

# Wavelet Descriptor of Planar Curves: Theory and Applications

Gene C.-H. Chuang and C.-C. Jay Kuo, *Senior Member, IEEE*

**Abstract**—By using the wavelet transform, we develop a hierarchical planar curve descriptor that decomposes a curve into components of different scales so that the coarsest scale components carry the global approximation information while the finer scale components contain the local detailed information. We show that the wavelet descriptor has many desirable properties such as multiresolution representation, invariance, uniqueness, stability, and spatial localization. A deformable wavelet descriptor is also proposed by interpreting the wavelet coefficients as random variables. The applications of the wavelet descriptor to character recognition and model-based contour extraction from low SNR images are examined. Numerical experiments are performed to illustrate the performance of the wavelet descriptor.

## I. INTRODUCTION

**B**OUNDARY representation is essential in shape description and recognition. It plays a key role in many applications such as image analysis, pattern recognition, computer graphics, and computer-aided animation. Many methods have been proposed to describe planar curves. The chain coding method [16] approximates a curve with a sequence of directional vectors lying on a square grid. Some quantitative measurements of the object, known as the shape factors [9], can be used to characterize shapes. Examples include moments, areas, perimeters, and corners. The spline technique [5] uses a set of piecewise low-order polynomials to approximate a curve so that the curve can be determined by a small number of parameters. The spline-generated boundary curve can be translated, scaled, and rotated by performing the transformations on a set of control points. The Fourier descriptor [27], [40] describes a curve with coefficients via Fourier analysis of a certain parametric representation such as the curvature or spatial coordinates of the curve.

The multiresolution (or multiscale) technique [3], [28], [31] for signal and image processing has grown very rapidly for the last several years. It was observed by study in psychophysics that the human visual system processes and analyzes image information at different resolutions. This motivated Rosenfeld [29] to develop a multiscale edge detection scheme. Marr [22] used the first- or second-order derivatives of a Gaussian

filter of different sizes for signal convolution and detected changes in the intensity or zero-crossing at different scales. More recently, Witkin [39] proposed a scale-space filtering approach that uses a set of parameterized Gaussian kernels to smooth a signal and extracts the descriptive primitives such as the locations of extrema or intervals bounded by the extrema. To characterize the features of local extremum, one can then detect the extrema at a certain coarse scale and track them across a couple of scales of finer resolutions.

For most curve matching or shape recognition tasks, it is important that the decision is made based on features that are insensitive to rotation, translation, and change in size. To match contours in a hierarchical manner is one of the main motivation for developing the scale-space methods so that the global alignment is conducted first, and local comparison is then performed at various resolutions. Fermeuller and Kropatsch [15] presented a multiresolution descriptor of planar curves using corners with a hierarchical structure. Based on the scale-space plot of the curvature of a planar curve, Mokhtarian and Mackworth [23] proposed a multiscale shape representation that locates points of inflection on the curve at varying scales. Moreover, Bajcsy and Kovacic [1] proposed a multiresolution elastic matching method in which they assumed that one of two objects was made of elastic material and the other served as a reference. By applying an external force to the elastic object, the shape of the elastic object was deformed to match the reference object along some scales. This matching method was shown to be effective in medical applications such as anatomical human brain atlas matching. For the scale-space methods, the Gaussian function is the optimal kernel for reducing noise with minimal delocalization [33], [21]. Nevertheless, a planar curve smoothed by the Gaussian kernel suffers from shrinkage [25], i.e., the perimeter becomes smaller after convolving with the Gaussian kernel. This may cause a problem in some pattern recognition applications.

In recent years, the wavelet transform became an active area of research for multiresolution signal and images analysis. In this paper, we consider another type of multiscale planar curve descriptor by using the wavelet transform. Our main objective is to demonstrate the nice properties and interesting applications of this new descriptor. We show that the wavelet descriptor has the properties of invariance, uniqueness, and stability. Features extracted from the wavelet approach can be normalized so that we can handle the effect of rotation, translation, and scaling easily. In contrast with the scale-space filtering approach, which serves primarily as an analytical

Manuscript received August 26, 1993; revised May 15, 1995. This work was supported by the National Science Foundation Presidential Faculty Fellow (PFF) Award ASC-9350309. The associate editor coordinating the review of this paper and approving it for publication was Dr. Michael Unser.

G. C.-H. Chuang is with Philips EBEI (Taiwan) Ltd., Taipei, Taiwan (e-mail: gene@rtptxt01.serigate.philips.nl).

C.-C. J. Kuo is with the Signal and Image Processing Institute and the Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: cckuo@sipi.usc.edu).

Publisher Item Identifier S 1057-7149(96)00138-8.

tool, the wavelet descriptor provides an effective synthesis tool as well. The application of the wavelet transform does not yield a coarse resolution curve smaller than the original one so that there is no shrinkage problem, even though some detailed variations of the curve are removed. Compared with the Fourier descriptor that uses global sinusoids as the basis functions, the wavelet descriptor is more efficient in representing and detecting local features of a curve due to the spatial and frequency localization property of wavelet bases. Furthermore, we may interpret the wavelet coefficients as random variables and use the deformable stochastic wavelet descriptor to model a group of shapes that have the same topological structure but may differ slightly due to local deformations so that this descriptor can be conveniently used in multiscale matching. Since the forward and inverse wavelet transforms can be implemented via the cascade of quadrature mirror filter banks, the wavelet descriptor is computationally efficient and has a great potential for real-time applications such as target recognition and detection.

This paper is organized as follows. We briefly review the wavelet theory and derive a wavelet representation for planar curves in Section II. In Section III, we examine several important features of the wavelet descriptor, including the effect due to scaling, translation, and rotation and the properties of invariance, uniqueness, and stability. These properties are essential for shape representation and recognition. The approximation capabilities of different wavelet bases with various vanishing moments and symmetry property are compared. In addition, we propose a procedure to normalize the wavelet coefficients contained in the feature set so that the recognition method is applicable to shapes with different sizes and variations. In Section IV, we study the application of the deformable wavelet descriptor to the model-based boundary extraction from noisy images using the maximum *a posteriori* (MAP) estimation method and perform an experiment to illustrate the performance of the wavelet descriptor. Some concluding remarks are given in Section V.

## II. PLANAR CURVE DESCRIPTOR USING WAVELET TRANSFORM

The parametrized closed curves can be represented by periodic sequences. Wavelets defined in  $L^2(R)$  are not suitable for this representation. Continuous wavelet analysis on the circle have been studied by Holschneider [17]. In this section, we will briefly review the theory of *periodized* wavelets [12]. Each periodized wavelet can be expressed as a sum of copies of periodically shifted continuous wavelets with reasonable decay. These functions constitute an orthonormal basis in the space  $L^2([0, 1])$ .

### A. Review of Periodized Wavelet Theory

We use  $\phi(t)$  to denote a scaling function such that for a certain  $m \in \mathbf{Z}$ , its translations

$$\phi_n^m(t) = 2^{-m/2} \phi(2^{-m}t - n), \quad n \in \mathbf{Z}$$

form an orthonormal basis for the wavelet subspaces  $V_m$  and that  $\{V_m\}_{m \in \mathbf{Z}}$  is a multiresolution approximation of the space

$L^2(R)$ . For each scaling function  $\phi(t)$ , one can determine the corresponding mother wavelet function  $\psi(t)$  such that the collection of its dilations and translations

$$\psi_n^m(t) = 2^{-m/2} \psi(2^{-m}t - n), \quad m, n \in \mathbf{Z}$$

form an orthonormal basis for  $L^2(R)$ . The functions  $\phi$  and  $\psi$  satisfy the following dilation equations:

$$\phi(t) = \sqrt{2} \sum_n h_n \phi(2t - n); \quad \psi(t) = \sqrt{2} \sum_n g_n \phi(2t - n).$$

The coefficients  $h_k$  and  $g_k$  are related via  $g_k = (-1)^k h_{1-k}$ .

The periodized wavelets in the space  $L^2([0, 1])$  can be examined based on the multiresolution analysis with the scaling function  $\phi$  and the wavelet  $\psi$  in  $L^2(R)$ . The periodic scaling and wavelet functions are defined as

$$\tilde{\phi}_n^m(t) = \sum_{l \in \mathbf{Z}} \phi_n^m(t + l); \quad \tilde{\psi}_n^m(t) = \sum_{l \in \mathbf{Z}} \psi_n^m(t + l). \quad (2.1)$$

The corresponding periodic multiresolution approximation spaces are

$$\tilde{V}^m = \overline{\text{Span}\{\tilde{\phi}_n^m; n \in \mathbf{Z}\}},$$

and

$$\tilde{W}^m = \overline{\text{Span}\{\tilde{\psi}_n^m; n \in \mathbf{Z}\}}. \quad (2.2)$$

In addition, as in the nonperiodic case, we have  $\tilde{V}^{m-1} = \tilde{V}^m \oplus \tilde{W}^m$ . It was proved in [12] that for negative integer  $m$ ,  $\tilde{V}^m$  is finite-dimensional  $\tilde{\phi}_{n+k2^{|m|}}^m(t) = \tilde{\phi}_n^m(t)$  for  $k \in \mathbf{Z}$ , and  $\tilde{V}^m$  is spanned by the  $2^{|m|}$  functions with  $n \in \mathbf{Z}_m \equiv \{0, 1, \dots, 2^{|m|} - 1\}$ . A similar result holds for  $\tilde{W}^m$  with  $\tilde{\phi}_n^m(t)$  replaced by  $\tilde{\psi}_n^m(t)$ .

For  $f(t) \in V_{M_f}$ , we can express its finite-scale orthogonal wavelet expansion

$$\begin{aligned} f(t) &= \sum_{n \in \mathbf{Z}_{M_f}} c_n^{M_f} \tilde{\phi}_n^{M_f}(t) \\ &= \sum_{n \in \mathbf{Z}_{M_c}} c_n^{M_c} \tilde{\phi}_n^{M_c}(t) + \sum_{m=M_f}^{M_c} \sum_{n \in \mathbf{Z}_m} d_n^m \tilde{\psi}_n^m(t) \end{aligned} \quad (2.3)$$

with

$$c_n^{M_c} = \int_0^1 f(t) \tilde{\phi}_n^{M_c}(t) dt; \quad d_n^m = \int_0^1 f(t) \tilde{\psi}_n^m(t) dt. \quad (2.4)$$

A fast algorithm to compute the finite-scale wavelet transform is given below. Let us define a pair of filter coefficients

$$\begin{aligned} \tilde{g}_l^m &\equiv \sum_{k \in \mathbf{Z}} g_{(l \bmod 2^{|m|}) + 2^{|m|}k}; \\ \tilde{h}_l^m &\equiv \sum_{k \in \mathbf{Z}} h_{(l \bmod 2^{|m|}) + 2^{|m|}k}, \quad l \in \mathbf{Z}. \end{aligned} \quad (2.5)$$

It is easy to verify that  $\tilde{g}_l^m$  and  $\tilde{h}_l^m$  are periodic sequences with period  $2^{|m|}$ . Then, the coefficients  $d_n^m$  and  $c_n^{M_c}$  can be computed from coefficients  $c_n^{M_f}$  via the following recursive formulas:

$$\begin{aligned} c_n^{m+1} &= \sum_{l \in \mathbf{Z}_m} \tilde{h}_{l-2n}^m c_l^m; \quad d_n^{m+1} = \sum_{l \in \mathbf{Z}_m} \tilde{g}_{l-2n}^m c_l^m, \\ m &= M_f, \dots, M_c - 2, M_c - 1. \end{aligned} \quad (2.6)$$

One can also obtain the coefficients  $c_n^{M_f}$  from  $d_n^m$  and  $c_n^{M_c}$  via the synthesis formula

$$c_n^m = \sum_{l \in \mathbf{Z}_{m+1}} (\tilde{h}_{n-2l}^m c_l^{m+1} + \tilde{g}_{n-2l}^m d_l^{m+1}), \quad m = M_c - 1, M_c - 2, \dots, M_f. \quad (2.7)$$

Equations (2.6) and (2.7) are called the forward and inverse discrete periodic wavelet transforms (DPWT), respectively. It can be easily shown that the DPWT is perfectly invertible when applied to sequences of finite length.

### B. Coordinate-Based Wavelet Descriptor

Let us denote a clockwise-oriented closed plane curve with parametric coordinates  $x(t)$  and  $y(t)$  by

$$\alpha(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad t(l) = l/L, \quad 0 \leq l \leq L$$

where

- $t$  normalized arc length
- $l$  arc length along the curve from a certain starting point  $t_0$
- $L$  total arc length.

By applying the wavelet transform to the parameterized coordinates, we obtain

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} x_a^M(t) \\ y_a^M(t) \end{bmatrix} + \sum_{m=M-m_0}^M \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix} \quad (2.8)$$

where

$$x_a^M(t) = \sum_n a_n^M \tilde{\phi}_n^M(t), \quad y_a^M(t) = \sum_n c_n^M \tilde{\phi}_n^M(t) \quad (2.9)$$

are called the approximation coefficients at scale  $M$  and

$$x_d^m(t) = \sum_n r_n^m \tilde{\psi}_n^m(t), \quad y_d^m(t) = \sum_n d_n^m \tilde{\psi}_n^m(t) \quad (2.10)$$

are called the detailed signals at scale  $m$  with  $m = M - m_0$  being the finest scale and  $m = M$  being the coarsest scale. Then, we can use the wavelet coefficients  $a_n^M$ ,  $c_n^M$ ,  $r_n^m$ , and  $d_n^m$  given in (2.9) and (2.10) as the planar curve descriptor.

By using only a subset of wavelet coefficients consisting of primarily coarser scale components, i.e., small values of  $|m|$ , we can have different multiresolution representation of the shape. In terms of mathematical expression, we can modify (2.8) to be

$$\begin{bmatrix} \hat{x}(k; t) \\ \hat{y}(k; t) \end{bmatrix} = \begin{bmatrix} x_a^M(t) \\ y_a^M(t) \end{bmatrix} + \sum_{m=k}^M \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix} \quad (2.11)$$

where  $M - m_0 \leq k \leq M + 1$ . The curves with coordinates  $\hat{x}(k; t)$  and  $\hat{y}(k; t)$  provide a sequence of multiresolution approximations to the original curve. This dyadic approximation sequence can be best depicted by its two extremes. On one hand, we obtain the original curve at  $k = M - m_0$ , and on the other hand, the approximating curve contains only the coarsest scale description, i.e.,  $x_a^M(t)$  and  $y_a^M(t)$  for  $k = M + 1$ . To give an example, we show the Koch's snowflake in Fig. 1 and its different approximations with a certain subset of wavelet coefficients. We perform an eight-level decomposition

of wavelet transform using the biorthogonal cubic B-spline wavelets [34], [35]. Fig. 1(1)–(8) are its approximations from the finest to the coarsest resolution, whereas the original curve with 3072 samples is given in Fig. 1(0). In practice, it is proper for the coarsest-scale in shape representation to contain between 4 to 16 sampled points in applications such as shape matching. One main advantage of the wavelet descriptor is that the hierarchical decomposition can lead to significant data compression. For example, Fig. 1(3) shows the approximation of the curve reconstructed from 384 points. We see that the approximation is very close to the original.

The detailed signals at scale  $m$  can also be represented by using the polar coordinates, i.e.

$$\begin{aligned} \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix} &= \sum_n \begin{bmatrix} r_n^m \\ d_n^m \end{bmatrix} \tilde{\psi}_n^m(t) \\ &= \sum_n \begin{bmatrix} \cos \theta_n^m & -\sin \theta_n^m \\ \sin \theta_n^m & \cos \theta_n^m \end{bmatrix} \begin{bmatrix} A_n^m \\ 0 \end{bmatrix} \tilde{\psi}_n^m(t) \end{aligned} \quad (2.12)$$

where

$$\theta_n^m = \arctan\left(\frac{d_n^m}{r_n^m}\right), \quad \text{and} \quad A_n^m = \sqrt{(d_n^m)^2 + (r_n^m)^2}. \quad (2.13)$$

With this polar coordinate representation, we can plot a vector function with amplitude  $A_n^m \cdot \tilde{\psi}_n^m(t)$  and phase  $\theta_n^m$ . The wavelet vectors are the building elements of curves from a scale to its next higher resolution scale. The representation of wavelet vector depends on the wavelet basis  $\tilde{\psi}(t)$  it uses. In general,  $\tilde{\psi}(t)$  with shorter support offers better spatial locality property at the expense of the smoothness of reconstructed curves.

We want to point out similarities and differences between the expression in (2.12) and elliptic Fourier features discussed in [20] and [30], where the sinusoidal orthogonal bases, i.e., sine and cosine functions, were adopted. If the coefficients of the sine and cosine functions are equal, the building elements are circles. Otherwise, they are ellipses. The building elements are flattened to vectors if only the sine or the cosine functions are used. In (2.12), the wavelet curve descriptor is reduced to a vector via the Cartesian-to-polar coordinate transformation.

It is also possible to represent a smooth curve using the chain-coded tangent, cumulative tangent, or curvature. These kinds of differential representations [18], [40], are invariant under translation, rotation, and scaling. The curvature  $\kappa$  of the planar curve is defined as the rate of change of the tangent angle with respect to arc length, i.e.

$$\kappa(t) = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}}$$

where  $\dot{x}$  and  $\ddot{x}$  (or  $\dot{y}$  and  $\ddot{y}$ ) denote, respectively, the first and second derivatives of the coordinate function  $x(t)$  (or  $y(t)$ ) with respect to the parameter  $t$  of the curve representation. We can apply the finite-scale orthogonal wavelet expansion to the curvature via

$$\kappa(t) = \sum_n \alpha_n^M \tilde{\phi}_n^M(t) + \sum_{m=M-m_0}^M \sum_n \delta_n^m \tilde{\psi}_n^m(t).$$

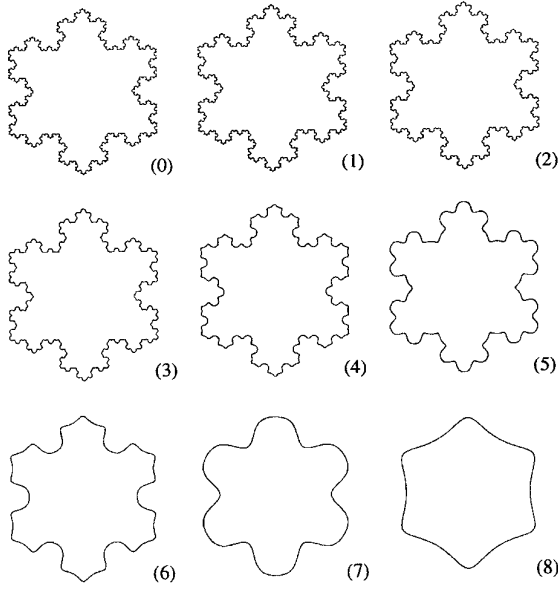


Fig. 1. Multiscale representation of Koch's snowflake using the biorthogonal spline wavelet.

However, for a closed curve with differential representation, its reconstruction based on a set of truncated wavelet coefficients may not be closed. Besides, it is difficult to compute the derivatives at sharp corners or under a noise-corrupted environment. Thus, we will focus on the coordinate-based wavelet descriptor in this paper.

It is worthwhile to review Fourier descriptor, which is a popular curve descriptor in many applications, for the comparison at the end of this section. The Fourier descriptor is a method of describing a closed planar curve by a set of Fourier coefficients. Let  $(x(n), y(n))$  be the  $N$ -point discrete parameterized coordinates of a curve. By assuming the contour to be traced repeatedly, the coordinate functions  $x(n)$  and  $y(n)$  are periodic functions and can be represented by using the Fourier series as

$$\begin{aligned} x(n) &= \sum_{k=0}^{N-1} X(k) e^{i2\pi nk/N} \\ y(n) &= \sum_{k=0}^{N-1} Y(k) e^{i2\pi nk/N}, \quad n = 0, 1, 2, \dots, N-1 \end{aligned} \quad (2.14)$$

where  $X(k)$  and  $Y(k)$  are the Fourier coefficients. To reconstruct the closed planar curve largely, one may only need a small set of coefficients of low-frequency components when the coefficients of higher frequency components are small.

### III. PROPERTIES OF WAVELET DESCRIPTOR

#### A. Scaling, Translation, and Rotation

The scaling, translation, and rotation of a planar curve can be described via a suitable transformation of wavelet coefficients. First, we examine the scaling effect. The scaling

of a curve by a factor  $\beta$  can be written as

$$\begin{aligned} \begin{bmatrix} \tilde{x}(t) \\ \tilde{y}(t) \end{bmatrix} &= \beta \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \beta \begin{bmatrix} x_a^M(t) \\ y_a^M(t) \end{bmatrix} \\ &+ \sum_{m=M-m_0}^M \beta \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix}. \end{aligned}$$

Thus, by using the linearity of the wavelet transform, it implies that we can scale the wavelet coefficients  $a_n^M$ ,  $c_n^M$ ,  $r_n^m$ , and  $d_n^m$  given in (2.9) and (2.10) by the same factor  $\beta$

$$\begin{aligned} \text{Scaling: } \begin{bmatrix} \tilde{a}_n^M \\ \tilde{c}_n^M \end{bmatrix} &= \beta \begin{bmatrix} a_n^M \\ c_n^M \end{bmatrix}, \\ \begin{bmatrix} \tilde{r}_n^m \\ \tilde{d}_n^m \end{bmatrix} &= \beta \begin{bmatrix} r_n^m \\ d_n^m \end{bmatrix}. \end{aligned} \quad (3.1)$$

As far as the polar coordinates (2.12) are concerned, we have

$$\tilde{\theta}_n^m = \arctan \left( \frac{\beta d_n^m}{\beta r_n^m} \right) = \arctan \left( \frac{d_n^m}{r_n^m} \right) = \theta_n^m$$

and

$$\tilde{A}_n^m = \sqrt{(\beta d_n^m)^2 + (\beta r_n^m)^2} = \beta \sqrt{(d_n^m)^2 + (r_n^m)^2} = \beta A_n^m.$$

Thus, the phase remains unchanged while the magnitude is scaled by the same factor  $\beta$ .

Next, we examine the translation of the curve by a distance  $(b_x, b_y)$ . By using (2.1) and the *admissibility* property of wavelet basis, we have

$$\int_0^1 \tilde{\psi}_n^m(t) dt = \sum_{l \in \mathbf{Z}} \int_l^{l+1} \psi(t') dt' = 0$$

and it is easy to see that

$$\int_0^1 [f(t) + b] \tilde{\psi}_n^m(t) dt = \int_0^1 f(t) \tilde{\psi}_n^m(t) dt, \quad m \in -\mathbf{N}; n \in \mathbf{Z}_m.$$

Thus, the wavelet coefficients  $r_n^m$  and  $d_n^m$  of the detailed signals  $x_d^m(t)$  and  $y_d^m(t)$  are invariant under translation. The displacement of the curve affects only the approximation coefficients  $x_a^M(t)$  and  $y_a^M(t)$ . Since

$$\int_0^1 \tilde{\phi}_n^M(t) dt = \sum_{l \in \mathbf{Z}} \int_0^1 \phi_n^M(t+l) dt = \int_{-\infty}^{\infty} \phi_n^M(t) dt$$

and  $\phi_n^M(t) = 2^{-M/2} \phi(2^{-M}t - n)$ , we have

$$\int_0^1 \tilde{\phi}_n^M(t) dt = 2^{-M/2} \int_{-\infty}^{+\infty} \phi(t) 2^M dt = 2^{M/2}$$

where the property that the integral of the mother scaling function is equal to unity is used. It follows that

$$\int_0^1 [f(t) + b] \tilde{\phi}_n^M(t) dt = \int_0^1 f(t) \tilde{\phi}_n^M(t) dt + 2^{M/2} \cdot b, \quad n \in \mathbf{Z}_m.$$

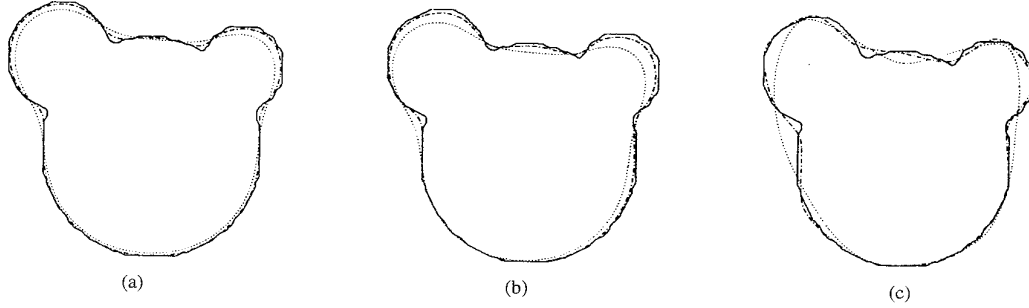


Fig. 2. Perimeter of a curve is reduced by applying multiple Gaussian and discrete diffusion smoothing steps as shown, respectively, in (a) and (b), whereas the wavelet approximation does not have the shrinkage problem, as shown in (c).

In addition, it can be summarized as

$$\text{Translation: } \begin{bmatrix} \tilde{a}_n^M \\ \tilde{c}_n^M \end{bmatrix} = \begin{bmatrix} a_n^M + 2^{M/2} \cdot b_x \\ c_n^M + 2^{M/2} \cdot b_y \end{bmatrix},$$

and

$$\begin{bmatrix} \tilde{r}_n^M \\ \tilde{d}_n^M \end{bmatrix} = \begin{bmatrix} r_n^M \\ d_n^M \end{bmatrix}.$$

Third, by rotating the curve by a counterclockwise angle  $\varphi$  with the centroid as the pivot point, we have

$$\text{Rotation: } \begin{bmatrix} \tilde{a}_n^m \\ \tilde{c}_n^m \end{bmatrix} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} a_n^m \\ c_n^m \end{bmatrix}$$

$$\begin{bmatrix} \tilde{r}_n^m \\ \tilde{d}_n^m \end{bmatrix} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} r_n^m \\ d_n^m \end{bmatrix}.$$

It is more convenient to describe the rotation in terms of polar coordinates and straightforward to derive that

$$\text{Rotation: } \tilde{\theta}_n^m = \theta_n^m + \varphi, \quad \text{and} \quad \tilde{A}_n^m = A_n^m$$

for the wavelet coefficients of the detailed signals. The same relationship also holds for the polar coordinate representation of the approximation coefficients.

In matching or recognition applications, it is known that the features selected as descriptors should be as insensitive as possible to the variation of changes in size, translation, and rotation. Curve normalization by using these properties will be described in detail in Section III-E.

### B. Nonshrinking Property

One common problem in multiresolution curve representation occurs when the curve convolves many times with some smoothing kernels. That is, the perimeter of a shape may become smaller. This is known as the shrinkage problem [25]. For example, let us consider the Gaussian and discrete diffusion smoothing kernels that satisfy analog and discrete scale-space conditions [21], respectively. Note that the family of discrete diffusion smoothing kernel is of the form  $T(n; t) = e^{-t} I_n(t)$ , where  $I_n(t)$  is the Bessel functions of integer order. The kernel  $T(n; t)$  is the solution of a discretized version of the diffusion equation and is proved to be the discrete counterpart of the Gaussian kernel for scale-space filtering. By comparing Fig. 2(a) and (b), we see that the perimeter of a curve is reduced by applying multiple times of Gaussian or discrete

diffusion smoothings. In Fig. 2, the solid lines are the original curves, whereas the dashdot and dotted lines represent two curves as the results of multiple smoothing. The shrinkage problem in the Gaussian and discrete diffusion smoothings is due to the reduction in low- as well as high-frequency components. In contrast, the wavelet approximations have no shrinkage problem, as shown in Fig. 2(c). For the proof of the nonshrinking property of the wavelet transform, we refer to work of Oliensis [25].

### C. Invariance, Uniqueness, and Stability

Mokhtarian and Mackworth [24] discussed some desired properties of a general shape descriptor as stated the following:

*Invariance:* For two curves with the same shape, they should have the same representation.

*Uniqueness:* For two curves with different shapes, they should have different representations.

*Stability:* Small differences in the shapes of curves correspond to small differences in their representations, and vice versa.

We can rephrase the three properties with commonly understood mathematical terminologies. First, the invariance property means that the mapping of a shape to its representation in fact defines a function. Second, the uniqueness property says that the function is one to one. Third, the stability property implies that the function is well posed. It is clear that the wavelet descriptor given by (2.8) is a one-to-one function so that it satisfies the invariance and uniqueness properties. The stability property will be examined below.

Since the  $x$  and  $y$  coordinates of a curve can be handled separately, we will focus on the scalar function rather than the vector function. Here, to present our explanation in a more general setting, we consider a class of square-integrable functions  $f \in L^2([0, 1])$  and their corresponding wavelet frame representations [7], [12]. (Note that the frame is a concept that is more general than the basis. A basis is a frame, but a frame may not be a basis.) Based on the property of a Hilbert space, we want to explain that if two functions  $f_1$  and  $f_2$  are close to each other, their wavelet frame representations  $W(f_1)$  and  $W(f_2)$  are also close, and vice versa.

By choosing  $\psi(x)$  such that functions

$$\{\tilde{\phi}_0^0\} \cup \{\tilde{\psi}_n^m(x)\}_{m \in -N; n \in \mathbb{Z}_m}$$

constitute a frame in  $L^2([0, 1])$ , we have [7], [12, pp. 55–56].

$$\begin{aligned} A\|f\|^2 &\leq |\langle f, \tilde{\phi}_0 \rangle|^2 \\ &+ \sum_{m \in -\mathbf{N}} \sum_{n \in \mathbf{Z}_m} |\langle f, \tilde{\psi}_n^m \rangle|^2 \leq B\|f\|^2, \\ f &\in L^2([0, 1]) \end{aligned} \quad (3.4)$$

for  $0 < A < B < \infty$ . Define the 2-norm of the wavelet representation

$$\begin{aligned} \|W(f)\| &\equiv (|\langle f, \tilde{\phi}_0 \rangle|^2 \\ &+ \sum_{m \in -\mathbf{N}} \sum_{n \in \mathbf{Z}_m} |\langle f, \tilde{\psi}_n^m \rangle|^2)^{1/2}. \end{aligned}$$

It is easy to see that if two representations are close, the curves that they represent should be close as well. In the case that the functions  $\tilde{\psi}_n^m(x)$  are the orthonormal set  $A = B = 1$ , and the computation of the transform is numerically stable. The limitation of these properties is that the parametrizations of curves must have the same starting point.

#### D. Approximation by Different Wavelet Bases

A good wavelet basis allows a signal to be approximated with a small error up to a certain resolution. “Which wavelet should we choose?” is a question with no quick answers. In this section, we compare the approximation capability of different wavelet bases in representing smooth curves. Following the idea of Daubechies [12], we investigate wavelet bases with different vanishing moments and symmetry property.

We will discuss the importance of vanishing moments of wavelets first. By definition, we say that a wavelet basis has  $L$  vanishing moments if the corresponding mother wavelet  $\tilde{\psi}(t)$  satisfies

$$\int t^l \tilde{\psi}(t) dt = 0, \quad l = 0, \dots, L-1.$$

By expanding a scalar function  $f(t)$  with Taylor series at  $t_0 = 2^m n$ , where  $m$  is a negative integer, we obtain

$$\begin{aligned} f(t) &= f(2^m n) + f^{(1)}(2^m n)(t - 2^m n) \\ &+ \frac{1}{2!} f^{(2)}(2^m n)(t - 2^m n)^2 + \dots \\ &+ \frac{1}{(L-1)!} f^{(L-1)}(2^m n)(t - 2^m n)^{L-1} \\ &+ (t - 2^m n)^L R(t) \end{aligned}$$

where  $|R(t)| < \infty$  if  $f \in C^{L-1}$ . Let us multiply both sides of the above equation by  $2^{-m/2} \tilde{\psi}(2^{-m}t - n)$  and integrate from 0 to 1. If the basis has  $L$  vanishing moments, the first  $L$  terms will become zero. Therefore, for a very fine scale (or a large value of  $|m|$ ), the wavelet coefficient  $|\langle f, \tilde{\psi}_n^m \rangle|$  will be small unless  $t = 2^m n$  is near one of the singularities of  $f$ . Suppose we decompose a function into wavelets and throw away all the detailed coefficients at finer scales. To this end, the wavelet  $\tilde{\psi}$  with higher vanishing moments tends to pack more information in the approximate coefficients. For some applications, the scaling function  $\tilde{\phi}(t)$  characterizes the coarsest scale approximation  $(x_a^M, y_a^M)$ . Thus, it is also

important to consider the vanishing moments of  $\tilde{\phi}(t)$ . Coiflets, which we will discuss later, fall into this category.

Physiological and psychological studies show that the human visual system is sensitive to the error induced by asymmetric reconstruction. It was proved that complete symmetry of  $\tilde{\psi}$  and  $\tilde{\phi}$  cannot be achieved by using compactly supported orthonormal wavelet bases except for the Haar basis [12], [37]. The orthonormal wavelets with noncompactly supported bases can only be implemented by the IIR filters. The implementation is generally recursive, which may lead to instability and limit cycles. Some other alternatives are biorthogonal wavelets [8] with different filters for decomposition and reconstruction.

To compare the performance of different wavelet bases in representing planar curves, we use the Daubechies [10], least asymmetric [12], Coiflets [4], [11], Battle-Lemarié [2], biorthogonal spline-variant [12], and biorthogonal cubic B-spline [34], [35] bases and show the results in Figs. 3 and 4. The previous four are orthogonal wavelets. Among them, the Daubechies and least asymmetric bases are both compactly supported wavelets with a maximum number of vanishing moments for a given support region. Coiflets are constructed by Daubechies and Coifman [11] from a class of orthonormal wavelet bases with an equal number of vanishing moments for  $\tilde{\psi}$  and  $\tilde{\phi}$ . It turns out that the mother wavelet  $\tilde{\psi}$  and the scaling function  $\tilde{\phi}$  of Coiflets are much more symmetric than their counterparts in Daubechies and least asymmetric bases. However, Coiflets require longer support of the basis function (and, hence, more computation efforts) to achieve the same order of vanishing moments. The Battle-Lemarié, biorthogonal spline-variant and cubic B-spline wavelets are constructed from spline functions and have perfectly symmetric scaling functions. We summarize the properties of these wavelets in Table I, where IIR/FIR means that the analysis and synthesis filters are IIR and FIR, respectively.

The approximations of a square by wavelet bases with various vanishing moments are plotted in Fig. 3. It appears that the biorthogonal spline-variant wavelet gives the best results for this symmetric test case since it represents both the corners and edges well. Furthermore, we compare the approximations of California’s outline by different basis with similar computational complexity (12 filter taps in most cases) in Fig. 4. At  $M = -7$ , the results are almost of no difference, except the Battle-Lemarié wavelets show some reconstruction errors due to the infinite convolution sum being replaced with a finite one. On the other hand, at  $M = -6$ , the Daubechies wavelets preserve sharp corners, whereas Coiflets nicely represent the right angle. The errors due to nonperfect reconstruction accumulate when more levels of decomposition/reconstruction are performed. Hence, higher noise appears in the Battle-Lemarié’s approximation at  $M = -6$  than at  $M = -7$ . This truncation error can be significantly reduced by increasing the number of taps of the IIR filters. Note also that it can be catastrophic if computation is done by the nonperiodic wavelet transform. We performed many other experiments and obtained the following three observations. First, in representing the symmetric pattern such as the square and the Koch’s curve, the wavelet bases with symmetric (or almost symmetric) scaling functions such as the biorthogonal

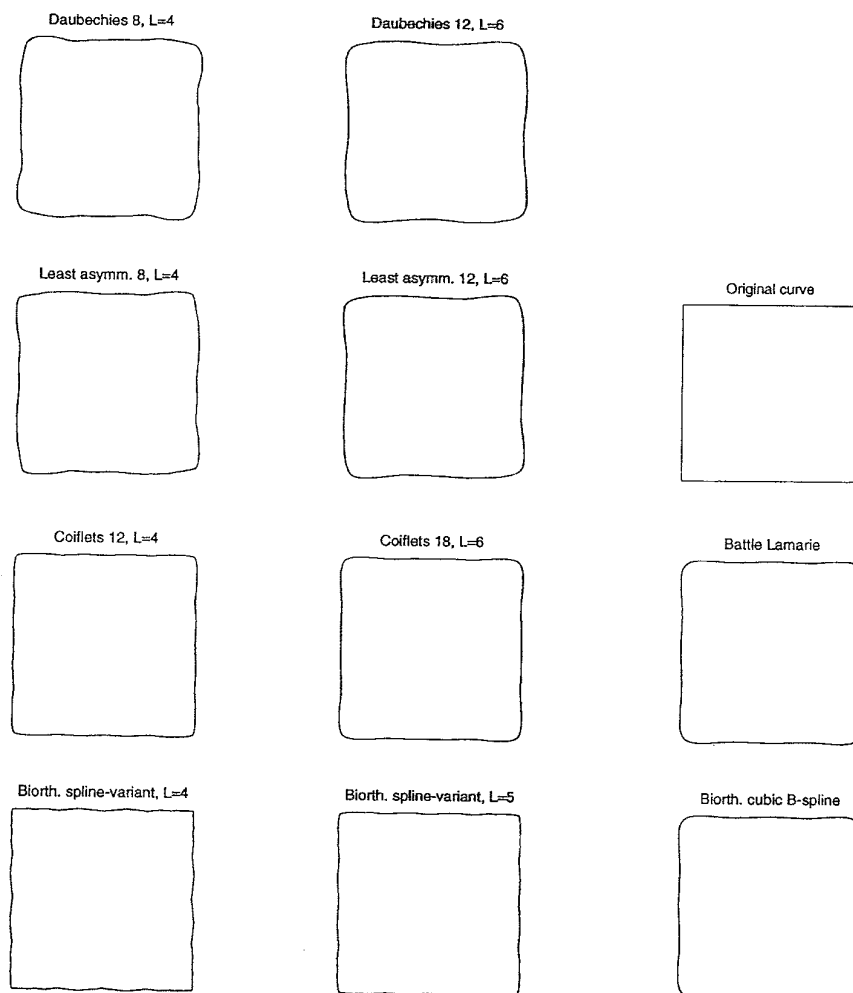


Fig. 3. Finite resolution approximations of a square with different wavelet bases.

TABLE I  
SYMMETRY AND OTHER PROPERTIES OF DIFFERENT WAVELETS

properties	wavelets					
	Daubechies	least asymmetric	Coiflets	Battle-Lemarié	spline-variant	cubic B-spline
symmetry	no	poor	almost	yes	yes	yes
filters	FIR/FIR	FIR/FIR	FIR/FIR	IIR/IIR	FIR/FIR	IIR/FIR
orthogonality	orthogonal	orthogonal	orthogonal	orthogonal	biorthogonal	biorthogonal

spline-variant or Coiflets give better results. Second, a basis with higher vanishing moments gives better approximations at the price of a higher computational cost. Third, with relatively light computations, the Daubechies wavelets preserve sharp corners in approximating not-so-smooth curves.

#### E. Normalization of Wavelet Coefficients

The invariance, uniqueness, and stability properties of the wavelet descriptor proved in Section III-C is important in the shape recognition application. Due to the invariance and uniqueness properties, it is convenient to perform clustering based on the wavelet representation rather than the spatial coordinates. The significance of the stability criterion is that

it guarantees that a small change in the shape of a curve will not cause a large change in its wavelet representation, and hence, the wavelet representation is stable with respect to noise. Thus, by measuring the distance between two boundary curves in terms of their wavelet coefficients, we can perform classification by the nearest-neighbor clustering rule. This procedure is detailed below.

Given a boundary curve and supposing that the starting point in traversing the curve is fixed, we calculate the wavelet transform and choose a set of significant coefficients as its features. Since the position, size, and orientation are not relevant in recognizing the shape, we have to normalize the contour so that the representation is invariant to these transformations. We now consider a normalization procedure

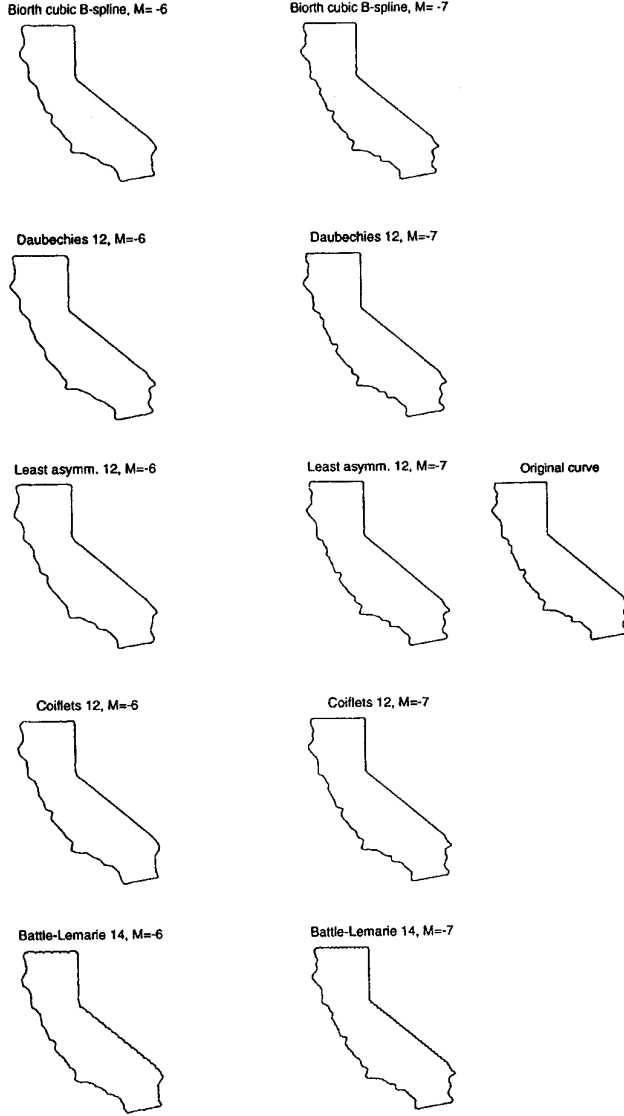


Fig. 4. Finite resolution approximations of the outline of California with different wavelet bases.

applied to the wavelet transformed data. It is clear from (3.2) that the wavelet coefficients  $r_n^m$  and  $d_n^m$  are invariant under translation. The centroid of the contour can be obtained by taking the average of the prime signals  $a_n^M$  and  $c_n^M$ , and the translation effect can be offset by setting the centroid of the contour to the origin. The scale normalization can be accomplished by dividing the magnitude  $A_n^m$  of wavelet coefficients with the averaged values  $\bar{A}$  over the significant set. By calculating the averaged phase  $\bar{\theta}$  of wavelet vectors over the significant set, we can offset the orientation effect by performing an inverse rotation.

In short, by performing the wavelet transform on the coordinate representation  $(x(l), y(l))$  of a contour via (2.6) and choosing a feature set, we can normalize the parameters, or wavelet coefficients, for each individual curve as follows:

- 1) For displacement normalization:  $(a_n^M, c_n^M) \leftarrow (a_n^M, c_n^M) - (b_x, b_y)$ , where  $(b_x, b_y) = \frac{1}{N_1} \sum (a_n^M, c_n^M)$

is the averaged displacement vector over the total number  $N_1$  of coefficients  $a_n^M$  (or  $c_n^M$ ) in the feature set.

- 2) For sizing normalization:  $A_n^m \leftarrow A_n^m / \bar{A}$ , where  $\bar{A} = \frac{1}{N_2} \sum_{n,m} A_n^m$  is the averaged magnitude over the total number  $N_2$  of coefficients  $r_n^m$  and  $d_n^m$  in the feature set.
- 3) For orientation normalization:  $\theta_n^m \leftarrow \theta_n^m - \bar{\theta}$ , where  $\bar{\theta} = \frac{1}{N_2} \sum_{n,m} \theta_n^m$  is the averaged angle over the total number  $N_2$  of coefficients  $r_n^m$  and  $d_n^m$  in the feature set.

This normalization sets up a scheme for silhouettes or signature recognition. It is worthwhile to point out that the above normalization procedure does not only normalize the shape for the purpose of classification but also removes noise in raw data.

#### F. An Example: Character Recognition

We use an example to demonstrate the the performance of the wavelet descriptor and will compare the result with that of the Fourier descriptor.

We work with the character set

$$\{2, Z, U, V, D, O, Q, 5, S, E, F, P, R, L, M, N\}$$

and represent each character by a binary image array of size  $24 \times 24$ . The reason for choosing such a character set is that elements in subsets  $\{2, Z\}$ ,  $\{U, V\}$ ,  $\{D, O, Q\}$ , and so on, are easily misclassified due to similar contour shapes. We trace the boundary of each character with a piecewise linear contour to represent the shape of the original image array. We have to mention that wavelet models depend on the choice of the starting points  $t_0$  of curves. Fourier descriptors, however, can be normalized [38] to global shift invariance, even though the normalization procedure is complicated. In this experiment, we suppose that the documents are aligned and scanned so that the choice of starting point is not a problem. Once the length of the contour is computed, we interpolate and resample the boundary contour so that the total number of samples is 256. (Note that both 128 and 512 points have also been tried, and we see any significant difference in the final results.) Then, the coefficients of Fourier and wavelet descriptors are computed by using (2.14) and (2.6), respectively. For the wavelet transform, we decompose the curve description into six levels with four coefficients at the coarsest scale.

Three fonts for each character, as shown in Fig. 5, were used in the training phase. It was observed that the regularity, vanishing moments, or even the choice of wavelets does not play a significant role in this character recognition experiment. In the following, we report the result based on the Daubechies D8 wavelet basis. Since only some lower resolution (or frequency) coefficients have substantially large values while the magnitudes of other coefficients are negligible, we choose a subset of the Fourier or wavelet coefficients as the feature vector. A reference feature vector for each character was obtained by averaging the feature vectors over all three fonts in the training stage. To perform nearest-neighbor rule clustering, we use the Euclidean distance. In the case of classification with only 4 features, which correspond to the coarsest scale wavelet coefficients, we found that the three characters 5, S,



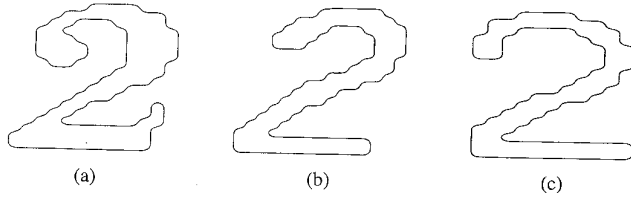


Fig. 5. Three fonts of the symbol "2": (a) New century schoolbook; (b) helvetica; (c) itc avant garde gothic.

and 2 were misclassified by the Fourier descriptor, whereas the two characters *Q* and *O* were misclassified by the wavelet descriptor. It is also possible to select the set of coefficients with the highest magnitudes as features. In the case of classification with seven features by using this approach, the character *S* was misclassified to 5 by the Fourier descriptor, whereas all characters were classified correctly by the wavelet descriptor. The above experiment is by no means complete, but the result supports that the wavelet descriptor can be an alternative to the Fourier representation that suffers from the drawback of not being sufficiently localized in space.

#### IV. DEFORMABLE WAVELET DESCRIPTOR

##### A. Basic Idea

To model a group of shapes that have the same topological structure but may differ slightly due to deformation, we can interpret the wavelet coefficients as random variables and propose a deformable stochastic wavelet descriptor to describe them. The possible applications of deformable wavelet descriptors include motion tracking, stereo matching, and computer animation.

Staib and Duncan [30] used coefficients of the Fourier descriptor as the parameters of deformable models. However, the basis functions of the Fourier descriptor are the sinusoids that are periodic and global (not sufficiently localized in space) so that a small perturbation of one parameter will affect the entire outline of a shape. This deformable model is not efficient in describing shapes with only local deformation. In contrast, the wavelet descriptor uses a set of basis functions with local support and multiscale dilations and, therefore, provides a scheme to model local as well as global deformation. In Fig. 6, we model the amplitude of a certain wavelet vector as a Gaussian random variable and assume that all other wavelet vectors (defined as in (2.13)) are deterministic. The deformation using the Fourier descriptor is also plotted for comparison. The superior local deformation property of the wavelet descriptor can be easily seen in this example.

Furthermore, we use the entropy as a measure of the averaged uncertainty of coefficients. The entropy  $H(\mathbf{x})$  of a random variable  $\mathbf{x}$  is defined as  $H(\mathbf{x}) = -\sum \Pr(x) \ln \Pr(x)$ , where  $\Pr(x)$  is the probability density function of  $\mathbf{x}$ . For a Gaussian random variable  $\mathbf{x}$ , the entropy is equal to  $H(\mathbf{x}) = \ln \sigma \sqrt{2\pi e}$ , where  $\sigma$  is the standard deviation of the Gaussian distribution of the random variable  $\mathbf{x}$  [26]. Now, consider a set of deformable curves with independent Gaussian-distributed spatial coordinates  $x(t)$  and  $y(t)$ . Then, the corresponding Fourier and wavelet coefficients are also Gaussian distributed.

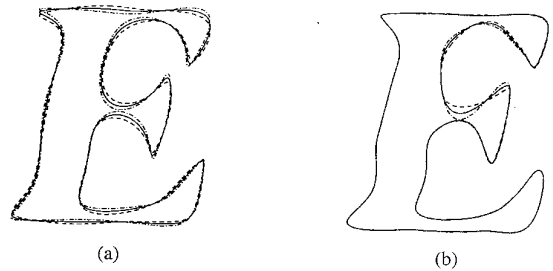


Fig. 6. Comparison of curves reconstructed by (a) Fourier and (b) wavelet descriptors.

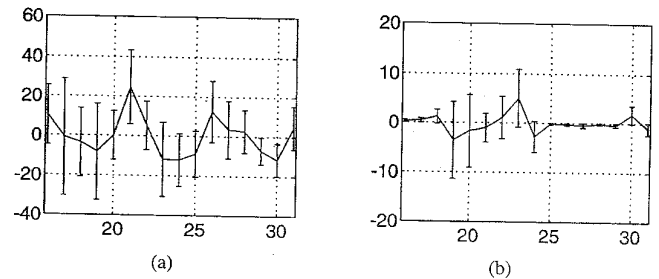


Fig. 7. Distributions of transform coefficients for a Gaussian deformed curve using (a) Fourier and (b) wavelet descriptors.

We plot the statistical distribution of the Fourier and wavelet transformed coefficients in Fig. 7, where the mean of the transform coefficient is indicated by the solid line, and the standard deviation for each transform coefficient is depicted with a vertical bar. Since the coefficient with a larger standard deviation has a larger entropy, we conclude that the neighborhood deformation of a curve can be more efficiently described by the wavelet descriptor than the Fourier descriptor.

For the case of a nonclosed curve, a straightforward representation of the curve would result in discontinuities; therefore, the boundary problem occurs when the wavelet transform is performed. This problem can be avoided by letting the parameter  $t$  start at one end of the curve, trace along the contour to the other end, and then trace back to the starting point to form a closed path. Fig. 8 shows an example of a nonclosed curve and its deformations, where the solid, dash, and dashdot lines represent the mean curves and curves with minus and plus one standard deviation, respectively. The curves are locally deformed in the  $x$  and  $y$  directions in Fig. 8(c) and (d). Note that these locally deformed curves cannot be easily achieved by using the Fourier descriptor deformable model. The locally deformable property makes the wavelet descriptor an attractive tool for handwritten signature generation and recognition.

##### B. Model-Based Contour Extraction

We consider one particular application of the deformable wavelet descriptor that extracts the boundary of an object from noisy images with a model-based/Bayesian approach. The idea of extracting the object boundaries from noisy images using the model-based Bayesian approach is inspired by the work of Staib and Duncan [30]. However, it is worthwhile to point out some differences. First, we propose a hierarchical matching

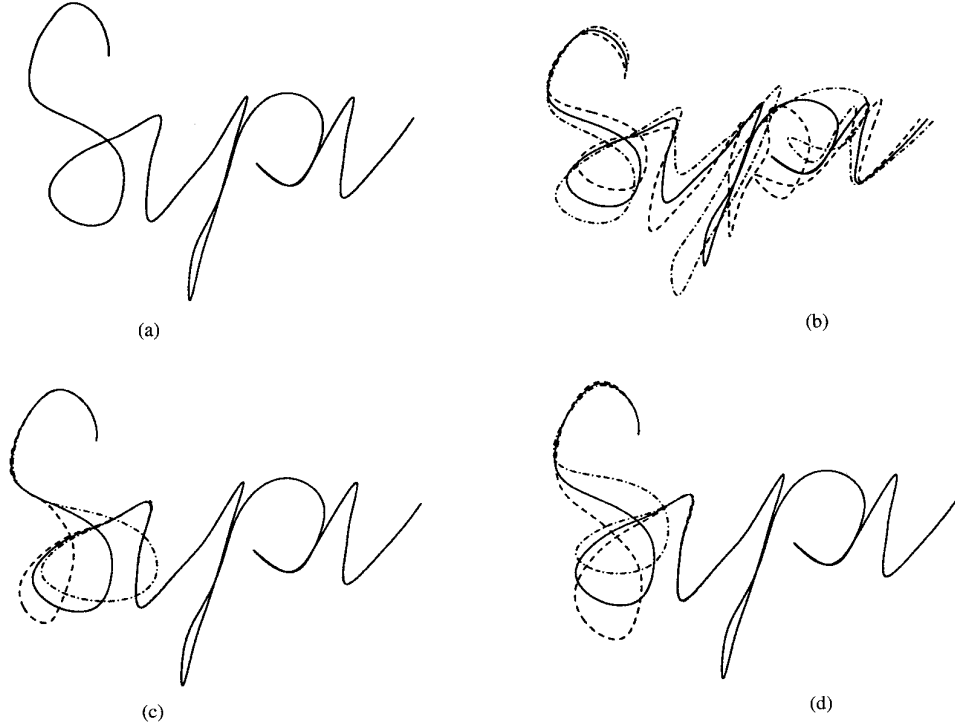


Fig. 8. Deformation of a nonclosed curve.

scheme by exploiting the multiscale representation capability of the wavelet descriptor. Second, to avoid exhaustive numerical line integrals for continuous gradient formulation in [30], we adopt the chamfer matching.

Let  $\vec{p}$  be the parameter vector of the wavelet descriptor of a contour where  $\vec{p}$  includes a set of  $N$  wavelet coefficients of the  $x, y$  coordinates as its elements. Given a contour template  $C_{\vec{p}}(j, k)$  corresponding to a particular value of the parameter vector  $\vec{p}$ , we want to detect the template  $C_{\vec{p}}$  from an image  $F(j, k)$ , where  $j$  and  $k$  are the indices of pixels in the image. To extract the contour of an object, some preprocessing of the image  $F(j, k)$  is needed. We apply the Laplacian of Gaussian (LoG) [22] operation to  $F(j, k)$  to obtain an estimate of the edge location. The LoG operator is defined as

$$I(j, k) \triangleq \text{LoG}(F)(j, k) = F(j, k) * (-\nabla^2 G)(j, k)$$

where  $\nabla^2$  is the Laplacian operator, and  $G$  is the Gaussian smoothing kernel. The lowpass Gaussian filter reduces the noise sensitivity of the Laplacian operator while preserving the location of zero crossings.

Based on the preprocessed information  $I(j, k)$ , we can formulate MAP estimation problem. Let  $C_{\text{MAP}}$  denote the MAP estimation of boundary curve  $C$ . Then, the MAP problem of our interest can be written as

$$\begin{aligned} \Pr(C_{\text{MAP}} | I) &= \max_{\vec{p}} \Pr(C_{\vec{p}} | I) \\ &= \max_{\vec{p}} \frac{\Pr(I | C_{\vec{p}}) \Pr(C_{\vec{p}})}{\Pr(I)} \end{aligned} \quad (4.1)$$

where  $\Pr(C_{\vec{p}})$  is some *a priori* knowledge of the contour, and  $\Pr(I | C_{\vec{p}})$  is the likelihood function of detected edge information  $I$  with given  $C_{\vec{p}}$ . Since the term  $\Pr(I)$  is independent of  $\vec{p}$ , we can take the logarithm on both sides and change the optimization problem (4.1) to be

$$\max_{\vec{p}} [\ln \Pr(I | C_{\vec{p}}) + \ln \Pr(C_{\vec{p}})]. \quad (4.2)$$

By assuming independent zero-mean Gaussian noise with standard deviation  $\sigma_n$ , we know that the likelihood function of  $I$  is of the form

$$\begin{aligned} \Pr(I | C_{\vec{p}}) &= \Pr(n = I - C_{\vec{p}}) \\ &= \prod_{j, k \in \mathcal{A}} \frac{1}{\sigma_n \sqrt{2\pi}} e^{-\frac{[I(j, k) - C_{\vec{p}}]^2}{2\sigma_n^2}}. \end{aligned}$$

In the above equation, the noise at each pixel  $n(j, k)$  is independently and identically distributed with the probability density  $\Pr(n)$ , and that probability for noise over the entire picture area  $\mathcal{A}$  is nothing but the product of noise on the individual pixel. The maximization problem (4.2) is equivalent to maximizing the objective function [28]

$$M(I, C_{\vec{p}}) \approx \max_{\vec{p}} \left[ \frac{k}{\sigma_n^2} \sum_{j, k \in \mathcal{A}} I(j, k) C_{\vec{p}}(j, k) + \ln \Pr(C_{\vec{p}}) \right]. \quad (4.3)$$

The second term of (4.3) is the natural logarithm of the *a priori* probability that confines the solution  $C_{\text{MAP}}$  to a group of curves within the prescribed model, whereas the first term

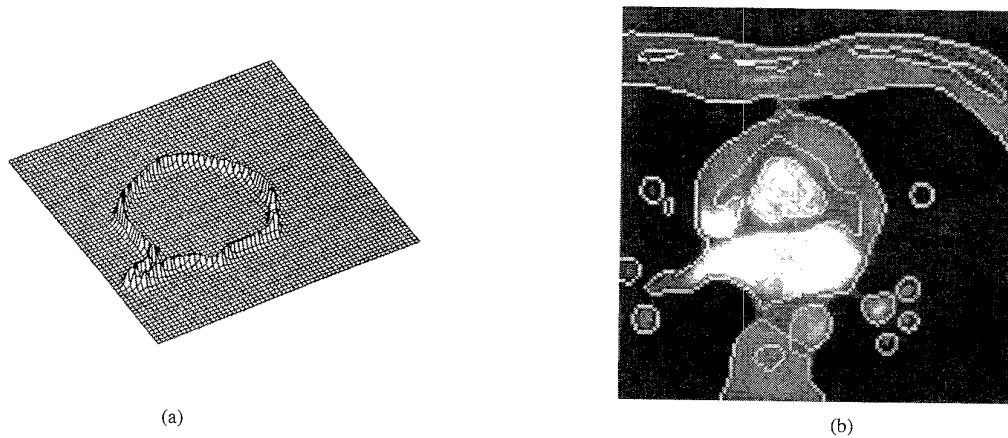


Fig. 9. Likelihood term approximated by the cross-correlation of (a) normalized chamfer correlation image of the contour template  $C_{\bar{p}}$  and (b) zero-crossing image  $I$ , where  $I$  is calculated by applying the Laplacian of Gaussian operator to the cardiac image.

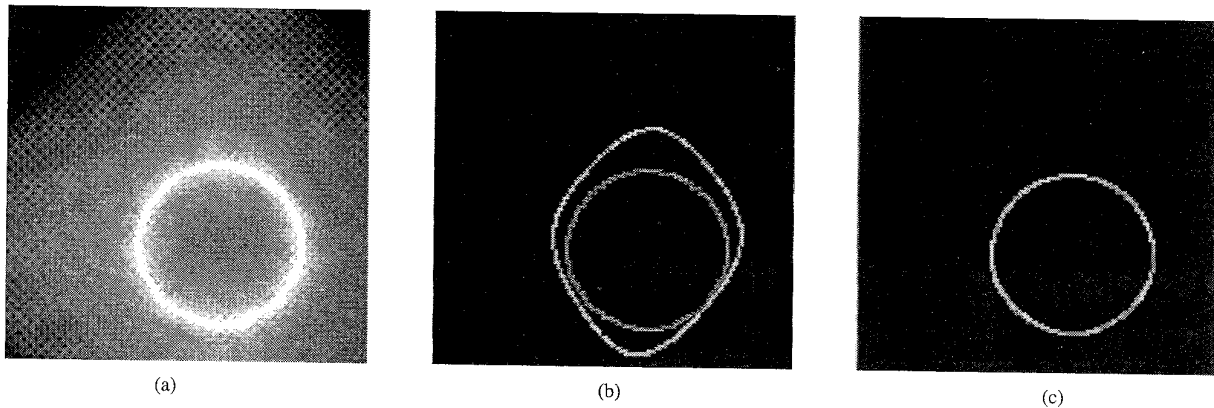


Fig. 10. Example of the normalized chamfer correlation image and its matching result using Powell's method: (a) Example of the normalized chamfer image of a synthesized image; (b) superimposed initial guess and the objective (the inner circle); (c) final result of multiscale elastic matching.

is the approximation of logarithm of the likelihood  $\Pr(I | C_{\bar{p}})$  and is, in some sense, a cross correlation of the contour  $C_{\bar{p}}$  and  $I(j, k)$ . Note also that the cross correlation is weighted by a factor  $k/\sigma_n^2$  that decreases as the energy of Gaussian noise increases. This implies that the estimation depends more heavily on the information of likelihood than the *a priori* knowledge when the noise is small and the *a priori* information prevails in the objective function when the noise corruption becomes more severe.

We will assume that all wavelet coefficients are Gaussian distributed. This assumption is in fact close to what we observed in the beating heart image sequence in the experiment described in Section IV-C. Let us choose

$$\ln \Pr(C_{\bar{p}}) = \sum_{i=1}^N \ln \frac{1}{\sigma_i \sqrt{2\pi}} - \frac{(p_i - m_i)^2}{2\sigma_i^2}$$

where the mean  $m_i$  and the standard deviation  $\sigma_i$  of each wavelet coefficient  $p_i$  are parameters of the deformation model. We assume that they are known in advance. Note that if no *a priori* information is available, the uniform probability yields a constant *a priori* term and therefore has no influence to the solution of the objective function.

The MAP estimation  $C_{\text{MAP}}$  can be calculated in a pyramidal structure. We first solve the MAP problem in a very coarse scale (by setting the wavelet coefficients in other finer scales to zero) and use the result as the initial guess at the next finer scale. That is, by fixing the coarse scale coefficients obtained from the coarse scale optimization, we solve the MAP problem by adjusting only the wavelet coefficients that are responsible for variations in this fine scale. The same procedure is applied recursively scale by scale until the highest resolution is reached. The hierarchical algorithm reduces the computational cost significantly.

It is worthwhile to compare our approach with a commonly used approach known as the snake [19]. The snake is an active contour model where a curve is deformed due to a certain external force. The proposed MAP estimate deforms the curve to maximize the probability of estimation given the *a priori* information and the *a posteriori* knowledge, whereas the snake scheme changes the curve shape to minimize the energy due to internal and external forces. The internal force acts as the smoothness constraint, whereas the external force guides the active contour toward image features. The smoothness constraint of snake is to minimize the first- and the second-

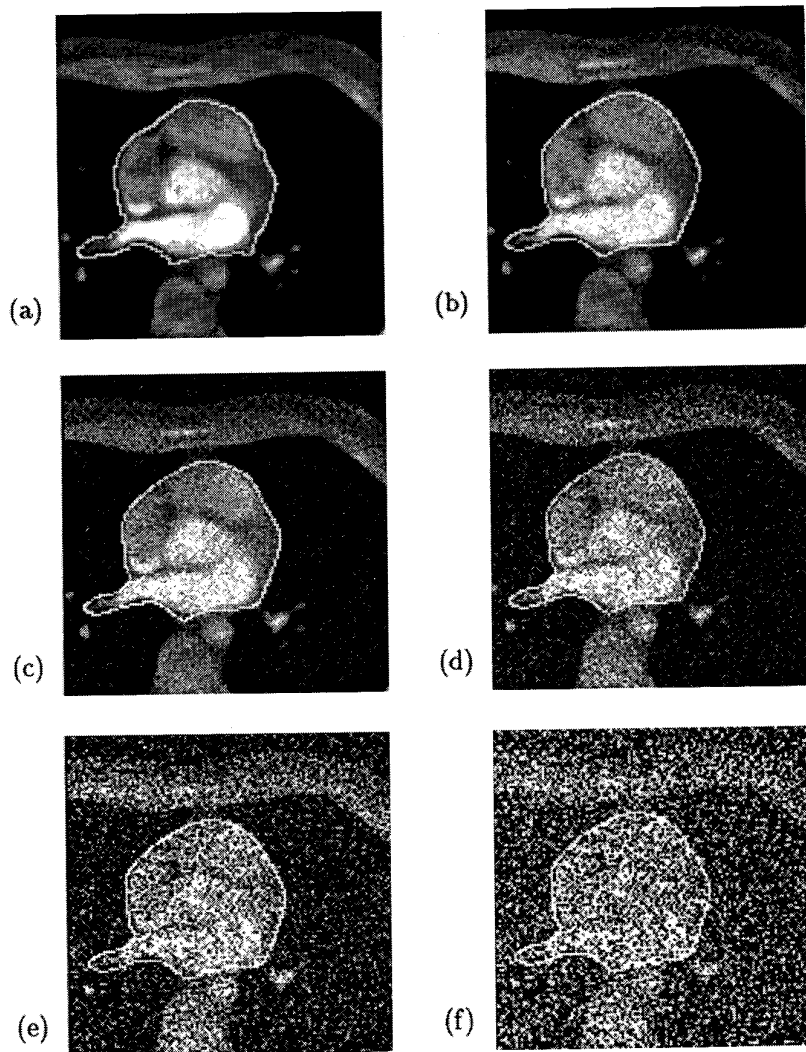


Fig. 11. Contour extraction by the MAP estimation using the deformable wavelet descriptor.

order derivatives of curves, but our scheme restricts the curve representation in the wavelet coefficient space to achieve multiscale matching. However, both the snake and wavelet deformable model share the advantage of having local control property so that the modification of the position of a control point results in only a local variation of the curve.

### C. Experiments

In this experiment, we use the deformable model to extract the contour of a heart from a certain computer tomography (CT) image. The extraction of the cardiac contour is important in assessing the regional diastolic function [14]. To collect the *a priori* information of the contour, we consider a set of 15-frame noiseless CT images of size  $128 \times 128$  that forms a full cycle of the heartbeat. We use the contours of these images to calculate the mean and the standard deviation of the deformable model. Note that the proposed method relies on other mechanisms to register the contours in the image sequence such that the starting points are identical.

By using the MAP estimation method, we show how to extract the contour of the heart from a certain frame, say, the second frame, with high-level additive white noise and compare the results of using the deformable Fourier and wavelet descriptors.

Since the gradient of the objective function provides information of the new search direction, most numerical optimization methods require it for fast convergence. However, it is sometimes difficult to obtain the analytic form of the gradient as well as its discrete approximation. In the current setting, to calculate the analytic form of a gradient vector involves a numerical line integral that is too tedious to be practical while its discrete approximation is sensitive to error resulting from an inadequate choice of step sizes. Thus, we propose to do the following.

In the experiment, we first ignore the first term in (4.3) by setting it to zero and focus on the optimization of the second term. To be more precise, we choose the quasi-Newton [13] method to maximize this term, which has a quadratic form

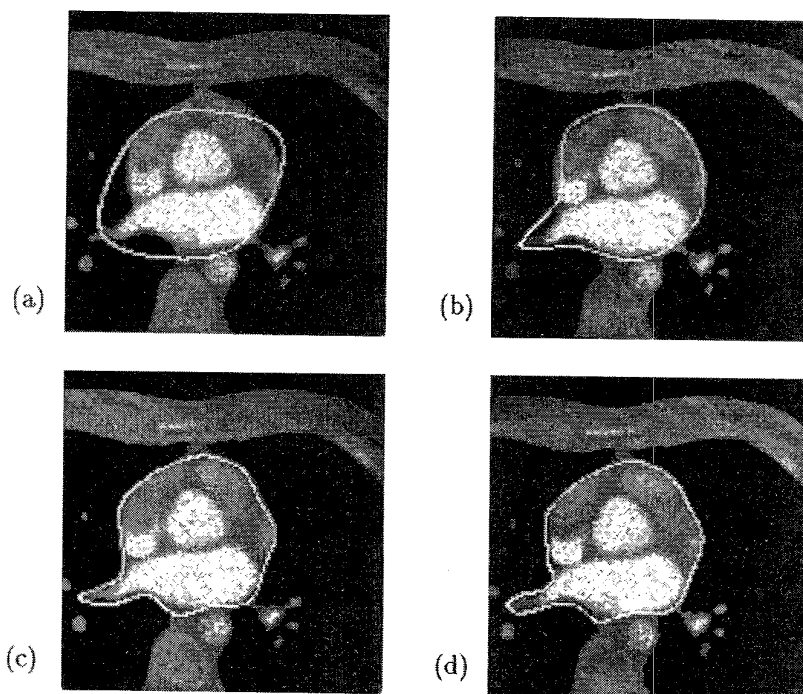


Fig. 12. Multiscale contour extraction by the MAP estimation.

as shown in (4.4), with the objective to match features in the coarsest scale. The interpolated solution obtained from the coarsest scale is then served as the initial guess for the optimization in the next finer scale.

For all optimization problems in the finer scales, we consider the objective function consisting of both terms as given in (4.3). We use Powell's method to solve the optimization problem and progress from the coarsest scale to the finest scale. Powell's method [13]—a special example of conjugate direction methods—does not involve explicit computation of the gradient of a function but takes more iterations to converge. To use Powell's method for optimization, we found that direct cross correlation fails to provide the direction that leads to optimum if a template  $C_{\bar{p}}$  has no coincident points with the object data  $I$ . Thus, we turn to a technique known as chamfer matching [6], which determines the best fit of imperfect edge data from two images. Chamfer matching offers a generalized correlation that shows the direction to achieve the maximum cross correlation. We calculate the normalized chamfer correlation image of the contour template  $C_{\bar{p}}$  and perform the cross correlation between the chamfer correlation image and the zero-crossing image  $I(j, k)$ . Fig. 9 gives an example of the normalized chamfer correlation image and the zero-crossing image (which is the Laplacian of Gaussian of the cardiac image).

The normalized chamfer correlation image is derived from the conventional chamfer distance by taking two to the power of  $-v_{j,k}$ , i.e.,  $2^{-v_{j,k}}$ , where  $v_{j,k}$  is the corresponding chamfer distance image, and  $j, k$  are the row and column indices of the image. We normalize the image so that chamfer correlation gives the value 1 if the estimated and the original curves are totally coincided, the value 0 if totally irrelevant, and

all values in between if they are proportional to the degree of coincidences. Fig. 10(a) is an example of the normalized chamfer image of a synthesized image. The binary edge image is a circle. The brightest pixels represent the value 1. The value decreases gradually to 0 as the distance from the circle increases. In the optimization process, the chamfer correlation is also helpful in improving the convergence rate, especially when the template is too far from the edge. We use Fig. 10 as an example to demonstrate how the chamfer matching works for a nonmodel-based (maximum likelihood) optimization with Powell's method. We first calculate the normalized chamfer correlation image in Fig. 10(a), where the circle of brightest grey level is the object to reach. In Fig. 10(b), we superimpose the initial guess and the objective (the inner circle), whereas Fig. 10(c) gives the final result of multiscale elastic matching.

Now, we are ready to show an example of contour extraction from a set of noisy images, which are obtained by adding white Gaussian noise of different levels to the second frame of the 15-frame image sequence. The noise level ranges from  $\sigma_n = 10$  as shown in Fig. 11(a) to  $\sigma_n = 160$  as given in Fig. 11(f), or the corresponding SNR values range from 21.3 dB to  $-2.8$  dB. The proposed MAP estimator extracts the contour accurately for these noise levels, but the estimation tends to be trapped to the local maxima for images with an even lower range of SNR. The applicable SNR region depends on the structure and the contrast of the image. Some external features tend to give erroneous results and hamper the estimator's ability to find the real contours. In addition, the applicability of the proposed scheme to structured (instead of Gaussian) noise remains an open problem.

TABLE II  
COMPARISON OF THE MAP ESTIMATION RESULTS IN TERMS OF NORMALIZED CHAMFER CORRELATION, WHERE 1  
MEANS THAT THE ORIGINAL AND ESTIMATED CURVES COMPLETELY COINCIDE, WHEREAS 0 MEANS NO CORRELATION

Descriptor	Estimation result of frame						
	2	4	6	8	10	12	14
Wavelet	0.7494	0.7917	0.7825	0.7861	0.8033	0.8222	0.8424
Fourier	0.7215	0.7201	0.7285	0.7330	0.7525	0.7436	0.8142

The multiscale matching process can be best illustrated in Fig. 12, where four levels of scales are used in this experiment. We show, in Fig. 12(a), the contour extraction done at the coarsest scale by using four wavelet coefficients as the feature vector. The result is then used as the initial curve for matching at the next finer scale. The results of intermediate scales are given in Fig. 12(b) and (c) by using eight and 16 wavelet coefficients as feature vectors, respectively. Finally, the optimal estimation is achieved at the finest scale as in Fig. 12(d), where the feature vector consists of 32 wavelet coefficients. This hierarchical process avoids the undesirable local minima resulting from noise or spurious details existing in finer scales.

We compare the results of the MAP estimation using wavelet and Fourier descriptors in Table II, where a quantitative comparison for even frames of the cardiac image sequences with additive noise  $\sigma_n = 80$  is reported, and the normalized chamfer correlation is chosen as the performance measure. The reason that the wavelet descriptor has better results in this case is that it is more effective in representing the local deformations of a group of deformable contours.

## V. CONCLUSION AND EXTENSIONS

By using the wavelet transform, we developed a planar curve descriptor that has a multiscale analysis capability and can be computed effectively. The effect of scaling, translation, and rotation of a planar curve on its wavelet coefficient vectors was explored. The invariance, uniqueness, and stability properties of the wavelet descriptor were derived. We also compared the performance of a class of wavelet bases with different vanishing moments and symmetry properties. The application of the wavelet descriptor to the modeling of deformable objects was studied.

The multiresolution bases provide a powerful tool for local-to-global shape description and will have impact in the analysis and synthesis of 2-D and 3-D shape deformation. We have recently found research using the wavelet transform for 3-D multiscale deformable modeling by Vemuri and Radisavljevic [36]. They made an improvement on the dynamic 3-D deformation model of Terzopoulos [32] with superquadrics and a global deformable model characterized by the coarsest level wavelet coefficients. Their work considered 2-D wavelets constructed from the tensor products of 1-D wavelets and examined the applications of surface deformation in a 3-D space. In this research, however, we focused on the planar curve descriptor, i.e., 1-D wavelet descriptor in a 2-D space. Since this is a relatively simpler mathematical problem, we

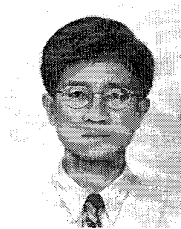
were able to discuss their properties in a more detailed way from a better perspective. Besides, most of the properties derived can be straightforwardly extended to the 3-D case, and the applications examined in this work also have their own practical values.

There are many interesting topics worth further study, including the application of wavelets to the description of object shapes in the 3-D space, dynamic shape warping in computer animation, and optimum surface reconstruction.

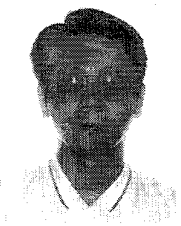
## REFERENCES

- [1] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Comput. Vision, Graphics, Image Processing*, vol. 46, pp. 1–21, 1989.
- [2] G. Battle, "A block spin construction of ondelettes. Part I: Lemarié functions," *Comm. Math. Phys.*, vol. 110, pp. 601–615, 1987.
- [3] A. Bengtsson and J. O. Eklundh, "Shape representation by multiscale contour approximation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, no. 1, pp. 85–93, 1991.
- [4] G. Beylkin, R. Coifman, and V. Rokhlin, "Fast wavelet transforms and numerical algorithms," *Comm. Pure Appl. Math.*, vol. 44, pp. 141–183, 1991.
- [5] C. D. Boor, *A Practical Guide to Splines*. New York: Springer, 1978.
- [6] G. Borgefors, "Hierarchical chamfer matching: A parametric edge matching algorithm," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 10, no. 6, pp. 849–865, 1988.
- [7] C. K. Chui, *An Introduction to Wavelets*. New York: Academic, 1991.
- [8] A. Cohen, "Biorthogonal wavelets," in *Wavelets: A Tutorial in Theory and Applications*, C. K. Chui, Ed. New York: Academic, 1992, pp. 123–152.
- [9] P. E. Danielsson, "A new shape factor," *Comput. Graphics Image Processing*, vol. 7, pp. 292–299, 1978.
- [10] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, 1988.
- [11] —, "Orthonormal bases of compactly supported wavelets II. Variations on a theme," submitted to *SIAM J. Math. Anal.*, 1990.
- [12] —, *Ten Lectures on Wavelets*, Philadelphia: SIAM, 1992.
- [13] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [14] E. L. Dove, K. P. Philip, D. D. McPherson, and K. B. Chandran, "Quantitative shape descriptors of left-ventricular cine-CT images," *IEEE Trans. Biomed. Eng.*, vol. 38, no. 12, pp. 1256–1261, 1991.
- [15] C. Fermeuller and W. Kropatsch, "Multi-resolution shape description by corners," in *Proc. Comput. Vision Patt. Recogn.*, 1992, pp. 271–276.
- [16] H. Freeman, "On the encoding of arbitrary geometric configurations," *IEEE Trans. Electron. Comput.*, vol. EC-10, no. 2, pp. 260–268, 1961.
- [17] M. Holschneider, "Wavelet analysis on the circle," *J. Math. Phys.*, vol. 31, no. 1, pp. 39–44, 1990.
- [18] B. K. P. Horn and E. J. Weldon, "Filtering closed curve," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 8, no. 5, pp. 665–668, 1986.
- [19] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: active contour models," *Int. J. Comput. Vision*, vol. 1, no. 4, pp. 321–331, 1988.
- [20] F. P. Kuhl and C. R. Giardina, "Elliptic Fourier features of a closed contour," *Comput. Graphics Image Processing*, vol. 18, pp. 236–258, 1982.

- [21] T. Lindeberg, "Scale-space for discrete signals," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 12, no. 3, pp. 234–254, 1990.
- [22] D. Marr, "Early processing of visual information," *Trans. Roy. Soc. London*, vol. 275, pp. 483–519, 1976.
- [23] F. Mokhtarian and A. K. Mackworth, "Scale-based description and recognition for planar curves and two-dimensional shapes," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-8, no. 1, pp. 34–43, 1986.
- [24] ———, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 14, no. 8, pp. 789–805, 1992.
- [25] J. Oliensis, "Local reproducible smoothing without shrinkage," *Proc. Comput. Vision Patt. Recogn.*, 1992, pp. 277–282.
- [26] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1984.
- [27] E. Persoon and K. S. Fu, "Shape discrimination using Fourier descriptor," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-7, pp. 170–179, Mar. 1977.
- [28] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*, 2nd ed. Orlando, FL: Academic, 1982.
- [29] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol. C-20, pp. 562–569, 1971.
- [30] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 14, no. 11, pp. 1061–1075, 1992.
- [31] D. Terzopoulos, "Image analysis using multigrid relaxation methods," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 8, no. 2, pp. 129–139, 1986.
- [32] D. Terzopoulos and D. Metaxas, "Dynamic 3-D models with local and global deformations: Deformable superquadrics," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, no. 7, pp. 703–714, 1991.
- [33] V. Torre and T. Poggio, "On edge detection," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-8, no. 2, pp. 147–163, 1986.
- [34] M. Unser, A. Aldroubi, and M. Eden, "A family of polynomial spline wavelet transforms," *Signal Processing*, vol. 30, pp. 141–162, 1993.
- [35] M. Unser and M. Eden, "FIR approximations of inverse filters and perfect reconstruction filter banks," *Signal Processing*, vol. 36, pp. 163–174, 1994.
- [36] B. C. Vemuri and A. Radisavljevi, "From global to local, a continuum of shape models with fractal priors," in *Proc. Comput. Vision Patt. Recogn.*, 1993, pp. 307–313.
- [37] M. Vetterli and C. Herley, "Wavelets and filter banks: relationships and new results," in *Proc. IEEE ICASSP*, 1990, pp. 1723–1726.
- [38] T. P. Wallace and P. A. Wintz, "An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors," *Comput. Graphics Image Processing*, vol. 13, pp. 99–126, 1980.
- [39] A. P. Witkin, "Scale space filtering," in *Proc. Int. Joint Conf. Artificial Intell.*, 1983, pp. 1019–1023.
- [40] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Comput.*, vol. C-21, no. 3, pp. 269–281, 1972.



TV/multimedia applications from 1988 to 1990. He was a research assistant with the Signal and Image Processing Institute of the University of Southern California, Los Angeles, USA, from 1992 to 1994. He is currently with the Philips (Taiwan) Ltd., where he is in charge of strategic planning for highly integrated Multimedia IC's. His main research interests are in digital video/image processing and computer graphics and their VLSI implementations.



**C.-C. Jay Kuo** (SM'92) received the B.S. degree from the National Taiwan University, Taipei, in 1980 and the M.S. and Ph.D. degrees from the Massachusetts Institute of Technology, Cambridge, USA, in 1985 and 1987, respectively, all in electrical engineering.

From October 1987 to December 1988, he was a computational and applied mathematics (CAM) research assistant professor in the Department of Mathematics at the University of California, Los Angeles, USA. Since January 1989, he has been with the Department of Electrical Engineering-Systems and the Signal and Image Processing Institute at the University of Southern California, Los Angeles, USA, where he is currently an Associate Professor. His research interests are in the areas of digital signal and image processing, wavelet theory and applications, video image compression, multimedia, and large-scale scientific computing. He has authored more than 160 technical publications that have appeared in journals and conference proceedings.

Dr. Kuo is a member of Sigma Xi, SIAM, SPIE, and ACM. He serves as an Associate Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, and he is on the editorial board of the *Journal of Visual Communication and Image Representation*. He received the National Science Foundation Young Investigator Award (NYI) and Presidential Faculty Fellow (PFF) Award in 1992 and 1993, respectively.