Contour-based Object Recognition using Wavelet-Transform

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 for shape recognition using the wavelet transform modulus maxima. And we apply it to the problem of content-based indexing and retrieval of fish contours. The description scheme and the similarity measure proposed here are simple and take into consideration the way our visual system perceives objects and compares them.

The proposed scheme is invariant to translation, rotation, scale change and to 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 good performance of the proposed technique.

1. Introduction

Content-based indexing and retrieval (CBIR) of digital images became a very active area of research; and both industrial and academic systems for image retrieval have been built. Most of these systems (e.g. *QBIC* [1] from IBM, *Netra* from UCSB, *Virage* from Virage Inc., *MUVIS* [2] from TUT) support one or more of the following options: browse, search by example, search based on a single or a combination of low level features. These features can be extracted from the image, such as color, shape, texture, spatial layout of objects in the scene or added to it after its capture, such as contextual information and keywords.

In this paper we will focus on one way of describing the shape feature of a given object. Therefore, our data consists of non-occluded object boundaries, single object boundary per image. We are not concerned by the way these boundaries were obtained. The problem of automatic object recognition and similarity estimation remains a hard problem even with such assumptions.

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 [3]. 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 [4]. Therefore, in this paper we preferred to represent shapes by their outer boundary and not by the regions they contain.

Hoffman et al. [5] 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.

In the case of shape, high curvature points are robust features in the sense that they are invariant under translation, rotation and scale change [6, 7]. Moreover, they provide reliable clues regarding objects even under occlusion and varying background levels [8]. Corner-based representation of objects reduces significantly the size of the feature vector representing the object-contour, while still keeping mach of the boundary information essential to object recognition [4]. Therefore, object recognition techniques based on corner point matching have been used in machine vision applications [8, 9]. It can be seen that complexity of the algorithms proposed in [8, 9] increases exponentially as the number of candidate objects increases. Therefore, these techniques are not suitable for large image databases where thousands of images are involved.

Hwang and Mallat [10] proved that there couldn't be a singularity without a local maximum of the wavelet transform at the finer scales. Therefore, the WTMM seem to be very appropriate for the description of contours. 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.

In this context we propose a robust wavelet-based matching algorithm, which is suitable for shape matching based on object contour. Moreover, it is not very sensitive to noise and is invariant to translation, rotation and scale change. The algorithm uses WTMM to detect the location of high curvature points and to estimate the degree of similarity between two shapes at these points. In this paper the performance of our approach is evaluated for estimating the similarity of natural objects. The retrieved images with the proposed approach are compared to those retrieved with the contour scale-space (CSS) technique [16] and to a set of images retrieved by human users.

2. Wavelet-based feature extraction and matching

Wavelet decomposition provides natural setting for the multi-level image contour analysis. Since wavelet transform modulus maxima (WTMM), provide useful information for curvature analysis [11, 12], we propose to use it here for fast feature extraction. It has been shown in [12, 13] that biquadratic wavelets, proposed by Mallat and Zhong [14], perform better than other wavelets for corner detection applications.

The boundary is tracked and its orientation profile is computed as in [15]. The orientation profile of each one of the two shapes is up-sampled 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 WTMM larger than a fixed threshold are considered important singularities. 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.

The search space is in general reduced by fast schemes for narrowing down the search space are becoming essential in content-based retrieval systems due to the increasing size of data sets under consideration. Here we used the aspect ratio γ to reduce the search space. Images with error on γ larger than 5% are discarded. The remaining candidates go through the second step of the retrieval process, where the wavelet features are extracted at the selected high curvature points and compared to those of the query image. These wavelet features are the locations and the magnitudes of the wavelet transform modulus maxima.

3. Algorithm

1. Select candidate objects with aspect ratios similar to the one of the query object (5% error is allowed),
2. Consider only WTMM greater than the threshold T_{WTMM},
3. Compare WTMM of both query image and the database,
4. Compute a similarity score, at each level,
5. The final similarity score is computed as the maximum of the scores at each level, not all the decomposition levels are used.

At each wavelet decomposition level l, we characterize the query image contour with two vectors, M_l^q and L_l^q, where, $M_l^q = \lfloor m_{l1}^q \quad m_{l2}^q \quad \cdots \quad m_{lm}^q \rfloor$ contains the m magnitudes of the WTMM and $L_l^q = \lfloor p_{l1}^q \quad p_{l2}^q \quad \cdots \quad p_{lm}^q \rfloor$ 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, feature vectors from the query and candidate contours are shifted in a way to have their largest magnitude values aligned. Which means that, at each decomposition level, we start the matching process from the boundary points having the highest curvature.

A valid match between two high curvature points is found if the differences between their locations and magnitudes are under the thresholds T_M and T_L respectively. Let K be the number of matched maxima. The similarity score at level l, for $l = j, \cdots, 6$, is computed as:

$$s_l = \left(\frac{2 \times (K - \xi)}{(m+n)} \right) \times 100 , \qquad (1)$$

where, $\xi = \sum_{i=1}^{K} \left(\frac{\delta m_i}{mean(m_{li}^q, m_{li}^c)} + \frac{2 \times |\delta p_i|}{L} \right)$, L is the length of the contours, δm_i and δp_i are the errors on magnitude and position for the i^{th} matched maxima. ξ gives an idea on how good the match is between the two sets of points.

The lower levels are not considered in the matching process in order to make our measure robust to minor changes in the boundary, which is a very important factor since we are dealing with natural objects. 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. [8], however the proposed approach is more effective since the locations of the HCP are accurately determined by tracking the WTMM through the decomposition levels until the original contour. Moreover it is faster, since few decomposition levels are needed, unlike the CSS where the full decomposition is required. Moreover, only the features considered important by our visual system are used to estimate the similarity of two contours and redundant information is discarded by ignoring smooth curves. The proposed descriptors preserve most of the shape information, since the object contour can be accurately reconstructed from its WTMM [15].

4. Experimental results

In these experiments we used 1130 fish contour images, (see Section 6). The boundary of each fish in the database is represented by a sequence of 1000 points. The results of querying the fish database with the "Query image" shown in Figure 1 are presented in Table 1. 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 T_L. The first $T_{WTMM} = 0.6$, separates high curvature points from insignificant singularities which could be present in the fish boundaries due to noise introduced during the acquisition or the contour extraction processes. While the last two thresholds, $T_M = 0.2$ and $T_L = 20$, are the tolerated errors on the magnitude and the location of the wavelet transform modulus maxima.

The shape of the neighborhood used is specified by the errors on the location and the magnitude and their relation. In our experiments we used a rectangular neighborhood, which imposes no further constraints on the possible values of the magnitude and the location. One intuitive relation could be $T_M + T_L \leq Const$, to allow a larger error on the location when the magnitude error is small and vise versa.

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 human users. And that the produced results with the proposed algorithm are clearly better than those obtained by CSS.

5. Conclusions

In this paper we proposed a fast and robust retrieval algorithm for contour images. This technique is invariant to translation, rotation, and scale change. Moreover it is robust to noise corruption. These characteristics and its simplicity make the proposed approach very suitable for object retrieval over large databases.

The performance of our algorithm was compared to that of the scale–space based technique "CSS" [16] and against results produced by human observers. And it was shown to be better than the results obtained by the CSS technique and very close to what the human observers have selected.

6. Acknowledgements

Authors acknowledge Prof. F. Mokhtarian and S. Abbasi, for providing the test image database. Authors also acknowledge Tampere Graduate School in Information Science and Engineering (TISE) and Tampere International Center for Signal Processing (TICSP) for providing the financial support for this research.

7. References

[1] Niblack, W. et al, "The QBIC project; querying images by content using color, texture and shape", *SPIE*, Vol. 1908, 1993.

[2] 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.

[3] Goldstein, E. B., "Sensation and Perception," 5th Edition University of Pittsburgh Published by Wadsworth Publishing, CB © 1999.

[4] Russ, J. C., "The Image Processing Handbook", 3rd edition, *CRC, Springer and IEEE Press inc.*, 1995.

[5] Hoffman, D. D., and Richards, W. A., "Parts of recognition", Cognition, Vol. 18, pp. 65-96, 1984.

[6] Attneave, F. 'Some informational aspects of visual perception', Psychological Rev., 61, (3), pp. 183-193, 1954.

[7] 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.

[8] 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.

[9] Koch, M. W. and Kashyap, R.L., " Using polygon to recognize and locate partially occluded objects", *IEEE Trans. PAMI*, vol. 9, pp. 483-494, 1987.

[10] Mallat, S., "A Wavelet Tour of Signal Processing", 2nd Edition, Academic Press, 1999.

[11] Quddus, A. and Fahmy, M. M., "Fast wavelet-based corner detection technique", *Electronics Letters*, Vol. 35, pp. 287-288, Feb. 1999.

[12] Quddus, A., "Curvature analysis using multi-resolution techniques", *Ph.D. Thesis, King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia*, 1998.

[13] Quddus, A. and Fahmy, M. M., "Corner detection using various wavelets", accepted for the ICICS'99, December 1999, Singapore.

[14] Mallat, S., G. and Zhong, S., "Characterization of signals from multiscale edges", *IEEE Trans. on PAMI*, Vol. 14, pp. 710-732, 1992.

[15] 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.

[16] Abbasi, S., F. Mokhtarian, and J. Kittler, "Curvature Scale Space image in Shape Similarity Retrieval," *Springer Journal of MultiMedia Systems*, 1999.

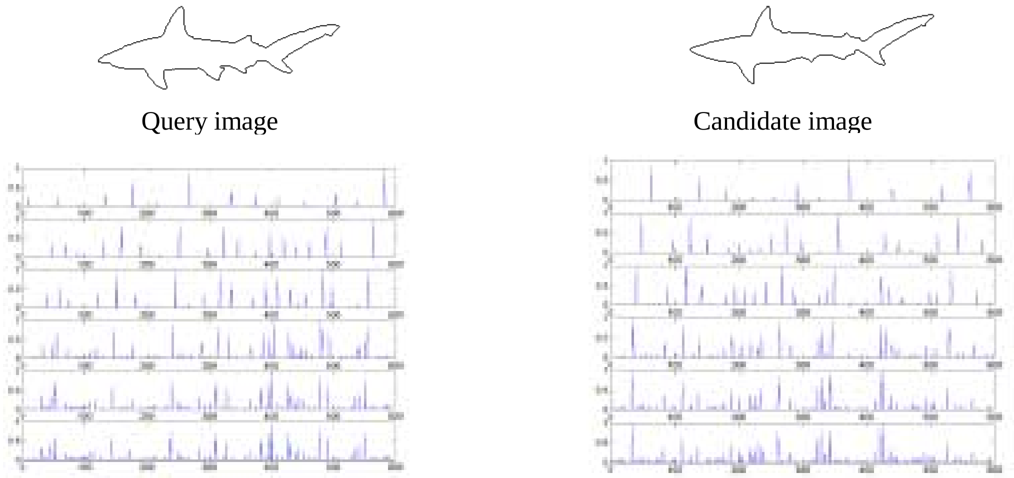
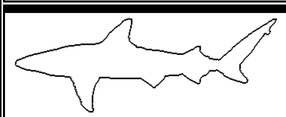
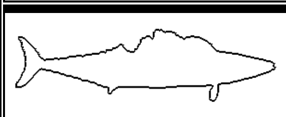
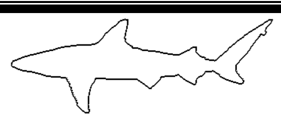


Figure 1. Two fish contours and their corresponding WTMM.

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

CSS	Human	Proposed algorithm	Similarity Scores with proposed algo.
			100
			83
			74
			67
			58