

# Shape Recognition based on Wavelet-Transform

## Modulus Maxima

Faouzi Alaya Cheikh, Azhar Quddus and Moncef Gabbouj

*Tampere University of Technology (TUT),  
Signal Processing Laboratory,  
P.O. Box 553, FIN-33101  
Tampere, Finland  
{faouzi, azhar, moncef}@cs.tut.fi*

**ABSTRACT:** In this paper we propose a new approach to object recognition based on the polygonal approximation of the object contour. The vertices of the polygon are the high curvature points of the contour, selected using the wavelet transform modulus maxima. We associate with this shape description a simple measure to estimate the similarity between objects. The description scheme and the similarity measure proposed take into consideration the way the human visual system perceives objects and compares them.

The description scheme and the similarity measure can be used in several applications such as content-based indexing and retrieval, evaluation of segmentation results in the context of MPEG-4, or in classical classification problems. The proposed scheme is invariant to translation, rotation, scale change and noise corruption. Moreover, this description scheme allows accurate reconstruction of the shape boundary from the feature vector used to describe it. The experimental results and comparisons show the performance of the proposed technique.

### 1. INTRODUCTION

Most people assume that what we see is exactly what our eyes see and report to our brain. This is not quite true in fact, our brain adds very substantially to the information it gets from the eye. It is even more interesting to know that the eye throws away much of the information it gets, leaving it to the brain to fill in additional information in its own way [1]. This capability is hard-wired into our retinas. Connected directly to the rods and cones of the retina are two layers of neurons that perform an operation similar to the Laplacian. This operation is called lateral inhibition and helps us to extract boundaries and edges [2]. Therefore, in this paper we opted for a boundary-based representation of shapes rather than a region-based one.

The object representation we are proposing in this paper simplifies the complex shape boundary using

polygonal approximation. Hoffman et al. [3] argued that when the human visual system decomposes objects it does so at points of high negative curvature. Therefore, approximating curves by straight lines joining these high curvature points (HCP) retain the maximal amount of information necessary for successful shape recognition. This can be explained by the fact that our visual system focuses on the singularities and ignores smooth curves thanks to the lateral inhibition mechanism.

Therefore, in the case of shape feature, high curvature points are robust descriptors in the sense that they are invariant under translation, rotation and scale change [4]. They provide reliable clues regarding objects even under occlusion and varying background levels [5]. Moreover, corner-based representation of object boundary reduces significantly the size of the feature vector, while still keeping much of the boundary information essential to object recognition [2]. Therefore, object recognition techniques based on corner point matching have been used in machine vision applications [5]. It can be seen, however, that the complexity of the algorithms proposed in [5] increases exponentially as the number of candidate objects increases. Therefore, these techniques are not suitable for applications such as retrieval in large image databases where online processing of thousands of images is required.

Hwang and Mallat [6] proved that there couldn't be a singularity in a signal without a local maximum of its wavelet transform at the finer scales. Therefore, the WTMM of the orientation profile of a given object contour seem to be very appropriate for the selection of the vertices for the polygonal approximation of the contour. Moreover, WTMM-based descriptors, unlike global contour descriptors such as the Fourier descriptors, provide precise local shape information. In this paper the importance of high curvature points and

their locations are estimated directly from the WTMM of the orientation profile of the contour.

Once the objects in the database are approximated by polygons, their similarity is measured by computing the Euclidean distance between their feature vectors. Each feature vector is composed of pairs of the magnitude of the WT and its location on the normalized contour, computed at each vertex of the polygonal approximation.

This paper is organized as follows: In section 2 the proposed method is described. Section 3 presents the shape recognition algorithm, the features used, the similarity measure and their properties. Experimental results are presented in section 4, and conclusions are drawn in section 5.

## 2. PROPOSED METHOD

Wavelet decomposition provides natural setting for the multi-level image contour analysis. Since wavelet transform modulus maxima (WTMM), provide useful information for curvature analysis, we propose to use it here for fast feature extraction. It has been shown in [7] that biquadratic wavelets, proposed by Mallat and Zhong [8], perform better than other wavelets for corner detection applications; therefore we will use it here.

The boundary is tracked and its orientation profile is computed as in [9]. The orientation profile of each one of the two shapes is upsampled and interpolated in a way to have the same number of points for each contour. The Wavelet transform of the orientation profile is computed for dyadic scales from  $2^1$  to  $2^6$ . WTMM are then computed, see Figure 1, and only those WTMMs larger than a certain threshold are considered important singularities (high curvature points). The shapes will be approximated by their polygonal approximation built from these high curvature points. And the vector feature describing a given shape will contain information extracted at each of the HCPs from the WT. In this paper we have used the most elementary information from the WT: specifically the location of the HCP on the normalized contour and the magnitude of the WT at that point. Similarity scores are then estimated at each level of the decomposition independently. And the overall similarity measure is computed as the maximum value of the single level similarity scores.

Speeding up the search in online applications, where processing large data sets is required in a limited time, such as content-based retrieval systems [11], is becoming a crucial factor in the usability of such systems. Therefore, fast schemes for narrowing down the search space are essential in content-based retrieval systems due to the size of the data sets under consideration. Here we used the aspect ratio  $\gamma$  to reduce

the search space. Images with error on  $\gamma$  larger than a fixed threshold are discarded. And the remaining candidates go through the second step of the retrieval process, where for each candidate image the set of low-level features is extracted at the high curvature points and compared to those of the query image.

## 3. SHAPE RECOGNITION ALGORITHM

1. Select candidate objects with aspect ratios similar to the one of the query object,
2. Consider only WTMM greater than certain threshold  $T_{WTMM}$  as vertices for the polygonal approximation,
3. Create a feature vector, for each object in the database, containing the location and magnitude of the WT at each one of the HCP,
4. Compute a similarity score, at each level, between the of query image and each of the candidate images,
5. The final similarity score is computed as the maximum of the scores at each level.

### 3.1 Feature Extraction

At each wavelet decomposition level  $l$ , we characterize the query image contour with two vectors,  $M_l^q$  and  $L_l^q$ , where,  $M_l^q = [m_{l1}^q, m_{l2}^q, \dots, m_{lm}^q]$  contains the  $m$  magnitudes of the WTMM and  $L_l^q = [p_{l1}^q, p_{l2}^q, \dots, p_{lm}^q]$  the  $m$  locations of the high curvature points on the normalized contour. Similarly, at each decomposition level  $l$ , the candidate image contour is characterized with two vectors,  $M_l^c$  and  $L_l^c$  of length  $n$ . Before starting the matching, circular shift is applied to the feature vector of the candidate object in a way to have its HCP associated with the largest WTMM, aligned with the HCP corresponding to the largest WTMM of the query image. Which means, that we start the matching process from the boundary points having the highest curvature at each decomposition level. This circular shift makes our description invariant to starting point change and to rotation of the objects.

### 3.2 Similarity Estimation

In this step we try to find correspondence between vertices from the polygonal approximation of the query and the candidate objects. A correspondence between two high curvature points ( $P_Q$  and  $P_c$  from the query image and a candidate image respectively) is found if

the Euclidean distance between their corresponding feature vectors is under a fixed threshold  $T$ . Let

$V_{i_0} = (m_{i_0}^q, p_{i_0}^q)$  be the feature vector extracted at the point  $P_Q$  and  $V_i = (m_i^c, p_i^c)$ ,

$p_i^c \in [p_{i_0}^q - \Delta p, p_{i_0}^q + \Delta p]$  be the feature vectors extracted at the candidate high curvature points located within the neighborhood centered around  $p_{i_0}^q$ . A valid match between the query image high curvature point  $P_Q$  and a high curvature point  $P_c$  from a candidate object contour is found if  $\exists V_i = (m_i^c, p_i^c)$ ,

$p_i^c \in [p_{i_0}^q - \Delta p, p_{i_0}^q + \Delta p]$  verifying that the distance  $\|V_{i_0} - V_i\|_2 \leq T$ . Therefore, several matches can be obtained for each high curvature point of the query contour. Only the best matching points are used for the estimation of the similarity between two shapes. The best matching points are those corresponding to local minima of the Euclidean distance mentioned before:

$$p_i^c, i = \arg \min_{p_i^c \in [p_{i_0}^q - \Delta p, p_{i_0}^q + \Delta p]} \|V_{i_0} - V_i\|_2.$$

Let  $K$  be the number of matched maxima. The similarity score at level  $l$ , for  $l = j, \dots, 6$ , is

$$\text{computed as: } s_l = \left( \frac{2 \times (K - \xi)}{m + n} \right) \times 100,$$

$$\text{where, } \xi = \sum_{i=1}^K \left( \frac{\delta m_i}{\text{mean}(m_{i_0}^q, m_{i_0}^c)} + 2 \times |\delta p_i| \right), \quad \delta m_i$$

and  $\delta p_i$  are the errors on the magnitude and the normalized position for the  $i^{\text{th}}$  matched maxima.  $\xi$  gives an idea on how good the correspondence is between the two sets of points.

The lower levels are not considered in the matching process in order to make our measure unaffected by noise presence. The overall similarity score is  $S = \max(s_l)$ , for  $l \in \{4, 5, 6\}$ .

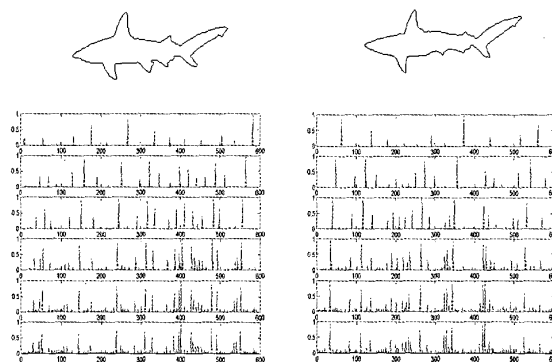
This approach is similar to the CSS technique proposed by Abbasi et al. [5], however our approach is more effective since the exact location of the HCP is determined with high precision by tracking the WTMM through the decomposition levels until the original contour. Moreover it is faster, since few dyadic scales are used, unlike the CSS where the full continuous scale decomposition is required. Moreover, by considering high curvature points only, just the visually most significant details are used to estimate the similarity of two contours and redundant information is discarded by

ignoring smooth curves. The proposed technique preserves most of the shape information, since the object contour can be accurately reconstructed from its WTMM [9].

### 3.3 Properties of the proposed scheme

The proposed description scheme is invariant to:

- **Translation:** since the orientation profile of the contour is independent of the location.
- **Rotation & starting point change:** the rotation of the shape or the change of the starting point induces circular shift of the maxima of the wavelet decomposition. Therefore, applying a circular shift to the feature vector until the largest maxima of the query and candidate shapes coincide solves these problems.
- **Scale:** achieved by using the same number of points to represent the orientation profile and the normalizing of the contour length to 1.
- **Noise corruption:** achieved by considering only the WTMM larger than a fixed threshold.
- **Flipping:** the problem of mirrored images can be solved by replacing the maxima position with the relative position from the largest maxima, however this may create problems if we have partial symmetry around the largest maxima.
- **Partial occlusion:** the proposed similarity measure is not very sensitive to the appearance or elimination of few HCP; therefore it is robust to partial occlusions.



**Figure 1.** Two fish contours and their corresponding WTMM.

## 4. EXPERIMENTAL RESULTS

In these experiments we used 1130 fish contour images, from the University of Surrey, UK and 1400

from the MPEG-7 CE1 test set B. The boundary of each fish in the database is represented by a sequence of 1000 points. Table 1 presents the results of a query with the fish in its first cell. Where, the five most similar images retrieved using the CSS algorithm, the matching results of human users and the proposed algorithm are shown. Three thresholds were used in these experiments,  $T_{WTMM}$ ,  $T_M$  and  $\Delta p$ .  $T_{WTMM} = 0.6$ , distinguishes high curvature points from insignificant singularities.  $\Delta p = 0.2$  and  $T_M = 0.2$  are the tolerated errors on the location of the WTMM and its magnitude. 5% error on the aspect ratio  $\gamma$  was allowed.

The CSS results shown in Table 1 were obtained from the web-based demo available at: "http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/dbase2.html".

From Table 1, one can easily see that the proposed algorithm often produces results, which are close to those selected by the human users. The produced results are clearly better than those obtained by CSS, (MPEG-7 has an optimized version of the CSS algorithm that we are planning to compare our results against in the future work). Figure 2 shows the 20 first images retrieved by our algorithm from the MPEG-7 test set B consisting of 1400 images (20 similar images in each category) of the Core Experimental data CE1, of the MPRG-7 group, when image "rat-01" is used as query.

### 5. CONCLUSIONS

In this paper we have proposed a new technique for object shape recognition based on a vector feature extracted from the WT of the orientation profile of a

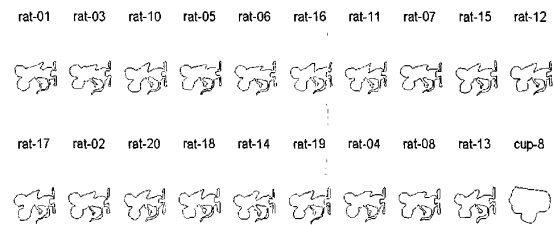
**Table 1.** Retrieval results obtained by the CSS, Human observers and our algorithm.

CSS	Human	Proposed algorithm

given object boundary and a simple similarity measure. The proposed algorithm is simple, fast and has invariance properties that make it very appropriate for applications such as object recognition and content-based retrieval in large databases.

### 6. REFERENCES

[1] Goldstein, E. B., "Sensation and Perception," 5th Edition University of Pittsburgh Published by Wadsworth Publishing, CB © 1999.  
 [2] Russ, J. C., "The Image Processing Handbook", 3rd edition, CRC, Springer and IEEE Press inc., 1995.  
 [3] Hoffman, D. D., and Richards, W. A., "Parts of recognition", *Cognition*, Vol. 18, pp. 65-96, 1984.  
 [4] Teh, C.H. and Chin, R. T., "On the detection of dominant points on digital curves", *IEEE Trans. PAMI*, vol. 11, pp. 859-872, 1989.  
 [5] Han, M. H. and Jang, D., "The use of maximum curvature points for the recognition of partially occluded objects", *Pattern Recognition*, vol. 23, pp. 21-23, 1990.  
 [6] Mallat, S., "A Wavelet Tour of Signal Processing", 2nd Edition, Academic Press, 1999.  
 [7] Quddus, A. and Fahmy, M. M., "Corner detection using various wavelets", accepted for the ICICS'99, December 1999, Singapore.  
 [8] Mallat, S., G. and Zhong, S., "Characterization of signals from multiscale edges", *IEEE Trans. on PAMI*, Vol. 14, pp. 710-732, 1992.  
 [9] Lee, J. S., Sun, Y. N. and Chen, C. H., "Multiscale corner detection by wavelet transform", *IEEE Trans. on Image Processing*, Vol. 4, pp. 100-104, 1995.  
 [10] Abbasi, S., F. Mokhtarian, and J. Kittler, "Curvature Scale Space image in Shape Similarity Retrieval," *Springer Journal of Multimedia Systems*, 1999.  
 [11] Alaya Cheikh, F., Cramariuc, B., Reynaud, C., Quinghong, M., Dragos-Adrian, B., Hnich, B., Gabbouj, M., Kerminen, P., Mäkinen, T. and Jaakkola, H., "MUVIS: A System for Content-Based Indexing and Retrieval in Large Image Databases," *Proceedings of the SPIE/EI'99 Conference on Storage and Retrieval for Image and Video Databases VII*, Vol. 3656, pp.98-106, San Jose, California, 26-29 January 1999.



**Figure 2.** Sample retrieval results from the MPEG-7 CE Shape/Motion test set B.