[6] A. Sajjanhar, G. Lu and J. Wright An experimental study of moment invariants and Fourier descriptors for shape based image retrieval *Proc, of the Second Australia Document Computing Symposium*, April 5 1997, Melbourne, Australia, pp. 46-54.

[7] J. Gary, and R. Mehrotra Similar-Shape Retrieval In Shape Data Management *IEEE, Computer* September 1995.

[8] F. Mokhtarian, S. Abbasi, and J. Kitter Efficient and Robust Retrieval by Shape Content through Curvature Scale Space *Proc. of International Workshop on Image Database and Multimedia Search*, pp. 35-42, Amesterdam, Netherlands, 1996.

[9] F. Korn, N. Sidiropoulos, C. Faloutsos, E. Siegel, and Z. Protopapas Fast Nearest Neighbor Search in Medical Image Databases *Proc. of 22nd VLDB Conference*, pp.215-226, Mumbai, India, 1996.