Shape Matching Using Wavelet Transform Modulus Maxima

Xiaojun Qi
Computer Science Department, Utah State University
4205 Old Main Hill, Logan, UT 84322-4205
Xiaojun.Qi@usu.edu

Abstract
Object contour matching has been prevalently used in pattern recognition and content-based indexing and retrieval of digital images. In this paper, the contour descriptors are extracted from a WTMM (Wavelet Transform Modulus Maxima) image, which is produced by a one level biorthogonal wavelet transform and contains curvature points of the contour. The most important WTMMs are further utilized to find the correspondences between points on the two shapes. In order to solve the correspondence problem, the centroid of the most important WTMMs and the line across the centroid and the point furthest away from the centroid are used respectively as the reference point and the reference line. Angle descriptors are constructed between the reference line and the lines connecting each important WTMM point and the reference point. The modulus maxima magnitudes, the angle descriptors, and the normalized distances from each important WTMM point to the centroid are used to construct the feature vector. Both the correspondence matching percentage and the similarity score are derived from the feature vector. They are invariant to the translation, rotation, and scale changes. The experimental results have demonstrated that the proposed approach can accurately and efficiently match the object contour under these three changes.

Keywords
Object contour, high curvature points, biorthogonal wavelet transform, wavelet transform modulus maxima, feature vector.

I Introduction

Content-based indexing and retrieval (CBIR) of digital images has become an active research area since the early 1990s. Most of these CBIR systems support the following one or several functionalities: browse, search by example, and search by a single or a combination of low level features (e.g., color, contour, texture, and spatial layout of objects) extracted from the image.

Hoffman *et al.* [1] propose to decompose objects at HCPs (High Curvature Points) and connect these HCPs via straight lines to approximate curves. This HCP-based approximation is proven to retain the maximal amount of information necessary for successful contour recognition. Teh and Chin [2] prove that HCP-based features are robust for recognizing objects due to their invariance to the translation, rotation, and scale changes. Russ [3] applies the corner-based representation of objects to reduce the size of the feature vector for recognizing objects. Koch and Kashyap [4] and Han and Jiang [5] independently propose an HCP-based approach to provide reliable clues about the objects under partial occlusion and varying background levels. Cheikh *et. al.* [6] present a robust wavelet-based shape recognition algorithm to construct two feature vectors to match the shapes. These two feature vectors contain the magnitude (importance) and the locations of the HCPs, which are estimated directly from the WTMM image at dyadic scales from 2^1 to 2^6. Khalil *et. al* [7] study the object recognition invariant functions calculated by different dyadic scales of wavelet transform. Belongie *et. al.* [8] propose a shape-context-based approach to measuring similarity between shapes and exploit it for object recognition. In general, either the complexity of the above algorithms increases exponentially as the number of candidate objects increases or the algorithms have long feature vectors at each dyadic scale for approximate comparison. Therefore, these techniques are not suitable for large image databases with thousands of images.

In this paper, we propose an efficient biorthogonal-wavelet-transform-based approach to describe the contour features of a given object. These contour features are represented by the outer boundary of the shape instead of the region encompassed by the shape. This type of shape representation precisely simulates the object detection process coordinated between our eyes and the brain. That is, our eyes pass the important information to our brain. This important information contains boundaries and edges extracted by two layers of neurons using an operation called lateral inhibition, which behaves like the highpass Laplacian operator [3]. Our brain then automatically fills in additional information to make an object detection decision [9]. Since we focus on a boundary-based representation of shapes rather than a region-based one, our digital image data exclusively contains one non-occluded object boundary per image.

The remainder is organized as follows:

- Section II briefly introduces the concept of the wavelet transformation maxima and its potential application.
- Section III proposes a practical implementation of biorthogonal-wavelet-based shape recognition.
- Section IV illustrates the experimental results of the shape recognition.
- Section V draws conclusions.

II Wavelet-Based Feature Extraction for Contour Analysis

Wavelet transforms can decompose images into elementary building blocks that are well localized both in space and frequency. This decomposition provides a natural approach for the multi-level image contour analysis. Hwang and Mallat [10] prove a local maximum of the wavelet transform (i.e., a strict local maximum of the modulus on either its right or left side) at the finer scales corresponds to a singularity. This singularity is a measure of the local regularity of the intensities and precisely identifies discontinuities in the intensities in images. Therefore, wavelet transforms are particularly useful in recognition of edges, boundaries and important features. The WTMM can further be utilized for contour description in the curvature analysis. The WTMM-based descriptors, unlike Fourier transformation-based global contour descriptors, provide precise local shape information [11].

III Object Contour Matching Method

This section details the step-by-step procedure to match object contours.

Step 1: Apply a one-level biorthogonal wavelet transform to each query/candidate image
To date, we have used the biorthogonal wavelet transform. The major reason to choose the biorthogonal wavelet transform is that the symmetry of the wavelets and scaling functions is maintained via biorthogonal wavelet bases by relaxing the orthogonality constraint and associating with short length filters [12]. The level one decomposition is chosen for contour identifying since our image contains non-occluded object boundaries with little or no noise involved. Moreover, such decomposition has the finest scale for fast computations as well as the least memory requirements.

Step 2: Construct a WTMM (Wavelet Transform Modulus Maxima) image for each query/candidate image
Identify the modulus maxima points at the low-high and high-low subbands, which respectively correspond to the wavelet decomposition of the object contour at the horizontal and vertical direction with the highest frequencies. The WTMM image is obtained by combining the identified modulus maxima points at the low-high and high-low subbands.

Step 3: Set up a threshold to determine the most important modulus maxima points (i.e., HCPs) for each query/candidate image
The threshold is empirically determined to be the one-tenth the maximum magnitude value of the WTMM image. Any point with a value greater than the determined threshold corresponds to sharper edges (salient features) in the original image. Therefore, such a point is considered as a HCP in the WTMM image and kept for further processing.

Step 4: Construct a feature vector for each query/candidate image
The contour centroid and the straight line for connecting the contour centroid and the furthest away HCP (This HCP has the longest distance to the centroid) are respectively considered as the reference point and the reference line. The feature vector contains the magnitude of each HCP, the normalized distance from each HCP to the reference point, and the angle between the reference line and the line across each HCP and the reference point.

Step 5: Compute the correspondence matching percentage and the similarity score between the query image and each candidate image.

Align the reference point and the reference line of the query image and each candidate image so the HCP with the longest distance to the centroid is aligned as well. A valid match between two HCPs is found if the differences between the relative angles with the reference line and the differences between the normalized distances to the reference point are under the thresholds of T_{ang} and $T_{dis,}$ respectively. The correspondence matching percentage is calculated as the ratio of the number of matching points (i.e., the points satisfy the two threshold conditions) and the number of the most important modulus maxima points (HCPs). The similarity score is calculated as:

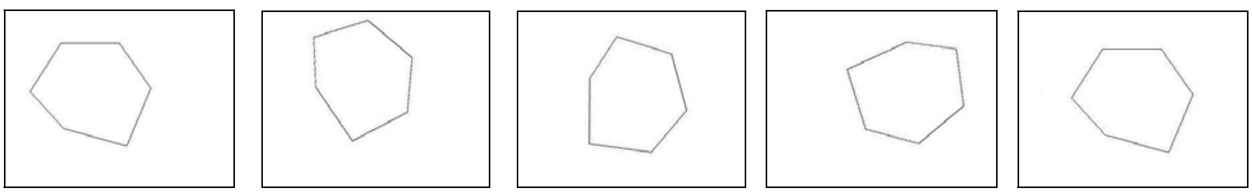
$$Value = \frac{\sum_{i=1}^{MatchNum}\left(abs(NDis_q(i) - NDis_c(i))\right)}{\sum_{i=1}^{MatchNum}\left(\frac{Magnitude_q(i) + Magnitude_c(i)}{2}\right)} + 2\times \sum_{i=1}^{MatchNum}\left(abs(RAngle_q(i) - RAngle_c(i))\right)$$

$$Score = \frac{(MatchNum - Value)}{MatchNum} \times 100.$$

where *MatchNum* is the total number of matching points, *NDis* is the normalized distance between each matching HCP and the reference point, *RAngle* is the relative angle between each matching HCP and the reference line, and *Magnitude* is the magnitude of each matching HCP in the WTMM image. *Value* measures how good the correspondence is between the two sets of HCPs by calculating the total differences between the matched normalized distance and the matched relative angle. *Score* measures the similarity between the two shapes with the maximum value of 100 indicating a perfect match.

IV Experimental Results

We have divided the experimental images into two groups. The first group is formed by geometrical drawn objects that change their location, size, and angle in an image. The second group consists of fish contour images [13].



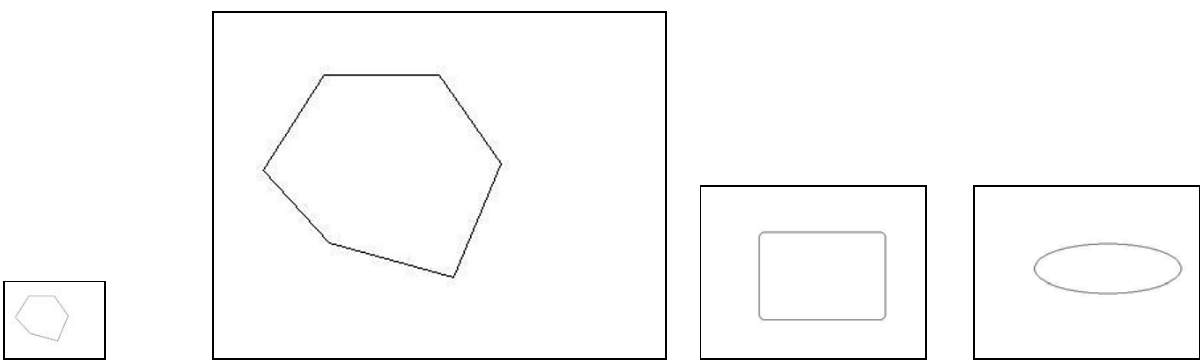
(a) Query image

(b) Similar candidate 1 -- Rotate 40° MP: 97.62% SS: 98.22

(c) Similar candidate 2 -- Rotate 75° MP: 96.36% SS: 98.04

(d) Similar candidate 3 -- Rotate 140° MP: 95.87% SS: 97.10

(e) Similar candidate 4 --Change Location MP: 95.12% SS: 97.06



(f) Similar candidate 5 -- Shrink to 1/3 MP: 98.02% SS: 98.44

(g) Similar candidate 6 -- Enlarge to 1.5 MP: 98.17% SS: 98.92

(h) Dissimilar candidate 1 MP: 51.63% SS: 65.01

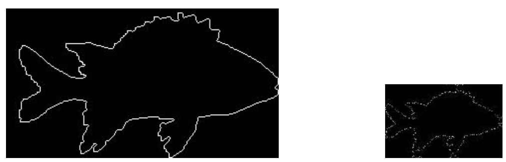
(i) Dissimilar candidate 2 MP: 47.75% SS: 58.19

Fig. 1: Objects in the First Group

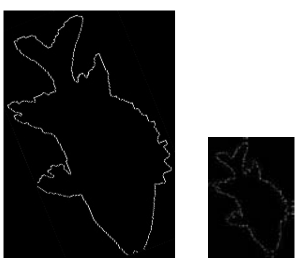
In Fig. 1, the query image Fig. 1(a) contains an object with a polygonal contour. The first six candidate images (Fig. 1(b), Fig. 1(c), Fig. 1(d), Fig. 1(e), Fig. 1(f), and Fig. 1(g)) contain an object with the same polygonal contour but either being rotated by 40°, 75°, or 140°, being located at different positions, or being scaled by a factor of 1/3 or 1.5. The last two candidate images (Fig. 1(h) and Fig. 1(i)) contain an object with different contours. The matching percentage and the similarity score for each image are listed below each image, where MP indicates matching percentage and SS indicates similarity score. Here the higher matching percentage and higher similarity score indicate the better shape matching.

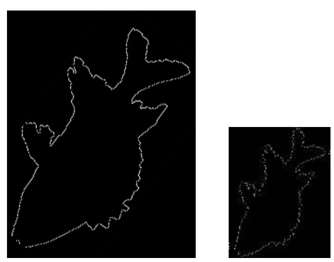
We observe that the first six similar candidate images have higher matching percentage and higher similarity score than the last two dissimilar candidate images when compared with the query image in Fig. 1(a). For instance, when query image is compared with the similar candidate image 2, which is being rotated 75° and translated to the center of the image, the matching percentage is 96.36% and the similarity score is 98.04. These two values are substantially larger than the values obtained when comparing the query image with the dissimilar image 1 as shown in Fig. 1(h). Therefore, the experimental results from the first group demonstrate our proposed feature-vector-based matching measurement (i.e., matching percentage and similarity score) is invariant to the changes in the translation, scale, and rotation.

Figure 2 shows several images from the second group. These images include one original image from the fish contour image database [13] and the corresponding rotated, translated, or scaled images. The WTMM images with the HCPs obtained by our proposed approach are illustrated at the right side of the corresponding original images. The feature vector with three elements (i.e., the magnitude of each HCP, the normalized distance from each HCP to the reference point, and the angle between the reference line and the line across each HCP and the reference point) for each HCP is constructed to calculate the matching percentage and the similarity score based on the conditions of two thresholds T_{ang} and $T_{dis.}$ These two measurements are listed below each candidate image to indicate the similarity with the original query image.



(a) Query Image and its WTMM image with the HCP





(b) Rotated Translated Image (68°) (c) Rotated Translated Image (126°) (d) Shrunk Image (1/4)
and Their WTMM Images

(b)	(c)	(d)
MP: 96.37%	MP: 96.13%	MP: 97.96%
SS: 97.07	SS: 97.03	SS: 98.38

Fig. 2: Query Image, Rotated Image, Shrunk Image, and their Corresponding WTMM Images with the HCPs

The experimental results from the second group indicate the proposed approach can precisely match object contours under any geometrical operations. As demonstrated by our experimental results, the accuracy for these two matches is high as well.

Fig. 3 shows one example of the retrieval results by using the proposed approach. The top three matched fish images of the query fish illustrated in Fig. 2(a) from the fish contour database are shown in Fig. 3. The results are comparable with those selected by the users. The two measurements are listed below each candidate image.

MP: 97.28% SS: 98.99

MP: 96.08% SS: 97.04

MP: 95.91% SS: 96.89

Fig. 3: Retrieval results by using the query image shown in Fig. 2(a)

V. Conclusions

We propose a fast and robust matching algorithm for contour images by using a biorthogonal wavelet transform. The derived WTMM image contains only the HCPs, which represent the most salient features in the contour. A feature vector is constructed by three elements, i.e., the magnitude of each HCP, the normalized distance from each HCP to the reference point, and the angle between the reference line and the line across each HCP and the reference point. The matching percentage and the similarity score are calculated by the matching feature vectors. This technique is invariant to the translation, rotation, and scale changes. It is very efficient to match the images within a large database due to the easy construction of the feature vector at only one dyadic scale of a wavelet transform.

VI. Acknowledgement

Author acknowledges professors S. Abbasi, F. Mokhtarian, and J. Kittler for providing the test image databases.

References

1. 1. D. D. Hoffman, and W. A. Richards, "Parts of recognition," *Cognition*, Vol. 18, pp. 65-96, 1984.
2. 2. C. H. Teh and R. T. Chin, "On the detection of dominant points on digital curves," *IEEE Trans. PAMI*, vol. 11, pp. 859-872, 1989.
3. J. C. Russ, The Image Processing Handbook, 3rd edition, *CRC, Springer and IEEE Press Inc.*, 1995.
4. M. W. Koch and R. L. Kashyap, "Using polygon to recognize and locate partially occluded objects," *IEEE Trans. PAMI*, vol. 9, pp. 483-494, 1987.
5. M. H. Han and D. Jiang, "The use of maximum curvature points for the recognition of partially occluded objects," *Pattern Recognition*, Vol. 23, pp. 21-23, 1990.
6. F. A. Cheikh, A Quddus, and M. Gabbouj, "Multi-level shape recognition based on wavelet-transform modulus maxima," *Proc. of Image Analysis and Interpretation*, pp. 8-12, April 2-4, 2000.
7. M. Khalil and M. Bayoumi, "Affine invariants for object recognition using the wavelet transform," *Pattern Recognition Letters*, Vol. 23, Issue1-3, pp. 57-72, 2002.
8. S. Belongie , J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. PAMI*, Vol. 24, No. 24, pp. 509-522, 2002.
9. E. B. Goldstein, "Sensation and Perception," 5th Edition University of Pittsburgh Published by Wadsworth Publishing, CB ©1999.
10. S. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," *IEEE Trans. Info. Theory*, Vol. 38, No. 2, pp. 617-643, March 1992.
11. A. Quddus and M. M. Fahmy, "Fast wavelet-based corner detection technique," *Electronics Letters*, Vol. 35, pp. 287-288, Feb. 1999.
12. D. Wei, J. Tian, R. O. Wells, C. S. Burrus, "A new class of biorthogonal wavelet systems for image transform coding," *IEEE Trans. Medical Imaging*, Vol. 7, No. 7, pp. 1000-1013, 1998.
13. Fish contour images, http://www.ee.surrey.ac.uk/Research/VSSP/imagedb/demo.html