Similarity retrieval of occluded shapes using wavelet-based shape features

Mejdi Trimeche[a], Faouzi Alaya Cheikh[b] and Moncef Gabbouj[b]

[a]Nokia Research Center, Visiokatu 1, Tampere, Finland
[b]Signal Processing Laboratory,Tampere University of Technology, P.O. Box 553, FIN-33101, Tampere, Finland

ABSTRACT

In this paper, we present a novel approach for describing and estimating similarity of shapes. The target application is content-based indexing and retrieval over large image databases. The shape feature vector is based on the efficient indexing of high curvature (HCP) points which are detected at different levels of resolution of the wavelet transform modulus maxima decomposition. The scale information, together with other topological information of those high curvature points are employed in a sophisticated similarity algorithm. The experimental results and comparisons show that the technique isolates efficiently similar shapes from a large database and reflects adequately the human similarity perception. The proposed algorithm also proved efficient in matching heavily occluded contours with their originals and with other shape contours in the database containing similar portions.

Keywords: Similarity retrieval, CBIR, occluded, wavelet, high curvature points, shape feature vector.

1. INTRODUCTION

The human brain is able to extract and abstract visual forms (shapes) from very complex scenes in order to recognize and understand their meaning. In fact, shape is one of the most important visual attributes in an image. For the purpose of retrieval by shape similarity in content based indexing and retrieval systems (CBIR), representations have to capture the salient perceptual aspects of a shape so that the human notion of closeness between shapes corresponds to the topological closeness in the representation space.[1] Therefore, we require that the shape descrition scheme to be:

- *invariant to simple geometric transformations (translation, rotation and scaling)*,
- *robust to noise,*
- *meaningfull in similarity retrieval algorithms,*
- *have a rich local support (robust to occlusions),*
- *cover the largest scope of forms,*
- *compress the data efficiently,*
- *be computationally feasible.*

In the case of contour-based shape representations, the high curvature points (HCP) are robust features in the sense that they provide reliable clues regarding objects under translations, rotations, scale changes, occlusions and varying backgroud levels.[2] Moreover, HCP-based descriptions reduce significantly the size of the feature vector representing the shape contour, while still keeping much of the boundary information essential for shape recognition. The techniques used to extract those points of interest can be broadly classified[3] into two categories: the ones that perform HCP's extraction at one scale and those that make use of different scales. Techniques working at just one level of resolution might suffer from finding a lot of unimportant details, while at the same time missing large rounded corners. On the other hand, multi-resolution techniques not only avoid these problems, but also provide additional information about the "structural" importance of the high curvature points.

In,[4] the curvature scale space representation (CSS) was used to extract the contour corners. However, since planar curves smoothed using Gaussian kernel suffer from shrinkage,[5] the tracking and the correct localization of high curvature points become more difficult. In contrast with the scale-space filtering approach, which serves

Further author information: (Send correspondence to Mejdi Trimeche)
Mejdi Trimeche: E-mail: mejdi.trimeche@nokia.com
Faouzi Alaya Cheikh: E-mail: faouzi@cs.tut.fi

primarily as an analysis tool, the wavelet decomposition can be used as a synthesis tool, which additionally, does not have any shrinkage problems. This makes the Wavelet decomposition very useful for detecting local features of a curve due to the spatial and frequency localization property of the wavelet bases.[6]

Hwang and Mallat[7] proved that there could not be a singularity without a local maximum of the wavelet transform at the finer scales, therefore, the WTMM (wavelet transform modulus maxima) seem to be very appropriate for the description of contours, enabling for the robust detection of high curvature (HCP) points at different resolutions.[8,9] We have used this approach for the determination of the HCP points used to generate the shape feature vector, this is explained in detail in section 2. In section 3, we define the shape feature vector used for indexing every contour in the database. In section 4, we present the matching algorithm that exploits the shape feature vector defined earlier to evaluate the similarity scores between different contours. The experimental results showing the performance of the algorithm in retrieving similar shapes from a large database are presented in section 5.

2. WAVELET-BASED HCP EXTRACTION

The Wavelet Transform (WT) decomposes a signal into a family of functions that are the translation and dilation of a unique function $\psi(x)$ whose average is zero. This function is called the mother wavelet. The WT of a signal $f(x)$ is given by

$$Wf(s,a) = f * \psi_s(a) \tag{1}$$

where $*$ denotes the convolution operator and $\psi_s(x) = \frac{1}{s}\psi(\frac{x}{s})$. The so called "dyadic scales" denote the scale that are powers of 2, that is $s = 2^j$, where $j = 1, 2, 3 \ldots$.

Mallat[7] has defined a class of basic wavelets that are regular and have compact supports. Additionally, their corresponding WT can be implemented by a fast algorithm. Mallat also proved that all singular points of a signal correspond necessarily to local maxima in the wavelet transform of that signal. This idea was adopted[8] to detect all the singularities of a contour by determing the local maxima of the wavelet transform of the contour's orientation profile. It was also confirmed[8] that the quadratic spline mother wavelet $\psi(x)$ given by equation (2) is very efficient for detecting singularities on the contour. $\psi(x)$ is the first derivative of the cubic spline smoothing function $\theta_s(x)$:

$$\psi(x) = \begin{cases} 24|x|^3/x - 16|x|^2/x & \text{if } |x| \le 0.5 \\ -8|x|^3/x + 16|x|^3/x - 8|x|/x & \text{if } 0.5 \le |x| \le 1 \\ 0 & \text{if } |x| \ge 1 \end{cases} \tag{2}$$

$$\theta(x) = \begin{cases} 8|x|^3 - 8|x|^2 & \text{if } |x| \le 0.5 \\ -8/3|x|^3 + 8|x|^2 - 8|x| + 8/3 & \text{if } 0.5 \le |x| \le 1 \\ 0 & \text{if } |x| \ge 1 \end{cases} \tag{3}$$

.

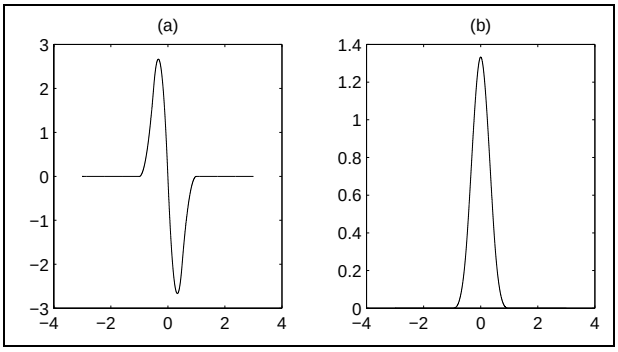

Figure 1. (a) Mother wavelet $\psi(x)$ used for HCP detection (quadratic spline), it is the derivative of the cubic spline $\theta(x)$ shown in (b).

Now, using the mother wavelet defined above, we try to detect and locate every singularity appearing on the contour $\Gamma(u) = (x(u), y(u))$, corresponds to a high curvature point (HCP). This is done by analysing the local maxima of the wavelet transform of the orientation profile $\phi(u)$:

$$W\phi(s,u) \quad = \quad \phi(u) * \psi_s(u) \tag{4}$$

where orientation function $\phi(u)$ is defined as:

$$\phi(u) = \tan^{-1}((dy(u)/du)/(dx(u)/du)). \tag{5}$$

Figure 2 shows the wavelet transform modulus maxima (WTMM) computed for dyadic scales 2^1 to 2^6. Note that as we increase the scale, the number of maxima decreases because of the smoothing process which happens as a larger filtering window is used each time. Only those maxima which correspond to important singularities survive the smoothing. For the purpose of extracting few important high curvature points that are useful for polygonal approximation, we select the WTMM maxima above a certain threshold ($T_j = 0.16$ was used) in each level at scale s^j, track them down to lower scales (to compensate for change of location across scales) and determine the exact location of these HCP on the contour. Figure 3 displays at different scales, those points that were extracted from the WTMM profile plotted in figure 2. On figure 4, the polygonal approximation using the high curvature points $(P_1, P_2, \ldots P_i, \ldots, P_n)$ extracted at level $j = 4$, is plotted against the original contour. It can be safely assumed that this approximation captures the salient properties of a shape and gives us a clear idea about its structure.

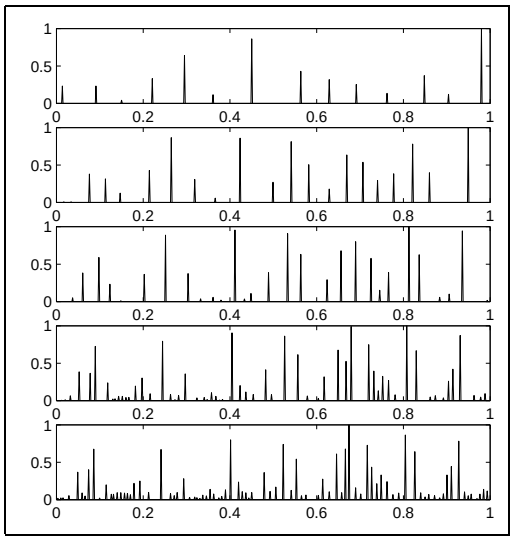

Figure 2. The wavelet transform modulus maxima (WTMM) at dyadic scales $s = 2^j, j = 2, 3, 4, 5, 6$, from bottom to top, respectively.

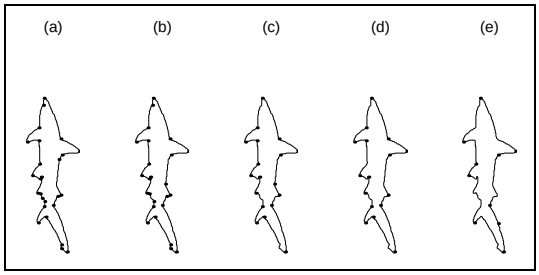

Figure 3. high curvature points detected for levels $j = 2, 3, 4, 5, 6$, from left to right, respectively.

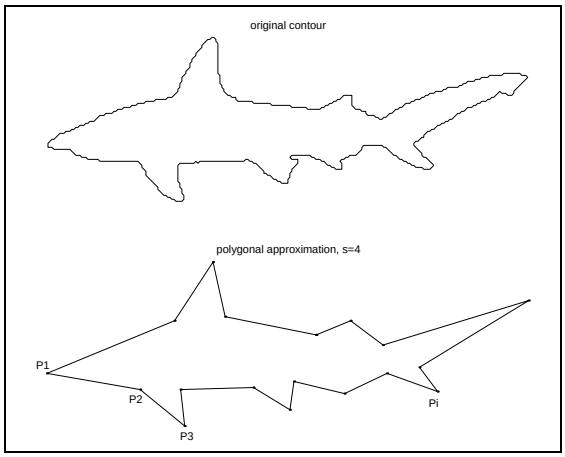

Figure 4. Polygonal approximation using high curvature points appearing at scales $s \geq 4$.

3. SHAPE FEATURE VECTOR

We used the vertices extracted at level $j = 4$, $(P_1, P_2, \ldots P_i, \ldots, P_n)$, ordered in counter clockwise direction, to represent the polygonal approximation of the shape contour. The shape feature vector was defined as:

$$F = \{(s_1, \theta_1, a_1, r_1) \ldots (s_i, \theta_i, a_i, r_i) \ldots (s_n, \theta_n, a_n, r_n)\} \tag{6}$$

where s_i is the largest scale at which the vertex P_i was detected, $(s_i \geq 2^4)$, θ_i is the angle at the vertex P_i, a_i is the logarithm of the ratio of consecutive segments and r_i is the logarithm of the ratio of the length of sides at each vertex to the overall perimeter length, (L):

$$a_i \stackrel{def}{=} \log\Big(\frac{dist(P_i, P_{i+1})}{dist(P_{i-1}, P_i)}\Big) \tag{7}$$

$$r_i \stackrel{def}{=} \log\Big(\frac{dist(P_i, P_{i+1}) + dist(P_{i-1}, P_i)}{L}\Big). \tag{8}$$

For closed contours, the feature vector F is stored in a linked list, thus when $i = 1$ in equations (7) and (8), P_{i-1} becomes P_n and P_{i+1} becomes P_1 when $i = n$. Note that a reconstruction of an approximation of the shape is possible using the shape feature vector defined above. As a matter of fact, for most applications, even the most demanding ones, we do not need the very fine details of the contour, this explains our choice to limit the verices used in the representation to only those HCP appearing at scales $s \geq 2^4$ (also $s \geq 2^3$ can be used), thus ignoring the fine details on the contour, which may be due to noise and could easily fool the matching algorithm discussed next. This also results in a significant reduction in the size of the shape feature vector. Moreover, it ensures that the shape representation is robust to noise and stable. Since the curvature function is used to extract the feature vector, translation invariance is achieved automatically. The representation is also robust to scaling and rotation. Change in starting point result in simple circular shift of the elements in the shape feature vector, this is efficiently dealt with in the matching algorithm discussed next.

4. MATCHING ALGORITHM

In this section, we present the matching algorithm that we designed and used with the shape feature vector described above. The basic idea behind the algorithm is the following: if we want to compare two polygons, we match their vertices based on their scale, angle, ratio of consecutive segments and ratio to the overall length, $(s_i, \theta_i, a_i, r_i)$ and then exploit the information about the succession of the vertices in each polygon to align the matched vertices such that the correct correspondence between the two sets of vertices is maximized. The allowed transformation from one set to another is circular shift of the feature vectors.

We will denote the feature vector of the query shape Q by

$$F^Q = \{(s_1^Q, \theta_1^Q, a_1^Q, r_1^Q) \dots (s_i^Q, \theta_i^Q, a_i^Q, r_i^Q) \dots (s_m^Q, \theta_m^Q, a_m^Q, r_m^Q)\}$$

and the feature vector of each candidate shape C by

$$F^C = \{(s_1^C, \theta_1^C, a_1^C, r_1^C) \dots (s_i^C, \theta_i^C, a_i^C, r_i^C) \dots (s_n^C, \theta_n^C, a_n^C, r_n^C)\}.$$

1. Create the match matrix M of size (m, n) such that it contains a list of all possible matches between the vertices in Q and in C. For convenience, we will always arrange M so that $m \leq n$, this can be done by simple transposition. Formally, every entry $M_{i,j}$ is calculated as follows:
   if $\{ (\theta_i^Q - \theta_j^C)_{mod(360)} \leq \Theta_{Min}) \ \& \ (|a_i^Q - a_j^C| \leq A_{Min}) \ \& \ (|r_i^Q - r_j^C| \leq R_{Min}) \ \}$
   $M_{i,j} = min(s_i^Q, s_j^C)$
   else $M_{i,j} = 0$.
   This means that if the angles at vertices P_i^Q and P_j^C are within a reasonable range (Θ_{min}) as well as the ratios of consecutive segments and those of the lengths of the corners are close enough (A_{min} and R_{min} respectively), then we assign the match value as the the minimum common scale at which the two vertices first appeared. The values of the "tolerated" deviations Θ_{min}, A_{min} and R_{min} can be adjusted to account for the size and nature of the database. That is, if our database consists of highly "similar" shapes, then those values can be made small enough to allow only tight matches, otherwise, if the database is small or contains shapes of hybrid structures then we can set those parameters to higher values and allow loose matches. If we want to account for the matching of heavily occluded shapes, then R_{min} should be set to a large value. In our experiments, we set the values of Θ_{min}, A_{min} and R_{min} to 15°, 0.45 and 0.3. With these values, we obtained good results when matching full shapes. For an enhanced performance over the matching of heavily occluded shapes, we changed the value of R_{min} to 0.9.

2. for two *identical* shapes, the match matrix M will contain a full diagonal of entries containing the scales s_i. The diagonal will be wrapped if the starting point is not the same:

$$M = \begin{bmatrix} \times & s_2 & \times & \times \\ \times & \times & s_3 & \times \\ \times & \times & \times & s_4 \\ s_1 & \times & \times & \times \end{bmatrix} \tag{9}$$

   The example above shows the match matrix M for two identical shapes with four vertices, with a shift of one vertex at the starting points. the $\times$'s denote "wrongly" matched or more usually, unmatched vertices, so as the value at $\times$ is usually 0.
   Now we want to find the best matching score among the diagonal entries of M. To simplify the approach, we construct a matrix M_D such that the "wrapped" diagonal entries of M are the columns of M_D. This is always possible since $m \leq n$. For the example matrix M in equation (9), M_D is given as

$$M_D = \begin{bmatrix} \times & s_2 & \times & \times \\ \times & s_3 & \times & \times \\ \times & s_4 & \times & \times \\ \times & s_1 & \times & \times \end{bmatrix} \tag{10}$$

3. For every column j in M_D calculate the score of match C_j of that column as follows:
   for every $i = 1 \dots n$
   if $M_D(i, j) \neq 0$ (contains a matched vertex)
   $C_j(i) = M_D(i, j)$
   else (does not contain matched vertex)
   $C_j(i) = R_i(j_{closest})/(|(j + n) - j_{closest}| + 1)$;
   where $R_i(j_{closest})$ is defined as the closest nonzero element (in both directions) in the row R_i,
   ($R_i = [M_D(i, :) \ M_D(i, :) \ M_D(i, :)]$) to the point $R_i(j + n)$. The idea here is when we do not find a match on a given vertex, we search for the consecutive or previous vertex that matches. The number of skipped vertices

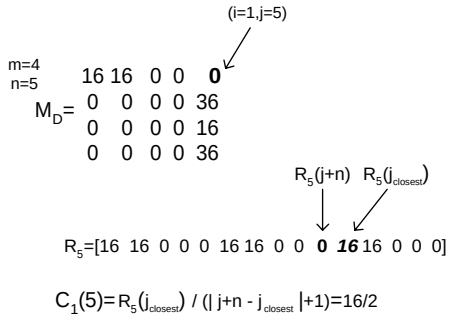

Figure 5. Calculation of $C_j(i)$

penelizes the match. Figure 5 explains more how $C_j(i)$ is calculated when $M_D(i,j) = 0$.

Finally, the matching score of each column j is calculated as:

$$C_j = \frac{\sum_{i=1}^{n} C_j(i)}{\sum_{i=1}^{m} s_i} \tag{11}$$

4. The column with the maximum score is selected as the best matched diagonal and the overall matching score will be given by:

$$Score = \max(C_j) \quad j = 1 \ldots n. \tag{12}$$

4.1. Experimental Results

We have implemented the above shape representation and matching algorithm, the software was developed using matlab©. The algorithm was tested on a database containing 1128 contours of marine animals. The feature database of all the shapes in the database was calculated offline. For each query image, we extracted the shape feature vector online and compared it to all candidate shapes in the database.

Figures 6 through 10 show the retrieval results on different fish contours. The image shown on the top left corner of each figure is the query image, its score is identically 1. The rest of the shapes are the best matches among all the shapes in the database. From these figures, it can be safely concluded that the retrieval results reflected similarity of shapes very sharply: most of the first matches are similar to the query image and their ranking according to their matching score "makes sense". This shows that a concise combination of the location and scale information in the matching algorithm can show very efficient in reflecting similarity of shapes. The system was insensitive to translations, rotations and scaling. Query images consisting of rotated or scaled fish contours always resulted in retrieving the same images as the those resulting from the search using the original images. The matching algorithm performed particularly well on heavily occluded shapes. In Figure 9, the query image represents a fish tail (more than 50% occlusion). The first retrieved image was the original fish to which the tail belonged, then the rest of the retrieved images contained tails that were similar to the query.

We have plotted in Figure 11 the normalized simlarity scores for the same query image achieved by the Modified Fourier Descriptors method,[10] the CSS technique[11,12] and by our technique. The scores were normalized between 0 and 1. The best match (not the exact match) was given the score 1. The scores performed by our algorithm had the steepest descent, which means that they isolated most efficiently similar shapes from the rest of the database, hence confirming the usefulness of our shape representation and matching algorithm in very large databases.

ACKNOWLEDGMENTS

We wish to thank Sadegh Abbasi and Farzin Mokhtarian[11] for making their fish database publicly available to the research community.

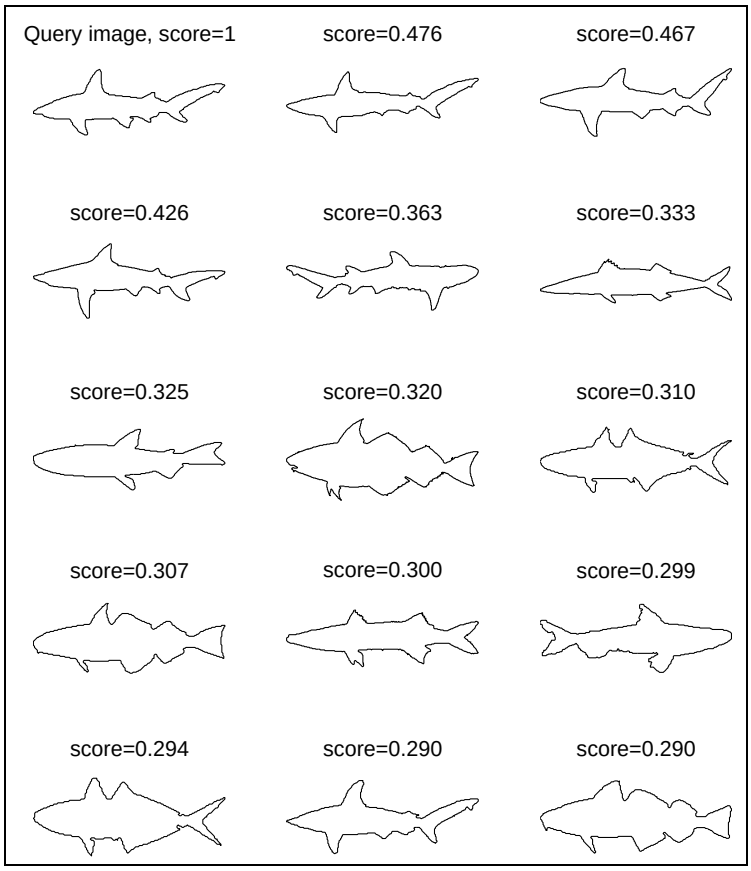

Figure 6. Example 1 of search results using WTMM-based technique.

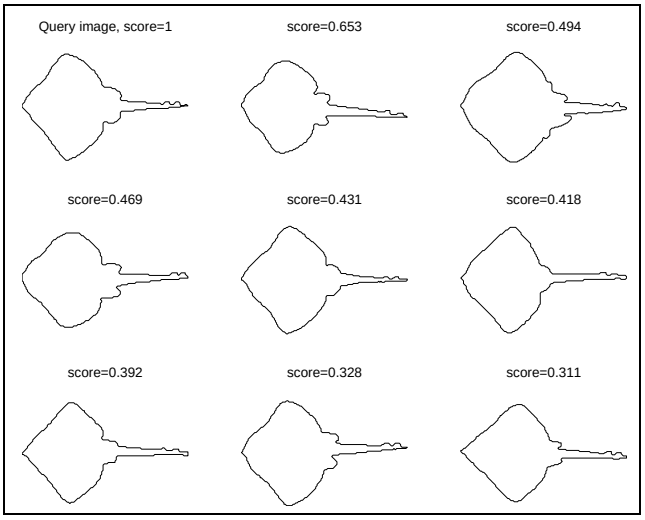

Figure 7. Example 2 of search results using WTMM-based technique.

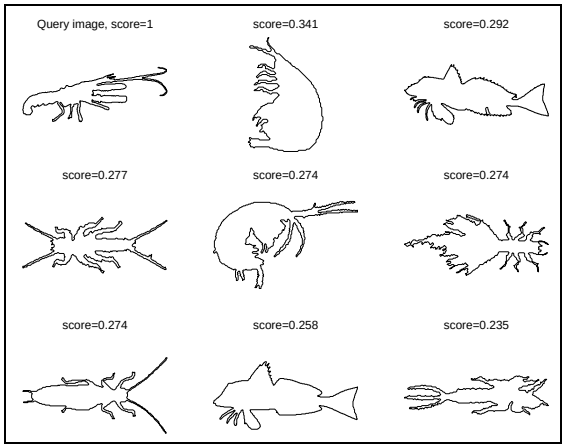

Figure 8. Example 3 of search results over a contour containing very sharp HCP's, most appearing at high scales ($s \geq 4$).

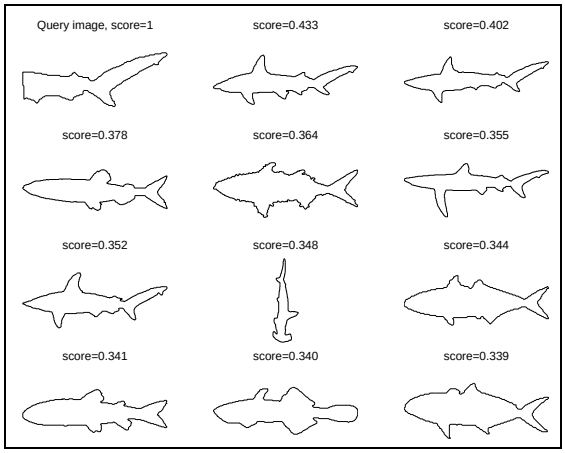

Figure 9. Example 4 of search results over a heavily occluded contour.

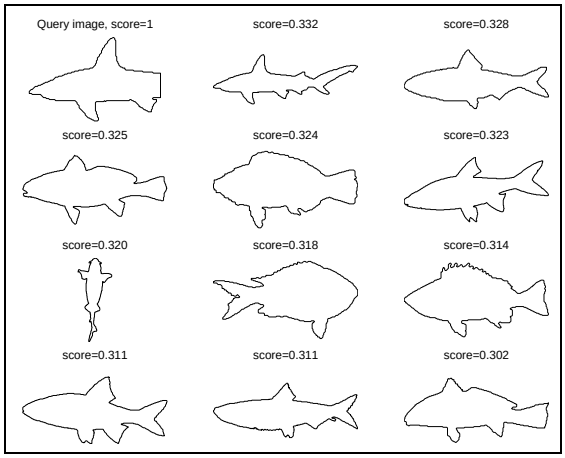

Figure 10. Example 5 of search results on a heavily occluded contour.

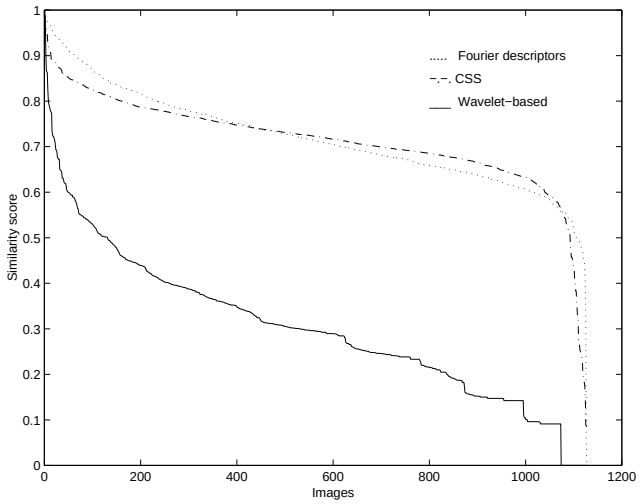

Figure 11. Plot of the ordered similarity scores for the Fourier disriptors method, CSS technique and our technique, respectively.

REFERENCES

1. A. Del Bimbo, *Visual Information Retrieval*, Academic Press; Morgan Kaufmann Publishers, June, 1999.
2. M. H. Han and D. Jang, " The use of maximum curvature points for the recognition of patially occluded objects," in *IEEE Transactions on Pattern Analysis and Machine Intellegence* Vol.23, pp. 21–23, 1990.
3. C. Fermuller and W. Kropatsch, "Multi-resolution shape description by corners", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 271–276, Champaign, IL, USA, 1992.
4. F. Mokhtarian and R. Suomela, "Robust Image Corner Detection Through Curvature Scale Space", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 20, No.12, pp. 1376–1381, December 1998.
5. T. Lindberg, "Scale-Space for Discrete Signals", *IEEE Transactions on Pattern Analysis and Machine Vision*, Vol. 12, No. 3, pp. 234–254, 1990.
6. G. C.-H. Chuang, C.-C. J. Kuo, "Wavelet Descriptor of Planar Curves : Theory and Applications", *IEEE Transactions on Image Processing*, Vol. 5, No. 1, pp. 56–70, May 1996.
7. S. Mallat, *A Wavelet Tour of Signal Processing, 2nd Edition*, Academic Press; 1999.
8. A. Quddus M. Fahmy, "Fast Wavelet-based Corner Detection Technique", *Electronics Letters*, Vol. 35, pp. 287–288, February 1999.
9. J. Lee, Y. Sun and C. Chen, "Multiscale Corner Detection by Using Wavelet Transform", *IEEE Transactions on Image Processing*, Vol. 4, No.1, pp. 100–104, January 1995.
10. Y. Rui, A. C. She and T. S. Huang, "Modified Fourier Descriptors for Shape Representation", *Proc. of First International Workshop on Image Databases and Multi Media Search*, Amsterdam, August, 1996.
11. F. Mokhtarian, S. Abbasi and J. Kittler, "Robust and Efficient Shape Indexing through Curvature Scale Space," *Proc. British Machine Vision Conference*, pp. 53–62, Edinburgh, UK, 1996.
12. F. Mokhtarian and A. K. Mackworth, "Scale-based Description and Recognition for Planar Curves and Two-Dimentional Shapes", *IEEE Transactions on Pattern Analysis and Machine Intellegence*, Vol. 8, No. 1, pp. 34-43, 1986.