USC–SIPI REPORT #238

A Wavelet Descriptor of Planar Curves: Theory and Applications

by

Chun-Hsiung Chuang and C.-C. Jay Kuo

August 1993

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 404
Los Angeles, CA 90089-2564 U.S.A.

A Wavelet Descriptor of Planar Curves: Theory and Applications *

Chun-Hsiung Chuang † and C.-C. Jay Kuo †

August 24, 1993

EDICS: IP 1.6

Abstract

By using the wavelet transform, we develop a hierarchical planar curve descriptor which 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 application of the wavelet descriptor to character recognition and model-based contour extraction from low SNR images is examined. Numerical experiments are performed to illustrate the performance of the wavelet descriptor.

1 Introduction

Boundary 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 [15] 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 [8], can be used to characterize the features of its shape. Examples include moments, areas, perimeters, and corners. The spline technique [4] 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

*This work was supported by the National Science Foundation Young Investigator (NYI) Award ASC-9258396 and Presidential Faculty Fellow (PFF) Award GER 93-50309.

†The authors are with the Signal and Image Processing Institute and the Department of Electrical Engineering-Systems, University of Southern California, Los Angeles, California 90089-2564. E-mail: chunhsiu@sipi.usc.edu and cckuo@sipi.usc.edu.

curve can be translated, scaled, and rotated by performing the transformations on a set of control points. The Fourier descriptor [24], [34] describes a curve with coefficients via Fourier analysis of a certain parametric representation of the curve such as the curvature or spatial coordinates.

One common drawback of the above methods is that they do not have a multiresolution representation capability. The multiresolution (or multiscale) technique [2], [25], [28] 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 [26] to develop an multiscale edge detection scheme. Marr [19] 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 [33] proposed a scale-space filtering approach which uses a set of parameterized Gaussian kernel functions to smooth a signal and extracts the descriptive primitives such as local extrema or intervals bounded by the extrema with a multiscale analysis. That is, to characterize the features of local extremum, one can 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 [14] 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 [20] 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 it was 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 altas matching. For the scale-space methods, the Gaussian function is the optimal kernel for reducing noise with minimal delocalization [30], [18]. Nevertheless, a planar curve smoothed by the Gaussian kernel suffers from shrinkage [22], 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 research, 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 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 to the Fourier descriptor with the global sinusoids as the basis functions, the wavelet descriptor is more efficient in representing and detecting local features of a curve due to the nice 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 which have the same topological structure but may differ slightly due to local deformation 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 wavelet theory and derive a wavelet representation for planar curves in Section 2. In Section 3, 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 orientations. In Section 4, 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 5.

2 Planar Curve Descriptor Using Wavelet Transform

2.1 Review of Wavelet Theory

Let us first review some basic results in wavelet theory [11], [6] and the biorthogonal wavelet transform [7]. We use $\phi(t)$ to denote a scaling function such that its dilation and translation

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

form an orthonormal basis for the wavelet subspaces V_m and that $\{V_m\}_{m\in Z}$ is a multiresolution approximation of the space $L^2(R)$. For each scaling function $\phi(t)$, we can determine the corresponding mother wavelet function $\psi(t)$ and such that its dilations and translations

$$\psi_n^m(t) = 2^{-m/2}\psi(2^{-m}t - n)$$

form an orthonormal basis for $L^2(R)$. Besides, we require that there exists a dual function $\tilde{\psi}(t)$ such that $\tilde{\psi}_n^m(t)$,$m, n \in Z$, with $< \psi_n^m, \tilde{\psi}_{n'}^{m'} >= \delta_{m,m'}\delta_{n,n'}$, and $\{\psi_n^m, \tilde{\psi}_n^m\}_{m,n,\in Z}$ are a pair of biorthogonal Riesz bases. Then, the wavelet coefficients d_n^m can be computed via

$$d_n^m = \int_{-\infty}^{+\infty} f(t)\psi_n^m(t)dt \equiv < f, \psi_n^m >, \tag{2.1}$$

and if $f(t) \in L^2(R)$, we have

$$f(t) = \sum_{m,n=-\infty}^{\infty} d_n^m \tilde{\psi}_n^m(t).$$

The bases $\{\psi_n^m(t)\}_{m,n}$ and $\{\tilde{\psi}_n^m(t)\}_{m,n}$ are known as the biorthogonal wavelets. In particular, if the mother wavelet $\psi(t)$ is equal to its dual $\tilde{\psi}(t)$, we call it the orthogonal wavelet.

For $f(t) \in V_0$, we have

$$f(t) = \sum_{n=-\infty}^{\infty} c_n^0 \tilde{\phi}_n^0(t) = \sum_{n=-\infty}^{\infty} c_n^M \tilde{\phi}_n^M(t) + \sum_{m=1}^{M}\sum_{n=-\infty}^{\infty} d_n^m \tilde{\psi}_n^m(t),$$

which is known as the finite-scale biorthogonal wavelet expansion. The coefficients d_n^m and c_n^M can be computed from the coefficients c_n^0 via the following recursive formulas:

$$\begin{array}{rcl} c_n^{m+1} & = & \sum_l h_{l-2n} c_l^m, \\ d_n^{m+1} & = & \sum_l g_{l-2n} c_l^m, \end{array} \quad m = 0, 1, ..., M-1, \tag{2.2}$$

where h_n is a lowpass filter and g_n is its mirror highpass filter. On the other hand, one can obtain the coefficients c_n^0 from coefficients d_n^m and c_n^M via the synthesis formula

$$c_n^{m-1} = \sum_l (\tilde{h}_{n-2l} c_l^m + \tilde{g}_{n-2l} d_l^m), \quad m = M, M-1, ..., 2, 1. \tag{2.3}$$

where $g_n = (-1)^{n+1}\tilde{h}_{-n+1}$ and $\tilde{g}_n = (-1)^{n+1}h_{-n+1}$, and the filters h_n, $\tilde{h}_n$, g_n and $\tilde{g}_n$ are called the biorthogonal filter bank. In particular, for the orthogonal filter bank, we choose $\tilde{h_n} = h_n$, $\tilde{g_n} = g_n$, and $g_n = (-1)^n h_{-n+1}$. Equations (2.2) and (2.3) are called the forward and inverse wavelet transforms, respectively.

2.2 Coordinate-Based Wavelet Descriptor

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

$$\alpha(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \quad t(l) = 2\pi l/L, \quad 0 \le l \le L,$$

where the parameter t corresponds to the normalized arc length, l is the arc length along the curve from a certain starting point, and L is the total arc length. By applying the wavelet transform to the parameterized coordinates, we obtain

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} x_a^M(t) \\ y_a^M(t) \end{bmatrix} + \sum_{m=1}^{M} \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix}, \tag{2.4}$$

where

$$x_a^M(t) = \sum_n a_n^M \tilde{\phi}_n^M(t), \quad y_a^M(t) = \sum_n c_n^M \tilde{\phi}_n^M(t) \tag{2.5}$$

are called the approximation coefficients at scale M and

$$x_d^m(t) \equiv \sum_n r_n^m \tilde{\psi}_n^m(t), \quad y_d^m(t) \equiv \sum_n d_n^m \tilde{\psi}_n^m(t), \tag{2.6}$$

are called the detail signals at scale m with $1 \le m \le M$. Then, we can use the wavelet coefficients a_n^M, c_n^M, r_n^m and d_n^m given in (2.5) and (2.6) as the planar curve descriptor.

By using only a subset of wavelet coefficients consisting of primarily low frequency components, i.e larger values of m, we can have different multiresolution representation of the shape. In terms of mathematical expression, we can modify (2.4) to be

$$\begin{bmatrix} \hat{x}(k;t) \\ \hat{y}(k;t) \end{bmatrix} = \begin{bmatrix} x_a^M(t) \\ y_a^M(t) \end{bmatrix} + \sum_{m=k}^{M} \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix}, \tag{2.7}$$

where $1 \le k \le M+1$. The curves with coordinates $\hat{x}(k;t)$ and $\hat{y}(k;t)$ provide a sequence of multiresolution approximations to the original curve. On one hand, we obtain the original curve for $k = 1$. 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

a biorthogonal spline wavelet transform with $M = 8$ as described by (2.2). Figs. 1(1)-(8) are its approximations from the resolution with $k = 2$ to the resolution with $k = 9$, while the original curve with 3072 samples is given in Fig. 1(0). One of the 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.

$$\begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix} = \sum_n \begin{bmatrix} r_n^m \\ d_n^m \end{bmatrix} \tilde{\psi}_n^m(t) = \sum_n \begin{bmatrix} \cos\theta_n^m & -\sin\theta_n^m \\ \sin\theta_n^m & \cos\theta_n^m \end{bmatrix} \begin{bmatrix} A_n^m \\ 0 \end{bmatrix} \tilde{\psi}_n^m(t), \tag{2.8}$$

where

$$\theta_n^m = \arctan(\frac{d_n^m}{r_n^m}), \quad \text{and} \quad A_n^m = \sqrt{(d_n^m)^2 + (r_n^m)^2}. \tag{2.9}$$

With this polar coordinate representation, we can plot a vector function with amplitude $A_n^m \cdot \tilde{\psi}_n^m(t)$ and phase θ_n^m. This wavelet vector can be viewed as the degenerated case of elliptic Fourier features [17], [27] with minor axis length equals to zero and hence the ellipses flattened to two coincident lines. 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.

It is also possible to represent a smooth curve using the chain-coded tangent, cumulative tangent or curvature. This kind of differential representation [16], [34] is invariant under translation, rotation and scaling. The curvature κ 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 $\ddot{x}$ ($\dot{y}$ and $\ddot{y}$) denote, respectively, the first and second derivatives of the coordinate function $x(t)$ ($y(t)$) with respect to the parameter t of the curve representation. We can apply the finite-scale biorthogonal wavelet expansion to the curvature via

$$\kappa(t) = \sum_n \alpha_n^M \tilde{\phi}_n^M(t) + \sum_{m=1}^{M} \sum_n \delta_n^m \tilde{\psi}_n^m(t).$$

However, for a closed curve with differential representation, its reconstruction based 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 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

$$\begin{array}{rcl} x(n) & = & \sum_{k=0}^{N-1} X(k) e^{i2\pi nk/N} \\ y(n) & = & \sum_{k=0}^{N-1} Y(k) e^{i2\pi nk/N}, \end{array} \qquad n = 0, 1, 2, \cdots, N-1, \tag{2.10}$$

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

3 Properties of Wavelet Descriptor

3.1 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 β can be written as

$$\begin{bmatrix} \breve{x}(t) \\ \breve{y}(t) \end{bmatrix} = \beta \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \beta \begin{bmatrix} x_a^M(t) \\ y_a^M(t) \end{bmatrix} + \sum_{m=1}^{M} \beta \begin{bmatrix} x_d^m(t) \\ y_d^m(t) \end{bmatrix}.$$

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 (2.5) and (2.6) by the same factor β,

$$\text{Scaling:} \quad \begin{bmatrix} \breve{a}_n^M \\ \breve{c}_n^M \end{bmatrix} = \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} a_n^M \\ c_n^M \end{bmatrix}, \quad \begin{bmatrix} \breve{r}_n^M \\ \breve{d}_n^M \end{bmatrix} = \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} r_n^M \\ d_n^M \end{bmatrix}. \tag{3.1}$$

As far as the polar coordinates (2.8) are concerned, we have

$$\breve{\theta}_n^m = \arctan(\frac{\beta d_n^m}{\beta r_n^m}) = \arctan(\frac{d_n^m}{r_n^m}) = \theta_n^m,$$

and

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

Thus, the phase remains unchanged while the magnitude is scaled by the same factor β.

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, i.e.

$$\int_{-\infty}^{+\infty} \psi_n^m(t)dt = 0,$$

it is easy to see that

$$\int_{-\infty}^{+\infty} [f(t)+b]\psi_n^m(t)dt = \int_{-\infty}^{+\infty} f(t)\psi_n^m(t)dt, \quad m,n \in Z.$$

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 $y_a^M(t)$. Now , using the property of *partition of unity* which leads to definite integral of the scaling function

$$\int_{-\infty}^{+\infty} \phi(t)dt = 1,$$

and it follows

$$\int_{-\infty}^{+\infty} \phi_n^M(t)dt = 2^{-M/2}\int_{-\infty}^{+\infty} \phi(t)2^M dt = 2^{M/2}.$$

We get evidently,

$$\int_{-\infty}^{+\infty} [f(t)+b]\phi_n^M(t)dt = \int_{-\infty}^{+\infty} f(t)\phi_n^M(t)dt + 2^{M/2}\cdot b, \quad n \in Z.$$

And they can then be written as

$$\text{Translation:} \quad \begin{bmatrix} \breve{a}_n^M \\ \breve{c}_n^M \end{bmatrix} = \begin{bmatrix} a_n^M + 2^{M/2}\cdot b_x \\ c_n^M + 2^{M/2}\cdot b_y \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} \breve{r}_n^M \\ \breve{d}_n^M \end{bmatrix} = \begin{bmatrix} r_n^M \\ d_n^M \end{bmatrix}. \tag{3.2}$$

Third, by rotating the curve by a counter-clockwise angle φ with the origin $(x,y) = (0,0)$ as the pivot point, we have

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

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

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

$$\text{Rotation:} \quad \breve{\theta}_n^m = \theta_n^m + \varphi, \quad \text{and} \quad \breve{A}_n^m = A_n^m, \tag{3.3}$$

for the wavelet coefficients of the detail 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 3.4.

3.2 Invariance, Uniqueness and Stability

Mokhtarian and Mackworth [21] discussed some desired properties of a general shape descriptor as stated as follows.

> *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 (2.4) 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 proof in a more general setting, we consider a class of square-integrable functions $f \in L^2(R)$ and their corresponding wavelet frame representations [6], [11]. (Note that the frame is a concept more general than the basis. A basis is a frame, but a frame may not be a basis.) We want to show 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 $\{\psi_n^m(x)\}_{m,n\in Z}$ constitute a frame in $L^2(R)$, we have [6], [11, pages 55-56].

$$A\|f\|^2 \leq \sum_{m,n\in Z} |< f, \psi_n^m >|^2 \leq B\|f\|^2, \quad f \in L^2(R), \tag{3.4}$$

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

$$\|W(f)\| \equiv (\sum_{m,n\in Z} |d_n^m|^2)^{1/2},$$

we can rewrite (3.4) as

$$A\|f\|^2 \leq \|W(f)\|^2 \leq B\|f\|^2, \quad f \in L^2(R).$$

Furthermore, by using the linearity of the operator W, we have

$$\|W(f_1) - W(f_2)\|^2 = \|W(f_1 - f_2)\|^2 \leq B\|f_1 - f_2\|^2.$$

Thus, if $\|f_1 - f_2\| < \delta = \varepsilon/B$, $\|W(f_1) - W(f_2)\| < \varepsilon$. In terms of words, the difference in the wavelet frame representations of two similar curves is small. On the other hand, since

$$\|f_1 - f_2\|^2 \leq \frac{1}{A}\|W(f_1) - W(f_2)\|^2,$$

we see that if two representations are close, the curves that they represent should be close as well.

For the biorthogonal case, we can choose the dual frames ψ_n^m and $\tilde{\psi}_n^m$ so that the following relationships [11, page 265]

$$\tilde{A}^{-1}\|f\|^2 \leq \sum_{m,n} | < f, \psi_n^m > |^2 \leq A\|f\|^2, \quad \text{and} \quad A^{-1}\|f\|^2 \leq \sum_{m,n} | < f, \tilde{\psi}_n^m > |^2 \leq \tilde{A}\|f\|^2$$

hold. The stability property can then be proved by using similar arguments.

3.3 Approximation by Different Wavelet Bases

A good wavelet basis allows a certain signal to be well represented at a lower resolution with a small approximation error. In this subsection, we compare the approximation capability of different wavelet bases in representing smooth curves. Following the idea of Daubechies [11], we investigate wavelet bases with different vanishing moments and symmetry property.

By definition, we say that a wavelet basis has L vanishing moments, if the corresponding mother wavelet $\psi(t)$ satisfies

$$\int t^l \psi(t) dt = 0, \quad l = 0, \cdots, L-1.$$

By expanding a scalar function $f(t)$, $f(t) \in V_{m_0-1}$ or any space includes V_{m_0-1}, with Taylor series at $t_0 = 2^{m_0}n$, we obtain

$$\begin{aligned} f(t) = \; & f(2^{m_0}n) + f^{(1)}(2^{m_0}n)(t - 2^{m_0}n) + \tfrac{1}{2!}f^{(2)}(2^{m_0}n)(t - 2^{m_0}n)^2 + \cdots + \\ & \tfrac{1}{(L-1)!}f^{(L-1)}(2^{m_0}n)(t - 2^{m_0}n)^{L-1} + (t - 2^{m_0}n)^L R(t), \end{aligned}$$

where $|R(t)| \leq K < \infty$, if $f \in C^{L-1}$. Let us multiply both sides of the above equation with $\psi(2^{-m_0}t - n)$ and integrate. If the basis has L vanishing moments, the first L terms will become zero so that we derive a bound on the wavelet coefficient, i.e.

$$\begin{aligned} |< f, \psi_n^{m_0} >| \; &= \; \left| 2^{-m_0/2} \int (t - 2^{m_0}n)^L R(t) \psi(2^{-m_0}t - n) \cdot dt \right| \\ &\leq \; K \cdot 2^{m_0(L+1/2)} \int |v|^L |\psi(v)| dv. \end{aligned}$$

For a very fine scale (or a large negative value of m_0), the term will be small unless $t = 2^{m_0}n$ is near one of the singularities of f.

For our current application, the scaling function $\phi(t)$ describes the coarsest scale approximation. Thus, it is also important to consider the vanishing moments of $\phi(t)$. Daubechies and Coifman [10] constructed a class of orthonormal wavelet bases with an equal number of vanishing moments for ψ and ϕ known as the Coiflets. 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 ψ and ϕ cannot be achieved by using compactly supported orthonormal wavelet bases except for the Haar basis. It turns out that the mother wavelet ψ and the scaling function ϕ of Coiflets are much more symmetric than their counterparts in Daubechies and least asymmetric bases. However, Coiflets requires longer support of the basis function to achieve the same order of vanishing moments.

To compare the performance of different wavelet bases in representing a shape, we use the Daubechies [9], least asymmetric [11], Coiflets [3], [10] and spline-variant [11] bases to represent a square, and show the results in Fig. 2. The Daubechies and least asymmetric bases are both compactly supported wavelets with a maximum number of vanishing moments for a given support region while the spline-variant wavelet is a biorthogonal basis. The approximations by wavelet bases with various vanishing moments are plotted.

We see from Fig. 2 that biorthogonal spline-variant wavelet gives the best results, since it represents both the corners and edges well. We perform many other experiments and obtain the following two 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 biorthogonal spline-variant or Coiflet give the better choices. Second, a basis with a higher vanishing moment gives a better approximation at the price of a higher computational cost.

3.4 Normalization of Wavelet Coefficients

The invariance, uniqueness and stability properties of the wavelet descriptor proved in Section 3.3 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 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, 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 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 be zero. 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 (2.2) 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 $(b_x, b_y) = N_1^{-1} \sum (a_n^M, c_n^M)$ 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 $\bar{A} = N_2^{-1} \sum_{n,m} A_n^m$ 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 $\bar{\theta} = N_2^{-1} \sum_{n,m} \theta_n^m$ 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 signatures recognition. It is worthwhile to point out that the above normalization procedure does not only normalize the shape for ease of classification but also remove noise in raw data.

3.5 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 which is the most widely used transform method in characters and silhouettes recognition.

To reconstruct the closed planar curve, it is often that only a small set of coefficients of low frequency components are needed since the remaining coefficients were small. As a result, the

Fourier descriptor provides a efficient tool to represent a closed curve. We refer to [32] for the normalization of the Fourier descriptor.

In this experiment, 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 24×24. 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 piece-wise linear contour to represent the shape of the original image array, and the starting point in traversing the boundary curves for each character is fixed. Once the length of the contour is computed, we interpolate and resample the boundary contour so that the total number of samples is of a power of two. Then, the coefficients of Fourier and wavelet descriptors are computed by using (2.10) and (2.2), respectively.

Three fonts for each character, as shown in Fig. 6, were used in the training phase. We found that only some lower resolution (or frequency) coefficients have substantial large values while the magnitudes of other coefficients are negligible. Thus, 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 four features, we found that the three characters 5, S, and 2 were misclassified by the Fourier descriptor while the two characters Q and O were misclassified by the wavelet descriptor. In the case of classification with seven features, the character S was misclassified to 5 by the Fourier descriptor while all characters were classified correctly by the wavelet descriptor.

4 Deformable Wavelet Descriptor

4.1 Basic Idea

To model a group of shapes which 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 [27] used coefficients of the Fourier descriptor as the parameters of deformable models. However, the basis functions of the Fourier descriptor are the sinusoids which are periodic

and global so that a small perturbation of one parameter will affect the entire boundary 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, provide a scheme to model local as well as global deformation. In Fig. 4, we model the amplitude of a certain wavelet vector as a Gaussian random variable and assume all other wavelet vectors (defined as in (2.9)) to be deterministic. The deformation using a 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 $H(\mathbf{x}) = -\sum p(x)\ln p(x)$, where $p(x)$ is the probability density function of $\mathbf{x}$. Therefore, for a Gaussian random variable $\mathbf{x}$ the entropy [23] is equal to $H(\mathbf{x}) = \ln\sigma\sqrt{2\pi e}$, where σ is the standard deviation of $p(x)$. In Fig. 3, we plot the distribution of coefficients of a group of deformed curves by applying the Gaussian deformation to the spatial coordinates, where the vertical bars depict the standard deviation σ of each wavelet coefficient. We see some coefficients have large entropy values while others have small ones in the case of wavelet descriptor. This confirms 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 straight forward representation of the curve would result in discontinuities so that 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. 5 shows an example of a nonclosed curve and its deformations. In Fig. 5 (c) and Fig. 5 (d) curves are locally deformed in x and y directions. In contrast, these locally deformed curves cannot be easily achieved by using the Fourier descriptor deformable model. The local deformable property makes the wavelet descriptor ideal for handwritten signature or symbol recognition.

4.2 Model-Based Contour Extraction

We consider one particular application of the deformable wavelet descriptor that extracts boundary of an object from noisy images with a model-based/Bayesian approach. 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. That is, we apply the Laplacian of Gaussian (LoG) [19] operation to $F(j,k)$ to obtain an estimate of the edge location. The LoG operator is defined as

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

where

$$(-\nabla^2 G)(j,k) = c[1 - \frac{j^2+k^2}{2\sigma^2}]\exp(-\frac{j^2+k^2}{2\sigma^2}).$$

and where c normalizes the sum of the elements of a given size mask to unity and the parameter σ controls the spread of the Gaussian smoothing kernel. The lowpass Gaussian filters reduce 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 a maximum *a posteriori* (MAP) estimation problem. Let C_{MAP} denote the maximum *a posteriori* estimation of boundary curve C. Then, the MAP problem of our interest can be written as

$$Pr(C_{MAP}|I) = \max_{\vec{p}} Pr(C_{\vec{p}}|I) = \max_{\vec{p}} \frac{Pr(I|C_{\vec{p}})Pr(C_{\vec{p}})}{Pr(I)}, \tag{4.1}$$

where $Pr(C_{\vec{p}})$ is the *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

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

By assuming independent zero-mean Gaussian noise with standard deviation σ_n, we know that the likelihood function of obtaining I given template $C_{\vec{p}}$ 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 above, 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 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

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

The second term of (4.3) is the natural logarithm of the *a priori* probability which confines the solution C_{MAP} to a group of curves within the prescribed model, while the first term is the approximation of logarithm of the likelihood $Pr(I|C_{\vec{p}})$ and is, in a sense, a cross-correlation [25] of the contour $C_{\vec{p}}$ and $I(j,k)$. Note also that the cross-correlation is weighted by a factor $\frac{k}{\sigma_n^2}$ which 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 severer.

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 4.3. Let us choose

$$\ln Pr(C_{\vec{p}}) = \sum_{i=1}^{N} \ln \frac{1}{\sigma_i \sqrt{2\pi}} - \frac{(p_i - m_i)^2}{2\sigma_i^2}, \tag{4.4}$$

where the mean m_i and the standard deviation σ_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 prior 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_{MAP} can be calculated in a pyramidal structure. i.e. searching the MAP estimation in a coarser resolution level and then switching to the higher resolution to meet the matching criteria. The hierarchical algorithm reduces the computational cost significantly and has the advantage that the result is very likely to be a global optimum.

4.3 Experiments

In this experiment, we use the deformable model to extract the contour of a heart from a certain CT (Computer Tomography) image. The extraction of the cardiac contour is important in assessing the regional diastolic function [13]. To obtain the *a priori* information of the contour of a heart, we consider a set of 15-frame noiseless CT images of size 128×128 which 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. By using the maximum *a posteriori* 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 which is too tedious to be practical while its discrete approximation is sensitive to errors resulting from an inadequate choice of step sizes. 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. We choose the quasi-Newton [12] method to maximize this term, which has a quadratic form as shown in (4.4), with the objective to match features in the coarsest scale.

The interpolated solution obtained from the coarsest scale serves as the initial guess for an optimization in the next finer scales. For all optimization problems in finer scales, we consider the objective function consisting of two terms as given in (4.3). To handle the first term of (4.3), it is useful to introduce a technique known as chamfer matching [5] which determines the best fit of edge points from two images. We calculate the normalized chamfer correlation image of the contour template $C_{\vec{p}}$ and perform the cross-correlation between the chamfer correlation image and the zero-crossing image of $F(j,k)$. Fig. 11 gives an example of the calculated cross-correlation between 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 can be calculated by fast numerical computation as follows. For a binary edge image, we first set each edge pixel to zero and each non-edge pixel to minus infinity. Then, we perform a two-pass converting on it, i.e. forward converting from upper-left to bottom-right followed by the backward converting from bottom-right to upper-left:

FORWARD CONVERTING:

for $j = 2, \cdots,$ row do

  for $k = 2, \cdots,$ column do

    $v_{j,k} = \text{maximum}(v_{j,k}, v_{j-1,k-1} - 2, v_{j-1,k+1} - 2, v_{j-1,k} - 1, v_{j,k-1} - 1).$

BACKWARD CONVERTING:

for $j =$ row $- 1, \cdots, 1$ do

  for $k =$ column $- 1, \cdots, 1$ do

    $v_{j,k} = \text{maximum}(v_{j,k}, v_{j+1,k+1} - 2, v_{j+1,k-1} - 2, v_{j+1,k} - 1, v_{j,k+1} - 1).$

In above, we use row and column to denote the dimensions of the image. By taking two to the power of $v_{j,k}$, i.e. $2^{(v_{j,k})}$, we finally normalize the image so that chamfer correlation gives the value 1 if the estimated and the original curves are totally coincided and the value 0 if totally irrelevant. Fig. 8 (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 also helps for convergence especially when the template is too far from the edge. We use Powell's method to solve the optimization problem, and progress from the coarsest scale to the finest scale. Powell's method [12] which is a special example of the conjugate direction (CD) methods, does not involve explicit computation of the gradient of a function but takes more iteration for optimization to converge. In our case, The Powell's method fits the local optimization purpose of finer scale matching which often starts from an initial curve close to the optimal solution at each scale. We show in Fig. 7 the convergence history of the objective function using the Powell's method. Note that since we perform a maximization process, the objective function grows as the optimization proceeds. Figs. 8 (b) and (c) give the corresponding result of non-model-based (maximum likelihood) optimization.

We show a typical 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. 9 (a) to $\sigma_n = 160$ as given in Fig. 9 (f), and the corresponding SNR values range from 21.3 dB to -2.8 dB. It is clear from these figures that the proposed MAP estimator using the wavelet deformable model extracts the contour accurately even when the SNR value becomes very low.

The multiscale matching process can be best illustrated in Fig. 10 where four levels of scales are used in this experiment. We show in Fig. 10 (a) that the contour extraction done at the coarsest scale by using 4 wavelet coefficients as the feature vector. The result is then used as the initial curve of matching at the next finer scale. The results of intermediate scales are given in Figs. 10 (b) and (c) by using 8 and 16 wavelet coefficients as feature vectors, respectively. Finally, the optimal estimation is achieved at the finest scale as in Fig. 10 (d) where the feature vector consists of 32 wavelet coefficients. This hierarchical process avoids the undesirable local minima resulted from noise or spurious details existing in finer scales.

We compare the results of the MAP estimation using wavelet and Fourier descriptors in Table 1, where a quantitative comparison for even frames of the cardiac image sequences with additive noise $\sigma_n = 80$ is reported and normalized chamfer correlation is chosen as the performance measure. The reason that the wavelet descriptor has better results is that it is more effective in representing the local deformations of a contour.

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

Table 1: Comparison of the MAP estimation results in terms of normalized chamfer correlation, where 1 means the original and estimated curves completely coincide while 0 means no correlation.

5 Conclusion and Extensions

By using the wavelet transform, we developed a planar curve descriptor which 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 derived. The invariance, uniqueness, and stability properties of the wavelet descriptor were proved. 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 2D and 3D shape deformation. We have recently found a work using the wavelet trasform for 3D multiscale deformable modeling by Vemuri and Radisavljevic [31]. They made an improvement on the dynamic 3D local and global deformation model of Terzopoulos [29] with superquadric parameters and a global deformable model characterized by the coarsest level wavelet coefficients . Their work considered 2D wavelets constructed from the tensor product of 1D wavelets and examined the applications of surface deformation in a 3D space. In this research, however, we focused on the planar curve descriptor, i.e. 1-D wavelet descriptor in a 2D space. Since this is a relatively simpler mathematical problem, we were able to discuss their properties in a more detail and perspective way. Besides, most of the properties derived can be straightforwardly extended to the 3D case, and the applications examined in this work also have their own practical values.

There are many interesting topics worth further studying, 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. In particular, we are currently conducting research on multiscale shape morphing. By using wavelet descriptor, we can manipulate curves and surfaces so that metamorphosis may occur along resolutions as well as spatial locations. The challenging problem is to perform shape morphing with non-self-intersecting intermediate contours.

References

[1] R. Bajcsy and S. Kovacic, "Multiresolution elastic matching," *Computer Vision, Graphics and Image Processing*, Vol. 46, pp. 1–21, 1989.

[2] A. Bengtsson and J. O. Eklundh, "Shape representation by multiscale contour approximation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 1, pp. 85–93, 1991.

[3] G. Beylkin, R. Coifman, and V. Rokhlin, "Fast wavelet transforms and numerical algorithms," *Comm. Pure Appl. Math.*, Vol. 44, pp. 141–183, 1991.

[4] C. D. Boor, *A Practical Guide to Splines*, Springer, 1978.

[5] G. Borgefors, "Hierarchical chamfer matching: a parametric edge matching algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-10, No. 6, pp. 849–865, 1988.

[6] C. K. Chui, *An Introduction to Wavelets*, Academic press, 1991.

[7] A. Cohen, "biorthogonal wavelets," in *Wavelets: A Tutorial in Theory and Applications* (C. K. Chui, ed.), pp. 123–152, Academic press, 1992.

[8] P. E. Danielsson, "A new shape factor," *Computer Graphics Image processing*, Vol. 7, pp. 292–299, 1978.

[9] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Comm. on Pure and Appl. Math.*, Vol. 41, pp. 909–996, 1988.

[10] I. Daubechies, "Orthonormal bases of compactly supported wavelets II. Variations on a theme," *submitted to SIAM J. Math. Anal.*, 1990.

[11] I. Daubechies, *Ten Lectures on Wavelets*, Philadephia: SIAM, 1992.

[12] J. E. Dennis and R. B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, 1983.

[13] E. L. Dove, K. P. Philip, D. D. McPherson, and K. B. Chandran, "Quantitaive shape descriptors of left-vectricular cine-CT images," *IEEE Trans. Biomedical Engineering*, Vol. 38, No. 12, pp. 1256–1261, 1991.

[14] C. Fermeuller and W. Kropatsch, "Multi-resolution shape description by corners," *Proc. Computer Vision and Pattern Recognition*, pp. 271–276, 1992.

[15] H. Freeman, "On the encoding of arbitrary geometric configurations," *IEEE Trans. Electronic Computers*, Vol. EC-10, No. 2, pp. 260–268, 1961.

[16] B. K. P. Horn and E. J. Weldon, "Filtering closed curve," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 8, No. 5, pp. 665–668, 1986.

[17] F. P. Kuhl and C. R. Giardina, "Elliptic Fourier features of a closed contour," *Computer Graphics Image processing*, Vol. 18, pp. 236–258, 1982.

[18] T. Lindeberg, "Scale-space for discrete signals," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 3, pp. 234–254, 1990.

[19] D. Marr, "Early processing of visual information," *Trans. Roy. Soc. London*, Vol. 275, pp. 483–519, 1976.

[20] F. Mokhtarian and A. K. Mackworth, "Scale-based description and recognition for planar curves and two-dimensional shapes," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 1, pp. 34–43, 1986.

[21] F. Mokhtarian and A. K. Mackworth, "A theory of multiscale, curvature-based shape representation for planar curves," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No. 8, pp. 789–805, 1992.

[22] J. Oliensis, "Local reproducible smoothing without shrinkage," *Proc. Computer Vision and Pattern Recognition*, pp. 277–282, 1992.

[23] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 1984.

[24] E. Persoon and K. S. Fu, "Shape Discrimination Using Fourier Descriptor," *IEEE Trans on Systems, Man, and Cybernetics*, Vol. Mar, pp. 170–179, 1977.

[25] A. Rosenfeld and A. C. Kak, *Digital Picture Processing (2nd Ed.)*, Orlando, FL: Academic, 1982.

[26] A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene anslysis," *IEEE Trans. Comput.*, Vol. C-20, pp. 562–569, 1971.

[27] L. H. Staib and J. S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 14, No. 11, pp. 1061–1075, 1992.

[28] D. Terzopoulos, "Image analysis using multigrid relaxation methods," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 8, No. 2, pp. 129–139, 1986.

[29] D. Terzopoulos, "Dynamic 3D models withe local and global deformations: Deformable superquadrics," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 703–714, 1991.

[30] V. Torre and T. Paggio, "On edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 2, pp. 147–163, 1986.

[31] B. C. Vemuri and A. Radisavljevi, "From global to local, a continuum of shape models with fractal priors," *Proc. Computer Vision and Pattern Recognition*, pp. 307–313, 1993.

[32] T. P. Wallace and P. A. Wintz, "An efficient three-dimensional aircraft recognition algorithm using normalized Fourier descriptors," *Computer Graphics Image processing*, Vol. 13, pp. 99–126, 1980.

[33] A. P. Witkin, "Scale space filtering," *Proc. Int. Joint Conf. Artificial Intell.*, pp. 1019–1023, 1983.

[34] C. T. Zahn and R. Z. Roskies, "Fourier descriptors for plane closed curves," *IEEE Trans. Computers*, Vol. c-21, No. 3, pp. 269–281, 1972.

Figure Captions

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

Figure 2: Finite resolution approximation of a square with different wavelet bases.

Figure 3: Distributions of coefficients for a Gaussian deformed curve using (a) Fourier and (b) wavelet descriptors.

Figure 4: Comparison of curves reconstructed by (a) Fourier and (b) wavelet descriptors.

Figure 5: Deformation of a nonclosed curve.

Figure 6: Three fonts of a set of characters and their normalized contours.

Figure 7: An example of the convergence history of Powell's method.

Figure 8: An example of the normalized chamfer correlation image and its matching result using Powell's method.

Figure 9: Contour extraction by the maximum *a posteriori* estimation using the deformable wavelet descriptor.

Figure 10: Multiscale contour extraction by the maximum *a posteriori* estimation.

Figure 11: Likelihood term is calculated by the cross-correlation of the normalized chamfer correlation image of the contour template and the Laplacian of Gaussian of the image.

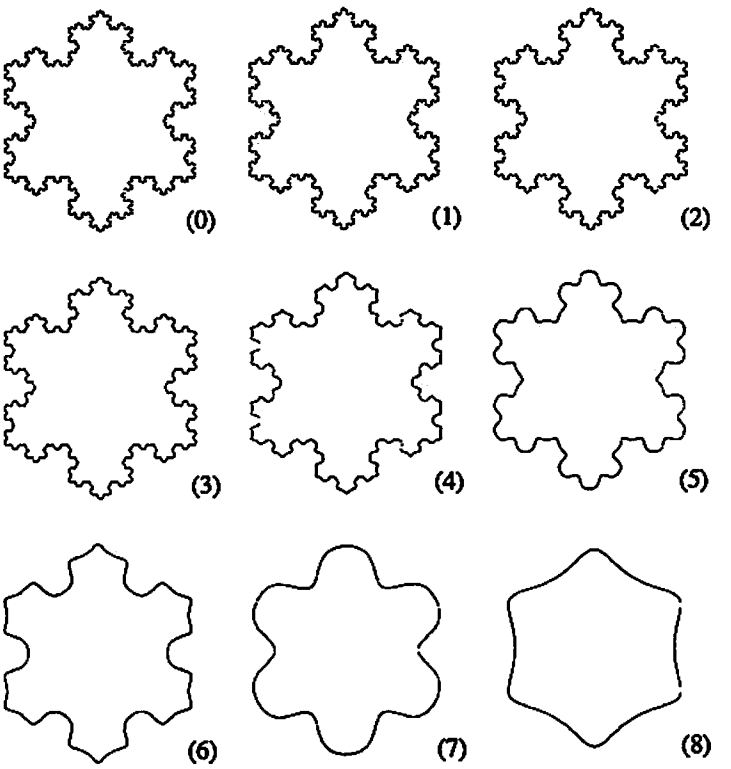


Figure 1: Multiscale representation of Koch's snowflake using the biorthogonal spline wavelet.

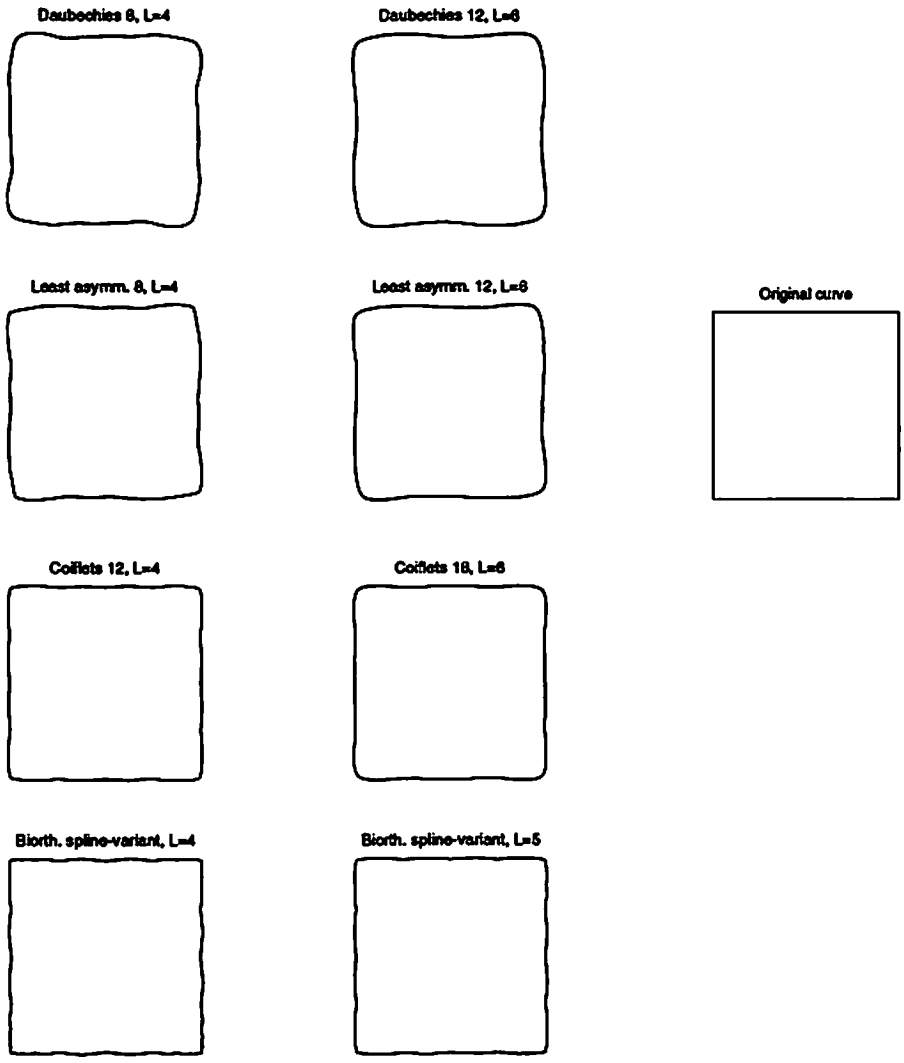


Figure 2: Finite resolution (M=5) approximation of a square with different wavelet bases, where *Daubechies n* means n-tap Daubechies filter is used and L denotes the order of vanishing moments.

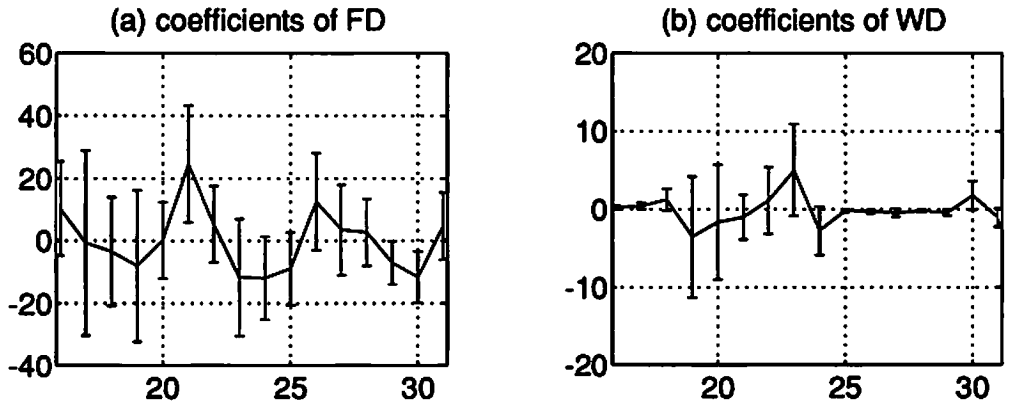
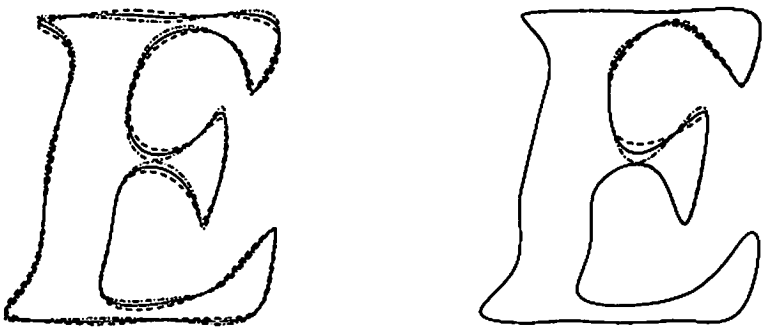


Figure 3: Distributions of coefficients for a Gaussian deformed curve using (a) Fourier and (b) wavelet descriptors.



(a) (b)

Figure 4: Comparison of curves reconstructed by (a) Fourier and (b) wavelet descriptors, where the solid, dash and dashdot lines represent the mean curves and curves with minus and plus one standard deviation, respectively.

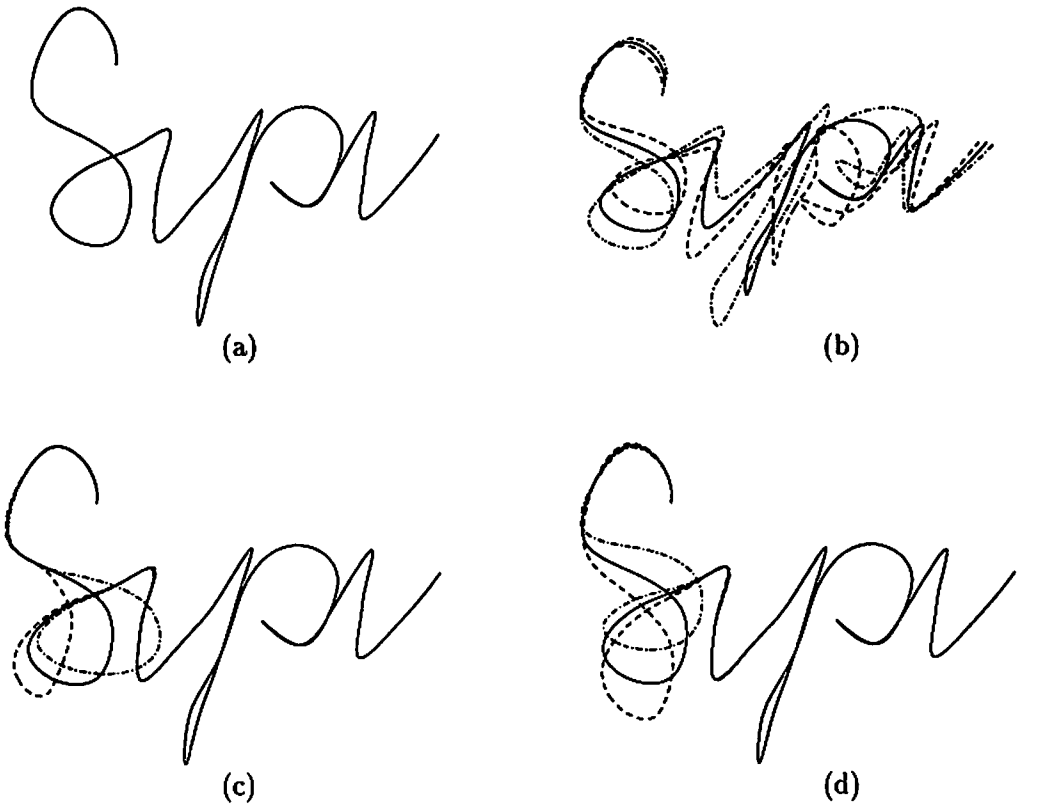


Figure 5: Deformation of a nonclosed curve: (a) the original curve; (b) the globally deformed curve; (c) locally deformed in x direction; and (d) locally deformed in y direction only. Note that the solid, dash and dashdot lines represent the mean curves, curves with minus and plus one standard deviation, respectively.

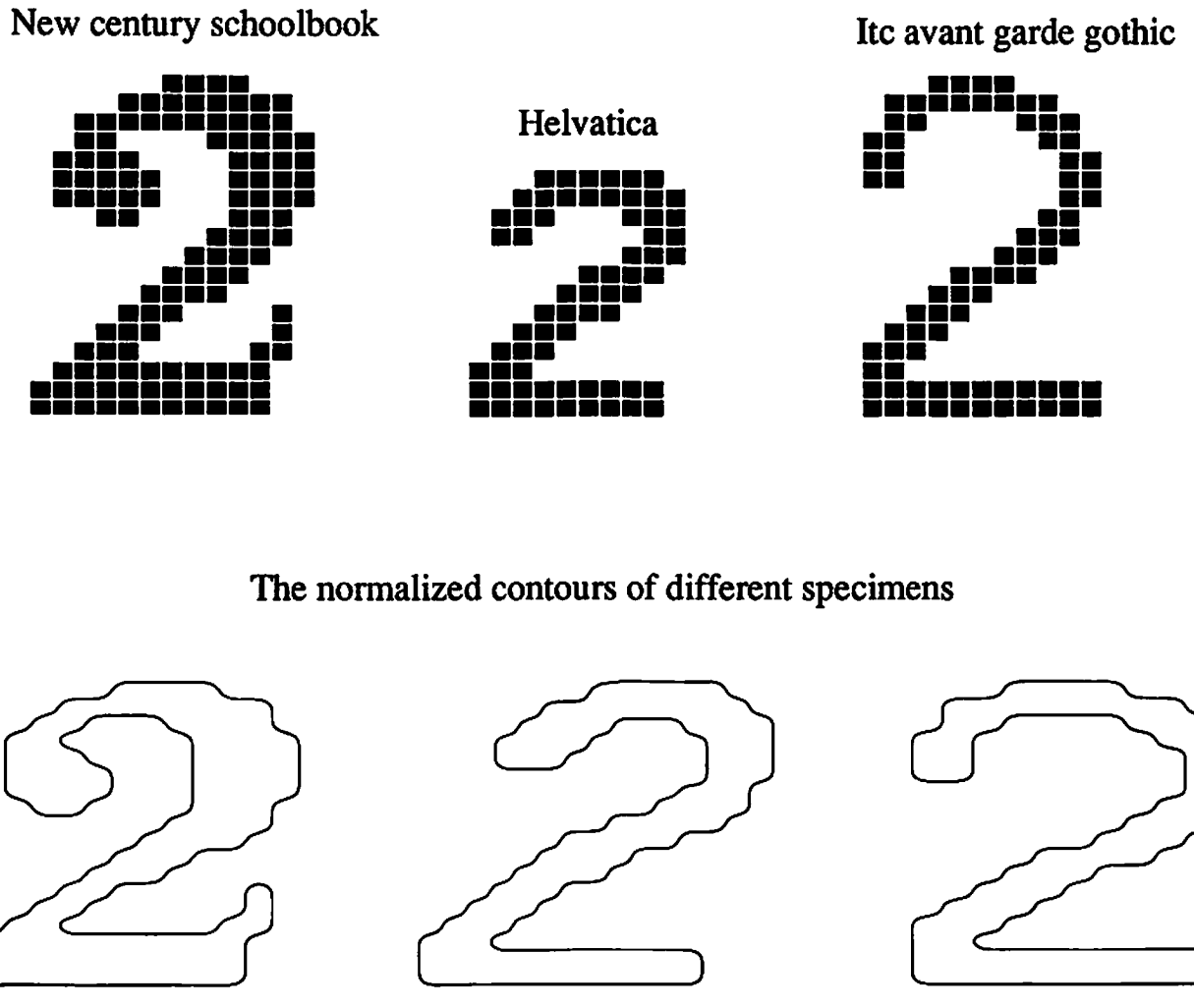


Figure 6: Three fonts of the symbol "2" and their normalized contours.

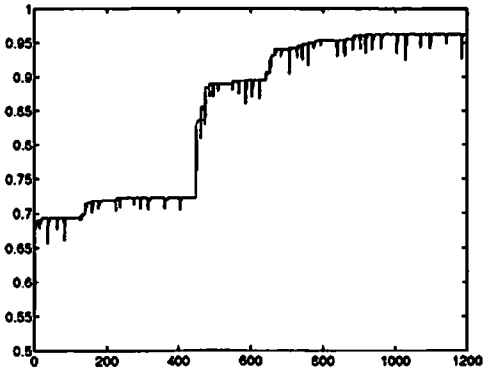
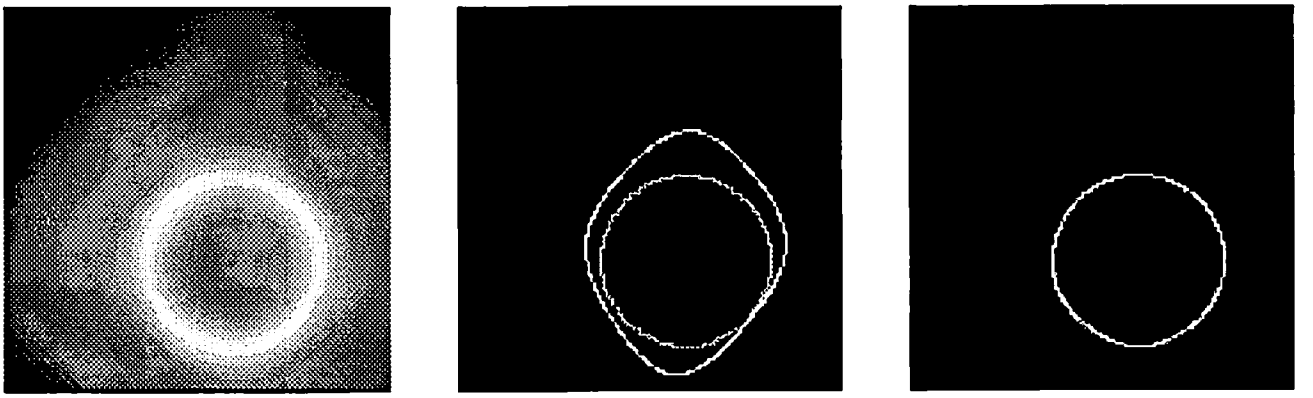


Figure 7: An example of the convergence history of Powell's method.



(a) (b) (c)

Figure 8: An example of the normalized chamfer correlation image and its matching result using Powell's method: (a) the normalized chamfer correlation image; (b) the initialization and (c) the final result of the multiscale elastic matching.

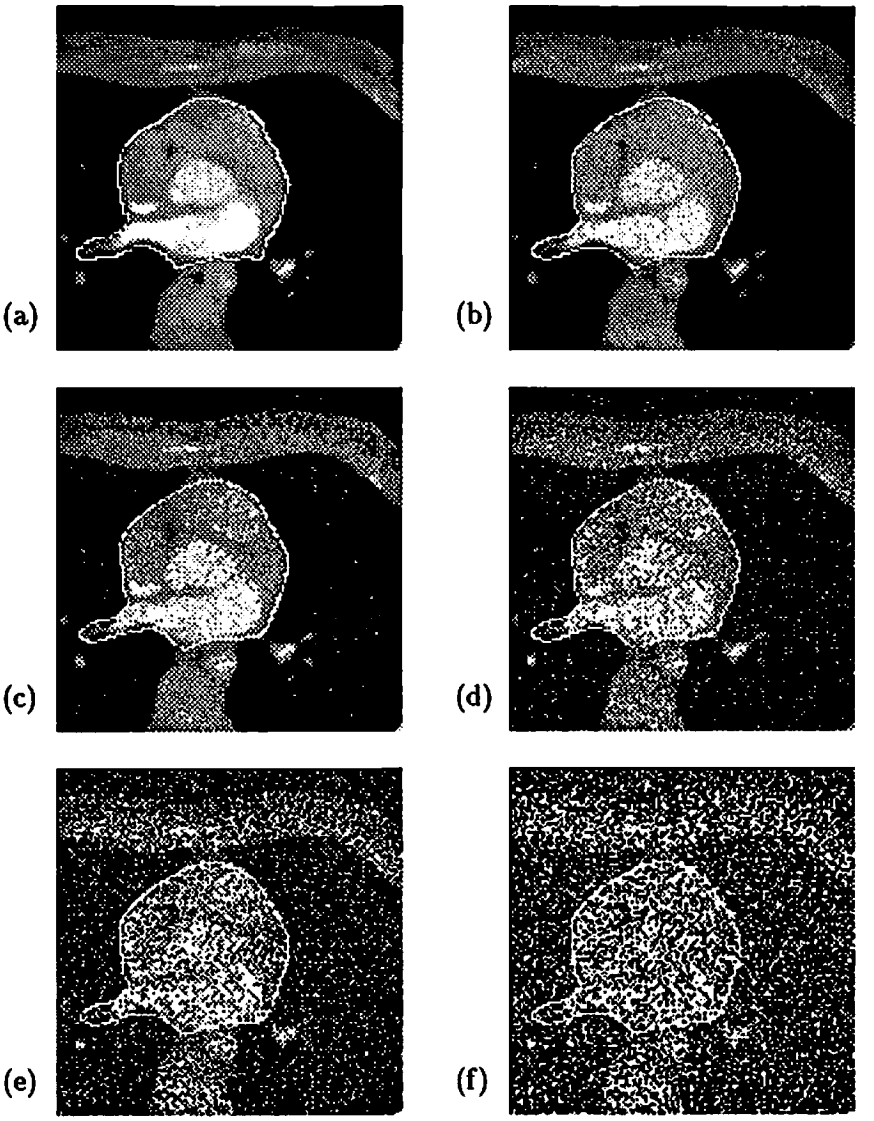


Figure 9: Contour extraction by the maximum *a posteriori* estimation using the deformable wavelet descriptor: (a) original picture (frame no. 2) with contours detected by the Laplacian of Gaussian and images corrupted by additive white Gaussian noise with the SNR equal to (b) 21.3 dB (c) 15.3 dB (d) 9.2 dB (e) 3.2 dB and (f) -2.8 dB.

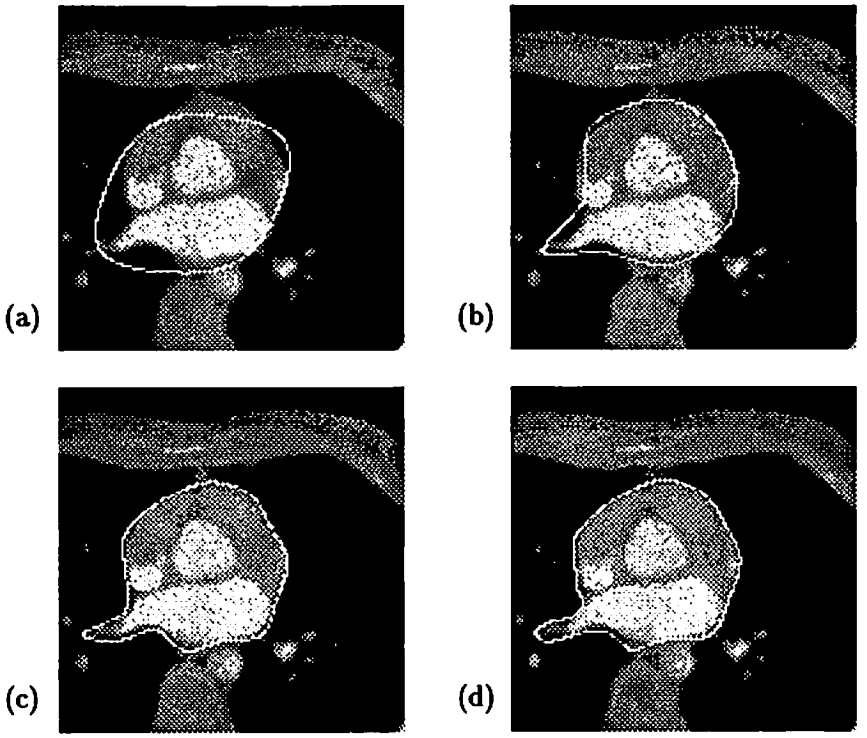


Figure 10: Multiscale contour extraction by the maximum *a posteriori* (MAP) estimation. The contour extraction, using 4 wavelet coefficients, was first done at the coarsest scale as shown in (a). The result is used as the initial curve for the next finer scale matching. Results of intermediate scales are given in (b) and (c). The optimal estimation obtained at finest scale with 32 wavelet coefficients is shown in (d).

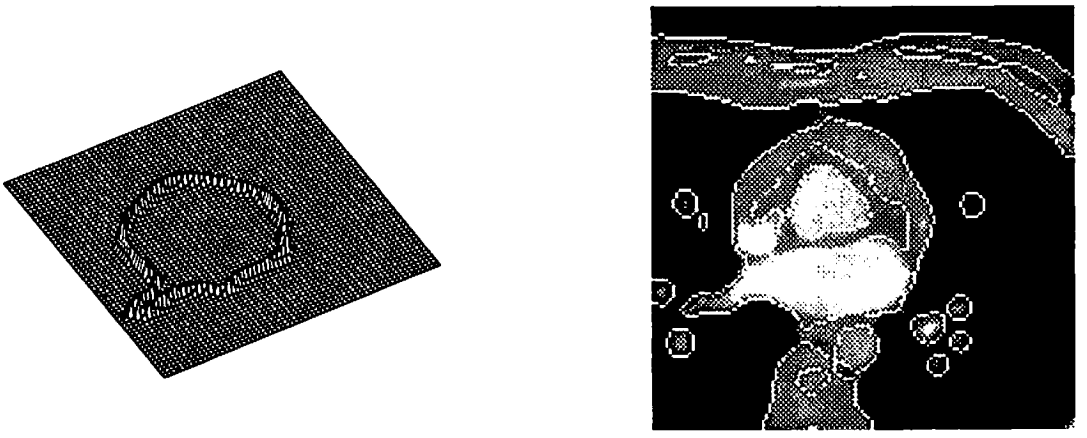


Figure 11: Likelihood term is approximated by the cross-correlation of (a) normalized chamfer correlation image of the contour template $C_{\vec{p}}$, and (b) zero-crossing image I. I is calculated by applying the Laplacian of Gaussian operator to the cardiac image.