Multimed Tools Appl
DOI 10.1007/s11042-006-0070-y

An experimental study of alternative shape-based image retrieval techniques

Cyrus Shahabi · Maytham Safar

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. These object databases are then queried and searched for different purposes. A sample query might be "find all the scenes that contain a certain object." Shape of an object is an important feature for image and multimedia similarity retrievals. Therefore, in this study we focus on shape-based object retrieval and conduct a comparison study on four of such techniques (i.e., Fourier descriptors, grid based, Delaunay triangulation, and our proposed *MBC-based* methods (e.g., $MBC\text{-}TPVAS$)). We measure the effectiveness of the similarity retrieval of the four different shape representation methods in terms of recall and precision. Our results show that the similarity retrieval accuracy of our method (MBC-TPVAS) is as good as that of the other methods, while it observes 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, MBC-TPVAS requires no normalization of the objects, and is the only method that

This research has been funded in part by NSF grants EEC-9529152 (IMSC ERC), IIS-0082826 (ITR), IIS-0238560 (PECASE), IIS-0324955 (ITR) and IIS-0307908, and unrestricted cash gifts from Okawa Foundation and Microsoft. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

C. Shahabi (✉)
Integrated Media Systems Center, Department of Computer Science, University of Southern California, Los Angeles, CA 90089–0781, USA
e-mail: shahabi@usc.edu

M. Safar
Computer Engineering Department, Kuwait University, Kuwait, Kuwait
e-mail: maytham@eng.kuniv.edu.kw

has direct support for $S\text{-}RST$ query types. In this paper, we also propose a new shape description taxonomy.

Keywords Shape representation · Shape similarity · Similarity measure · Image retrieval

1 Introduction

Several applications in the areas of CAD/CAM and computer graphics require to store and access large databases. A major data type stored and managed by these applications is representation of two dimensional (2D) objects. Objects contain many features (e.g., color, texture, shape, etc.) that have meaningful semantics. From those, shape is an important feature that conforms with the way human beings interpret and interact with the real world objects. The shape representation of objects can therefore be used for their indexing, retrieval, and as a similarity measure.

The object databases can be queried and searched for different purposes. For example, a CAD application [2] for manufacturing might intend to reduce the cost of building new industrial parts by searching for reusable existing parts in a database. For an alternative trade mark registry application [3], one might need to ensure that a new registered trademark is sufficiently distinctive from the existing marks by searching the database. Meanwhile, new multimedia applications such as structured video [7], animation, and MPEG-7 standard [23] define specific objects that consist in different scenes of a continuous presentation. These scenes and their objects can be stored in a database for future queries. A sample query might be "find all the scenes that contain a certain object." Therefore, one of the important functionalities required by all these applications is the capability to find objects in a database that match a given object.

This study is different from the work in computational geometry, pattern recognition and computer vision. In computational geometry, the main focus of the research is in developing algorithms to compare two given objects to see if they match or not. Although, we eventually employ variations of these comparison algorithms to eliminate false hits, the major focus is on filtering techniques to minimize false hits. In pattern recognition and computer vision the goal is to recognize objects in a given scene (e.g., by using edge detection algorithms). In our work, the set of objects in the database is already well defined by their vector representations either manually as a result of a drawing software or automatically as a result of an object recognition algorithm.

There is a variety of techniques that has been proposed in the literature for shape representation [13, 24]. In [13], 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 method FD [21] (based on objects' shape radii), grid-based method GB [19] (based on chain codes), Delaunay triangulation method DT [24, 25] (based on corner points), and MBC-based methods (based on minimum bounding circles and angle sequences). Three MBC-based approaches to shape representation and similarity measure were proposed by the authors in [17, 22]. We compare the retrieval performance of one of them (MBC-TPVAS) with that of the other two more established methods (i.e., FD, and GB) that were used in some commercial systems [13] and as bases for different comparison studies in [13, 19]. In addition, MBC-TPVAS is compared to a new method based on a new indexing technique (i.e., DT) [24]. Moreover, we tested MBC-TPVAS method for robustness to noise by performing queries on a database of noisy shapes.

Although MBC-TPVAS 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 [22]. Furthermore, MBC-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 MBC-TPVAS is comparable to the other methods.

In some application domains, the support of S-RST query types in which the matching objects could be within a specified rotation angle (R), scaling factor (S), translation vector (T), or any combination of the three is essential (see [22] for further details). An example is, searching for similar tumor shapes in a medical image database [11]. A tumor is represented by a set of 2D images, each corresponding to a slice cut through its 3D representation. 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. With TPVAS, a subset of the MBC features (i.e., radius r, center C, and start point SP) could be utilized to support S-RST query types. On the other hand, the other three methods cannot support S-RST queries. This is due to either the use of normalization, or lack of extracted features that could support such queries. In addition, this study introduces a new shape description taxonomy where we classify techniques that were not considered in previous studies [13, 24, 25] (e.g., turning angle, collinearity, ... etc.).

The remainder of this paper is organized as follows. Section 2, provides some background on related work. We present our proposed shape description taxonomy in Section 3. In Section 4, we describe different types of shape retrieval methods considered in this study. Section 5 briefly describes index structures previously proposed by the authors to support shape retrieval using the MBC-based methods. Section 6 discusses some of the drawbacks of the alternative shape description techniques. Section 7 reports on the performance results obtained from a set of experiments. Finally, Section 8 concludes the paper and provides an overview on our future plans.

2 Related work

As mentioned in Section 1, this work is different from the studies in the area of pattern and object recognition as well as the area of computational geometry. Therefore, in this section, we focus on the related works in the database area in order to distinguish this study from the others. Towards this end, we categorize those works into two groups. The first group investigates techniques to speed up the similarity between 2D objects based on shape, while the second group focuses on 1D sequences.

The first group of studies [5, 6, 8, 10, 13, 15], investigates techniques to speed up the search of 2D "similar" objects to a target object based on their shapes. There is a variety of techniques that has been proposed in the multimedia information systems area for shape representation [8, 13, 15]. Examples of those techniques are Fourier descriptors method (FD) [8, 15, 21] (based on objects' shape radii), grid-based method (GB) [12, 19] (based on chain codes). They were used in some commercial systems [13] and as bases for different comparison studies in [12, 13, 19, 20]. In addition, a new method based on a new indexing technique (i.e., Delaunay triangulation-DT) was proposed in [24]. With Delaunay triangulation, the shape features used are invariant under uniform translation, scaling, and rotation. However, with Delaunay triangulation, histogram-based representation is not very discriminative and is not unique, i.e., objects of different shapes may have the same feature point histogram representation. In addition, choosing a histogram with a small number of bins provides less discriminating ability. Both FD and GB are variant to rotation, and scaling. Therefore, shape signatures obtained for the same object with different orientations or with different scales are different. Hence, given the shape signature of an object, a normalization procedure is required to normalize the object such that its boundary has a standard size, and orientation. In addition, GB, FD, and DT are unable to support S-RST query type. With GB and FD, normalization is used so that the objects fit into a prespecified mesh, or force the boundaries to have a standard size and orientation, respectively. While the DT representation provides no information about the size, and orientation of the original objects.

The second group [1, 4] of approaches in the database literature focuses on speeding up the similarity search for one dimensional sequences. These techniques try to reduce the dimensionality of the problem. First, they map the vectors from the original space into a space of the same dimensionality using a transformation that preserves all information available about the original vectors. Second, they select the most significant components from the transformed vectors and build up an index on the reduced size vectors. These studies can be considered as a complement to our work. This is because we use object features in order to reduce the dimensionality into single dimension sequences and then we can utilize their proposed techniques to speed up the search further.

3 Shape description taxonomy

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 (see [13] for an overview). There are various methods for shape representation. For example, the study in [13], provides an overview of shape description techniques. It categorizes the techniques into boundary based and region based methods. Boundary based methods use only the contour of the objects' shape, while the region based methods use the internal details (e.g., holes) in addition to the contour. The region-based methods are further broken into spatial and transform domain subcategories 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 technique in transform domain, while chain codes can be considered as a 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, under spatial domain techniques. However, we believe that structural techniques such 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 [25], 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

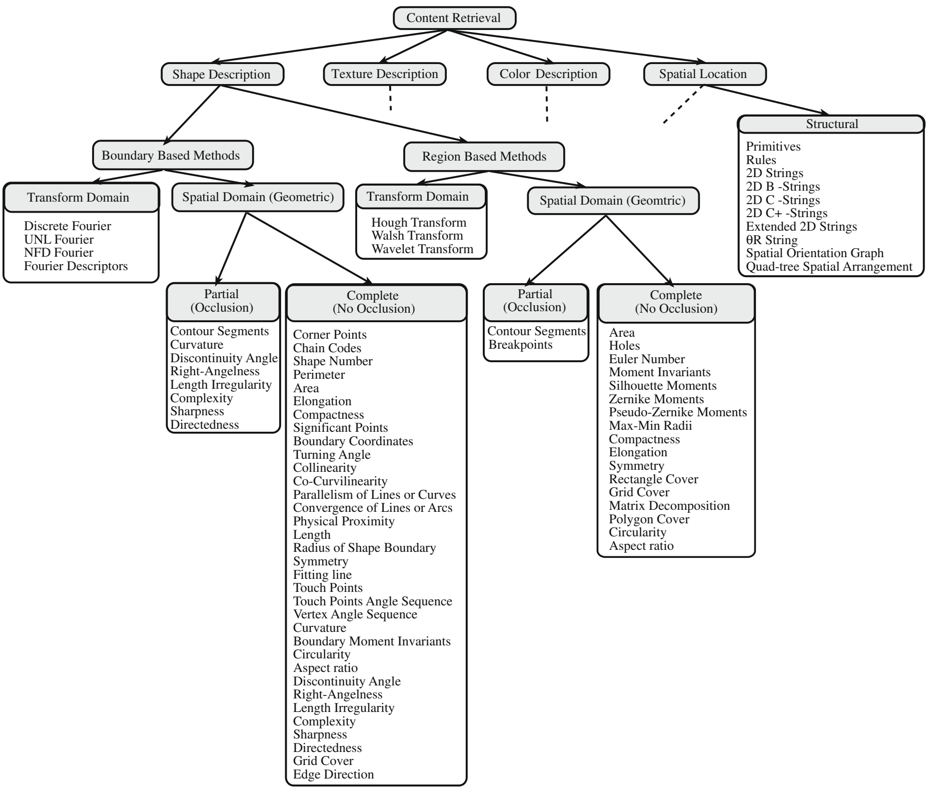


Fig. 1 A taxonomy of shape description techniques

Fig. 1. We also included additional techniques that were not considered before in [13, 25] (e.g., turning angle, collinearity, . . . etc.).

4 Alternative boundary-based shape description techniques

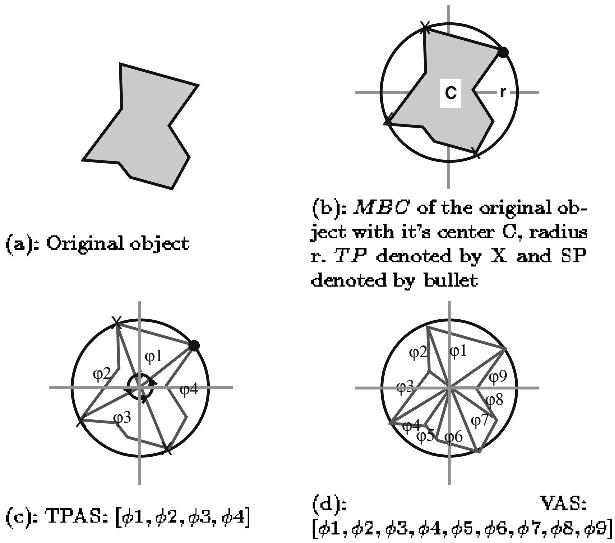
This study compares four boundary based methods for object shape descriptions: Fourier descriptors method (FD) [21], grid-based method (GB) [19], Delaunay triangulation method (DT) [24], and MBC-based methods (MBC-TPVAS, MBC-TPAS, and MBC-VAS) [22].

4.1 MBC-based methods

This approach for shape representation was proposed by the authors in another study [22]. We proposed three MBC-based shape representation methods (MBC-TPAS, MBC-VAS, and MBC-TPVAS) based on features that are extracted from the object's minimum bounding circle (MBC). Those features are the center coordinates of MBC, the radius of MBC, the set of touch points on MBC (TPs), touch points angle sequence (TPAS), vertex angle sequence (VAS), and the start point of the angle sequence (SP). Figure 2 depicts an example of an object with its six identified features. In Fig. 2b the MBC of an object is computed and its touch points are identified as the vertices that lay on its MBC (denoted by X). Figures 2c and d show the touch points and vertex angle sequences, respectively. For the algorithms to extract the MBC features as well as their formal definitions see [22].

With the MBC-based methods, the shape descriptor of an object depends on a subset of the six MBC features. Each method starts by obtaining a feature function of the object called shape signature $f(k)$. To obtain $f(k)$, four steps are required. The minimum bounding circles of the objects are first computed. Second, MBC features such as angle sequences (TPAS, and VAS) and number of touch

Fig. 2 MBC features



points (TP) are identified. Third, the object's shape signature is identified as the sequence combination of TP, TPAS and/or VAS where: MBC-TPAS uses a sequence combination of TP and TPAS; MBC-VAS uses only VAS; and MBC-TPVAS uses a sequence combination of TP and VAS. Finally, a unique object representation is obtained as the discrete Fourier series expansion of the object's shape signature. The Fourier coefficients obtained are used for shape representation, index key, and shape similarity computation. The discrete Fourier transformation (DFT) of a shape signature is given by:

$$F(u) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f(k) \exp(-j2\pi uk/N) \quad u = 0, 1, \ldots, N-1 \tag{1}$$

where N is the number of vertices of the polygon representation of an object.

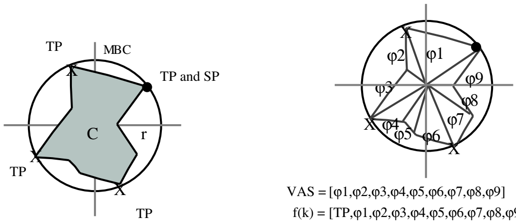
Although the MBC-based methods use simple attributes such as minimum bounding circles and angle sequences, these attributes were shown to be translation, rotation, and scale invariant in [22]. This is because, first when angle sequences are computed, only the angle between two vectors connecting the center of the MBC and the object's points are recorded. Thus, angle sequences does not depend on the exact coordinates of the points making the shape signatures of the objects translation invariant. Second, rotating two intersecting vectors in space does not change the angle between them. In addition, by taking the magnitude of $F(u)$ (i.e., $|F(u)|$), we are insuring that the angle sequence is independent of the choice of the start point. Hence, the MBC-based methods are also invariant to rotation. Finally, when applying uniform scaling to an object (assuming that the origin of the coordinate system is the center of the MBC), we are only changing its size and not its shape. Changing the size means that only the proportion of the distances between the points of the object and the center of its MBC are changing. Therefore, the new object's AS is equivalent to the old object's AS. Hence, MBC-based methods are also scaling invariant. In sum, the similarity measure is independent of object's position, scale, and rotation (i.e., objects does not need to be normalized).

The difference between two objects is the Euclidean distance between their corresponding feature vectors. Hence, the similarity measure of the MBC-based methods is that two objects are similar in shape, if and only if the Euclidean distance between their feature vectors is less than a prespecified threshold.

An example is illustrated in Fig. 3. Figure 3a shows an object with nine vertices with its minimum bounding circle, and touch points. Figure 3b shows the vertex angle sequence of the object and its shape descriptors $f(k)$ using MBC-TPVAS method.

A major observation is that MBC features are unique per object. In addition, a subset of the MBC features (e.g., TPAS, VAS, and TPs) are translation, scaling and rotation independent. This subset of features could be utilized to support query types in which the matching objects could be invariant with respect to translation, rotation and scaling (I-RST). It could also be utilized to support query types in which the matching objects could be within a specified rotation angle, scaling factor, translation vector, or any combination of the three (S-RST). We use object's SP to check if (and how much) it is rotated, object's r to check if (and how much) it is scaled, and the vertex coordinates of the object's C to check if (and how much) it is translated (see [22] for further details).

Fig. 3 MBC-TPVAS method



(a): An object with its MBC, center C, radius r, TP's, and SP

(b): VAS and shape descriptor $f(k)$

4.2 Fourier descriptors method

The FD method [9, 12, 13, 15, 19–21] obtains an object representation in the frequency domain as complex coefficients of the Fourier series expansion of the object's shape signature. The method starts by obtaining a feature function of the object called shape signature $f(k)$; which could be curvature based, radius based, or boundary coordinates based. $f(k)$ is also called the Fourier descriptors (FDs) of the boundary. In the next step, a discrete Fourier transform of the shape signature is obtained. The Fourier coefficients obtained are then used for shape representation, index key, and shape similarity computation. The discrete Fourier transformation (DFT) of a shape signature is given by:

$$F(u) = \frac{1}{N}\sum_{k=0}^{N-1} f(k)\exp(-j2\pi uk/N) \quad u = 0, 1, \ldots, N-1 \tag{2}$$

where N is the number of samples of $f(k)$.

Direct representation (e.g., radii) captures each individual detail of a shape, however, it is very sensitive to small changes and noise. As a consequence, a small change in the coordinates of the objects' boundary points may lead to a very different shape signature, hence, very poor retrieval performance. On the other hand, Fourier transformation captures the general feature of a shape by converting the sensitive direct representation measures into frequency domain. As a result, the data is more robust to small changes and noise. Therefore, Fourier transformation is used as shape representation instead of the direct representation (see [12] for further details).

A popular shape signature is the shape radii (or centroidal distance) which computes the distance between points uniformly sampled along the object boundary and its centroid (its center of mass). FD method is translation invariant, however, it is rotation and scaling variant. This is because, first when shape radii are used, only the distances between points are recorded and not the exact coordinates of the points. Thus, the shape signatures of the objects are translation invariant. Second, shape radii computes the distance between points uniformly sampled along the object boundary and its centroid. Thus, rotating an object would change the distances making the shape signatures of the objects rotation variance. Finally, in order to scale normalize the objects' FDs, the magnitude of all $F(u)$s are divided by the magnitude of $F(0)$. As a result of normalization a new feature vector, which is

Fig. 4 Fourier descriptors method

(a): An object with 8 sampling points (b): Fourier descriptor of the object

invariant to translation, rotation and scale, can be generated as follows: $FN = [|F(1)|/|F(0)||F(2)|/|F(0)|\ldots|F(N)|/|F(0)|]^T$ [12].

The difference between two objects is defined as the Euclidean distance between their corresponding feature vectors. Hence, the similarity measure of FD is that two objects are similar in shape, if and only if the Euclidean distance between their feature vectors is less than a prespecified threshold. In other words, they have the same set of distances between their centroids and their sampled boundary points.

An example is illustrated in Fig. 4, where Fig. 4a shows an object with its eight sampling points and its centroid, and Fig. 4b shows the radii distances and the Fourier descriptors $f(k)$ of the boundary.

4.3 Grid based method

An alternative approach for shape representation is the GB method [12, 13, 18–20]. With this method, an object is first normalized for rotation and scale. Then, the object is mapped on a grid of fixed cell size. Subsequently, the grid is scanned and a 1 or 0 is assigned to the cells depending on whether the number of pixels in the cell which are inside the object are greater than or less than a predetermined threshold. Finally, a unique binary number is obtained as the shape representation by reading the 1s and 0s assigned to the cells from left to right and top to bottom. To improve the efficiency of this method, another shape feature, called eccentricity [19], was used. Eccentricity of shape is the ratio of the number of cells used in x-direction to the number of cells used in the y-direction to represent a shape. Therefore, for two objects to be similar, their shape signatures and their eccentricities values should be similar.

GB method is translation invariant, however, it is rotation and scaling variant. This is because, first the shape signature is independent of object's vertices. Thus, the shape signatures of the objects are translation invariant. Second, the binary number of an object depends on the spatial relationship between the grid cells and the objects' boundaries, and also on both the grid and the object sizes. Hence, binary numbers obtained for the same object with a different orientations in space or with different scales will be different. As a consequence, normalization (rotation and scale) of the object boundaries prior to indexing is crucial. Rotation normalization is achieved by rotating the shape to make its major axis[1] parallel to the x-axis. Scaling normalization is achieved by first choosing a fixed length of the major axis, called the *standardized major axis*. Subsequently, the shape is scaled along the major axis

[1] Major axis of a shape is the line connecting the two points on the shape boundary that are farthest away from each other.

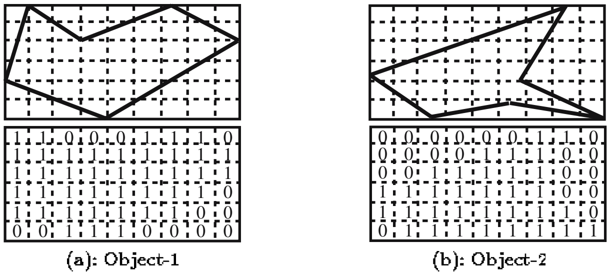


Fig. 5 Grid based method—object mapping and representation

such that its major axis becomes equal to the length of the standardized major axis. To maintain the perceptual similarity of the object, it is also scaled along the *minor axis* proportionally. The minor axis is perpendicular to the major axis and of such length that a rectangle with sides of major axis and minor axis defines the minimum bounding rectangle of the object.

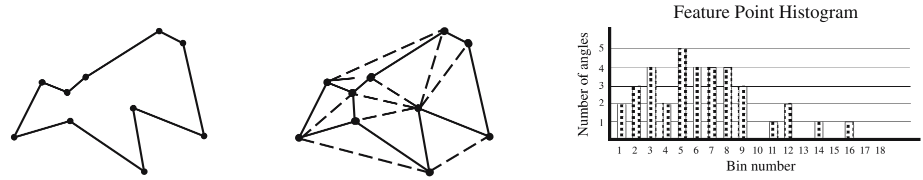
The difference between two objects is the number of cells in the grids which are covered by one shape and not the other, which is the same as the sum of 1s in the result of the exclusive OR of their binary numbers. Hence, the similarity measure of GB is that two objects are similar in shape, if and only if the difference between their binary representations is less than a prespecified threshold, and they have similar eccentricities.

An example is illustrated in Fig. 5, where the objects are mapped on to a grid of fixed cell size in a manner such that the objects are justified to the top left corner (i.e., assuming that the objects do not need to be normalized). The binary numbers obtained for the objects in Fig. 5a and b are “110001110 111111111 111111111 111111110 111111100 001110000” and “000000110 000011100 001111100 111111100 111111110 011111111,” respectively. Hence, the difference between the objects is 20.

4.4 Delaunay triangulation method

DT shape representation is a histogram-based approach [24, 25]. Given an object, *corner points* are used as the feature points of the object. Corner points are generally high-curvature points located along the crossings of an object’s edges or boundaries. Then a Delaunay triangulation of these feature points is constructed. Consequently, a feature point histogram is obtained by discretizing the angles produced by this triangulation into a set of bins and counting the number of times each discrete angle occurs in the triangulation. In building the feature point histogram, a selection criteria of which angles will contribute to the final feature point histogram is set. For example, the selection criteria could be counting the two largest angles, the two smallest angles, or all three angles of each individual Delaunay triangle.

The difference between two objects is the Euclidean distance between the corresponding bins of the objects’ feature point histograms. Hence, the similarity measure of DT is that two objects are similar in shape, if and only if the euclidean distance between their feature point histograms is less than a prespecified threshold. In other words, they have the same set of feature points. Thus, each pair of the corresponding Delaunay triangles in the two resulting Delaunay triangulations must be similar to each other. The angles of the resulting Delaunay triangles of a set of points is



(a): An object with 10 points

(b): Delaunay triangulation

(c): Feature point histogram

Fig. 6 Delaunay triangulation method

invariant under uniform translations, scaling, and rotations. Therefore, the similarity measure is independent of object's position, scale, and rotation (i.e., objects do not have to be normalized). This approach also supports an incremental approach to image object matching, from coarse to fine, by varying the bin sizes.

An example is illustrated in Fig. 6, where Fig. 6a shows an object that consist of ten feature points (corner points). Figure 6b shows the resulting Delaunay triangulation. While Fig. 6c shows the resulting feature point histogram built by counting all three angles of each individual Delaunay triangle with bin size of 10 degrees.

4.4.1 Modified Delaunay triangulations method

Optimally, a straight line could be represented by its two end points. However, typical edge detection algorithms used to generate shape files introduce more than two points to represent a straight line. The extra points make the straight line looks like a jagged line (like saw edge). With DT, *corner points* are used to reduce the number of vertices to represent the shapes while maintaining its general characteristics. However, in our *modified Delaunay triangulation* method (M-DT), we employ a three pass algorithm. Wherein each pass, we try to identify the straight lines on the boundary of the shape and eliminate the extra points representing the line. Our results show that DT is very sensitive to the technique used to find the representative vertices of the objects. Therefore, the accuracy of our modified DT (M-DT) is higher than the original DT.

5 Index structures and spatial queries

In [22], we presented a multi-step query processing method that efficiently supports the retrieval of 2D objects. We showed how our proposed MBC-based methods can be used for efficient retrieval of 2D objects by utilizing three different index structures I_{TPAS}, I_{VAS}, and I_{TPVAS} based on features that are extracted from the objects MBC. These features can help us in both reducing the complexity of comparison algorithms and building index structures to search only a subset of the candidate objects for match queries. We focused on three variations of match queries (*EM, RST, and SIM*) to cover cases where the database is searched for exactly identical objects; identical objects that are rotation, scaling, and translation invariant; or *similar* objects per our definition of similarity. Our performance results showed the superiority of our techniques as compared to a *naive* method with a minimum margin of 40% improvement in the I/O cost, and orders of magnitude improvement in CPU cost. We also identified that the index structure based of the MBC-TPVAS is the superior one among the three, independent of the size of the database and the number of vertices of the objects. Therefore, in this study, we will only focus on the performance on MBC-TPVAS method.

Besides similarity retrieval, in [16] we describe the support of spatial queries (i.e., topological and direction queries) using our MBC features. We investigated the usefulness and feasibility of utilizing the same MBC approximations and identified a subset of MBC features that could be utilized to support such queries. As a result, there is no need to maintain two approximations (i.e., MBR and MBC) per object for efficient support of different queries (i.e., similarity and spatial).

6 Discussion

This section describes general comparison criteria that can be used to distinguish between different shape retrieval techniques, and points out some of the drawbacks of the alternative shape description techniques.

Normalization With GB and FD, the shape representations obtained for the same object with different orientations in space or with different scales are different. Therefore, the normalization of the object boundaries prior to indexing is crucial to meet the uniqueness criteria of the shape descriptor. DT and MBC-TPVAS use simple shape attributes such as the angles of the resulting Delaunay triangles of a set of points, MBC and AS. These attributes were shown to be translation, rotation, and scale invariant. Therefore, no normalization is required.

Support for S-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 (termed I-RST), or with a specified size and/or orientation (termed S-RST). With MBC-TPVAS, a subset of the MBC features (i.e., r, C, and SP) could be utilized to support RST queries (see [22] for further details). On the other hand, the GB, FD, and DT methods are unable to support bot I-RST and S-RST. With GB and FD, normalization is used so that the objects fit into a prespecified mesh, or force the boundaries to have a standard size, and orientation, respectively. Therefore, they can support I-RST but not S-RST. DT representation provides no information about the size, and orientation of the original objects, and hence can also support I-RST but not S-RST.

Different shape signatures per object GB performs rotation normalization by rotating the original object until its *major axis* is parallel with the *x-axis*. The object could be rotated either clockwise or counterclockwise. Hence, two different shape signatures are maintained for each object. Furthermore, an object may have multiple major axes. Therefore, two more shape signatures should be maintained for the object per each additional major axis. With FD, a shape signature is constructed using shape radii of points uniformly sampled along the object boundary. These points are not the exact coordinates of the object vertices, therefore, using different sampled points leads to different shape signatures.*Size of shape signature* With GB, there is a tradeoff between the number of grid cells and both the time required to create object's signatures and similarity computation time during retrieval. The smaller the cells the greater the accuracy with which the shapes can be indexed but the more computationally complex the signature generating task and the similarity measure computation. In addition, depending on the minor axis, the binary number obtained for each shape might be of different lengths. Consequently, it would be more difficult to index these numbers using spatial indices that require fixed dimension (e.g., R-tree).

Shape representation inaccuracy With GB, the binary numbers obtained for shapes when using a grid of fixed cell size becomes increasingly inaccurate for shapes with larger eccentricities. A proposed solution in [19] is to use an adaptive grid size which will vary the grid layout in a manner that the shape is most effectively indexed while generating a binary number of fixed length. A drawback is that it is no longer possible to cluster/classify objects according to their eccentricities.

Choice of parameters With GB, both the shape and the grid size affect the derivation of the binary number. Hence, an optimum standardized major axis size, and the number of pixels inside a grid should be selected carefully. With DT, choosing a histogram with a small number of bins is less discriminating. The coarser the bin size of the feature point histogram is, the worse the overall effectiveness of the shape matching becomes. Therefore, the size and the number of histogram bins should be selected carefully. In addition, the choice of the angle selection criterion influences the ability of object shape discrimination. Moreover, its performance depends on the shape representation scheme, which in turn depends on the quality of the technique used to find the corner points (high-curvature points).

7 Comparison study

Section 7.1 provides cost analysis for the different methods and compares their efficiency in terms of computation and storage requirements. In Section 7.2, we describe the implementation of our prototype shape retrieval system. Finally, Section 7.3 reports the experimental results of the retrieval accuracy of the four different shape representation methods (i.e., FD, GB, M-DT, and MBC-TPVAS).

7.1 Analytical comparison

This section provides cost analysis for the different methods and compares their efficiency in terms of computation and storage requirements. From Table 1(a) we can conclude that both MBC-TPVAS and FD have the lowest storage requirement, while M-DT requires the largest storage requirement. GB requires larger storage than M-DT if the objects have multiple major axes. Both GB and M-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 MBC-TPVAS has the lowest computation cost and GB the highest. Both M-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.

Table 1 Costs

GB	M-DT	FD	MBC-TPVAS
a) Storage cost (n objects)			
$2*n*SS+8*n$ and *temp* $px/8$	$4*n*nb$ and *temp* $3*N$	$8*n*fn$	$8*n*fn$
b) Shape signature generation cost			
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			
$8*SS$ *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 $fn2-1$ float add

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

7.2 Graphical comparison

We implemented a prototype shape retrieval system in Java, on a Sun Ultra 60 workstation with 512 MB of RAM. Figure 7 shows an example of the shape signatures of an object using four alternative shape representation techniques.

7.3 Experiments

In many applications, an important criterion for testing the efficacy of the shape retrieval methods is that for each query object the relevant items (similar shapes) in the database should be retrieved. Therefore, an experiment was designed to measure the effectiveness of the similarity retrieval of the four different shape representation methods (i.e., FD, GB, M-DT, and MBC-TPVAS) in terms of recall and precision that are commonly used in the literature. Recall measures the ability of retrieving relevant shapes in a database and is defined as the ratio between the number of relevant shapes retrieved and the total number of relevant shapes in the database. While, precision measures the retrieval accuracy and is defined as the ratio between the number of relevant shapes retrieved and the number of total shapes retrieved.

For our experiments, boundary features (based on shape) are first extracted from the objects during the population of the database. At query time, the user enters a query into the system and features are extracted from the query object. The query object is compared with the database of objects by the object retrieval system on the basis of the closeness of the extracted features and the best match objects are presented to the user. The accuracy of the shape retrieval methods was measured in terms of recall and precision. In the following, we briefly describe the experimental setup, and the experimental results.

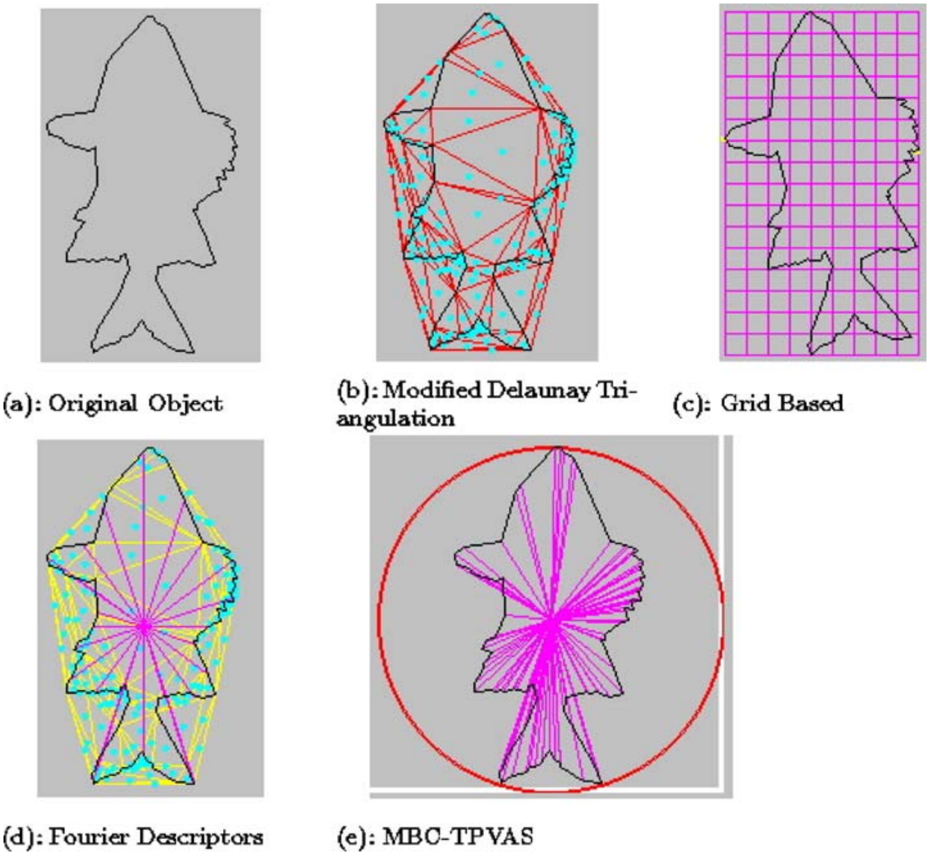


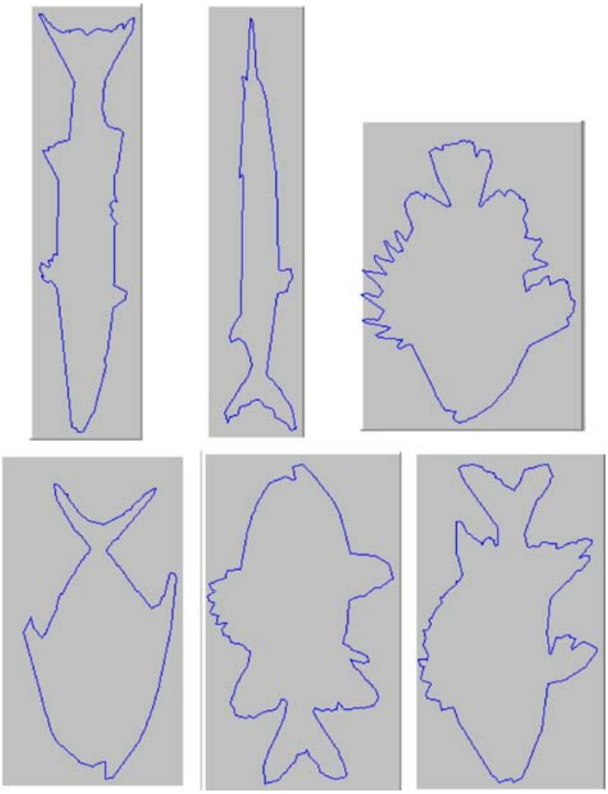
Fig. 7 An object and its four-shape representations

7.3.1 Experimental setup

The shape database of the system consists of 296 fish and tool shapes provided to us by [6, 14]. For each of the 296 objects, ten similar objects' variants were constructed as follows: three rotation variants, three rotation scaled up variants, three rotation scaled down variants, and one translated variant. Those the total number of the shapes in the database was increased to 2,960 objects. Some examples of shapes inside the database are shown in Fig. 8. Fifteen query shapes were selected randomly from the database, where each query object has ten 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. Figure 9 shows a query object and some of its variants. For our experiments, we simply assume that each shape is relevant only to itself and to its ten variants.

As advised in [13, 19, 22, 24], for GB we employed a grid cell size of 12×12 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

Fig. 8 Examples of objects in the dataset



both FD and MBC-TPVAS, we kept eight low frequency coefficients of the Fourier transform. Finally, for M-DT, we used two largest angle histogram with 18 bins of size 10 degrees.

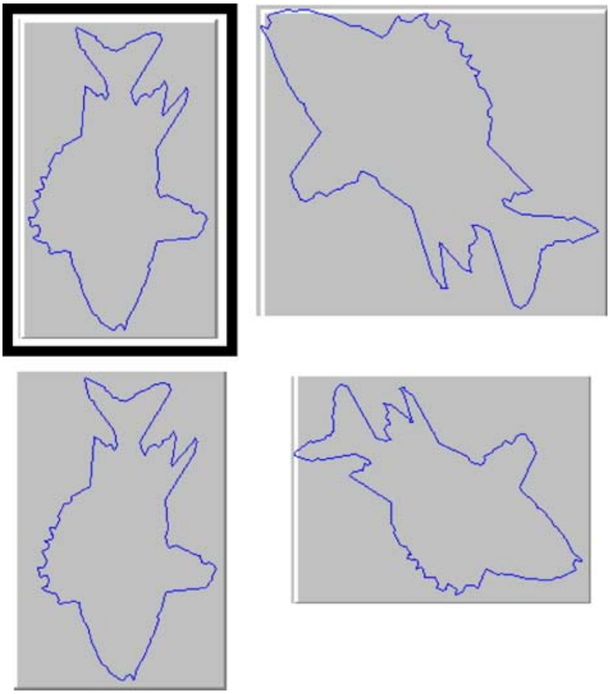
7.3.2 Experimental results

Given a query object, the system retrieves relevant shapes from the database, in decreasing order of similarity to the query shapes. First, the shape features of the query object using the four different methods are extracted. Second, the Euclidean distances between the shape signature of the query object and all other objects' shape signatures are computed.[2] Third, the Euclidean distances are ordered in an ascending order for all the methods. Finally, the accuracy of the retrieval methods was calculated as precision-recall. Towards this end, we identify the ranks of the relevant objects to the query object within the ordered list (according to Euclidean distance).

The recall and precision results for 15 query shapes are averaged and shown in Fig. 10. From Fig. 10a, it is observed that FD was significantly outperformed by all

[2]The retrieval can be done more efficiently using index structures, but since this is not the focus of this study we assumed the sequential scan of the database for all the methods.

Fig. 9 A query object and its three variants



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 from the Fourier coefficients ($F(u)$) and ignoring its phase information as explained in [19].

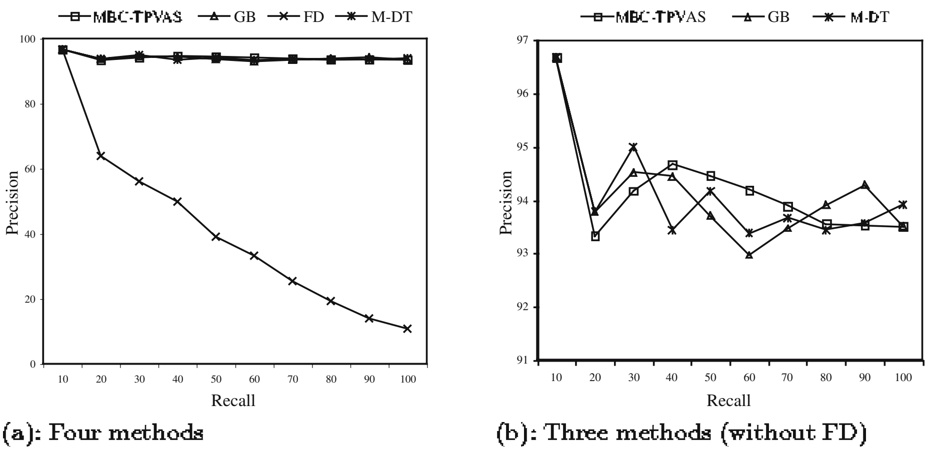


Fig. 10 Recall-precision curves

Figure 10b, shows the recall-precision curves of the methods excluding FD. From Fig. 10b, we make the following observations. First, that the accuracy of our proposed method (MBC-TPVAS) is as good as that of other methods. Second, unlike what was expected, the results obtained for our modified DT (M-DT) were better than what was reported in [24]. 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*. Instead, in our experiments, we used the three pass algorithm described in Section 4.4.1. 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, the 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.

8 Conclusion and future work

This study compared four shape-based object retrieval techniques (FD, GB, DT, and MBC-TPVAS). The similarity retrieval accuracy of our method (MBC-TVPAS) was as good as 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, MBC-TPVAS requires no normalization of the objects, and is the only method that has direct support for S-RST query types. Finally, the results of GB and DT methods are sensitive to human decisions, choice of parameters, and algorithms.

We intend to extend this work in several directions. First, we plan to investigate how to handle objects with curved parts, holes, and open curves or lines. Second, partial similarity match is not supported by all the systems, therefore we intend to identify better features to be used for partial similarity match queries. Finally, we want to extend this work to support three dimensional objects. Our preliminary investigations show that analogous to MBC features for 2D objects, we can extract features from 3D objects by using their minimum bounding spheres.

Acknowledgements The authors would like to thank Geoff Leach for providing us with the Java code for implementing Delaunay triangulation algorithm (available at http://goanna.cs.rmit.edu.au/gl/research/comp_geom/delaunay/delaunay.html). Special thanks also go to F. Mokhtarian, X. Ding, and S. Abbasi for providing us with the fish shapes data set (available at ftp://ftp.ee.surrey.ac.uk/pub/vision/misc/fish_contours.tar.Z). We would also like to thank J. Gary for kindly providing us with the tool shapes data set.

References

1. Agrawal R, Faloutsos C, Swami A (1993, October) Efficient similarity search in sequence databases. In: Proceedings of the international conference of FODO, Chicago, IL
2. Berchtold S, Keim D, Kriegel HP (1997) Using extended feature objects for partial similarity retrieval. In: Proceedings of 23rd very large databases (VLDB) conference, Springer, Berlin Heidelberg New York, pp 333–348

3. Eakins JP (1994) Retrieval of trade mark images by shape feature. In: Proceedings first international conference on electronic library and visual information system research, de Montfort University, Milton Keynes, UK, pp 101–109
4. Faloutsos C, Ranganathan M, Manolopoulos Y (1994) Fast subsequence matching in time–series databases. In: Proceedings of the ACM SIGMOD international conference on management of data, Minneapolis, MN
5. Gary J, Mehrotra R (1993) Feature-based retrieval of similar shapes. In: Proceedings of international conference on data engineering (ICDE), Vienna, Austria, pp 108–115
6. Gary J, Mehrotra R (1995) Similar-shape retrieval in shape data management. IEEE Comput Mag 28:57–62
7. Ghandeharizadeh S (1995) Stream-based versus structured video objects: issues, solutions, and challenges. In: Jajodia S, Subrahmanian V (eds) Multimedia DB systems: issues and res. direct. Springer, Berlin Heidelberg New York
8. Gonzalez RC, Wintz P (1987) Digital image processing 2nd edn. Addison-Wesley, Reading, MA
9. Granlund JH (1972) Fourier preprocessing for hand print character recognition. IEEE Trans Comput 21:195–201
10. Jagadish HV (1991) A retrieval technique for similarity shapes. In: Proceedings ACM SIGMOD international conference on management of data, Denver, CO, pp 208–217
11. Korn F, Sidiropoulos N, Faloutsos C, Siegel E, Protopapas Z (1996) Fast nearest neighbor search in medical image databases. In: Proceedings 22nd VLDB conference, Mumbai, India, pp 215–226
12. Lu G, Sajjanhar A (1999) Region-based shape representation and similarity measure suitable for content-based image retrieval. Multimedia Syst 7(2):165–174
13. Mehtre BM, Kankanhalli MS, Lee WF (1997) Shape measures for content based image retrieval: a comparison. Inf Process Manag, 33(3):319–337
14. Mokhtarian F, Abbasi S, Kitter J (1996) Efficient and robust retrieval by shape content through curvature scale space. In: Proceedings of international workshop on image database and multimedia search, Amesterdam, Netherlands, pp 35–42
15. Pitas I (1993) Digital image processing algorithms. Prentice Hall, Englewood Cliffs, NJ
16. Safar M, Shahabi C (1999) 2D topological and direction relations in the world of minimum bounding circles. In: Proceedings of IEEE international database engineering and applications symposium (IDEAS), Montreal, Canada, pp 239–247, 2–4 August
17. Safar M, Shahabi C, Tan, C-H (2000) Resiliency and robustness of alternative shape-based image retrieval techniques. In: Proceedings of IEEE international database engineering and applications symposium (IDEAS), Yokohama, Japan
18. Sajjanhar A, Lu G (1997a) Indexing 2D non-occluded shape for similarity retrieval. In: Proceedings of SPIE conference on applications of digital image processing XX, vol 3164, San Diego, CA, pp 188–197, 30 July–1 August
19. Sajjanhar A, Lu G (1997b) A grid based shape indexing and retrieval method. Aust Comput J 29(4):131–140
20. Sajjanhar A, Lu G (1998) A comparison of techniques for shape retrieval. In: International conference on computational intelligence and multimedia applications, Monash University, Gippsland Campus, Australia, pp 854–859, 9–11 February
21. Sajjanhar A, Lu G, Wright J (1997) An experimental study of moment invariants and Fourier descriptors for shape based image retrieval. In: Proceedings of the second Australia document computing symposium, Melbourne, Australia, pp 46–54, 5 April
22. Shahabi C, Safar M (1999) Efficient retrieval and spatial querying of 2D objects. In: Proceedings of IEEE international conference on multimedia computing and systems (ICMCS), Florence, Italy, pp 611–617, 7–11 June
23. Sul C, Lee K, Wohn K (1998) Virtual stage: a location-based karaoke system. IEEE Multimed 5:42–52
24. Tao Y, Grosky WI (1999, January) Delaunay triangulation for image object indexing: a novel method for shape representation. In: Proceedings of the 7th SPIE symposium on storage and retrieval for image and video databases, San Jose, CA, pp 631–642
25. Tao Y, Grosky WI (1999, January) Object-based image retrieval using point feature maps. In: Proceedings of the international conference on database semantics (DS-8), Rotorua, New Zealand, pp 59–73

Cyrus Shahabi is currently an Associate Professor and the Director of the Information Laboratory (InfoLAB) at the Computer Science Department and also a Research Area Director at the NSF's Integrated Media Systems Center (IMSC) at the University of Southern California. He received his B.S. degree in Computer Engineering from Sharif University of Technology in 1989 and his M.S. and Ph.D. degree in Computer Science from the University of Southern California in 1993 and 1996, respectively. He has two books and more than hundred articles, book chapters, and conference papers in the areas of databases, GIS and multimedia. Dr. Shahabi's current research interests include Geospatial and Multidimensional Data Analysis, Peer-to-Peer Systems and Streaming Architectures. He is currently an associate editor of the IEEE Transactions on Parallel and Distributed Systems (TPDS) and on the editorial board of ACM Computers in Entertainment magazine. He is also in the steering committee of IEEE NetDB and ACM GIS. He serves on many conference program committees such as ACM SIGKDD 2006, IEEE ICDE 2006, ACM CIKM 2005, SSTD 2005 and ACM SIGMOD 2004. Dr. Shahabi is the recipient of the 2002 National Science Foundation CAREER Award and 2003 Presidential Early Career Awards for Scientists and Engineers (PECASE). In 2001, he also received an award from the Okawa Foundations.

Maytham Safar is currently an Assistant Professor at the Computer Engineering Department at Kuwait University. He received his Ph.D. degree in Computer Science from the University of Southern California in 2000. He has one book and more than 25 articles, book chapters, and conference/journal papers in the areas of databases and multimedia. Dr. Safar's current research interests include Peer-to-Peer Networks, Spatial Databases, Multidimensional Databases, and Geographic Information Systems. He served on many conferences as a reviewer and/or a scientific program committee member such as ICDCS, EURASIA-ICT, WWW/Internet, ICWI, ICME, AINA, WEBIST, IPSI, HPC&S, ICICT, i-Society, ET-WBC, ICDIM and iiWAS. He also served as a member on the editorial board or a reviewer for many journals such as IEEE TMJ, ACM Computing Reviews, JDIM, MTAP, IEEE TPAMI, and ACM MSJ.