INVARIANT 2D OBJECT RECOGNITION USING THE MODULUS MAXIMA OF A CONTINUOUS WAVELET TRANSFORM

Mahmoud I. Khalil and Mohamed M. Bayoumi

Department of Electrical and Computer Engineering, Queen's University
Kingston, ON, Canada
khalil@eleceng.ee.queensu.ca

ABSTRACT

In this paper, a technique is proposed to recognize 2D objects under translation, rotation, and scale transformation. The technique is based on the continuous wavelet transform and neural networks. Experimental results are presented and compared with traditional methods. The experimental results showed that this refined technique was successfully capable of classifying the objects and that it outperformed some traditional methods especially in the presence of noise.

1. INTRODUCTION

A great deal of research in computer vision focuses on invariant 2D object recognition, yet the available methods suffer from various disadvantages that limit their applicability. In this paper, a new technique is presented for recognizing 2D objects under orientation, location, and size transformation. The developed technique is based on analyzing the boundary of objects using the decay of the continuous wavelet transform and neural networks. Mallat [1] stated that the wavelet mathematical theory is reaching a mature stage but it is not clear how to use these wavelet descriptors to solve computer vision problems.

The paper is organized as follows. Section 2 describes the $\Theta - S$ diagram and how to eliminate the wrap around error. Section 3 investigates how to calculate the position and the regularity of each singularity. Then, in Section 4, the experimental results are presented. Section 5 provides a comparison among certain traditional methods and our proposed technique. Finally, Section 6 contains the conclusion of the given work.

2. THE WRAP AROUND ERROR-FREE $\Theta - S$ DIAGRAM

In our recognition system, the first step is to extract the object boundary. One of the edge detection techniques can be used to extract the boundary of the object [2]. The boundary is recovered by one of the edge-following techniques [3]. The object boundary will be represented by the $\Theta - S$ diagram, where Θ is the angle made between a fixed line and the tangent to the boundary of the object. This angle is plotted against S, the arc of the boundary traversed. A boundary is considered as a sequence of successive boundary pixel coordinates (x_i, y_j).

The $\Theta - S$ diagram at any boundary point i can then be expressed as the tangent value calculated in the window w:

$$\Theta_i = tan^{-1}\frac{y_i - y_{i-w}}{x_i - x_{i-w}} \qquad i = 0, 1, \ldots, N-1 \tag{1}$$

Here, N is the number of the boundary points.

As an example, Figure 1(a) shows the image of an object, and Figure 1(b) shows its boundary. The range of equation (1) is bounded between $-\pi/2$ and $\pi/2$. As a result, the wrap around error occurs when $|\Theta|$ exceeds $\pi/2$ and this error creates additive discontinuities which do not represent any meaningful features of the object. In the next section, the continuous wavelet transform is used to detect the discontinuities in the $\Theta - S$ diagram. The discontinuities introduced by this wrap around error will be misinterpreted by the wavelet transform. To eliminate this error, an offset is added to the Θ_i values to get the compensated ones:

$$\Theta_{compensated} = \Theta_i + \text{offset} \tag{2}$$

This offset is initialized to zero, then it is calculated at each boundary point as:

$$\begin{aligned} \text{offset} &= \text{offset} + \pi \;\; \text{if} \;\; (\Theta_{i-1} - \Theta_i) \geq \pi * \text{factor} \\ &= \text{offset} - \pi \;\; \text{if} \;\; (\Theta_i - \Theta_{i-1}) \geq \pi * \text{factor} \end{aligned} \tag{3}$$

We have selected the factor to be equal to 0.95. Figure 2(a) shows the wrap around error $\Theta - S$ diagram for the plane object shown in Figure 1(a), and Figure 2(b) shows the wrap around error-free $\Theta - S$ diagram. Each $\Theta - S$ diagram will be normalized such that its length is equal to 256 (i.e. $N = 256$). The $\Theta - S$ diagram itself is invariant to rotation. To make it invariant to translation, the starting point is chosen to be the point on the boundary that intersects with the major principal axis [4].

3. THE MODULUS MAXIMA OF A CONTINUOUS WAVELET TRANSFORM

In the previous section, it was shown how each object may be represented by its $\Theta - S$ diagram. The $\Theta - S$ diagram consists of singularities that carry the most important information that represents each object. In this section, the

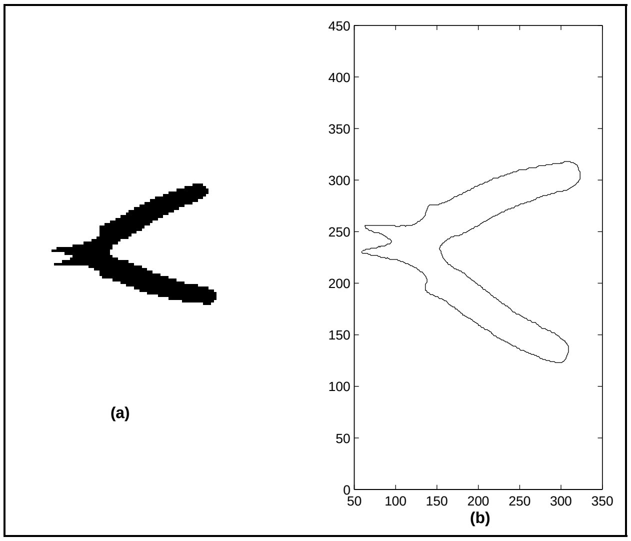

Figure 1: (a) A tool object. (b) The tool boundary.

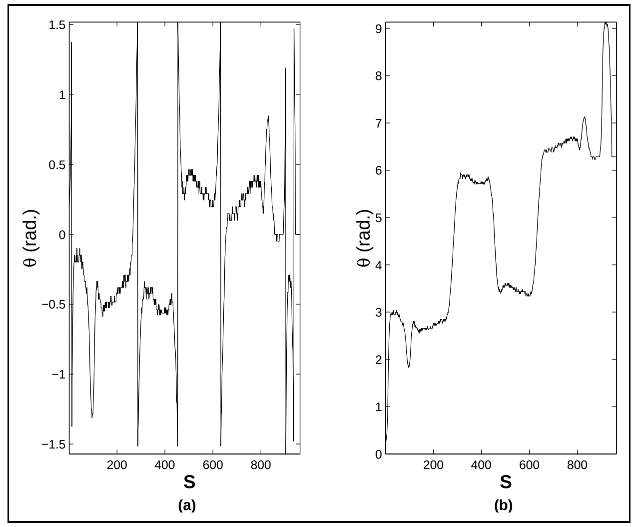

Figure 2: (a) Wrap around error $\Theta - S$ diagram. (b) Wrap around error free $\Theta - S$ diagram.

evolution across the scales of the continuous wavelet transform is used to specify the location of these singularities and to calculate their Lipschitz exponents.

If we denote the $\Theta - S$ diagram as $f(t)$, then $Wf(u,s)$ denotes the continuous wavelet transform in the scale space diagram where u is the position parameter and s is the scale. The function $Wf(u,s)$ can be calculated by convolving $f(t)$ with the wavelet function $\psi(t)$ as follows

$$Wf(u,s) = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{(s)}} \psi\left(\frac{t-u}{s}\right) dt \tag{4}$$

The decay of $|Wf(u,s)|$ can be calculated by measuring the decay of its modulus maxima values. The modulus maxima line consists of points which are locally maximum.

The kind of wavelet function used is

$$\psi(t) = (-1)^n \frac{d^n \theta(t)}{dt^n} = (-1)^n \theta^{(n)} \tag{5}$$

where θ is a smoothing function satisfying $\int_{-\infty}^{+\infty} \theta(t) dt \neq 0$.

The smoothing function θ is chosen to be a Gaussian function such that the modulus maxima of $Wf(u,s)$ are not interrupted when the scale decreases [5].

The following algorithm is used to extract the singularity position and its regularity.

1. Choose $n = 1$ in equation (5) and get the continuous wavelet transform. The scale is chosen to start from $s = 1$ to $s = 8$.
2. Extract the maxima line which propagates more than a certain scale value. In our experiment, this value equals 2. In this way, we neglect maxima lines that are smaller than a given small length which either does not carry a significant information or it may occur due to noise.
3. Get the position of each singularity where the maxima line intersects with the finest possible scale level ($s = 1$)
4. The singularity exponent $\alpha + \frac{1}{2}$ is calculated by measuring the decay slope of $\log_2 |Wf(u,s)|$ as a function of $log_2 s$ [6].

   If all the singularities have a Lipschitz exponent α less than 1 then the algorithm will terminate.
5. If the singularity has a Lipschitz exponent α that is almost equal to 1, then the actual singularity exponents may be larger than 1. In this case, the same procedure is repeated again by choosing the wavelet which has $n = 2$ in equation (5).

   We are not going to check all the maxima lines in the new wavelet transform. We need to calculate the actual Lipschitz exponent of the singularity with $\alpha > 1$ but can not be calculated when the wavelet used has $n = 1$. Knowing the position of the singularity u_0 from step (3), the cone of influence of this singularity as the region in the scale space is formed where $|u - u_0| \leq Cs$ and the wavelet ψ has a compact support equal to $[-C, C]$. The largest maxima line in this cone of influence is used to calculate the Lipschitz exponent of this singularity.

   The procedure can be repeated if the Lipschitz exponent α is greater than 2.

If the number of singularities detected is denoted by M, then the feature vector for each object is:

$$V = V_A \cup V_B \tag{6}$$

whereby

$$V_A = \{d_1, d_2, \ldots, d_M\}, \tag{7}$$

$$V_B = \{\alpha_1, \alpha_2, \ldots, \alpha_M\}, \tag{8}$$

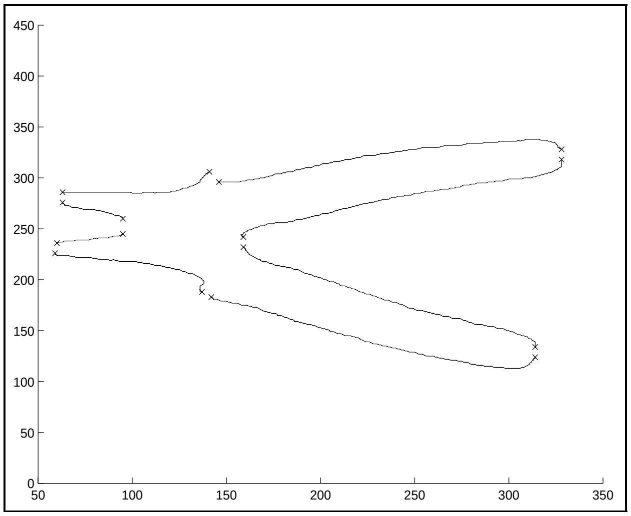

Figure 3: The segmented tool.

and where

d_i	is the distance between singularity i and $i+1$,$(i = 1, \ldots, M-1)$
d_M	is equal to the distance between singularity M to the end of the transformed signal and the beginning of the transformed signal and the first singularity, and
α_i	is the Lipschitz exponent for the singularity number i, $(i = 1, \ldots, M)$

Figure 3 shows the tool shown in Figure 1(a) after it has been segmented in the positions where the singularities are detected by the previous algorithm.

Neural networks are used to solve the supervised classification problems [7]. The classifier used is the multilayer feedforward network which is trained by means of the back-propagation algorithm [8]. The number of input nodes is equal to the maximum number of feature vector length M_{max} (V in equation (6)). The feature vectors for other training models are padded with zeros to have the same length M_{max}. The number of output nodes is equal to the number of training models.

4. EXPERIMENTAL RESULTS

In order to evaluate the performance of the recognition system, the images of eight different tools [9] were used. Figure 4 shows the tool objects used in the experiments. Some of these images are different objects, others are similar objects (e.g. model2 and model3 or model6 and model7). The former case shows the discrimination power of the features in general, and the latter shows the ability to describe small variations.

The recognition system was tested by adding random Gaussian noise to the shape boundary points. This kind of distortion is used in You and Jain [10], and Kauppinen *et. al.* [11]. If the coordinates of the ith boundary point are

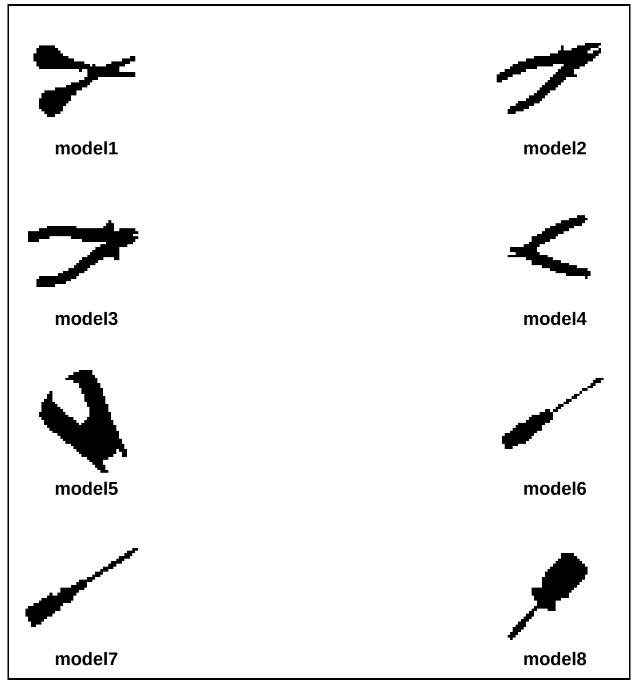

Figure 4: The tool objects.

$(x(i), y(i))$, then the coordinates of the corresponding point on the noisy boundary are given by:

$$x_{noisy}(i) = x(i) + drc\cos(\Theta_i)$$

$$y_{noisy}(i) = y(i) + drc\sin(\Theta_i)$$

where

r	is a sample from Gaussian distribution $N(0,1)$,
d	is the distance between successive boundary points,
Θ_i	is the tangent angle at boundary point i, and
c	is a parameter which controls the amount of noise.

The points that caused crossover on the boundary will be omitted. The measure of noise in an object contour is determined by the percentage of the contour points which are corrupted with noise. Figures 5(a) and (b) show the model4 object boundary with 50% and 75% noise levels. The size of the test images is varying by $\pm 50\%$. The simulation results are illustrated in table 1.

5. COMPARISON WITH TRADITIONAL METHODS

In this section, the algebraic invariant moments and the Fourier descriptors techniques are used to evaluate the performance of our refined technique.

Hu [12] reported the mathematical foundation of the two-dimensional moment invariant method and its application to visual information processing. He used the six absolute orthogonal invariant functions calculated for the second and third moments.

Table 1: Recognition accuracy for the tool objects.

Noise (%)	model1	model2	model3	model4	model5	model6	model7	model8
0	100	100	100	100	100	100	100	100
10	100	100	100	100	100	100	100	100
20	100	100	100	100	100	100	100	100
30	100	100	100	100	100	100	100	100
40	100	95	100	100	100	100	95	100

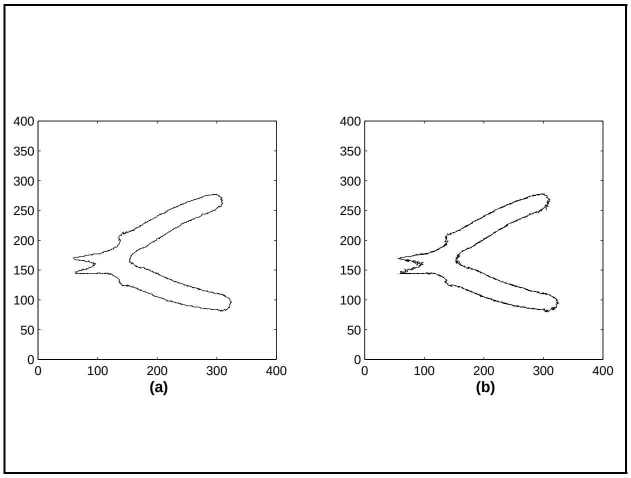

Figure 5: (a) A tool object boundary with %50 noise. (b) A tool object boundary with %75 noise.

Table 2 shows the experimental results obtained with algebraic invariant moments. The classifier used is the multilayer feedforward network with one hidden layer. The number of input nodes is taken to be equal to 6 and the number of output nodes is 8.

Zahn and Roskies [13] developed a method for the analysis and synthesis of closed curves using the Fourier descriptors method. These descriptors represent the shape of the object in the frequency domain. A subset of the Fourier descriptors is often enough to discriminate between different shapes. The Fourier descriptors used is based on the normalized $\Theta - S$ diagram. A subset containing the twelve lowest frequency coefficients was used.

Table 3 shows the experimental results obtained with the Fourier descriptors method for the two data sets. The classifier used is the multilayer feedforward network with one hidden layer. The number of input nodes is twelve and the number of output nodes is chosen to be equal to 8.

It should be noted that although the performance of the Fourier descriptors method and that of invariant moments method are comparable with our recognition system in the case of noiseless objects, their performance was much worse in the case of heavy noisy objects.

6. CONCLUSION

This paper has proposed a recognition system based on the continuous wavelet representation and neural networks approach to recognize 2D objects under translation, rotation and scale transformation. The technique is based on extracting the singularity position and its regularity by measuring the decay of modulus maxima lines of the wavelet transform. The experimental results show that the refined technique gives better results than some traditional methods, such as the Fourier descriptors method and the invariant moments method especially in the presence of noise.

7. REFERENCES

[1] S. Mallat, "Wavelets for a vision", *Proceedings of the IEEE*, vol. 84, no. 4, pp. 604–614, 1996.

[2] J.F. Canny, "A computational approach to edge detection.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[3] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley and Sons, New York, 1973.

[4] R.J. Prokop and A.P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition", *CVGIP: Graphical Models and Image Processing*, vol. 54, no. 5, pp. 438–460, 1992.

[5] S. Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, San Diego, 1998.

[6] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges.", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, 1992.

[7] R.P. Lippmann, "An introduction to computing with neural nets", *IEEE Acoustics, Speech and Signal Processing Magazine*, vol. 61, pp. 4–22, 1987.

[8] D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1:Foundations.*, MIT Press, Cambridge, MA, 1986.

[9] G.N. Bebis and G.M. Papadourakis, "Object recognition using invariant object boundary representations and neural network models", *Pattern Recognition*, vol. 25, no. 1, pp. 25–44, 1992.

Table 2: Recognition accuracy for the tool objects by algebraic moment invariants.

Noise (%)	model1	model2	model3	model4	model5	model6	model7	model8
0	100	100	100	100	100	100	100	100
10	100	80	100	100	100	100	85	100
20	100	65	100	100	100	100	70	100
30	100	40	80	100	100	95	55	100
40	100	25	70	100	100	85	30	100

Table 3: Recognition accuracy for the tool objects by Fourier descriptors.

Noise (%)	model1	model2	model3	model4	model5	model6	model7	model8
0	100	100	100	100	100	100	100	100
10	100	95	100	100	100	100	95	100
20	100	85	100	100	100	100	85	100
30	100	75	100	100	100	100	80	100
40	100	60	95	95	100	95	75	100

[10] Z. You and A.K. Jain, "Performance evaluation of shape matching via chord length distribution", *Computer Vision, Graphics and Image Processing*, vol. 28, pp. 185–198, 1984.

[11] H. Kauppinen, T. Seppanen, and M. Pietikainen, "An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 201–207, 1995.

[12] M. Hu, "Visual pattern recognition by moment invariants", *IRE Transactions on Information Theory*, vol. 8, pp. 179–187, 1962.

[13] C.T. Zahn and R.Z. Roskies, "Fourier descriptors for plane closed curves", *IEEE Transactions on Computers*, vol. C-21, no. 3, pp. 269–281, 1972.