A Consensus-based Fusion Algorithm in Shape-based Image Retrieval

(1) Akrem El-ghazal, (2) Otman Basir
Department of Electrical and Computer Engineering
University of Waterloo
Waterloo, On., Canada
(1) akrem@pami.uwaterloo.ca
(2) obasir@uwaterloo.ca

Saeid Belkasim
Department of Computer Science
Georgia State University
Atlanta, GA., USA
sbelkasim@cs.gst.edu

Abstract - Shape-based image retrieval techniques are among the most successful Content-Based Image Retrieval (CBIR) approaches. In recent years, the number of shape-based image retrieval techniques has dramatically increased; however, each technique has both advantages and shortcomings. This paper proposes a consensus-based fusion algorithm to integrate several shape-based image retrieval techniques so as to enhance the performance of the image retrieval process. In this algorithm, several techniques work as a team: they exchange their ranking information based on pair-wise co-ranking to reach a consensus that will improve their final ranking decisions. Although the proposed algorithm handles any number of CBIR techniques, only three common techniques are used to demonstrate its effectiveness. Several experiments were conducted on the widely used MPEG-7 database. The results indicate that the proposed fusion algorithm significantly improves the retrieval process.

Keywords: Consensus, fusion, image retrieval, shape.

1 Introduction

In recent years, there has been an enormous increase in the size of digital image databases. Contributing to this increased interest are the ease and convenience of capturing, transmitting, and storing digital images in personal and public image databases. The internet is already rich in image depositories that cover a wide range of applications, including biomedical, multimedia, geological, and astronomy-related. Although the format for storing image data is relatively standardized, the process of retrieving images from an image depository tends to be quite complex and has thus become a limiting factor that needs to be addressed.

Typically, images in a database are retrieved based on either textual information or content information. Early retrieval techniques were based on textual annotation of images. Images were first annotated with text and then searched based on their textual tags. However, text-based techniques have many limitations, including their reliance on manual annotation, which can be difficult and tedious for large image sets. Furthermore, the rich content typically found in images and the subjectivity of human perception make the task of describing images using words a difficult if not an impossible task.

To overcome these difficulties, Content-Based Image Retrieval (CBIR) was proposed [1]. This approach relies on the visual content of images rather than on textual annotations to search for images and hence has the potential to be more effective in responding to specific user queries. CBIR techniques use visual content such as color, texture, and shape to represent and index the image. Visual color and texture content have been explored more thoroughly than shape content. The increasing interest in using the shape features of objects for CBIR is not surprising, given the considerable evidence that natural objects are recognized primarily by their shape [2, 3]. A survey of users with respect to the cognition aspects of image retrieval confirmed that users prefer retrieving images based on shape rather than color or texture [4]. However, retrieval based on shape content remains more difficult than retrieval based on other visual features [2]. During the last decade, significant progress has been made in both the theoretical and practical research aspects of shape-based image retrieval [6,5]. Approaches to representing shapes can be divided into two main categories: region-based approaches and boundary-based approaches (also known as contour-based approaches). Region-based approaches often use moment descriptors to describe shapes. These descriptors include geometrical moments [7], Zernike moments [8-11], pseudo-Zernike moments [12], and Legendre moments [8]. Although region-based approaches are global in nature and can be applied to generic shapes, they fail to distinguish between similar objects [13].

In many applications, the internal content of the shape is not as important as its boundary. Boundary-based techniques tend to be more efficient for handling shapes that are describable by their object contours [13]. Many boundary-based techniques have been proposed in the literature, including Fourier descriptors [14-16], Curvature Scale Space (CSS) [17, 18], wavelet descriptors [19], and contour displacement [20].
Although the number of available shape-based image retrieval techniques has been increasing rapidly, such techniques still exhibit shortcomings. Techniques that have demonstrated reasonable robustness often tend to be computationally complex [20, 21]. In this paper, we propose to integrate several of the existing techniques in a pair-wise co-ranking scheme so as to obtain superior shape-based image retrieval.

The remainder of the paper is organized as follows: Section 2 introduces the problem formulation, and Section 3 discusses the proposed pair-wise co-ranking scheme. The process of formulating the consensus about the final ranking decision is presented in Section 4. Section 5 describes experiments for evaluating the performance of the proposed fusion algorithm. Conclusions that can be drawn from this study and suggestions for future work are presented in Section 6.

2 Problem Formulation

The study considers a group of image retrieval techniques that have complementary image representation capabilities. These techniques are viewed as agents that cooperate to determine the best candidate image \( x \) from an ensemble of images \( \Theta \) that matches a query image. This group of image retrieval agents is indexed by the set \( \text{IRA} = \{\text{IRA}_1, \text{IRA}_2, \ldots, \text{IRA}_M\} \). Each agent \( \text{IRA}_i \) uses a feature extraction scheme \( F_i \) and a matching strategy \( \Gamma_i \) to determine a similarity measure \( S_i \) between query image \( x \) and all images \( y \in \Theta \); that is,

\[
S_i(x, y) = \Gamma_i(F_i(y), F_i(x)), \forall y \in \Theta
\]  

(1)

Each agent \( \text{IRA}_i \) establishes a ranking \( R_i(y \mid F_i(y)) \in \{1, 2, \ldots, N\}, \forall y \in \Theta \) such that \( R_i(y \mid F_i(y)) \leq R_i(z \mid F_i(z)) \) implies \( S_i(x, y) \geq S_i(x, z) \).

Without loss of generality, \( R_i(y \mid F_i(y)) \) can be viewed as an index set from 1 to \( N \), where index 1, for example, points to the closest candidate image to the query image \( x \), and index \( N \) points to the most dissimilar candidate image to the query image \( x \). In general, index \( l \) in this set points to the image that is preceded by \( l-1 \) candidates; these candidates are viewed by \( \text{IRA}_i \) as better candidates for the query image \( x \) than the candidate given the rank \( l \). Since each agent uses a different feature extraction scheme, it is expected that the agents will reach different ranking decisions for each different image in the set. Each of the CBIR techniques must have demonstrated reasonable performance before it was selected to be a member of the group. It is thus reasonable to expect that good image candidates will cluster at the top of the ranking for all agents.

3 Pair-wise Co-ranking Scheme

The following discussion proposes an information exchange scheme between the image retrieval agents so as to assist each in determining the best image as a candidate for matching the user query. This scheme allows the relative advantages and disadvantages of each agent to be exploited.

Pair-wise co-ranking revolves around the hypothesis that each agent \( \text{IRA}_i \) may re-adjust its ranking of a candidate image if it is exposed to the ranking produced by another agent \( \text{IRA}_j \). The communication process between two agents is depicted in Figure 1. The initial ranking of an agent is referred to as marginal ranking. This ranking reflects the assessment of each image by the agent with respect to how close the image is to the query image. On the other hand, the revised ranking of an agent is referred to as conditional ranking. This ranking reflects how the assessment by an agent is influenced by the assessments produced by other agents.

To set up the process of information exchange among the agents, the ranking set of each agent \( \text{IRA}_i \) is divided into two partitions: an Elite Candidates Partition (\( \text{ECP}_i \)) and a Potential Candidates Partition (\( \text{PCP}_i \)). It is expected that good matches to the query will cluster in the elite partitions. The \( \text{ECP} \) contains the first \( m \) candidates; the \( \text{PCP} \) contains the last \( N-m \) candidates. Thus, the marginal ranking of agent \( \text{IRA}_i \) can be viewed as a concatenation of two ranking sets; that is, \( R_i = \{R_i^{\text{ECP}}, R_i^{\text{PCP}}\} \), where \( R_i^{\text{ECP}} \) is the ranking for \( R_i \leq m \), and \( R_i^{\text{PCP}} \) is the ranking for \( R_i > m \).

We need to introduce the conditional ranking like we did with the marginal ranking. Figure 2 depicts the process of forming conditional elites, in which agent \( \text{IRA}_i \) revises its conditional elite partition ranking \( R_i^{\text{ECP}} \) based on the marginal ranking information provided by agent \( \text{IRA}_j \). Here, image \( z \) is seen as an elite image by agent \( \text{IRA}_j \). Agent \( \text{IRA}_i \) uses its own feature extraction scheme \( F_i \) to determine the rank of image \( z \) in its conditional elite candidates partition; that is,

\[
R_i^{\text{ECP}}(z \mid F_i(z), R_i(z \mid F_i(z)) \leq m), z \in \text{ECP}_i
\]  

(2)

This formula can be read as follows: agent \( j \) re-ranks image \( z \) based on its feature extraction scheme \( F_j(z) \) which takes into consideration that image \( z \) has been ranked as an elite candidate by agent \( \text{IRA}_i \) based on its feature extraction scheme \( F_i(z) \).
Figure 2: Illustration of the pair-wise co-ranking scheme.

The fact that image $z$ is placed in the conditional elite partition of agent $IRA_j$ does not necessarily imply that image $z$ is in the marginal elite partition of $IRA_j$. Similarly, the conditional rankings of the potential candidate partitions are computed:

$$R^{ECP}_i(z | F_i(z), R_i(z | F_i(z)) > m), z \in PCP$$

(3)

The conditional ranking of agent $IRA_j$, based on information received from agent $IRA_i$, is viewed as a concatenation of two ranking sets; that is,

$$R_j = \{R^{ECP}_j, R^{PCP}_j\}$$

(4)

The results of the above process are summarized in Table 1.

Table 1: Marginal and Pair-wise Conditional Rankings

$$R^{ECP}_i(z | F_i(z)) \forall R_i \leq m$$

$$R^{PCP}_i(z | F_i(z)) \forall R_i > m$$

$$R^{ECP}_j(y | F_j(z), R_j(y | F_j(z)) \leq m) \forall R_j < m$$

$$R^{PCP}_j(y | F_j(z), R_j(y | F_j(z)) > m) \forall R_j > m$$

$$R^{ECP}_{ji}(y | F_i(z), R_i(y | F_i(z)) \leq m) \forall R_{ji} \leq m$$

$$R^{PCP}_{ji}(y | F_i(z), R_i(y | F_i(z)) > m) \forall R_{ji} > m$$

$$R^{ECP}_{ji}(y | F_j(z), R_j(y | F_j(z)) \leq m) \forall R_{ji} \leq m$$

$$R^{PCP}_{ji}(y | F_j(z), R_j(y | F_j(z)) > m) \forall R_{ji} > m$$

$$R^{ECP}_{ji} = \{R^{ECP}_{ji}, R^{PCP}_{ji}\}$$

Illustrative Example:

The database described in Section 5 consists of 70 classes, each class having 20 objects. The target is to have the relevant images be ranked in the top 20 positions. Figure 3 and 4 display the results produced by agents $IRA_1$ and $IRA_2$, respectively.

In both figures, the top left shape is the query shape (Frog); the retrieved shapes are arranged in descending order according to each shape’s similarity to the query shape. Figure 3 indicates that the first agent $IRA_1$ has managed to correctly rank 9 of 20 shapes (45%). Furthermore, most of the irrelevant shapes, indicated by the dashed frame, are objects that belong to the same class (bell). In Figure 4, it is evident that the irrelevant shapes retrieved by agent $IRA_2$, indicated by the dashed frame, differ from those retrieved by agent $IRA_1$. Figure 5 shows the conditional ranking of agent $IRA_1$, based on the ranking information it has received from the second agent $IRA_2$. As can be seen in Figure 5, the irrelevant shapes of the first agent, indicated by the dashed frame in Figure 3, do not appear in top-ranked positions. The relevant shapes, enclosed in ellipses in Figure 3, as ranked by the first agent $IRA_1$, are repositioned to be the top 20 positions in $R_{21}$, resulting in an accuracy rate of 75% (15 out of 20). It is clear that the pair-wise conditional ranking $R_{21}$ is much better than the marginal rankings $R_1$ and $R_2$. 

Figure 3: Retrieved results of $IRA_1$ ($R_1$)

Figure 4: Retrieved results of $IRA_2$ ($R_2$).

Figure 5: Pair-wise conditional ranking of $IRA_1$ based on $IRA_2$ ($R_{21}$).
4 Consensus-based Fusion

The main motivation for this research is to identify image retrieval techniques that capture different characteristics of an image and combine them in a way that enhances retrieval performance through consensus about the rankings. These techniques are considered to be a team of agents who cooperate in order to determine the image in the database that best matches the query image. This collaboration is accomplished through an exchange of candidate ranking information. Each agent uses its feature extraction scheme to compute a ranking $R_{i}$. Double indexing is used to simplify the mathematical notations. The ranking will reflect the technique’s preference for the candidate that best matches the query image (i.e., marginal ranking). Furthermore, each agent is exposed to the results of the ranking by the other agents so as to compute a conditional ranking for each candidate image, as explained above. Therefore, $M$ retrieval agents yield $M^2$ rankings: $M$ marginal rankings, plus $M(M-1)$ conditional rankings. Figure 6 depicts the ranking set of a three-agent system.

![Figure 6: Information exchange for three image retrieval agents.](image)

For $M$ agents, the ranking sets can be organized in a matrix format:

$$
R = \begin{bmatrix}
R_{11} & R_{12} & \cdots & R_{1M} \\
R_{21} & R_{22} & \cdots & R_{2M} \\
\vdots & \vdots & \ddots & \vdots \\
R_{M1} & R_{M2} & \cdots & R_{MM}
\end{bmatrix}
$$

where $R_{ij}$ is the conditional ranking of the $i^{th}$ agent, given the ranking of the $j^{th}$ agent, and $R_{ii}$ is the marginal ranking of the $i^{th}$ agent.

Figure 7 portrays the rank information exchange that yields the pair-wise conditional rankings.

![Figure 7: Steps in applying the pair-wise co-ranking.](image)

To obtain a final decision about rankings, a pair-wise consensus algorithm is proposed. In this algorithm, each set of pair-wise conditional rankings, obtained from the cooperation of each pair of the techniques used, is first combined by averaging their rankings:

$$
R_{cc} = (R_{cc1} + R_{cc2})/2
$$

where $c=1,2,\ldots M(M-1)/2$, and $M$ is the number of the used techniques.

Then the pair-wise co-ranking scheme is applied to the combined conditional rankings $R_{cc}$ to produce a new set of pair-wise conditional rankings $R_{ij}^t$ ($t$: represents the number of the iterations). At each iteration $t$, the algorithm produces set of $(M(M-1)/2)$ combined conditional rankings $R_{cc}$. The initial rankings of the individual agents are represented by $R_{ii}^0 (i=j, i=1,2,\ldots M)$.

In this manner, each technique influences the ranking information of the other techniques until all combined pair-wise conditional rankings $R_{cc}$ reach a consensus about their rankings. Figure 8 shows the main steps in the proposed algorithm.

![Figure 8: The main steps in the proposed algorithm.](image)

**Illustrative Example:**

This section introduces an example using the MPEG-7 database that illustrates the effectiveness of the proposed algorithm. For the sake of simplicity, numbers are used to refer to all images in the database.

An image that has an index of 746 is considered to be a query image. In this example, three techniques are used to retrieve similar images from part B of the MPEG-7 database for the given query image.

The proposed algorithm is applied and the rankings after different iterations are shown in Table 1 and Table 2. The first three columns of Table 1 show the rankings of the three individual techniques. Only the first top 20
images retrieved by each technique are listed, and star signs are used to point out the irrelevant retrieved images. The results for whose rankings all the techniques reached consensus are highlighted.

From the first three columns of Table 1, it can be seen that the three individual techniques reached consensus about the rankings of only two images. Moreover, many images were irrelevant, as indicated by the star signs. After the fourth iteration, most of the irrelevant images, indicated by star signs, from the individual rankings, do not appear in the top 20 ranked positions. Furthermore, the irrelevant images are replaced with relevant images.

Table 1: The rankings of the three individual techniques and the proposed algorithm after the fourth iteration.

<table>
<thead>
<tr>
<th>$R_{1}^{4}$</th>
<th>$R_{2}^{4}$</th>
<th>$R_{3}^{4}$</th>
<th>$R_{1}^{4}$</th>
<th>$R_{2}^{4}$</th>
<th>$R_{3}^{4}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>748</td>
<td>748</td>
<td>748</td>
<td>748</td>
<td>748</td>
<td>748</td>
</tr>
<tr>
<td>749</td>
<td>749</td>
<td>749</td>
<td>749</td>
<td>749</td>
<td>749</td>
</tr>
<tr>
<td>746</td>
<td>746</td>
<td>759</td>
<td>757</td>
<td>741</td>
<td>741</td>
</tr>
<tr>
<td>747</td>
<td>747</td>
<td>758</td>
<td>747</td>
<td>747</td>
<td>757</td>
</tr>
<tr>
<td>757</td>
<td>756</td>
<td>741</td>
<td>741</td>
<td>757</td>
<td>747</td>
</tr>
<tr>
<td>756</td>
<td>757</td>
<td>760</td>
<td>760</td>
<td>760</td>
<td>760</td>
</tr>
<tr>
<td>745</td>
<td>760</td>
<td>753</td>
<td>744</td>
<td>744</td>
<td>744</td>
</tr>
<tr>
<td>744</td>
<td>593*</td>
<td>757</td>
<td>746</td>
<td>746</td>
<td>758</td>
</tr>
<tr>
<td>750</td>
<td>592*</td>
<td>431*</td>
<td>758</td>
<td>754</td>
<td>754</td>
</tr>
<tr>
<td>751</td>
<td>741</td>
<td>702</td>
<td>754</td>
<td>756</td>
<td>746</td>
</tr>
<tr>
<td>742</td>
<td>565*</td>
<td>742</td>
<td>756</td>
<td>758</td>
<td>756</td>
</tr>
<tr>
<td>743</td>
<td>507*</td>
<td>430*</td>
<td>43</td>
<td>43</td>
<td>753</td>
</tr>
<tr>
<td>760</td>
<td>754</td>
<td>745</td>
<td>745</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>741</td>
<td>744</td>
<td>720</td>
<td>753</td>
<td>753</td>
<td>745</td>
</tr>
<tr>
<td>1312</td>
<td>561*</td>
<td>752</td>
<td>750</td>
<td>759</td>
<td>759</td>
</tr>
<tr>
<td>1319</td>
<td>579</td>
<td>291</td>
<td>743</td>
<td>750</td>
<td>742</td>
</tr>
<tr>
<td>13176</td>
<td>600*</td>
<td>717</td>
<td>742</td>
<td>743</td>
<td>743</td>
</tr>
<tr>
<td>1314*</td>
<td>599*</td>
<td>743</td>
<td>751</td>
<td>752</td>
<td>752</td>
</tr>
<tr>
<td>1313*</td>
<td>745</td>
<td>191*</td>
<td>759</td>
<td>742</td>
<td>750</td>
</tr>
<tr>
<td>678</td>
<td>576*</td>
<td>57</td>
<td>752</td>
<td>751</td>
<td>191*</td>
</tr>
</tbody>
</table>

Table 2: The rankings of the proposed algorithm after the sixth and ninth iterations.

<table>
<thead>
<tr>
<th>$R_{1}^{6}$</th>
<th>$R_{2}^{6}$</th>
<th>$R_{3}^{6}$</th>
<th>$R_{1}^{9}$</th>
<th>$R_{2}^{9}$</th>
<th>$R_{3}^{9}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>748</td>
<td>748</td>
<td>748</td>
<td>748</td>
<td>748</td>
<td>748</td>
</tr>
<tr>
<td>749</td>
<td>749</td>
<td>749</td>
<td>749</td>
<td>749</td>
<td>749</td>
</tr>
<tr>
<td>741</td>
<td>741</td>
<td>741</td>
<td>741</td>
<td>741</td>
<td>741</td>
</tr>
<tr>
<td>757</td>
<td>747</td>
<td>757</td>
<td>747</td>
<td>747</td>
<td>747</td>
</tr>
<tr>
<td>747</td>
<td>757</td>
<td>747</td>
<td>757</td>
<td>757</td>
<td>757</td>
</tr>
<tr>
<td>760</td>
<td>760</td>
<td>760</td>
<td>760</td>
<td>760</td>
<td>760</td>
</tr>
<tr>
<td>746</td>
<td>746</td>
<td>746</td>
<td>746</td>
<td>746</td>
<td>746</td>
</tr>
<tr>
<td>758</td>
<td>754</td>
<td>758</td>
<td>754</td>
<td>754</td>
<td>754</td>
</tr>
<tr>
<td>754</td>
<td>758</td>
<td>754</td>
<td>758</td>
<td>758</td>
<td>758</td>
</tr>
<tr>
<td>756</td>
<td>756</td>
<td>756</td>
<td>756</td>
<td>756</td>
<td>756</td>
</tr>
<tr>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>745</td>
<td>745</td>
<td>753</td>
<td>745</td>
<td>745</td>
<td>745</td>
</tr>
<tr>
<td>753</td>
<td>753</td>
<td>745</td>
<td>753</td>
<td>753</td>
<td>753</td>
</tr>
<tr>
<td>743</td>
<td>743</td>
<td>742</td>
<td>743</td>
<td>743</td>
<td>743</td>
</tr>
<tr>
<td>750</td>
<td>759</td>
<td>743</td>
<td>742</td>
<td>742</td>
<td>742</td>
</tr>
<tr>
<td>742</td>
<td>750</td>
<td>759</td>
<td>750</td>
<td>750</td>
<td>750</td>
</tr>
<tr>
<td>759</td>
<td>742</td>
<td>750</td>
<td>759</td>
<td>759</td>
<td>759</td>
</tr>
<tr>
<td>751</td>
<td>751</td>
<td>752</td>
<td>751</td>
<td>751</td>
<td>751</td>
</tr>
<tr>
<td>752</td>
<td>752</td>
<td>751</td>
<td>752</td>
<td>752</td>
<td>752</td>
</tr>
</tbody>
</table>

Figure 9: (a) Retrieval results from the first technique ($R_{1}^{4}$). (b) Retrieval results from the second technique ($R_{2}^{4}$). (c) Retrieval results from the third technique ($R_{3}^{4}$). (d) Retrieval results from the proposed algorithm after the ninth iteration.
Table 2 shows the rankings of the proposed algorithm after the sixth and ninth iterations. From this table, it is clear that the number of irrelevant images has dropped off and that a consensus about the ranking has been reached for 50% of the images. After the ninth iteration, a 100% consensus about all the rankings has been obtained. Moreover, not only has a consensus been achieved after the ninth iteration, but an accuracy rate of 95% has also been achieved.

Figure 9 shows the retrieval results from the three selected techniques and the proposed algorithm after the ninth iteration. In Figure 9, the top left shape is the query image, and the irrelevant retrieved images are marked by cross signs. It is clear from the figure that the proposed algorithm guides the three techniques used so that they achieve better accuracy (95%) than any of the three individual techniques. The proposed algorithm achieves an accuracy of 95% with 55% accuracy and 60% accuracy.

5 Experimental Results

The lack of a standard database for evaluating the different shape-based image retrieval techniques is recognized. Researchers in this field tend to develop their own databases, which are often limited in size and/or application. The MPEG-7 developers have built a database of a reasonable size and generality [13]. This database consists of three main sets: set A, set B, and set C. Set A consists of two subsets, A1 and A2, each of which includes 420 images. A1 is used for testing scale invariance; A2 is used to test for rotation invariance. Set B consists of 1400 images that are classified into 70 classes of 20 images each. Set B is used to test for similarity-based retrieval performance and to test the shape descriptors for robustness with respect to arbitrary shape distortions, including rotation, scaling, arbitrary skew, stretching, deflection, and indentation. Because of these attributes, set B has been selected for evaluating the performance of the proposed algorithm. Set C consists of 200 affine transformed Bream fish and 1100 marine fish that are unclassified. The 200 bream fish are designated as queries. This set is used to test the shape descriptors for robustness with respect to non-rigid object distortions.

Samples of the images from this database are depicted in Figure 11.

To evaluate the performance of the different techniques with respect to image retrieval, a performance measure is required. Precision and recall are the most commonly used measures. Precision measures the retrieval accuracy, whereas recall measures the capability of retrieving relevant items from the database [16]. To evaluate the performance of the proposed algorithm, experiments were conducted using set B of the MPEG-7 database. All the contours of the images were resampled so that each contour consists of 128 points. To implement and test the newly developed algorithm, three techniques were selected.

The first technique is the Invariant Zernike Moment (IZM) technique [9], a region-based technique that provides the global characteristics of an image. Zernike moments are effective both for image representation and as a global shape descriptor [10]. This technique has been adopted by MPEG-7 as a standard region technique.

The second technique, Multi-Triangular Area Representation (MTAR) [22], is a boundary-based technique used to provide the local characteristics of an image for the proposed algorithm. This technique is designed to capture local information using triangular areas. El-Rube et al. [22] have demonstrated that, in terms of accuracy, the MTAR technique outperforms the curvature scale space (CSS) technique, which has been adopted by MPEG-7 as a standard boundary technique.

The third technique is the Fourier Descriptor (FD) technique, based on the Centroid Distance (CD) signature. According to Zhang [16], the CD signature outperforms the Complex Coordinate (CC), Triangular Centroid Area (TCA), and the Chord-Length Distance (CLD) signatures. The FD technique was selected because it is a boundary-based technique and can capture both local and global characteristics. The low frequency provides global characteristics, and the high frequency symbolizes the local characteristics of an image [16].

To fuse the three techniques, the first 50 (m=50) retrieved results from each technique are considered to be an elite candidate partition in the co-ranking stage. Figure 11 shows the precision-recall curves of the combined pair-wise conditional rankings $R_{i,j}$ after different iterations (t=0 to 9). The precision value at a specific recall is the average of the precision values of all the database shapes at that recall. From Figure 11, it is clear that the combined pair-wise conditional rankings $R_{i,j}$ have comparable recall and precision curves after the fourth iteration. Based on the experimental results, all of the combined pair-wise conditional rankings ($R_{i,j}$, $R_{i,k}$, $R_{j,k}$) reach a consensus about their rankings and yield the same precision-recall curve after the ninth iteration.

Figure 10: Samples from set B of the MPEG-7 database.
6 Conclusions

This paper presents a consensus-based fusion algorithm that improves the overall retrieval performance of a group of shape-based image retrieval techniques. A co-ranking procedure is used to obtain a consensus that improves the individual rankings. Experimental results have been presented in order to investigate the performance of the proposed algorithm using set B of the widely used MPEG-7 database.

The experimental results demonstrate that the proposed algorithm yields better results than any one of the three fused techniques. Furthermore, the performance of the proposed algorithm is superior to that produced by pooling the three individual rankings. This improvement is the result of allowing the individual retrieval techniques to recursively exchange their ranking information until all the techniques reach a consensus about their rankings. The overall reported experimental results demonstrate the effectiveness of the proposed algorithm for image retrieval applications. In future work, the complexity of the proposed algorithm will be investigated and compared to other techniques.

7 References


