A New Scheme for Planar Shape Recognition Using Wavelets

JIANN-DER LEE
Department of Electrical Engineering
Chang Gung University, Tao-Yuan, Taiwan 333, R.O.C.

(Received July 1998; revised and accepted March 1999)

Abstract—This paper presents a novel approach to planar shape recognition using wavelets. There are two stages, called representation stage and recognition stage, in the proposed method. In the representation stage, in order to extract the features of a shape, a set of wavelet basis are investigated, and wavelet decomposition strategy from the orientation function of the boundary curve of the shape are then performed. The representation of a shape is achieved with a collection of multiscale feature set (MFS) which consists of the scale parameters of wavelet function, the positions where the dominant feature take place, a similarity measure, etc. In the recognition stage, the test shape is compared with various model shapes stored in a database by computing the distance of their MFSs, and the one with the minimum distance is chosen as the correct matching of the test shape. Experimental results obtained with the proposed scheme are encouraging which demonstrates the effectiveness and robustness of the approach. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords—Shape recognition, Wavelet transform, Multiscale features.

1. INTRODUCTION

Planar shape representation and recognition are essential tasks in the field of computer vision and pattern recognition. Usually planar shapes are represented in terms of their bounding contour, which can be obtained by various boundary detection algorithms [1–10]. A general-purpose shape representation method in computational vision should satisfy a number of criteria for further recognition process. The essential criteria are invariance, uniqueness, and stability. Two planar contours are considered to be the same shape if there exists a transformation consisting of uniform scaling, rotation, and translation which will cause one of those contours to overlap the other. The invariance criterion, if two contours have the same shape, they should also have the same representation, guarantees that all contours with the same shape will have the same representation. The uniqueness criterion, if two contours do not have the same shape, they should have different representations, guarantees that two contours with different shapes will have entirely different representations. The stability criterion, if two contours have a small
shape difference, their representations should also have a small difference, guarantees that a small change of a contour will not cause a large variation in its representation. When this criterion is satisfied, the representation of a contour can be considered to be stable with respect to noise. In addition, other criteria are usually required for the applications of practical shape recognition. For example, the criteria may include efficiency, ease of implementation, and computation of shape properties. The former is necessary for an object recognition system to perform real-time recognition. The later is often useful while determining the geometric properties of a contour using its representation. The methods to represent a contour rely on the extracted structure features. Hence, the features extracted from the contour must be efficient for further processing. In other words, the efficient features must satisfy the following criteria.

1. They can represent, as complete as possible, the characteristics of a contour.
2. The properties of these features should be invariant under scaling, rotation, and translation for the same contour.
3. The number of feature representing a contour should be minimal.
4. They should be robust in the presence of noise.

Generally, there are two broad classes of techniques in planar shape recognition: single-scale approaches and multiscale approaches. For single-scale approaches, numerous schemes in a variety of domains can be found in literature, such as chain codes [1], splines [2], invariant moment [3], normalized rapid descriptor [4], normalized Fourier descriptors [5], the circular autogressive (CAR) models [6], and polygonal approximation. For example, the chain coding method approximates a shape boundary with a sequence of directional vectors lying on a square grid. The spline technique uses a set of piecewise low-order polynomials to approximate a boundary contour so that the contour can be characterized by a small number of parameters. This spline technique generates a boundary contour which satisfies the properties of translated, scaled, and rotated invariant. The Fourier descriptor describes a boundary contour with the coefficients obtained by Fourier transform on a certain parametric representation of the contour such as the curvature or spatial coordinates. However, since the boundary contours usually contain detailed information at various scales, it is often difficult to represent the contours with the significant features by the single-scale approaches. Moreover, the performance of this kind of approach is easily affected by the noise from the environment. In other words, to extract the boundary information of the shape effectively for further processing, not only the global information for the qualitative description of the overall shape but also the local information for the detailed variation of the shape should be considered at the same times.

Recent works on the analysis of shape and image structure have addressed these problems by investigating the feasibility of multiscale representation. Based on the concept that human visual system can process and analyze signal information at different resolutions, a lot of multiscale techniques [7–9] have been proposed. Shape representation using coarse-to-fine structure can be produced via the multiscale analysis of a set of parameterized contours. For instance, Rosenfeld and Thurston [7], Witkin [8] introduced the notation of scale-space representation for an object shape. To smooth a one-dimensional signal, Witkin convolved it with the Gaussian kernel and treated the parameter of the kernel as a continuous scale parameter. Then, the structures, which are stable over various scales, are considered as the significant features for representing the boundary contour. Mokhtarian and Mackworth [9] extended this work to two-dimensional shapes. They located the inflection points on the contours and constructed a scale-space image showing the movements of these inflection points over various scales. Bengtsson and Eklundh [10] selected the inflection points of a polygon as features. The positions of these inflection points are estimated by using the coordinates of the midpoints of the inflection segments as features. A shape preserving representation is then derived. The concept of scale-space filtering is also used to extract primitives. Nattarangri and Chin [11] used it to detect and localize corner points on the digital contours. They locate maximum curvature points on digital contours by convolving
it with the Gaussian kernel. For the scale methods, the Gaussian function is an optimal kernel for reducing noise with minimum delocalization [12,13]. Nevertheless, a planar contour smoothed by the Gaussian kernel suffers from shrinkage [14], i.e., the perimeter becomes smaller after convolving it with the Gaussian kernel.

Most of the previous methods for shape (contour) representation only consider either the structure features of the shape which include corners, inflections, and its evolution over scales or qualitative description of a certain segmented part of the contours. In this paper, we propose a new algorithm to represent the contour of an object shape using a collection of multiscale feature set (MFS) which consists of the scale parameters of wavelet function, the positions where the extrema take place, the value of the similarity measure, the amplitude of alternative component of the signal function and the offset value of the signal function. Using the extracted MFSs, one statistical classifier based on the nearest-neighbor clustering rule is then utilized to recognize the unknown object shape. The flow chart of the proposed method is shown in Figure 1.

![Flow chart of the proposed method](image)

**Figure 1.** The flow chart of the proposed method.

![Contour of a pair of pliers](image)

**Figure 2.** The contour of a pair of pliers.

The remainder of this paper is organized as follows. The proposed algorithm for contour representation using wavelet decomposition strategy is presented in Section 2. The MFS extraction
Figure 3. The orientation function of the contour of the pliers.

Figure 4. The evolution of SM(s, t) value of the pliers at different scales.

Figure 5. The position of the MFSs on the shape of pliers.
and shape matching are described in Section 3. Experimental results to show the effectiveness of this method are presented in Section 4. Finally, conclusion is included in Section 5.

2. SHAPE REPRESENTATION

Planar shapes usually are the perspective projection of real 3D objects under arbitrary orientation. The aim of shape representation is to approximate it with less memory such that the amount of data to be processed in the analysis of the shape can be reduced drastically. That is the reason why objects are essentially three-dimensional, two-dimensional representation schemes are commonly used in practical computer vision systems for economic considerations. Since the representation scheme is based on wavelet transformation, its property is then first illustrated below.

2.1. Brief Review of the Wavelet Transform

In the recent years, the wavelet transform became an active area of research for multiresolution signal and image analysis [15-20]. It analyses image information at different resolutions by dilating the scale of the wavelet function and constructing a time-scale representation of a signal, which relates the local properties of the signal to the evolution of wavelet transform coefficients when the scale varies. Through the time-scale representation the local regularity of the input signal can be characterized.

A function \( \psi(x) \) is defined as a basic wavelet, also termed mother wavelet, if it satisfies the admissibility condition

\[
\int_{-\infty}^{\infty} \left| \widehat{\psi}(\omega) \right|^2 \frac{d\omega}{|\omega|} < \infty,
\]

where \( \widehat{\psi}(\omega) \) denotes the Fourier transform of \( \psi(x) \). The dilation of the \( \psi(x) \) by a factor \( s \) is denoted as

\[
\psi_s(x) = \frac{1}{s} \psi \left( \frac{x}{s} \right)
\]

and the wavelet transform of a function \( f(x) \) at scale \( s \) and position \( t \) is defined by

\[
Wf(s, t) = f^* \psi_s(x),
\]

where \( * \) denotes the convolution operator. The wavelet transform depends on two parameters \( s \) and \( t \) that vary continually over the set of real numbers. For practical applications these parameters must be discretized. Since the scale \( s \) decreases with the support of \( \psi_s(x) \), the wavelet transform \( Wf(s, t) \) is sensitive to the finer details of a contour [21]. That is, the scale \( s \) characterizes the size and the regularity of the signal features extracted by the wavelet transform.

As we know, the convolution of a function \( f(x) \) with a smoothing function will attenuate the higher frequencies part without modifying the lower frequencies part. In the past, various smoothing functions have been proposed in the analysis of signal and images. Based on the superior property of localizing the extrema of a signal profile, the first derivative of the Gaussian function is used as the mother wavelet \( \psi(x) \),

\[
\psi(x) = -x \exp \left( -\frac{x^2}{2} \right).
\]
2.2. Multiscale Contour Representation

A planar contour \( C(t) \) is usually represented in terms of \( X(t) \) and \( Y(t) \) coordinates, i.e., \( C(t) = \{ X(t), Y(t) \} \), where \( t \) is the arc length along the contour from a certain starting point. The orientation is defined as

\[
\phi(t) = \tan^{-1} \frac{dy}{dx}.
\]  

(5)

Due to the quantization error in discretization and the limited orientation resolution, equation (5) is not suitable for obtaining the orientation parameter at point \( P_i \). To reduce the quantization error and increase the orientation resolution, a smoothing level \( q \), \( q > 1 \), is used. Then, the orientation at point \( P_i \) can be expressed as

\[
\phi(t) = \tan^{-1} \frac{y_{i+q} - y_{i-q}}{x_{i+q} - x_{i-q}}.
\]  

(6)

Here, \( q \) is set to 5 to obtain the smoothed orientation function \( \phi(t) \), which has lower quantization error and higher orientation resolution of the object contour. The sampling length of the curvature function convolved with the wavelet function is determined by the support of wavelet function at a certain scale. Each part selected from the curvature function exists an offset, which is useful for constructing the MFS. The offset of the part corresponding to the convolution with wavelet function at the scale \( s \) and position \( t \) is denoted as \( \text{Off}(s, t) \),

\[
\text{Off}(s, t) = \int_{t-w_s/2}^{t+w_s/2} \phi(u) \, du,
\]  

(7)

where \( w_s \) is the support length of the wavelet function at the scale \( s \). When performing the wavelet transform on the curvature function, the offset should be calculated and subtracted from the curvature function to get the fluctuation of it.

Let \( A \) and \( B \) be two nonzero vectors; the scalar product of them can be obtained by

\[
A \cdot B = |A||B| \cos \theta,
\]  

(8)

where \( |A|, |B| \) are the magnitude of \( A \) and \( B \), respectively, \( \theta \) is the interior angle between \( A \) and \( B \). Rewriting equation (8), the term \( \cos \theta \) can be derived by

\[
\cos \theta = \frac{A \cdot B}{|A||B|}.
\]  

(9)

From equation (9), it is clear that \( A = B \) if and only if \( |A| = |B| \) and \( \cos \theta = 1 \). For \( |A| \leq |B| \), we define the similarity measure (SM) of two vectors as following:

\[
\text{SM} = \frac{|A|}{|B|} \cos \theta
\]  

(10)

or

\[
\text{SM} = \frac{A \cdot B}{|B|^2}.
\]  

(11)

The maximum value of the similarity is 1 corresponding to the condition \( A = B \). In other words, the similarity can be regarded as the degree of one vector fitting to the reference vector and \( \text{SM} = 1 \) is the unique complete fitting.

The wavelet transform \( W \phi(s, t) \) of the curvature of a curve at the scale \( s \) and position \( t \) can be expressed by

\[
W \phi(s, t) = \phi^* \psi_s(t)
\]

\[
= \int_{-\infty}^{\infty} (\phi(u) - \text{Off}(s, t)) \psi_s(u - t) \, du
\]

\[
= \int_{-\infty}^{\infty} (\phi(u) - \text{Off}(s, t)) \frac{(u - t)}{s} \exp \left( -\frac{(u - t)^2}{s^2} \right) \, du.
\]  

(12)
Since the wavelet function selected is antisymmetric, $\psi(-x) = -\psi(x)$. Hence, the absolute value of the wavelet transform is equal to the absolute value of scalar product of two functions, i.e.,

$$|W\phi(s,t)| = |\langle \phi, \psi_s \rangle|$$

$$= \int_{t-w_s/2}^{t+w_s/2} (\phi(u) - \text{Off}(s,t))\psi_s(u)\,du. \quad (13)$$

Rearranging equation (13), the similarity measure $SM(s,t)$ is then obtained by

$$SM(s,t) = \frac{|W\phi(s,t)|}{|\phi||\psi_s|}, \quad (14)$$

where $|\phi|$ is the amplitude of signal function extracted from the orientation function which the offset value is subtracted and calculated by

$$|\phi| = \left[ \int_{t-w_s/2}^{t+w_s/2} (\phi(u) - \text{Off}(s,t))^2\psi_s(u)\,du \right]^{1/2} \quad (15)$$

and $|\psi_s|$, the amplitude of support of wavelet function at scale $s$, is calculated by

$$|\psi_s| = \left[ \int_{-w_s/2}^{w_s/2} \psi_s(u)^2\,du \right]^{1/2}. \quad (16)$$

In this approach, the scale $s$ is used as a main feature from the evolution of the similarity measure $SM(s,t)$ if it exceeds a threshold value. In addition, other parameters, including position $t$, the amplitude $|\phi|$, the $\text{Off}(s,t)$ value of the sampling part of the curvature function which is used to convolve with wavelet function at the scale $s$ and position $t$, are also recorded.

### 3. FEATURE EXTRACTION ALGORITHM AND SHAPE MATCHING

Most of the previous methods for shape representation are based on the dominant points (corner points) which contain important information of the object shape, and have been successfully used in some applications, such as pattern recognition. Nevertheless, the information of the contour does not only exist a set of feature points. In other words, to represent a contour, one should consider both the feature points and other significant signatures derived from the contour. The more information the selected features consist of, the better representation of the contour obtained. Based on this concept, other features that simultaneously contain dominant points and the shape information neighbor on it are used. These features used in this approach consist of a set of multiscale features. The following is the details about the MFS extraction algorithm.

#### 3.1. MFS Algorithm

**STEP 1.** Parameterize the boundary contour $C(t)$ of an object shape by using the path length $t$ and calculate the smoothed orientation function $\phi(t)$ using equation (6).

**STEP 2.** Give the scale $s$ to sample the wavelet function $\psi_s(x)$ and determine the support $w_s$ with a preset value

$$\psi_s(x) = \sum_{n=-w_s/2}^{w_s/2} \psi(n) - \frac{w_s}{2} < n < \frac{w_s}{2}.$$  

**STEP 3.** Capture the orientation data by setting the range from $t - w_s/2$ to $t + w_s/2$, and obtain alternative component $\phi_{al,s}(t)$ by subtracting its offset, $\text{Off}(s,t)$. The alternative component $\phi_{al,s}(t)$ can be expressed by

$$\phi_{al,s}(t) = \sum_{n=-w_s/2}^{w_s/2} \phi(n - t) - \text{Off}(s,t),$$
where \( \text{Off}(s, t) \) can be described by

\[
\text{Off}(s, t) = \sum_{n=-w_s/2}^{w_s/2} \frac{\phi(n-t)}{w_s}.
\]

**STEP 4.** Compute wavelet transform of the alternative component \( \phi_{al,s}(t) \)

\[
\phi_{al,s}(s, t) = \phi_{al,s}(t) \ast \psi_s(t) = \sum_{n=-w_s/2}^{w_s/2} \phi_{al,s}(n-t) \times \psi_s(n).
\]

**STEP 5.** Obtain the similarity measure \( SM(s, t) \), i.e.,

\[
SM(s, t) = \frac{W \phi_{al,s}(s, t)}{|\phi_{al,s}| |\psi_s|},
\]

where \( |\phi_{al,s}| \) and \( |\psi_s| \) are defined as

\[
|\phi_{al,s}| = \left[ \sum_{n=-w_s/2}^{w_s/2} (\phi_{al,s}(n))^2 \right]^{1/2}
\]

and

\[
|\psi_s| = \left[ \sum_{n=-w_s/2}^{w_s/2} (\psi_s(n))^2 \right]^{1/2},
\]

respectively.

**STEP 6.** Increase \( s \) and repeat Steps 2–5 to obtain the evolution of the similarity measure \( SM(s, t) \) until \( s \) reach to a certain value. (In our case, \( s < 20 \), it is obtained from the experiment.)

**STEP 7.** Find the local extrema of those \( SM(s, t) \) and extract the scale \( s \) as the main feature. The corresponding parameters including \( SM(s, t) \) value, position \( t \), the \( \text{Off}(s, t) \), and the amplitude \( |\phi_{al,s}| \) are also recorded.

When the parameters are extracted by using the above algorithm, a hierarchical structure of these parameters, namely MFS, is constructed for a sequence of scales \( s \), the value of similarity measure \( SM(s, t) \), position \( t \), the amplitude \( |\phi_{al,s}| \), and the \( \text{Off}(s, t) \). Then the representation of a contour is completed.

### 3.2. Shape Matching Strategy

The performance of any shape recognition methods relies on the quality of the feature measurements provided. The best way to improve the recognition capability is by extending the feature set and selecting the best feature out of this set. Based on this concept, we propose a multiresolution object recognition strategy using the MFSs of each object shape. Since each component of an MFS has different significance in representing the shape, it is necessary to multiply a weighting factor with each component of the MFS. That is, an MFS can be denoted as

\[
\text{MFS} = \{k_1s, k_2SM(s, t), k_3t, k_4|\phi_{al,s}|, k_5\text{Off}(s, t)\},
\]

where \( k_i, i = 1, \ldots, 5 \), are the weighting factors, respectively. After obtaining the overall MFSs of each test object shape, a statistical classifier based on the nearest-neighbor clustering rule is used.
to shape matching. Here, we utilize ten classes for the training shapes and each class consists of thirty templates. The strategy of the classifier is summarized as following.

Let the $r$th MFS of an object shape $u$ in class $a$ be denoted as $MFS_{r,a}^{a,u} = (f_{r_1}^{a,u}, f_{r_2}^{a,u}, f_{r_3}^{a,u}, f_{r_4}^{a,u}, f_{r_5}^{a,u})^t$, where $f_{r_1}^{a,u} = k_1 s_{r_1}^{a,u}$, $f_{r_2}^{a,u} = k_2 (SM)_{r_2}^{a,u}$, $f_{r_3}^{a,u} = k_3 t_{r_3}^{a,u}$, $f_{r_4}^{a,u} = k_4 |\psi_{a,u}|_{a,u}$, and $f_{r_5}^{a,u} =
Figure 7. The orientation function of the object shapes shown in Figure 6.
Denote the mean and the standard deviation of the $r^{th}$ MFS in class $a$ as $\overline{\text{MFS}}_r$ and $\sigma_r^a$, respectively. The distance between a test shape $T = (t_1, t_2, \ldots, t_m)^T$ and $\text{MFS}^a$ can be expressed as

$$D(T, \text{MFS}^a) = \sum_{i,j=1}^{m,n} \frac{|t_i - \overline{\text{MFS}}_j|}{\sigma^a_j},$$

where $m, n$ are the total numbers of the MFSs for an object shape in class $a$ and a test shape, respectively.
Decision criteria: if the overall MFSs of an object shape in class \( a \) have the minimum distance from \( T \) among all the training shapes, we declare that the unknown shape belongs to class \( a \), i.e.,

\[ T \in \text{class} \ a, \quad \text{if} \ D(T, \text{MFS}^a) \leq D(T, \text{MFS}^b), \quad \text{for} \ 1 \leq a, b \leq 10. \]

4. EXPERIMENTAL RESULTS

Many experiments have been carried out on different object shapes. We implement the algorithm using C programming language and run it on a personal computer-PC586. These objects

Figure 8. The evolution of SM(s, t) values of the object shapes shown in Figure 6 at different scales.
are captured with a CCD camera and digitized to 8-bits gray level and stored as 512×512 bytes. Next, by choosing appropriate threshold, we obtain the binary images, which contain only 0 and 1 value in each pixel where 1 denotes the object and 0 denotes the background, refer to Figure 6. Before obtaining the boundary image of each object, the techniques including edge detection and thinning [23], are executed in order. Finally, the boundary of the object is extracted by using a boundary tracing algorithm [24].

For example, Figure 2 illustrates the thinned image of a pair of pliers. The orientation function of Figure 2 is shown in Figure 3. In this figure, the abscissa is the parameter, path length, and the vertical is the angle parameter ranging from $-\pi$ to $\pi$.

Figure 4 shows the evolution of the similarity measure $SM(s, t)$ value of Figure 3 over some different scales. If the local extrema of $SM(s, t)$ value exceeds a preset threshold, an MFS
Figure 9. The positions of the MFSs on the text object shapes.

is recorded at this local extrema point. More specifically, a real extrema should appear on
the evolution map such as Figure 3 cross various scales, but a false extrema does not. After
obtaining all the MFSs using the MFS extraction algorithm, their positions are marked on the origin boundary (see Figure 5). In Figure 5, we observe that the positions of these MFSs are exactly on the positions of the dominant points of the shape. The MFSs of a pair of pliers is listed in Table 1.

<table>
<thead>
<tr>
<th>Scale (s)</th>
<th>Position (t)</th>
<th>Similarity Measure SM(s,t)</th>
<th>Amplitude</th>
<th>Offset Off(s,t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>414</td>
<td>0.859</td>
<td>0.254</td>
<td>2.816</td>
</tr>
<tr>
<td>7</td>
<td>862</td>
<td>0.823</td>
<td>0.288</td>
<td>2.812</td>
</tr>
<tr>
<td>8</td>
<td>160</td>
<td>0.929</td>
<td>0.317</td>
<td>2.809</td>
</tr>
<tr>
<td>8</td>
<td>678</td>
<td>0.833</td>
<td>0.317</td>
<td>2.809</td>
</tr>
<tr>
<td>8</td>
<td>844</td>
<td>0.881</td>
<td>0.317</td>
<td>2.809</td>
</tr>
<tr>
<td>9</td>
<td>131</td>
<td>0.864</td>
<td>0.343</td>
<td>2.806</td>
</tr>
<tr>
<td>9</td>
<td>541</td>
<td>0.836</td>
<td>0.365</td>
<td>2.803</td>
</tr>
<tr>
<td>10</td>
<td>229</td>
<td>0.864</td>
<td>0.406</td>
<td>2.799</td>
</tr>
<tr>
<td>10</td>
<td>335</td>
<td>0.872</td>
<td>0.406</td>
<td>2.799</td>
</tr>
<tr>
<td>14</td>
<td>283</td>
<td>0.852</td>
<td>0.570</td>
<td>2.776</td>
</tr>
<tr>
<td>14</td>
<td>372</td>
<td>0.856</td>
<td>0.570</td>
<td>2.776</td>
</tr>
<tr>
<td>14</td>
<td>1076</td>
<td>0.814</td>
<td>0.570</td>
<td>2.776</td>
</tr>
<tr>
<td>16</td>
<td>125</td>
<td>0.862</td>
<td>0.648</td>
<td>2.771</td>
</tr>
<tr>
<td>16</td>
<td>642</td>
<td>0.817</td>
<td>0.648</td>
<td>2.771</td>
</tr>
<tr>
<td>19</td>
<td>868</td>
<td>0.859</td>
<td>1.050</td>
<td>2.778</td>
</tr>
</tbody>
</table>

To evaluate the performance of this scheme, 300 images including ten kinds of objects under various locations, orientations and scales are used in the experiment. Some of them are shown in Figure 6. Their orientation function and the evolution of the similarity measure SM(s,t) value over several scales are then shown in Figures 7 and 8, respectively. Figure 9 shows the positions of the MFS on the origin shapes. In the experiments, all the test patterns are recognized correctly and this proves the feasibility of the proposed method. The details of the experimental results are included in [25].

5. CONCLUSIONS

In this paper, we propose a new algorithm to represent the two-dimensional object contour using the wavelet transform. The representation of a contour is achieved with a collection of MFS, which are extracted by using the MFS extraction algorithm. It is an essential hierarchy presentation. The MFSs of a contour occurs when the local extrema of the similarity measure SM(s,t) evolution exceeds a preset threshold value. A MFS consists of the scale parameters of wavelet function, the positions where the MFSs take place, the similarity measure SM(s,t) which describes the degree of similarity, the amplitude of alternative component of the signal function and the offset value of the signal function. One statistical classifier based on the nearest-neighbor clustering rule is then utilized to shape recognition. Experimental results that prove the feasibility of the proposed algorithm are also included.

REFERENCES