# OBJECT IDENTIFICATION USING THE DYADIC WAVELET TRANSFORM AND INDEXING TECHNIQUES

*Quang Minh Tieng and W. W. Boles*

Signal Processing Research Centre
Queensland University of Technology
GPO. Box 2434, Brisbane, QLD 4001, Australia
quang@lime.eese.qut.edu.au, w.boles@qut.edu.au

## ABSTRACT

A wavelet based representation of planar objects is introduced. Based on this representation, a matching algorithm using indexing techniques is also developed for identifying unknown objects. Rather than considering all points on the representation, only extrema are used in constructing the look-up table and for matching. Simulations demonstrate that the proposed algorithm is effective and accurate in classifying objects under similarity transformations and in a noisy environment.

## 1. INTRODUCTION

Singularities and irregular structures often carry the most important information in object contours and it has been suggested, from the view point of the human visual system, that some points along these contours are rich in information content and are sufficient to characterise the shapes of the objects. Examples of such critical boundary points (feature points) are minima, maxima, zero-crossings of curvature, and discontinuities (or singularities) of curvature or tangent angle, etc.

By decomposing signals into elementary building blocks that are well localised in both space and frequency domains, the wavelet transform can detect and characterise the local regularity of signals [1]. In this paper, we introduce a method for recognising planar objects based on the singularity points on their contours. The proposed algorithm consists of two steps : (i) constructing the object representation in the form of discontinuities of the tangent angle (discussed in section 2) and (ii) classifying unknown objects into different classes represented by given models in a database (given in section 3). The discontinuities are detected using the dyadic wavelet transform, and the classification method is based on the indexing technique which has been shown to be an efficient technique to deal with large size databases [2]. Some experimental results will be given in section 4 followed by the conclusions in section 5.

## 2. WAVELET BASED REPRESENTATION

Our proposed object identification system uses shape in an image as an important factor to recognise and associate 2D unknown objects. The system is designed to handle planar closed boundary objects of general shape with or without noise. Translation, uniform scale and rotation are permissible. Our starting point for digital shape analysis will be the segmented image, in which the individual connected components have been identified and labelled. This can be accomplished by means of an edge detector followed by a tracing algorithm [3].

The proposed representation is constructed based on the *tangent angle function* which is defined as follows :

$$\theta(t) = \arctan\left(\frac{\dot{y}(t)}{\dot{x}(t)}\right) \qquad t \in [0, T]$$

where $T$ is the perimeter of the object, the dot denotes differentiation with respect the arc length parameter, $t$, and $(x, y)$ are the horizontal and vertical coordinates of the contour points, respectively. In the discrete case, the derivatives are approximated by the finite differences and the *discrete* tangent function can be implemented as :

$$\theta(n) = \arctan\left(\frac{y(n+1) - y(n)}{x(n+1) - x(n)}\right) \tag{1}$$

with $n = 0, \ldots, N-1$, and $N$ is the number of equi-spaced segments along the object contour.

This expression of the tangent angle bounds the gradient into the region $[-\pi, \pi]$. Hence, values which fall outside this range will be wrapped around and abrupt changes are observed at locations where the difference of the gradient magnitude between two successive points exceeds $\pi$ radians. These abrupt changes will form artificial discontinuities which will be mis-interpreted as critical boundary points when the representation is used for object recognition. The wrap around error can be eliminated with the addition of a compensation factor, $F(\cdot)$, as follows :

$$\theta_c(n) = \theta(n) + 2\pi F(\theta(n) - \theta(n-1)) \tag{2}$$

with

$$F(z) = \begin{cases} 1 & \text{for} \quad z \leq -\pi \\ -1 & \text{for} \quad z \geq \pi \\ 0 & \text{otherwise} \end{cases}$$

In relation to shape representation and analysis, $\theta_c$ has one serious drawback : it is not a periodic function. For simple closed curves $\theta_c$ has the property $|\theta_c(0) - \theta_c(N)| = 2\pi$. In order to overcome this problem, the compensated tangent angle function, $\theta_c$, needs to be periodised as follows :

$$\theta_p(n) = \theta_c(n) + \text{sgn}(\theta_c(0) - \theta_c(N)) \cdot \frac{2\pi}{N} n \tag{3}$$



0-7803-2431-5/95 $4.00 © 1995 IEEE

with

$$\text{sgn}(z) = \begin{cases} 1 & \text{for} \quad z > 0 \\ -1 & \text{for} \quad z < 0 \\ 0 & \text{for} \quad z = 0 \end{cases}$$

This tangent angle function $\theta_p$ is, in general, dependent on the size of object contours. In order to make the representation size (or scale) invariant, a normalisation process on object contour arc lengths is necessary to be carried out before finding the discontinuities on the function. Since the proposed algorithm makes use of the dyadic wavelet transform in constructing the object representation, it is necessary that the arc length be normalised to a power of two integer in order to extract all the information contained in the tangent function without retaining the approximated tangent at the coarsest resolution level. For matching purposes, the number of data points on object and model contours must be the same. The normalised contour thus is re-sampled such that the number of data points equals the normalisation constant. There is a tradeoff between the normalisation constant value, the classification result and the computational time [4].

The next step of constructing the object representation is locating feature points on the object contour which correspond to discontinuities on the periodised tangent angle function $\theta_p$. Adapting the technique in [1], and using a wavelet which is the first derivative of a smoothing function, the irregular structure of the tangent function is detected. Let $\varphi$ be a smoothing function, then the wavelet function is defined as follows :

$$\psi(t) = \frac{d\varphi(t)}{dt} \tag{4}$$

The dyadic wavelet transforms of a signal $f$ with respect to this wavelet is defined as :

$$W_j f(t) = f * \psi_j(t) = 2^j \frac{d}{dt} (f * \varphi_j)(t) \tag{5}$$

where $\varphi_j(t) = \frac{1}{2^j}\varphi(t/2^j)$.

With the wavelet as the first derivative of a smoothing function, we will show that the wavelet transform of the periodised tangent angle function, $\theta_p$, will be proportional to the curvature of the object contour. For simplicity and to be able to extend Mallat and Hwang's results, we will consider the problem in the continuous domain.

In the continuous domain, the periodised tangent angle function (3) can be rewritten as :

$$\theta_p(t) = \theta_c(t) + \frac{2\pi}{T} t \tag{6}$$

here we assume that the tangent angle at the starting point on the contour is larger than that at the end point. Taking the dyadic wavelet transform of function (6) using a wavelet which is the first derivative of a smoothing function, we obtain

$$\begin{aligned} W_j\theta_p(t) &= \theta_p * \psi_j(t) \\ &= 2^j \frac{d}{dt} (\theta_p * \varphi_j)(t) \\ &= 2^j \frac{d}{dt} (\theta_c * \varphi_j)(t) + 2^j \cdot \frac{2\pi}{T} \frac{d}{dt} (t * \varphi_j)(t) \\ &= 2^j \kappa * \varphi_j(t) + 2^j \cdot \frac{2\pi}{T} \int_{-\infty}^{\infty} \varphi_j(t) dt \end{aligned} \tag{7}$$

where

$$\kappa(t) = \frac{d\theta_c(t)}{dt} \tag{8}$$

is the definition of the curvature of a 2D curve.

If $\varphi$ is normalised such that;

$$\int_{-\infty}^{\infty} \varphi(t) dt = \int_{-\infty}^{\infty} \varphi_j(t) dt = 1$$

then the dyadic wavelet transform of the periodised tangent angle function with respect to $\psi$ is proportional to the curvature of the contour smoothed by the smoothing function $\varphi$ with an offset factor which depends on values of $j$ and $T$. Curvature descriptions of planar curves are intrinsic, that is, rotation and translation invariant, and with an appropriate normalisation, they are even scale invariant. Curvature also uniquely represents planar curves such that a contour can be reconstructed from its curvature description.

## 3. MATCHING ALGORITHM

Having represented an object contour by a wavelet based representation, we now wish to match an unknown object contour with model contours. Rather than considering every point on the representation, only extrema will be used in our matching algorithm. The use of extrema in matching reduces the computational effort and also makes the classification process less sensitive to variations on the representation due to the effect of uncontrolled sources.

In order to classify unknown objects into given classes represented by models in the database, one can repeatedly match the unknown object representations with model ones with the help of a similarity (or dissimilarity) function such as the one in [4]. This technique is referred to as search-based matching since the representation of the unknown objects must be matched with all available model representations. This means that it requires all feature combinations to be explored. As a result, the matching is computationally equivalent to an exponential search. In applications with large databases, this computational burden has to be borne every time an image is processed.

Recently, an alternate technique called *indexing* or *hashing* has been proposed for matching visual shapes [2]. In this technique, the feature correspondence and search of model database are replaced by a table look-up mechanism. The invariant features extracted from unknown objects are used as indexes to look-up a table containing references to the object models. The look-up table then returns a list of candidate models with associated weights indicating their likelihood. The advantages of indexing over traditional search-based matching schemes are especially evident in applications involving large collections of object models. Indexing does not require considering each model separately and is thus less dependent on the database size. A general computational framework of indexing can be found in [2].

Index-based recognition systems, in general, compute invariant features from an image which are then used to index a look-up table. In our case, there are four features used as indexes for the look-up table. They are the resolution levels, and the magnitude, type and location of extrema of the representation at each level. The first three features are invariant to similarity transformations. The locations of extrema, however, vary since the representation is not shift invariant. Thus some techniques must be used to overcome this problem.

*Resolution Level :* Not all resolution levels are used as indexes for the look-up table. Only the levels which contain most of the representation energy are considered. By using the representation at these levels, the effects of noise and quantisation errors are significantly reduced. This is because, at these levels, the signal to noise ratio (SNR) is high if we assume that the noise and errors have a uniform distribution. In practice, these levels must be determined for each model in the database. Knowing the range of these levels for each model, common levels of these ranges, denoted by $\mathcal{L}$, are used as one of the indexes for the look-up table.

*Extrema Magnitude :* The second index, denoted by $\mathcal{M}$, is the extrema magnitude. The magnitude has a real and unbounded value. Since each index of the table must be a finite number of discrete values, the magnitude will be normalised and quantised. The normalisation process is performed to make the largest absolute value of the magnitude of the extrema in the common levels become 1. As a result, the extrema magnitudes have values in the range of $[-1, 1]$. This range will be digitised to generate discrete indexes. Different quantisation methods can be applied to the normalised magnitude. The simplest scheme is uniform quantisation. If the range of the magnitude index is chosen as $M$, then using uniform quantisation, the $i^{th}$ index will be indicated by the normalised magnitude in the interval of $(-1 + \Delta m(i - 0.5) \pm \Delta m/2)$ with $\Delta m = 2/M$.

*Extrema Position :* The extrema positions are used as the third index which is denoted by $\mathcal{D}$. The maximum value of the position is defined based on the normalisation constant, $S$, used in constructing the representation. Since the proposed representation is shift variant, the extrema positions vary depending on the starting point on the object contour. However, the relative distance between them is considerately stable. Thus, in order to use the extrema positions as an index, a reference position must be defined such that the positions of extrema with respect to it will be fixed regardless of different starting points. By relating the wavelet and subband theories, it can be shown that the number of equivalent frequency components in detail signals, which are the representation in our case, will be halved when one moves from one resolution level to the next lower level. As a result, the corresponding number of extrema on the detail signals (or the representation) will be reduced. At the coarsest resolution level, the detail signal contains only the first harmonic which has two extrema, one maximum and one minimum, in the spatial domain. This means that at this level, the maximum point is unique and its position could be used as a reference. At the coarsest resolution level, however, most of object representations contain a small amount of energy. Thus they are easily disturbed by uncontrolled sources and the position of the extrema at this level will be unstable. In order to obtain a stable reference point, one must find it at the common resolution levels where SNR is high. At these levels, however, there is more than one extreme point. Consequently, it is hard to choose a unique reference point. For reducing the number of extrema and the difficulty in selecting the reference point, the coarsest resolution level among the common ones will be considered. When building the model representation, only one of the extrema (usually the one with the largest magnitude) will be used as the reference point. During the recognition phase, each extreme point will be used in turn as the reference point. Although the extrema positions have integer values, and as a result the relative distances also have integer values, a quantisation process must be applied to limit them to a finite number of discrete values. If a uniform quantisation scheme is used with the range of quantisation levels $D$, then the $i^{th}$ extreme position index will include all relative distances in the interval of $(\Delta d(i - 0.5) \pm \Delta d/2)$ with $\Delta d = S/D$

*Extrema Type :* The last index is the type of extrema and is denoted by $\mathcal{T}$. Two kinds of extrema exist. These are the local maxima and minima whose types have values of 1 and $-1$, respectively. The extrema type can be used directly without any modification.

## 4. EXPERIMENTAL RESULTS

To demonstrate the reliability and robustness of the proposed representation and matching technique, the proposed algorithm was tested on different silhouettes of aircraft shown in figure 1(a) with different sizes, orientations and locations as well as corrupted by uniformly distributed white noise (see figure 2). In figure 1(b), the proposed representation of one of the aircrafts in the database (aircraft (a)) is given. The normalisation constant used in this case is 512. Consequently, there is a total of 9 resolution levels in this representation (the last five low resolution levels are displayed in the figure).

During the experiment, the representation at the four resolution levels 4, 5, 6 and 7 is used. The numbers of magnitude index $M$ and position index $D$ are chosen as 10 and 30, respectively. Table 1 shows some of our experimental results. It consists of two parts. The first one gives the overall test results using the proposed (indexing) technique after 100 tests. The second one gives the results when the same representations are used with a search-based matching algorithm [4]. The experimental results show that the algorithm is very sensitive to shape differences while being invariant to translation, rotation and scaling. The average computational time for 100 tests of the proposed algorithm is 560 seconds. This consists of the time required for constructing the representation, the look-up table, and matching. This was more than 3 times faster than the algorithm using the search-based matching technique. The results show that the proposed algorithm is not only faster but also more accurate when compared with the matching algorithm in [4]. In 100 tests, the proposed algorithm was successful in classifying objects in a noisy environment with 100% correct classification at low and moderate amounts of noise (from 50 to 30dB SNR) and only 2% mis-classification occuring between two similar aircrafts (aircrafts (a) and (c)) with 20dB SNR.

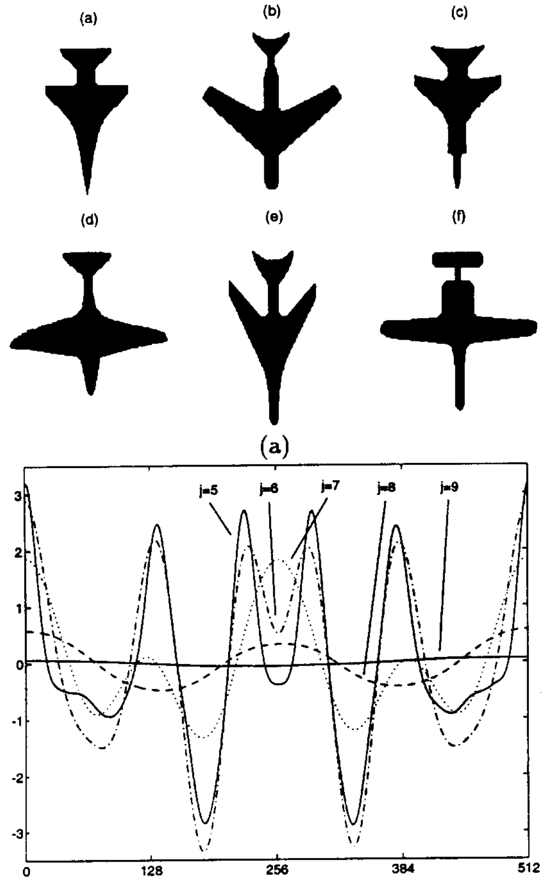


Figure 1: Test objects and the proposed representation. (a) : Aircraft models and (b) : The representation of aircraft model (a).

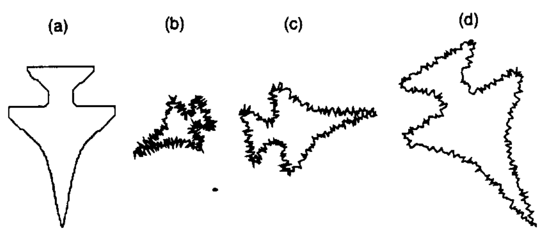


Figure 2: Some typical distorted versions of aircraft (a). (a) : the original contour, (b) : resulting contour at 20dB SNR, reduced with a scale of 0.5 and rotated 300°, (c) : resulting contour at 25dB SNR, reduced with a scale of 0.8 and rotated 100° and (d) : resulting contour at 30dB SNR, enlarged with a scale of 1.2 and rotated 45°.

Table 1: Overall Experimental Results at 20dB SNR after 100 tests.

| Proposed Technique | | | | | | |
|---|---|---|---|---|---|---|
| 20dB | M(a) | M(b) | M(c) | M(d) | M(e) | M(f) |
| O(a) | 98 | 0 | 2 | 0 | 0 | 0 |
| O(b) | 0 | 100 | 0 | 0 | 0 | 0 |
| O(c) | 0 | 0 | 100 | 0 | 0 | 0 |
| O(d) | 0 | 0 | 0 | 100 | 0 | 0 |
| O(e) | 0 | 0 | 0 | 0 | 100 | 0 |
| O(f) | 0 | 0 | 0 | 0 | 0 | 100 |

| Search–Based Technique | | | | | | |
|---|---|---|---|---|---|---|
| 20dB | M(a) | M(b) | M(c) | M(d) | M(e) | M(f) |
| O(a) | 100 | 0 | 0 | 0 | 0 | 0 |
| O(b) | 0 | 96 | 0 | 3 | 1 | 0 |
| O(c) | 2 | 0 | 98 | 0 | 0 | 0 |
| O(d) | 0 | 5 | 0 | 59 | 19 | 17 |
| O(e) | 3 | 0 | 26 | 0 | 71 | 0 |
| O(f) | 0 | 0 | 0 | 0 | 0 | 100 |

## 5. CONCLUSION

In this paper, we presented a technique for constructing a wavelet based representation for planar objects. This representation is translation, rotation and scaling invariant. Based on this representation, a matching algorithm was also developed. Our matching algorithm is based on indexing techniques which use a table look–up mechanism in the classification process. Rather than using every point on the representation, only extrema are considered in constructing the look–up table and matching. From these extrema, three invariant features - magnitude, type and location - are extracted and incorporated with the resolution levels to address the locations of a look–up table. Experimental results show that the proposed algorithm is computationally efficient and that it has successfully classified different objects under the similarity transformation and in a noisy environment.

## 6. REFERENCES

[1] S. G. Mallat and W. L. Hwang, "Singularity detection and processing with wavelets," tech. rep., Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, Mar. 1991.

[2] A. Califano and R. Mohan, "Multidimensional indexing for recognising visual shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, pp. 373–392, Apr. 1994.

[3] T. D. Haig, Y. Attikiouzel, and M. D. Alder, "Border following : new definition gives improved borders," *IEE Proceeding I*, vol. 139, pp. 206–211, Apr. 1992.

[4] Q. M. Tieng and W. W. Boles, "Complex Daubechies wavelet based affine invariant representation for object recognition," *Proc. First IEEE Inter. Conf. Image Processing, Austin, Texas, USA*, vol. 1, pp. 198–202, 13-16 Nov. 1994.