

# LV Volume Quantification via Spatiotemporal Analysis of Real-Time 3-D Echocardiography

Elsa D. Angelini, Andrew F. Laine\*, Shin Takuma, Jeffrey W. Holmes, and Shunichi Homma

**Abstract**—This paper presents a method of four-dimensional (4-D) (3-D + Time) space–frequency analysis for directional denoising and enhancement of real-time three-dimensional (RT3D) ultrasound and quantitative measures in diagnostic cardiac ultrasound. Expansion of echocardiographic volumes is performed with complex exponential wavelet-like basis functions called brushlets. These functions offer good localization in time and frequency and decompose a signal into distinct patterns of oriented harmonics, which are invariant to intensity and contrast range. Deformable-model segmentation is carried out on denoised data after thresholding of transform coefficients. This process attenuates speckle noise while preserving cardiac structure location. The superiority of 4-D over 3-D analysis for decorrelating additive white noise and multiplicative speckle noise on a 4-D phantom volume expanding in time is demonstrated. Quantitative validation, computed for contours and volumes, is performed on *in vitro* balloon phantoms. Clinical applications of this spatiotemporal analysis tool are reported for six patient cases providing measures of left ventricular volumes and ejection fraction.

**Index Terms**—Echocardiography, LV volume, spatiotemporal analysis, speckle denoising.

## I. INTRODUCTION

CARDIAC three-dimensional (3-D) imaging was conceived as a method to circumvent the shortcomings of two-dimensional (2-D) echoplanar imaging. Volume quantification and performance assessment of irregularly shaped 3-D cardiac chambers or the description of valve morphology using 2-D images is inherently problematic because of the dynamic nature of the heart. There has been a tremendous effort within the ultrasound community over the last decade dedicated to the development of new cardiac 3-D echo technology. In the first generation of 3-D echo systems (3D FreeScan), a spatial locator was incorporated on the ultrasound beam for the reconstruction of a 3-D representation of the heart. A sound or magnetic signal-emitting device was positioned in a *free-hand* ultrasound transducer, allowing the position of any particular

2-D plane to be registered by an external locator and used for later reconstruction. Since five to ten planes through the heart were required at a preset point in each cardiac cycle, this system was cumbersome to use and only a static 3-D image could be reconstructed [1], [2]. The second-generation 3-D imaging system, originally developed by Hewlett-Packard, took the form of a transducer that rotated around a fixed axis or moved along a fixed path, obtaining images of the heart at regularly spaced intervals. The 2-D images were later reconstructed into a 3-D volume. By obtaining 2-D images at different time points in the cardiac cycle, these axial slices could be later reconstructed to show the motion of the heart. However, the acquisition remained cumbersome and reconstruction was a time-consuming, off-line process. Nevertheless, these *rotational probes* have yielded accurate results for mass and volume calculations of cardiac chambers [3], [4] and described well a variety of cardiac lesions [5]. However, because of inherent difficulties these systems have not become clinically applicable. In order to circumvent the problems associated with the previous forms of 3-D imaging, a third-generation of volumetric imaging (Volumetrics Medical Imaging, Durham, NC) was developed [6]–[10] introducing real-time 3-D ultrasound (RT3D). This modality is fundamentally different from the previous generations of 3-D systems as a true volume rather than a summation or interpolation of 2-D planes is acquired. This system uses a 2.5-MHz matrix-array transducer that consists of 2-D phased arrays and offers steering in both the azimuth and elevation of the beam, permitting interrogation of a pyramidal volume through the body. In each of these 2-D arrays, 512 elements connect to the system, 256 of the 512 elements are used for transmission, and 256 are used for reception. Because the RT3D system has 16:1 parallel processing, the volumetric scanner consists of a 4096 channel-receive system (256 elements  $\times$  16 parallel processing). The RT3D transducer can accomplish a typical cardiac scan at a rate of 20 volumes/s at a depth setting of 15 cm. Therefore, for a typical heart rate of 60 beats/min, 1 s is required to obtain an entire cardiac cycle with 20 volumes captured from systole to diastole. During a clinical examination, the data is displayed as two adjustable orthogonal B-scans, which are perpendicular to the transducer, and two or three C-scans that are parallel to the face of the transducer [7], [11]. The transducer acquisition volume and clinical display interface tool (FreeScan from Echotech 3D Imaging Systems) are presented in Fig. 1. All scan angles are steerable so that various planes within each of the volume sets can be reviewed on screen while the data is recorded on a computer.

Clinical studies on cardiac function require an accurate method to extract myocardial boundaries from 3-D echo data sets. Several approaches for epicardial and endocardial border

Manuscript received June 16, 2000; revised March 24, 2001. This work was supported in part by the Withaker Foundation. The Associate Editor responsible for coordinating the review of this paper and recommending its publication was M. Insana. *Asterisk indicates corresponding author.*

E. D. Angelini and J. W. Holmes are with the Department of Biomedical Engineering, Fu Foundation School of Engineering and Applied Science, Columbia University, New York, NY 10027 USA.

\*A. F. Laine is with the Department of Biomedical Engineering and Radiology, Fu Foundation School of Engineering and Applied Science, Columbia University, 416 CEPSR, MC8904, 530 West 120th Street, New York, NY 10027 USA (e-mail: laine@columbia.edu).

S. Takuma and S. Homma are with the Department of Medicine, Echocardiography Laboratories, College of Physicians and Surgeons, Columbia-Presbyterian Medical Center, Columbia University, New York, NY 10032 USA.

Publisher Item Identifier S 0278-0062(01)04691-2.

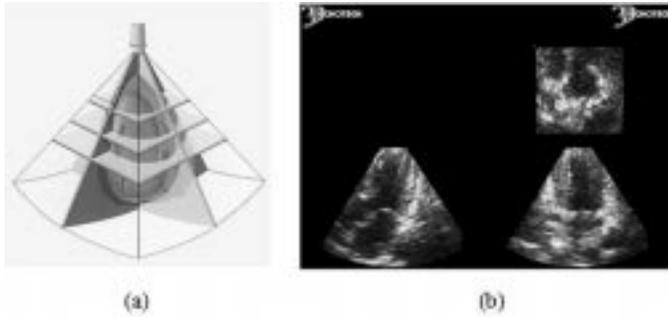


Fig. 1. RT3D acquisition and clinical interface. (a) Volumetric scanning of the left ventricle with steerable sectors. Two adjustable orthogonal B-scans perpendicular to the transducer and three adjustable C-scans parallel to the face of the transducer are displayed. (b) Initial views of RT3D echocardiography with two parallel C-scans and two orthogonal B-scans obtained from an apical window.

detection from ultrasound data have been reported during the past decade with partial success. Recent studies include methods based on statistical Markov random field models [12]–[15], fuzzy logic [16], [17], neural networks [18], [19], morphological filters [18], [20], active contours, and level sets [21]–[25]. A common motivation for these efforts, which have focused on the development of new methods of volume extraction, is that existing segmentation tools are not adapted to this type of data and do not meet the accuracy requirements of clinical applications. Existing segmentation methods perform poorly on RT3D ultrasound data because of the low spatial resolution of this modality, attenuation artifacts that lead to ill-defined myocardial borders and the addition of frequency-dependent speckle noise that corrupts the specificity of gray level values to characterize the presence of an interface or a particular tissue type. These characteristics of ultrasonic data, inherent to the physics of acquisition, impede simple region-based or boundary-based methods from performing correctly. Indeed, the underlying physics and mathematical assumptions supporting these methods are often violated during the acquisition of *in vivo* RT3D ultrasound data. This motivated our development of a new framework to recover information within a nontraditional domain where speckle noise is decorrelated.

The recent development of a RT3D imaging modality that captures an entire left-ventricular (LV) volume instantaneously with fixed geometric parameters for an entire data set raises new issues and challenges for denoising and volume extraction. On one hand, resolution of RT3D is even lower than in previous 3-D generations and the level of speckle noise is very high. But, on the other hand, the amount of information recorded per cardiac cycle is much greater than with any other ultrasonic modality. And there exists a strong coherence of surfaces in 3-D space and time for echos recorded from cardiac tissue. Also, RT3D eliminates the need for slice registration and reconstruction, leaving segmentation as the only barrier to a rapid, accurate and, therefore, clinically applicable automated calculation of LV cavity volume and quantification of LV function.

Pedagogically, our approach was to construct a denoising and enhancement method tailored to the nature of RT3D ultrasound that allows a *simple* segmentation tool to extract cardiac volumes in real-time and construct a computer dynamic model of the beating heart.

Recent publications in the literature on 3-D segmentation of 4-D freehand (first-generation) and rotational (second-generation) ultrasound include the work of Salustri *et al.* [26], Rohling *et al.* [27], Treece *et al.* [28], Ofili *et al.* [29], Noble *et al.* [30]. Stetten *et al.* developed a medial-node model for rapid identification and measurement of objects in RT3D ultrasound data [31].

Our spatiotemporal analysis tool is based on phase information rather than intensity. This strategy is best suited for denoising and cardiac boundary enhancement as first suggested by Noble *et al.* [32] with directional Gabor filters. In this paper, we use a new set of basis functions called brushlets, introduced by Coifman and Meyer in 1997 [33].

Spatiotemporal analysis of 4-D cardiac volumes with brushlet basis functions provides projected coefficients that are associated with distinct “brush strokes” of a particular size and orientation in four dimensions (3-D volume + Time). Brushlet coefficients may be modified in the transform domain via resetting of high frequency and nonlinear enhancement of low frequency coefficients.

In 1996, McInerney and Terzopoulos published an extensive survey of the use of deformable models in medical image analysis [34]. More recently, several papers on cardiac echocardiography with deformable models have been published. Chalani *et al.* [22] used active contours to detect both epicardium and endocardium borders on echoplanar ultrasound images, Papademetris *et al.* [35] used integrated deformable models to segment *in vitro* ultrasound of dog hearts acquired with a rotational-probe; Montagnat *et al.* [36] used 3-D deformable models with constrained deformations on cylindrical echocardiograms. In this study, segmentation was performed on denoised volumes, after enhancement and denoising in the transform domain. A 2-D deformable model performed segmentation of the endocardial borders and LV volumes were then reconstructed for quantitative measures. This segmentation was tested and validated on phantom and clinical RT3D data sets. Quantitative measures of the extracted volumes are reported for both validation studies.

The remainder of this paper is organized as follows. Section II describes the methodology for multidimensional space–frequency analysis with brushlet functions, nonlinear denoising–enhancement schemes and a deformable-model segmentation algorithm. Section III first presents results on denoising performance of the brushlet analysis and then describes quantitative segmentation results obtained from phantom and clinical RT3D ultrasound data sets. Section IV discusses the quality of analysis and segmentation. Section V concludes with an emphasis on future work for validation and possible enhancements of this research.

## II. METHODS

### A. Multidimensional Space–Frequency Analysis

Speckle noise corrupts ultrasonic data by introducing sharp changes in an image intensity profile, while attenuation alters the intensity of equally significant cardiac structures. These properties introduce inhomogeneity in the spatial domain and suggest that measures based on phase information rather than intensity profiles are more suited for analysis of cardiac volumes. Brushlet functions are a new family of steerable wavelet

packets based on the expansion of the Fourier transform (FT) of a signal onto windowed complex exponential functions. These functions, first introduced by Coifman and Meyer [33] for compression of highly textured images, are well localized in both time and frequency. However, we point out that the goals of compression are completely opposite to this application. We modified this original expansion to provide an overcomplete (redundant) representation. Such intentional redundancy has been shown beneficial for denoising applications [37]–[41].

1) *Definition of Brushlet Basis Functions in One Dimension:* An initial goal of the construction is the desire to build an orthonormal basis of transient functions with good time–frequency localization. For this purpose, windowed complex exponential functions for Lapped orthogonal transforms have been used for many years in the context of sine-cosine transforms [42], [43] and Gabor functions. Following the Balian–Low theorem there is no differentiable and compactly supported window  $g$  such that:  $\forall u_0$  and  $\zeta_0$ ,  $\{g(t - nu_0)e^{ik\zeta_0 t}\}_{(n,k) \in \mathbb{Z}^2}$  forms an orthonormal basis of  $L^2(\mathbb{R})$ . In order to overcome this limitation while working with complex exponentials, Meyer and Coifman introduced a division of the real axis into subintervals  $[a_n, a_{n+1}]$ , along with a new set of analysis basis functions

$$u_{j,n}(x) = b_n(x - c_n)e_{j,n}(x) + v(x - a_n)e_{j,n}(2a_n - x) - v(x - a_{n+1})e_{j,n}(2a_{n+1} - x). \quad (1)$$

The two window functions  $b_n$  and  $v$  are defined from the ramp function

$$r(t) = \begin{cases} 0, & \text{if } t \leq -1 \\ 1, & \text{if } t \geq 1 \end{cases} \quad (2)$$

and

$$r^2(t) + r^2(-t) = 1, \quad \forall t \in \mathbb{R}. \quad (3)$$

The function  $v$  is defined as

$$v(t) = r\left(\frac{t}{\varepsilon}\right)r\left(\frac{-t}{\varepsilon}\right), \quad t \in [-\varepsilon, \varepsilon]. \quad (4)$$

The function  $b$  is defined on  $[-l_n/2 - \varepsilon, l_n/2 + \varepsilon]$  as

$$b(t) = \begin{cases} r^2\left(\frac{t+l_n/2}{\varepsilon}\right), & \text{if } t \in \left[-\frac{l_n}{2} - \varepsilon, -\frac{l_n}{2} + \varepsilon\right] \\ 1, & \text{if } t \in \left[-\frac{l_n}{2} + \varepsilon, \frac{l_n}{2} - \varepsilon\right] \\ r^2\left(\frac{l_n/2-t}{\varepsilon}\right), & \text{if } t \in \left[\frac{l_n}{2} - \varepsilon, \frac{l_n}{2} + \varepsilon\right] \end{cases} \quad (5)$$

where  $l_n = a_{n+1} - a_n$  and  $c_n = l_n/2$ .

The complex-valued exponential  $e_{j,n}$  is defined as

$$e_{j,n}(x) = \frac{1}{\sqrt{l_n}} \exp\left[-2i\pi j \frac{(x - a_n)}{l_n}\right]. \quad (6)$$

The windowing functions  $b_n$  defined on  $[-l_n/2 - \varepsilon, l_n/2 + \varepsilon]$  and  $v$  defined on  $[-\varepsilon, \varepsilon]$  are displayed in Fig. 2. The parameter  $\varepsilon$  controls the degree of localization of the brushlet function and its inverse FT in time and frequency. As illustrated in Fig. 3, the smaller the value of  $\varepsilon$ , the better the localization in frequency (smaller second peak) but localization in time becomes less (spread of the central peak). This tradeoff in time–frequency resolution is analogous to the Heisenberg uncertainty principle

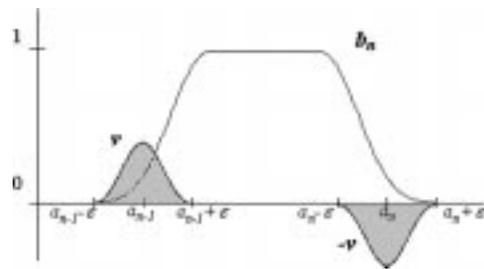


Fig. 2. Windowing functions  $b_n$  defined on  $[-l_n/2 - \varepsilon, l_n/2 + \varepsilon]$  and  $v$  defined on  $[-\varepsilon, \varepsilon]$ .

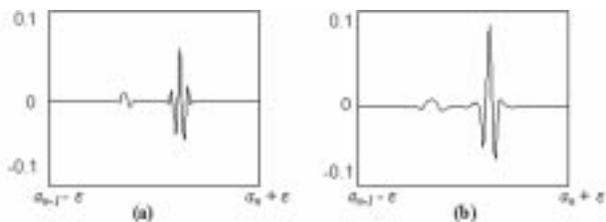


Fig. 3. Brushlet synthesis function  $W_{n,j}$  for  $l_n = 32$ ,  $j = 4$ , and (a)  $\varepsilon = 16$ , (b)  $\varepsilon = 8$ .

that constrains wavelet packet resolution and all other methods of time–frequency analysis.

Let us call a given one-dimensional (1-D) signal  $f$ , and its FT  $\hat{f}$ . We can project  $\hat{f}$  on a brushlet basis

$$\hat{f} = \sum_n \sum_j \hat{f}_{n,j} u_{n,j} \quad (7)$$

with  $u_{n,j}$  the brushlet analysis functions and  $\hat{f}_{n,j}$  the brushlet coefficients. By doing so, the FT of the signal is divided into subintervals. Each interval, indexed by  $n$ , and of size  $l_n$  is projected onto  $u_{n,j}$ , with  $j = 0, 1/l_n, \dots, (l_n - 1)/l_n$ . By applying an inverse FT, we can then compute a decomposition of  $f$

$$f = \sum_n \sum_j \hat{f}_{n,j} w_{n,j} \quad (8)$$

on the orthonormal basis  $w_{n,j}$ , inverse FT of  $u_{n,j}$ . The basis functions  $w_{n,j}$  are expressed as

$$w_{n,j}(x) = \sqrt{l_n} \exp\left[2i\pi a_n - \frac{l_n}{2}x\right] \times \left\{ (-1)^j \hat{b}_\sigma(l_n x - j) - 2i \sin(\pi l_n x) \hat{v}_\sigma(l_n x + j) \right\}. \quad (9)$$

We observe here the wavelet-like structure of the  $w_{n,j}$  synthesis functions with scaling factor  $l_n$  and translation factor  $j$ . The major difference between the brushlet basis and wavelet packets is the arbitrary tiling of the time–frequency plane and the perfect localization of a single frequency in one coefficient.

The projection of  $\hat{f}$  on the analysis functions  $u_{n,j}$  can be implemented efficiently by a folding technique and fast Fourier transform (FFT). The folding technique, described by Wickerauser [44], consists of folding the overlapping parts of the window function  $b_n$  and the bump function  $v$  across the end points of the interval, back inside itself.

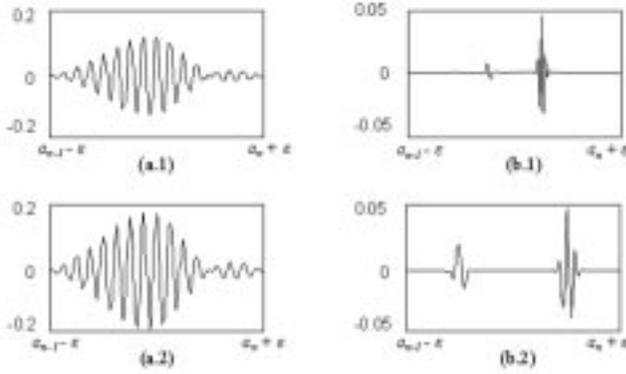


Fig. 4. Pairs of (a) analysis and (b) synthesis brushlet functions associated with intervals of different length but centered at the same frequency. (a.1–b.1):  $l_n = 64$ ,  $\varepsilon = 32$ , and  $j = 16$ ; (a.2–b.2):  $l_n = 32$ ,  $\varepsilon = 16$ , and  $j = 8$ .

If we call  $T_{a_n, a_{n+1}}$  a folding smooth periodic restriction operator and  $T_{a_n, a_{n+1}}^*$  its adjoint unfolding operator, we have the following properties for the projection of a signal  $\hat{f}$  on the functions  $u_{j,n}$

$$\langle \hat{f}, u_{j,n} \rangle = \langle \hat{f}, T_{a_n, a_{n+1}}^* e_{j,n} \rangle = \langle T_{a_n, a_{n+1}} \hat{f}, e_{j,n} \rangle. \quad (10)$$

This shows how folding and unfolding operators can be computed efficiently to project a signal on the complex exponential functions  $e_{j,n}$  with an FFT instead of computing the inner product of the original signal with the brushlet basis functions. The reconstruction can be simply computed in the same manner by an inverse FFT and an unfolding operation.

Brushlet functions constitute an orthogonal basis of a domain divided into sub intervals. There are as many basis functions as there are subintervals defining brushstrokes associated with the center frequency of each interval. The resolution of each brushstroke is inversely proportional to the size of the interval. This property is illustrated in Fig. 4, where two examples of pairs of analysis and synthesis brushlet functions associated with two intervals of different length but centered at the same frequency are plotted.

In order to control the oscillations of the brushlet synthesis function  $w_{n,j}$ , it is desirable to have a positive FT of the windowing function  $b$ . This condition is not compatible with the original construction of the  $u_{n,j}$  functions. Meyer and Coifman then introduced two biorthogonal windowed Fourier bases: synthesis function  $w_{n,j}$  and dual analysis functions  $\tilde{u}_{n,j}$ . In order to have  $\hat{b}$ , the FT of  $b$ , positive, they relaxed conditions on the ramp function defined in (2)–(3) and defined a new ramp function with the following properties:

$$r(t) = 0, \quad \text{if } t \leq 1 \quad (11)$$

and

$$r^2(t) + r^2(-t) > 0, \quad \forall t \in \mathbb{R}. \quad (12)$$

It is important to point out that since the second derivative of  $b$  is:  $b''(t) = -4\pi^2 \int_{-\infty}^{\infty} \xi^2 \hat{b}(\xi) e^{2i\pi\xi t} d\xi$ , the constraint  $\hat{b} \geq 0$  implies that  $b''(0) = -4\pi^2 \int_{-\infty}^{\infty} \xi^2 \hat{b}(\xi) d\xi$ . This verifies that  $b''(0) < 0$  and, thus, the bell function  $b$  is not flat around zero. This condition further imposes that  $\varepsilon = l_n/2$ . In our implemen-

tation, we followed the suggestion of Coifman and Meyer and defined

$$b(t) = \chi_{[-l/4, l/4]} * \chi_{[-l/4, l/4]} * \chi_{[-l/4, l/4]} * \chi_{[-l/4, l/4]} \quad (13)$$

with  $\chi_{[-l/4, l/4]}$  the characteristic function defined on the interval  $[-l/4, l/4]$ . In this case,  $b$  is compactly supported on  $[-l, l]$  and we write

$$\hat{b}(\xi) = \left( \frac{\sin(\pi\xi)}{\pi\xi} \right)^4 \geq 0. \quad (14)$$

The synthesis functions  $w_{n,j}$  are defined as stated in (9) with the cubic spline window function  $b$  and associated bump function  $v$ . The dual analysis functions  $\tilde{u}_{n,j}$  are defined in (1) by replacing  $b$  and  $v$  by  $\hat{b}$  and  $\tilde{v}$  defined with the new ramp function

$$\tilde{r}(t) = \frac{r(t)}{r^2(t) + r^2(-t)}. \quad (15)$$

2) *Brushlet Basis Functions in  $n$ -D*: We can extend brushlet analysis to  $n$ -D via separable tensor products. For a  $n$ -D signal the following steps are involved in the analysis (expansion).

- a) Computation of the  $n$ -dimensional ( $n$ -D) FFT of the signal.
- b) Tiling of the Fourier domain with arbitrary subvolume sizes along each dimension.
- c) Folding along each dimension,
- d)  $n$ -D FFT on each subvolume to project a folded signal on complex exponential terms of the brushlet functions.

Synthesis (reconstruction) of the signal includes:

- a)  $n$ -D inverse FFT on each subvolume;
- b) unfolding along each dimension;
- c) reassemble the subvolumes according to the original tiling of the Fourier domain;
- d)  $n$ -D inverse FFT of the signal.

3) *Tiling of Fourier Domain in Multiscale Analysis*: The brushlet basis offers the possibility to decompose and represent  $n$ -D signals in terms of texture patterns with different orientations and resolutions. The original work of Meyer and Coifman focused on image compression and showed that brushlets are a powerful tool for efficiently representing richly textured images [45], [46]. In this paper, we exploit this oriented-frequency representation to accomplish 4-D volume denoising prior to segmentation. The tiling of the Fourier domain determines the center frequencies of the decomposition and the resolution of the analysis. In other words, the tiling of the Fourier domain determines entirely the structure of the transform domain and can be viewed as a hyper-dimension of the analysis. The choice of the tiling determines at which frequencies the original signal is analyzed and at which resolution it is expanded. This additional dimension in the analysis process is illustrated in Fig. 5. In this figure, we illustrate in 3-D tiling in the FT domain with arbitrary sampling along each direction and the corresponding structure of brushlet coefficients in the transform domain. This flexibility in partitioning the transform domain is extremely powerful as it allows us to accommodate and precisely match the sampling rates of each dimension obtained during acquisition (sample spacing in  $x$ ,  $y$ ,  $z$ , and  $time$ ). By comparison,

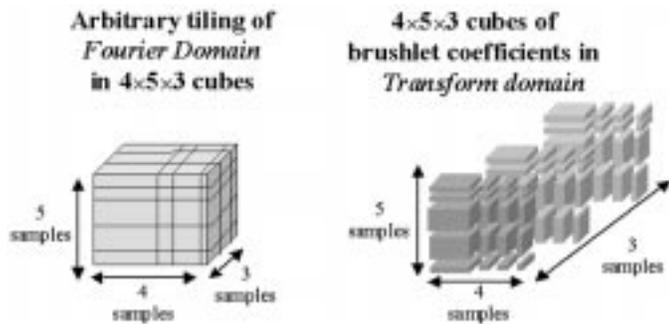


Fig. 5. Tiling of 3-D Fourier domain with arbitrary sampling along each direction and corresponding structure of brushlet coefficients in 3-D transform domain.

less optimal approaches, usually based on some interpolation scheme are required to accommodate the nonisotropic sampling rates needed in each dimension of ultrasound acquisition.

4) *An Overcomplete Framework*: A problem of particular interest for us is the manipulation of an overcomplete decomposition on brushlet basis. Overcomplete multiscale representations are well suited for image analysis and enhancement, because they avoid aliasing effects introduced by critical sampling [47] and yield a translation-invariant representation. This property is crucial for feature analysis such as edge (surface or volume) enhancement. Indeed, without this translation-invariant property, manipulation of the coefficients for feature extraction in the transform domain or thresholding for selective reconstruction could not be accurately carried out. In our case, the aliasing effects arise from the selection of overlapping subintervals of the Fourier plane expanded into a local Fourier basis. To avoid this and at the same time increase the number of coefficients for the same subinterval size along each dimension, we project onto an extended Fourier basis. The overcomplete projection is efficiently implemented by padding the folded signals with zeros along each dimension and computing its FT. Since padding a signal will increase the resolution of the FT, overcomplete projections increase the number of coefficients for the same interval and, therefore, increase resolution in the transform (coefficient) plane. For an overcomplete projection, each subinterval of the Fourier domain is projected onto a brushstroke of dimension equal to the original number of elements. The orientation and size of the original brushstroke are preserved, as the phase of the brushstroke is not modified. However, the number of points defining each brushstroke in  $n$ -D is increased to match the original size of the volumetric data. Inside a subvolume of the FT, the coefficients are stored in the same manner as the data points in the original signal. Therefore, in the overcomplete case, there is a perfect homomorphism between the location of data points in the original set and the position of the coefficients in each projected subvolume. This bijection (1:1 mapping) introduces redundancy in the transform domain, which results in a dramatic improvement in feature representation in each dimension, as observed in Fig. 6.

We point out that extension of brushlet analysis in four dimensions and overcompleteness of the expansion are two innovations from the initial 2-D decimated implementation. The computational cost added is not prohibitive since the  $n$ -D FT

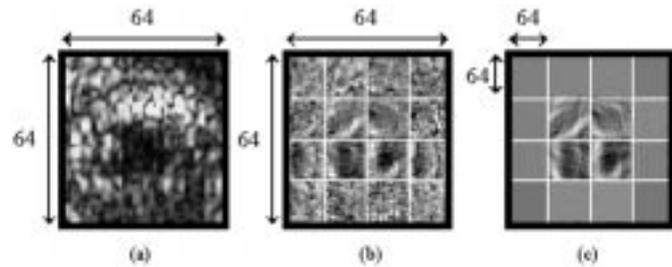


Fig. 6. Decimated and overcomplete brushlet analysis with  $(4 \times 4)$  tiling of the Fourier plane. (a) Original slice of clinical RT3D ultrasound data. (b) The 16 coefficient planes for a decimated brushlet analysis. Each coefficient quadrant is of dimension  $16 \times 16$ , four times smaller than the original data. (c) The 16 coefficient planes for an overcomplete brushlet analysis. Each coefficient quadrant is of dimension  $64 \times 64$ , the same size as the original data.

can be efficiently implemented with fast algorithms similar to the 2-D FFT.

#### B. Nonlinear Modification of Brushlet Coefficients in the Transform Domain

In order to extract features of interest and eliminate speckle noise components, we applied nonlinear thresholding on coefficients in the transform domain, in a similar fashion as previous works in cardiac denoising [48]–[50]. Recall that overcompleteness of the analysis guarantees that no distortion is introduced and a direct correspondence of feature location in both the spatial and transform domains exists. Because of the high frequency nature of the speckle noise and the absence of correlation in space and time we chose to decompose each dimension into four equal length intervals to divide the frequency spectrum of the signal into low frequencies (two inner intervals centered around the zero frequency) and high frequencies (two outer intervals). For an  $n$ -D signal we apply a  $4^n$  tiling of the Fourier domain (four subintervals/dimension) and extract only the  $2^n$  inner subvolumes centered around zero frequency. These subvolumes encode only the low frequency components of the signal along  $45^\circ$  diagonal directions in  $n$ -D. Since we wanted to detect rather isotropic endocardial and epicardial borders, we did not privilege any specific direction with tiling of the Fourier domain. Denoising in each subquadrant was performed via hard thresholding of the coefficients at 25% of maxima in the spatial domain and soft-thresholding at 15% in the time direction. These levels were selected via experimentation with a judicious trade off between attenuation of noise components and enhancement of directional features. Reconstruction was then performed with these thresholded coefficients and segmentation, described next, was then applied on the resulting reconstructed volumes.

#### C. Deformable-Model Segmentation

Segmentation of anatomical (and functional) structures remains in practice the most time consuming part of modeling a dynamic process. Automatic segmentation of anatomical structures in medical images remains an open problem. In clinical practice, manual segmentation is often considered the most reliable technique (i.e., the gold standard). Efforts by researchers in this field have shown the advantage of *deformable models* and templates coupled with object parameterization, such as those reported by McNerney and Terzopoulos in a recent survey [51].

Along these lines, techniques based on the use of deformable models for reconstructing 3-D surface shape and motion of the left ventricle from computed tomography and magnetic resonance imaging data have been developed [51], [52]. Our approach more closely follows the work of Metaxas [53] and uses a balloon deformable model first introduced by L. Cohen [54].

The class of deformable models originated with the introduction of the “snake” by Kass, Witkin, and Terzopoulos [55] and has been widely used to detect smooth contours in 2-D images. Deformable models were then applied for 2 1/2 D and 3-D object extraction from 2-D slice images [56].

A “snake” can be viewed as a curve that deforms under the influence of internal and external forces. Internal forces depend on the snake’s intrinsic properties such as its elasticity and curvature. External forces are derived from image properties (intensity gradient and edge maps).

Let a curve  $v$  be defined on  $\Omega = [0, 1]$  through the following mapping:

$$\begin{aligned} \Omega &\rightarrow \mathbb{R}^2 \\ s &\rightarrow v(s) = (x(s), y(s)). \end{aligned} \quad (16)$$

We define a snake as the space of admissible deformations that minimize the functional  $E$

$$E(v) = \int_{\Omega} \alpha \|v'(s)\|^2 + \beta \|v''(s)\|^2 + P_{\text{ext}}(v(s)) ds \quad (17)$$

with  $\alpha$  and  $\beta$  weighting parameters that control respectively the elasticity and rigidity of the snake and  $P_{\text{ext}}$  the potential of external forces, derived from the image edges. Let  $v$  be a local minimum of  $E$ , the associated Euler–Lagrange equation verified for  $v$  is

$$\begin{cases} -(\alpha v')' + (\beta v'')'' + \nabla P_{\text{ext}}(v) = 0 \\ \text{for some } v(0), v'(0), v(1), v'(1). \end{cases} \quad (18)$$

The first two terms represent the internal force that impose regularity to the curve, and the last term represents the potential of the external force that attracts the curve to features of interest. There are several approaches for defining this potential. In order to attract the snake to minima of the potential force, a simple model is to define an edge map

$$P(v) = -\|\nabla I(v)\|^2 \quad (19)$$

on the image  $I$ , or

$$P(v) = -\|\nabla (G_{\sigma} * I(v))\|^2 \quad (20)$$

with  $G_{\sigma}$  a Gaussian kernel of standard deviation  $\sigma$ . Even though the convolution of the image with a Gaussian kernel blurs the edges it has the effect of increasing the capture range of an active contour so that the initialization does not need to be close to the correct position to converge. Cohen and Cohen [57] modified the traditional gradient-based potential force to produce more stable results. They proposed to define an external force as

$$F = k_1 \vec{n}(s) - k \frac{\nabla P(v(s))}{\|\nabla P\|} \quad (21)$$

where  $\vec{n}(s)$  is the unit vector normal to the snake curve at point  $v(s)$ . The first term represents an internal pressure force that

inflates the snake as a *balloon*. This prevents the curve from shrinking or being trapped by spurious isolated edges and makes the minimization process less sensitive to the initialization. The second term normalizes the potential force to simulate a local time step. This applies the same speed to every point in the snake to overcome issues related to the selection of a common time step for every snake point. Coefficients  $k$  and  $k_1$  are chosen of the same subpixel order with  $k > k_1$  to let the snake stop at edge locations. In our work, in order to keep implementation simple, the *balloon* snake was programmed using a finite difference approximation scheme as in [58] and [59] with similar parameter values as suggested in [59]. However, rigidity was increased to diminish the time step of the dynamic equations and ensure the smoothness of a contour in situations where part of the LV wall was missing (as was often the case in our data). The edge map was defined as the gradient of the image blurred with 2-D Gaussian filter of standard deviation  $\sigma = 2$ . Three-hundred iterations were run for each short axis slice in the transform domain.

### III. RESULTS

#### A. Comparison of Denoising Performance in Three and Four Dimensions

To quantitatively evaluate the performance of dynamic 4-D analysis in decorrelating noise components in low spatial frequencies, we first experimented with mathematical phantoms corrupted with white or speckle noise. The phantom consisted of an ovoid volume growing in time that schematically mimicked the aspect of the left ventricle with an inner gray cavity surrounded by a thick white wall on a black background. The size of a single volume matrix was  $64 \times 64 \times 64$  and there were 16 volume “snap shots” growing in time. The volume size increased by 70% over 16 time frames, similar to the average ejection fraction in normal patients. We corrupted the volumes with 1) additive uniform white noise; 2) multiplicative speckle noise derived from uniform distribution; and 3) multiplicative speckle noise derived from Rayleigh distribution. All noisy signals were constructed with a SNR equal to  $-15$  dB. Corrupting noise is commonly modeled as Gaussian white noise. Ultrasound noise has been extensively studied in the literature. Because the ultrasound signal acquired by a transducer is not the raw wave signal itself but the envelope of the complex radio frequency signal, it can be shown that the initial Gaussian random noise is transformed into a signal dependent noise that follows a Rayleigh distribution in the case of fully formed speckle noise [60]. This Rayleigh distribution tends to a Gaussian white noise distribution when spatial resolution is increased. We have tested this hypothesis of Rayleigh distributed noise in data acquired with RT3D transducer and we found that it was indeed verified with good agreement. Brushlet analysis was performed with four subintervals in each direction. Denoising was performed via resetting of the higher frequency components and hard-thresholding of the lower frequency coefficients at 15% of coefficient maxima in each 3-D or 4-D subvolumes. In order to compare denoising performance to methods of reference we also performed denoising with Wiener filtering. Wiener filtering is optimal for denoising of additive white noise in the mean square error sense. We further adapted the filter to use variance

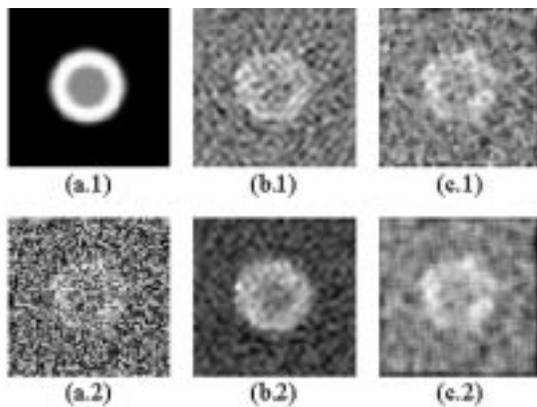


Fig. 7. Comparison of denoising performance for white noise. (a.1) Original slice, (a.2) noisy slice (SNR =  $-15$  dB). (b.1) Denoised slice with 3-D brushlet analysis (SNR =  $-14.16$  dB), (b.2) denoised slice with 4-D brushlet analysis (SNR =  $-8.5$  dB). (c.1) Denoised slice with Wiener filtering for white noise model (SNR =  $-14.84$  dB), and (c.2) denoised slice with Wiener filtering for Rayleigh noise model (SNR =  $-14.7$  dB).

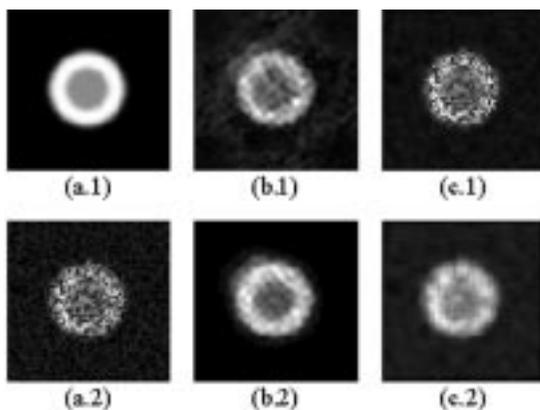


Fig. 8. Comparison of denoising performance for speckle noise with uniform distribution. (a.1) Original slice, (a.2) noisy slice (SNR =  $-15$  dB). (b.1) Denoised slice with 3-D brushlet analysis (SNR =  $-13.21$  dB), (b.2) denoised slice with 4-D brushlet analysis (SNR =  $-6.85$  dB). (c.1) Denoised slice with Wiener filtering for white noise model (SNR =  $-14.97$  dB), and (c.2) denoised slice with Wiener filtering for Rayleigh noise model (SNR =  $-14.81$  dB).

estimated from a Rayleigh noise model as presented in [61] to improve the performance on noisy data corrupted with speckle noise. A single slice from the original volumes, noisy volumes and denoised volumes processed with the four different methods is displayed in Fig. 7 for white noise, Fig. 8 for speckle noise derived from uniform distribution and Fig. 9 for speckle noise derived from Rayleigh distribution. Detailed results for the three different types of noise are summarized in Table I and commented below.

- In the case of additive white noise, the SNR of the noisy volume was  $-15$  dB. In Fig. 7, denoised results are displayed after brushlet analysis in 3-D and 4-D, and Wiener filtering with white and Rayleigh noise models. Visually, 4-D brushlet analysis clearly performed better at denoising the data. Quantitatively, we measured a SNR improvement of 51% with Wiener filtering using white noise model and 43% with 4-D brushlet analysis. Quantitative SNR improvement with 3-D brushlet analysis and Wiener filtering for speckle noise was negligible.

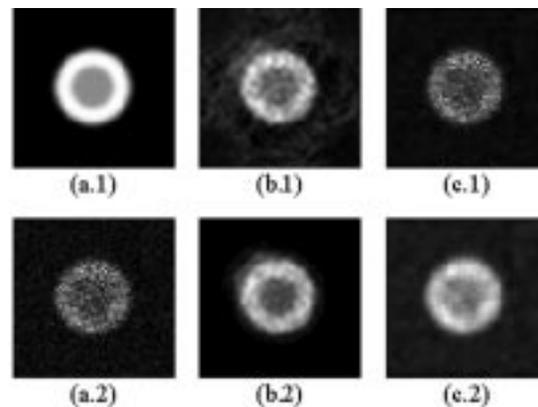


Fig. 9. Comparison of denoising performance for speckle noise with Rayleigh distribution. (a.1) Original slice, (a.2) noisy slice (SNR =  $-15$  dB). (b.1) Denoised slice with 3-D brushlet analysis (SNR =  $-13.35$  dB), (b.2) denoised slice with 4-D brushlet analysis (SNR =  $-6.75$  dB). (c.1) Denoised slice with Wiener filtering for white noise model (SNR =  $-14.98$  dB), (c.2) denoised slice with Wiener filtering for Rayleigh noise model (SNR =  $-14.84$  dB).

- In the case of multiplicative speckle noise derived from a uniform distribution, the SNR of the noisy volume was again  $-15$  dB. In Fig. 8, denoised results are displayed after brushlet analysis in three and four dimensions, and Wiener filtering with white and Rayleigh noise models. Visually, 4-D brushlet analysis clearly performed better at denoising the data. Quantitatively, significant SNR improvement of 55% was achieved by 4-D brushlet analysis while the other methods only achieved negligible improvements.
- In the case of multiplicative speckle noise derived from a Rayleigh distribution, the SNR of the noisy volume was again  $-15$  dB. In Fig. 9, denoised results are displayed after brushlet analysis in three and four dimensions, and Wiener filtering with white and Rayleigh noise models. Visually, 4-D brushlet analysis clearly performed better at denoising the data. Quantitatively, significant SNR improvement of 55% was achieved by 4-D brushlet analysis while the other methods only achieved negligible improvements.

### B. Segmentation for Volume Extraction

Segmentation was executed on two types of data: *in vitro* RT3D contrast echocardiogram phantoms, and *in vivo* clinical RT3D volume data sets.

1) *In Vitro Contrast Echocardiogram Phantoms*: *In vitro* contrast echocardiogram phantoms consisted of balloons filled with human albumin. A layer of bubbles was artificially created on the surface of the internal balloon wall to mimic the appearance of the white myocardium and increase the contrast at the wall interface. Ultrasonic RT3D data was acquired with a set up identical to clinical conditions. The phantom database used in this study consisted of two balloons of 35.0 ml and 65.0 ml. Each RT3D volume had a size of  $64 \times 64 \times 258$  voxels. Since the volume spanned by the transducer beam is conic, voxel position is determined in cylindrical coordinates. Each sample slice in the  $z$  direction (depth) was separated by 0.308 mm. For each slice,  $64 \times 64$  points were recorded spanning a total

TABLE I

QUANTITATIVE COMPARISON OF DENOISING PERFORMANCE ON 4-D MATHEMATICAL PHANTOM CORRUPTED WITH 1) WHITE NOISE; 2) SPECKLE NOISE DERIVED FROM UNIFORM DISTRIBUTION; AND 3) SPECKLE NOISE DERIVED FROM RAYLEIGH DISTRIBUTION. MEASURES OF SNR ARE REPORTED FOR A SINGLE SLICE OF THE PHANTOM DATA FOR THE ORIGINAL NOISY SLICE AND AFTER DENOISING WITH 1) 3-D BRUSHLET ANALYSIS; 2) 4-D BRUSHLET ANALYSIS; AND 3) WIENER FILTERING FOR WHITE NOISE, AND 4) WIENER FILTERING FOR RAYLEIGH NOISE

	Noisy	3D brushlet analysis	4D brushlet analysis	Wiener filtering (White noise)	Wiener filtering (Rayleigh noise)
White noise	-15 dB	-14.16 dB	-8.5 dB	14.8 dB	-7.3 dB
Speckle noise (Uniform)	-15 dB	-13.2 dB	-6.8 dB	-14.9 dB	-14.8 dB
Speckle noise (Rayleigh)	-15 dB	-13.3 dB	-6.7 dB	-14.9 dB	-14.8 dB

azimuth and elevation angles of  $64^\circ$  with  $1^\circ$  increment. With a total depth of acquisition of 7.92 cm, the optimal resolution in the short axis direction (perpendicular to the transducer) was 1.736 mm/pixel.

Brushlet analysis was performed in three dimensions with a tiling of the Fourier domain partitioned by  $4 \times 4 \times 4$  cubes as described previously in Section II. Denoising was performed in space via hard-thresholding of the lower-frequency coefficients as described in the previous paragraph. Segmentation was carried out using deformable-model on each slice. The snake was initialized with a four-pixel-radius circle inside the cavity to segment. The center of the circle was identified by a circular Hough transform applied on the edges extracted with a Prewitt filter for every ten slices of the volume. For each inspected slice, if more than one circle was detected, no center point was selected and the next slice inspected. In this fashion, for every group of ten slices we either selected one center point or none. We then computed parameters of the best linear fit in the least square sense over the entire set of center points detected. We, thus, obtained two linear models  $x = az + b$  and  $y = az + b$  to determine the  $x$  and  $y$  coordinates of the center points for each slice  $z$ . The circular Hough transform has been applied previously to RT3D data by Stetten *et al.* [62], [63] for volume measurements on balloon phantoms. Because of the high incidence of spurious edges in ultrasound data, a simple Hough transform on single slices might detect several different circles. Missing edges on the other hand can lead the Hough transform to shift the position of the center of the circular cavity. The linear best fit on the series of center points was observed to be well suited to handle these errors and provided a more robust initialization than a single slice Hough transform.

Parameters values for the active contour were set to:  $\alpha = 0.2$ ,  $\beta = 1$ ,  $\gamma = 2$ ,  $K = 3$ ,  $k_1 = 0.2$ , and  $\sigma = 2$ . An expert cardiologist performed manual segmentation of the inner balloon wall on the original volumes. Table II reports volume measures for manually traced balloon and deformable model segmentation. Examples of segmentation of the inside cavity of the two balloons are displayed in Fig. 10 for two slices.

In both cases, there is a consistent underestimation of volume estimation when using the deformable model while manual tracing estimations were within 5% of accuracy. This difference can be explained by the fact that the expert cardiologist knew that the real balloon wall borders were actually located inside the thin layer of bubbles that was inserted to increase the con-

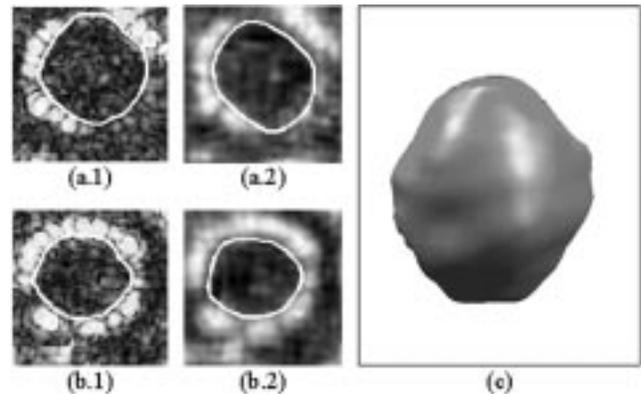


Fig. 10. Segmentation on balloon phantom data for the two data sets. (a.1–b.1) Segmentation on original data via manual tracing. (a.2–b.2) Corresponding slices after denoising via 3-D brushlet analysis using deformable-model. (c) One balloon volume reconstructed from deformable-model contours.

TABLE II  
MEASURES OF BALLOON VOLUMES WITH MANUAL SEGMENTATION ON ORIGINAL DATA AND DEFORMABLE-MODEL SEGMENTATION ON FILTERED DATA. TRUE BALLOON VOLUMES ARE REPORTED FOR COMPARISON

	Manual volume (ml)	Deformable-model volume (ml)	Real Volume (ml)
Balloon1	32.85	27.31	35
Balloon2	63.56	55.32	65

trast. From its construction and the constraints applied to it, the active contour is forbidden to go inside the white-appearance bubble layer, so that the final contours always underestimated the real balloon cavity. Testing on the balloon data sets allowed us to tune the deformable-model parameters on denoised data leading to a stable behavior in situation where part of the wall borders is missing. This situation is illustrated in Fig. 10.

2) *In Vivo Clinical RT3D Volume Data Sets*: A database of six patients co-screened with cardiac functional MRI was used for validation of LV volume segmentation and measure. Clinical data sets consisted of either: 21 time frames of size  $(64 \times 64 \times 373)$  voxels (two cases), 26 time frames of size  $(64 \times 64 \times 438)$  voxels (one case), or 21 time frames containing  $(64 \times 64 \times 438)$  voxels (three cases).

Brushlet analysis was performed in three dimensions for spatial denoising followed by 1-D analysis along time for enhance-

ment of cardiac structures. Tiling of the Fourier domain was set to four subintervals in each direction. Denoising in space was performed via hard thresholding of the lower frequency components at 25% of coefficient maxima. Enhancement was then performed in time via soft-thresholding at 15% of coefficient maxima. Threshold values in space and time were empirically set via testing on clinical data sets to achieve a good tradeoff between removal of noise artifacts and enhancement of cardiac structures.

Rescaling of the volumes from spherical coordinates to Cartesian coordinates via linear interpolation was performed prior to segmentation. Rescaling produces isotropic volumes with identical resolution in each of the three dimensions. The best spatial resolution in the acquired volumes is along the long-axis with pixel size of 0.308 mm with over 400 slices. Because of the computational time needed for the segmentation process, whether manually or with the deformable model, we decided to downsample the rescaled volumes. We initially downsampled by a factor of four but manual tracing ended up too far off from the correct values. We finally opted for a downsampling by a factor of two, reaching an acceptable compromise between accuracy of tracing and processing time. On average, 100 slices were segmented for each volume with a voxel resolution of  $(2 \times 0.308)^3 = 0.23 \mu\text{l}$ .

End-diastole and end-systole time frames were identified and manually segmented by an expert cardiologist. Manual tracing was performed on the unprocessed data while deformable-model segmentation was performed on the denoised data. The deformable model was initialized with a five-pixel-radius circle inside the cavity to segment. An operator manually selected the centers of the circle for every ten slices since the circular Hough transform applied to the balloon data sets failed to produce accurate center estimations in this case. Center locations for in-between slices were determined via linear best fit. By identifying the best linear fit to the center points detected inside the LV cavity we defined an approximation of the LV medial axis. Parameters values for the active contour were set to:  $\alpha = 0.07$ ,  $\beta = 2$ ,  $\gamma = 1$ ,  $K = 2$ ,  $k_1 = 0.09$ , and  $\sigma = 2$ . Examples of LV endocardium segmentation via manual tracing and deformable model are displayed in Fig. 11 for one short-axis and one long-axis view. After segmentation, the LV cavity was reconstructed to compute end-diastolic volume (EDV) and end-systolic volume (ESV). An example of LV volume reconstruction is also provided in Fig. 11.

Ejection-fraction (EF) was then computed as

$$EF = \left( \frac{EDV - ESV}{EDV} \right) \times 100. \quad (22)$$

The six patients were co-screened with cardiac functional MRI, which is considered as the method of reference in this study. An expert cardiologist performed manual tracing on the MRI data. Quantitative measures and errors for the six clinical cases are reported in Tables III–V.

Absolute errors of quantitative measures were computed for volumes and ejection-fraction. We evaluated mean-error values and standard deviation over the six cases for the three quantitative measures. These results are reported in Table IV. Maximum and minimum error values are provided to better assess the range

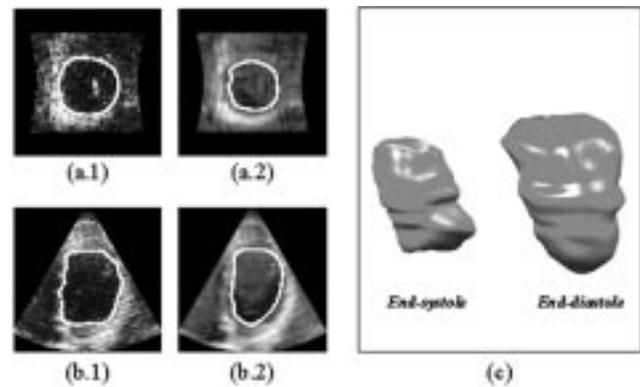


Fig. 11. Segmentation of the LV cavity in short-axis and long-axis views for one clinical data set at end-diastole. Segmentation was performed on short-axis slices. (a.1–b.1) LV segmentation on original clinical data via manual tracing. (b.1–b.2) Corresponding slices after denoising via 4-D brushlet analysis and segmented contour using deformable-model. (c) The LV cavity reconstructed from deformable-model contours at end-systole (left) and end-diastole (right).

TABLE III  
QUANTIFICATION OF EDV, ESV, AND EF FOR SIX CLINICAL CASES WITH:  
1) MANUAL TRACING ON ORIGINAL DATA AND 2) DEFORMABLE-MODEL  
SEGMENTATION ON DENOISED DATA

	Manual tracing			2D Deformable model			MRI		
	EDV (ml)	ESV (ml)	EF (%)	EDV (ml)	ESV (ml)	EF (%)	EDV (ml)	ESV (ml)	EF (%)
Volume1	87.0	43.1	50.5	99.6	41.1	58.8	97.4	41.6	57.3
Volume2	77.9	35.6	54.3	91.1	37.8	58.5	91.2	36.3	60.2
Volume3	75.8	29.5	61.1	80.0	17.0	78.7	76.1	22.9	69.9
Volume4	135.1	61.5	54.5	93.8	39.1	58.3	111.2	47.3	54.5
Volume5	79.4	37.5	52.7	75.1	25.6	66.0	77	26.2	66.0
Volume6	88.6	41.8	52.9	80.7	33.6	58.4	73.6	32.8	55.4

TABLE IV  
ABSOLUTE ERRORS IN QUANTIFICATION OF VOLUMES AND EJECTION FRACTION FOR THE SIX CLINICAL CASES REPORTED IN TABLE III. “MANUAL” REFERS TO MANUAL TRACING BY AN EXPERT CARDIOLOGIST ON ORIGINAL RT3D ECHO DATA, “MRI” REFERS TO MANUAL TRACING BY AN EXPERT CARDIOLOGIST ON MRI DATA AND “DEFORMABLE-MODEL” REFERS TO DEFORMABLE-MODEL SEGMENTATION ON DENOISED RT3D DATA. ERROR MEASURES WERE COMPUTED FOR THE SIX CASES. THE MEAN ERROR (MEAN), STANDARD DEVIATION (STD) AND MAXIMUM AND MINIMUM ERROR VALUES (MAX-MIN) ARE REPORTED FOR EDV, ESV, AND EF

Error	Manual vs. MRI			Deformable-model vs. MRI		
	EDV (ml)	ESV (ml)	EF (%)	EDV (ml)	ESV (ml)	EF (%)
Mean	10.9	7.2	6.7	5.4	2.9	2.6
Std	8.7	5.4	3.9	6.3	3.3	3.2
Max	23.9	14.2	13.2	17.4	8.2	8.8
Min	0.3	0.7	2.6	0.1	0.6	0.0

of variability achieved by the different segmentation methods. We also report in Table V percentage mean error for volumes and ejection-fraction measures over the six cases. We observed

TABLE V  
MEAN PERCENTAGE ERROR IN QUANTIFICATION OF END-DIASTOLIC VOLUME (EDV), END-SYSTOLIC VOLUME (ESV) AND EJECTION FRACTION (EF) FOR THE SIX CLINICAL CASES PRESENTED IN TABLE III

Error	Manual vs. MRI			Deformable-model vs. MRI		
	EDV (ml)	ESV (ml)	EF (%)	EDV (ml)	ESV (ml)	EF (%)
Mean (%)	11.8	22.5	10.7	5.9	8.9	4.1

with these results a significant improvement of quantitative accuracy with the deformable model when comparing to manual tracing. EDV accuracy was improved by 50.1%, ESV accuracy was improved by 60.6%, and ejection fraction measure accuracy improved by 61.4%.

#### IV. DISCUSSION

This paper presented a method for feature extraction from RT3D ultrasound that combines 4-D directional denoising and model-based segmentation.

Denoising of ultrasound data was performed via space-frequency brushlet analysis of the  $n$ -D Fourier domain and thresholding of the coefficients to remove speckle noise and enhance coherent anatomical structures in space and time. Brushlet analysis identifies efficient tilings of the Fourier domain, along each dimension of a signal. Through characterization of coefficients within sets of redundant articulated (orientation rich) bases we can separate signal and noise components. The performance of brushlet analysis to decorrelate signal from white additive and multiplicative speckle noise components was demonstrated on a 4-D mathematical phantom. These examples also showed the superiority of 4-D brushlet denoising over 3-D for 4-D data sets. We have previously shown that we can characterize and isolate features of interest in LV volumes by selection of specific brushlet coefficients [64]. Since decomposition on a brushlet basis can efficiently isolate directional features at specific frequencies, preprocessing of RT3D volumes via thresholding of lower frequency brushlet coefficient can assist segmentation by removing noise components and enhancing anatomical features. The RT3D clinical data sets suffer from very low spatial resolution due to subsampling by the transducer during acquisition of the echo, the high level of speckle noise and the motion of the cardiac muscle. These three factors diminish the resolution of the acquisition and create artifacts that can either remove parts of the myocardial wall or introduce "myocardium-like" bright signals inside the LV cavity or outside the ventricle. Since the temporal resolution is much better compared to the three spatial dimensions, the inclusion of time in the analysis of the data was a tremendous advantage. It helped to remove artifacts not persistent in adjacent time frames and added myocardial components (fill in holes) when the cardiac surface was present in adjacent time frames. We qualitatively compared denoising performance on clinical data using 3-D or 4-D brushlet analysis and Wiener filtering adapted to speckle noise. Results for the different denoising schemes are illustrated in Fig. 12 for four long-axis and four short-axis slices extracted from one clin-

ical data set. It is interesting to note that Wiener filtering produced a high quality denoising in long-axis views but introduced significant blurring that corrupted the data when displayed in short-axis views. Blurring of the data, which is a common artifact with denoising filters, prevents segmentation from realizing accurate volume quantification. Such blurring was not observed with brushlet expansions. When comparing 3-D and 4-D brushlet denoising we observed that the introduction of the temporal dimension improved the contrast of the denoised data with enhanced myocardial wall and a more homogeneous inside LV cavity. We observed that segmentation performance was enhanced by such contrast improvement as the deformable model was not trapped in local minima inside the cavity, moved faster toward the myocardial border, and localized border edges more accurately. The main power of brushlet analysis is its flexibility in decomposing  $n$ -D signals and its ability to accommodate and precisely match nonuniform sampling rates of each dimension obtained during RT3D acquisition (sample spacing in  $x, y, z$  and  $time$ ). The tiling of the Fourier domain determined the orientation and the resolution of the brush strokes on which the signal was projected. In different terms, the tiling selected the textural patterns used for the analysis of the original signal. From this point of view, tiling can be considered as an added dimension of the analysis. As described in Section II, brushlet analysis was performed for a fixed tiling of the FT that divided each dimension into four intervals. This choice allowed us to extract and denoise only the lower frequency components without selecting any specific direction. This approach was originally motivated by the isotropic nature of the LV myocardium shape in space and identification of low-frequency movements in time (i.e., slowly moving structures). The next step is to investigate mathematical and empirical methods for identifying the most "efficient" brush stroke sizes and orientations for decomposition and reconstruction. From an empirical point of view, the shape of the left ventricle is highly isotropic in short axis planes (circular wall) while rather elongated in long-axis views. This observation suggests that a selective tiling for horizontal features might be more appropriate in the  $z$  dimension. A mathematical framework for identifying an optimal basis was introduced by Coifman and Wickerhauser [65] for the wavelet packets and cosine transforms. By measuring a predefined cost function, such as entropy, the algorithm selects an optimal tiling of the Fourier domain where the energy of the signal is "best" decomposed. Meyer and Coifman [33] applied this algorithm on brushlet functions for compression of highly textured images. In the current application, optimal speckle noise removal is desired and a cost function adapted to denoising performance should be designed to carry out an optimal-basis search. We presented some preliminary work in that direction [61] with qualitative evaluation of denoising quality as a function of the tiling. However, optimization of the tiling needs to be studied further and would enable us to systematically adjust specific brushlet coefficients for noise reduction and selective reconstruction of more salient cardiac temporal features. Computation time for the 4-D brushlet analysis applied to clinical data sets will also need to be optimized for future real-time applications. Currently, the whole process of analysis and segmentation takes less than one hour, running on a Pentium II 400 MHz under Matlab.

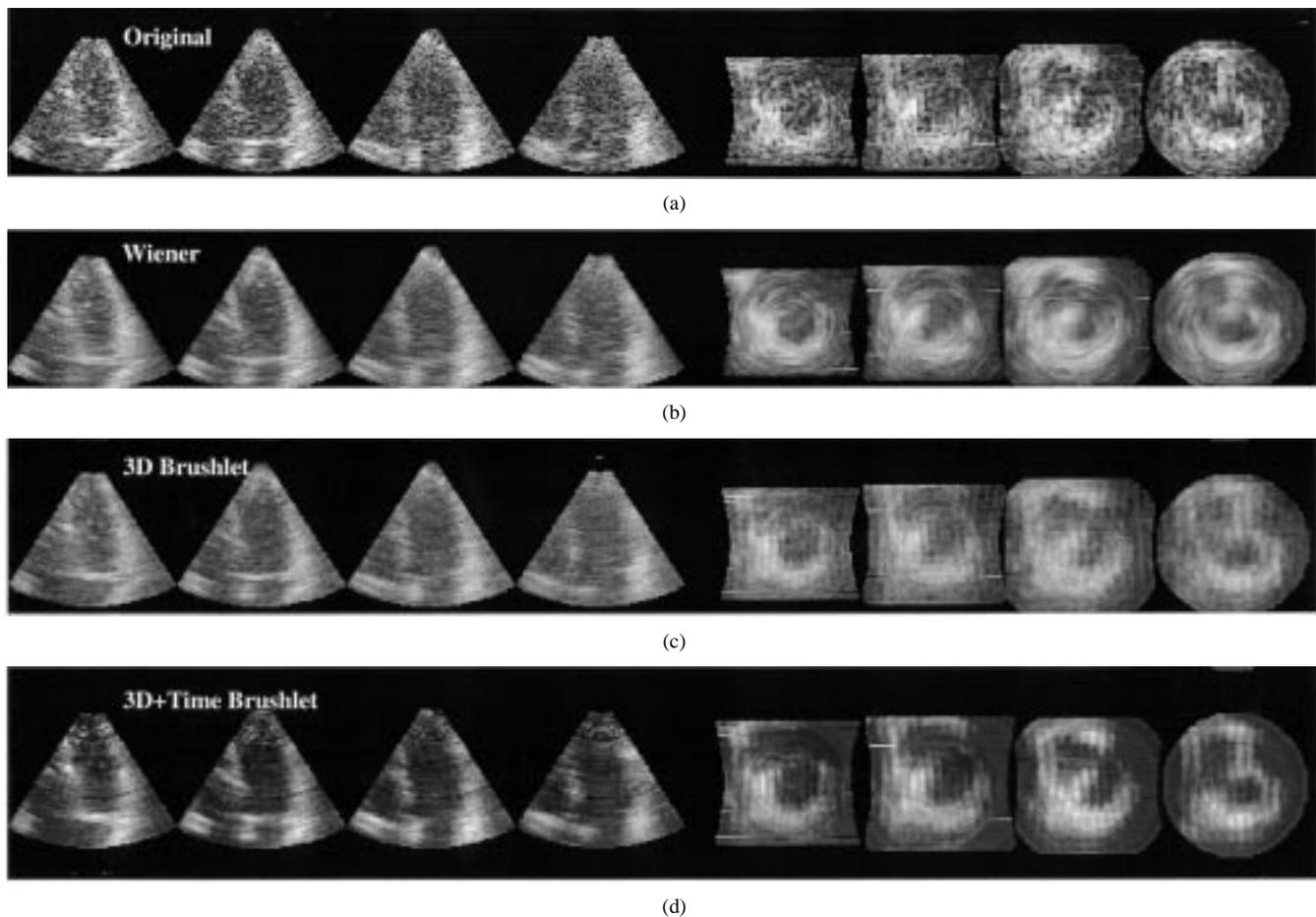


Fig. 12. Qualitative comparison of denoising and enhancement performance on RT3D clinical data. Denoising/enhancement results are presented for four long-axis and four short-axis views from a single clinical data set. (a) Original data, (b) denoising with Wiener filtering adapted to Rayleigh speckle noise model, (c) denoising with 3-D brushlet analysis, and (d) denoising with 4-D (3-D + Time) brushlet analysis.

Regarding volume quantification using RT3D, we showed that the deformable model could accurately segment balloon phantom data with consistency in observed behavior. Quantitative measures obtained from manual tracing were biased as detailed in Section III such that validation of volume measurements could not be obtained on this data.

Testing of volume quantification using a database of six patients showed that simple 2-D deformable-model segmentation applied on denoised data improved by over 50% the accuracy of volume and ejection fraction measures when compared to manual tracing by an expert cardiologist. The gold standard used in this clinical study was functional cardiac MRI. Takuma *et al.* [66] reported an inter-observer variability of 8.3% and intra-observer variability of 3.7% for LV volume estimation with parallel-plane-disks summation on RT3D echocardiograms. These values set the variability limits for acceptability of automatic volume extraction. Any segmentation method should achieve this range of accuracy in order to be used in clinical settings. As reported in Table V, mean percentage errors for volumes and ejection fraction measures are higher than these values for manual tracing while they fall below 8.7% with deformable-model-based segmentation.

These results are quite encouraging and suggest that a deformable model applied to denoised RT3D ultrasound could be used in clinical practice for rapid quantification of cardiac func-

tion. To fully validate our methods we need to carry out further testing on a larger database of patients to evaluate the clinical significance of these results. We are currently working on the implementation of a deformable model in three dimensions following the work of Jones and Metaxas [67]. Three-dimensional deformable surfaces applied to medical image segmentation have drawn a lot of interest in recent years. A good survey of the different approaches and implementations developed for these surfaces is provided in [68]. We believe that extending the segmentation algorithm to true three dimensions will have a great impact on the quality and the robustness of the extracted volumes. When using 2-D deformable models there are cases where the model is stopped by false edges inside the cavity and does not fully expand to reach the actual myocardium wall border. This can be avoided with 3-D deformable models whose deformations are controlled in 3-D space, ensuring smoothness of the contours in every dimension.

After building a 3-D deformable model, we plan to further adapt the method for segmentation of the epicardium borders. A long-term goal of this research project is to extract the myocardium wall borders and analyze wall deformation in time as studied during stress-echo examinations. As true volumetric data is acquired in time, usual post processing tasks such as interpolation for volume reconstruction and registration of volumes from different times is not required, leaving segmentation

as the final step to realize a patient-specific dynamic model of the left ventricle. This property makes RT3D ultrasound a very strong candidate for becoming a powerful echocardiographic screening modality that would enable fast and accurate computation of dynamic computer heart models for standard screening, patient monitoring and stress-echo testing.

## V. CONCLUSION

This paper presented a spatiotemporal analysis method for feature extraction using a third-generation RT3D ultrasound system. The challenge of developing a segmentation tool for quantification of cardiac function from RT3D ultrasound lies in the novelty of the data itself and the wealth of dynamic information that can be extracted despite low spatial resolution and high levels of speckle noise. Overcomplete wavelet-like brushlet functions were used to analyze the multidimensional FT of echocardiographic data and perform denoising and enhancement on low-frequency directional components. Speckle noise was decorrelated well from the original signal and anatomical structures were enhanced at the same time via nonlinear thresholding of analysis coefficients. The deformable-model based segmentation method was performed on the denoised data in two dimensions.

We illustrated the superiority of 4-D (3-D + Time) over 3-D analysis in decorrelating white and speckle noise from 4-D phantom and clinical data. We have shown that we can extract LV endocardial borders using 2-D deformable models and quantify volumes of interest with a higher degree of accuracy than those achieved with manual tracing. Future work will extend the segmentation process itself to higher dimensions to integrate the continuity of the echo data in space and time. Finally, we plan to investigate mathematical and empirical methods for identifying the most "efficient" brushlet analysis tiling for denoising and enhancement.

## ACKNOWLEDGMENT

The authors would like to thank Dr. J. Donis, Department of Medicine, Echocardiography Laboratories, Columbia-Presbyterian Medical Center, for performing manual tracing, Volumetrics Medical Imaging Inc. (Durham, NC) and Echotech 3D Imaging Systems (Hallbergmoos, Germany) for assisting in reading and visualizing RT3D data, and Dr. M. Teragaki, Department of Internal Medicine and Cardiology, Graduate School of Medicine, Osaka City University for providing some RT3D clinical cases with cardiac functional MRI data.

## REFERENCES

- [1] A. S. Gopal, M. J. Schnellbaecher, and Z. Shen *et al.*, "Freehand three-dimensional echocardiography for determination of left ventricular volume and mass in patients with abnormal ventricles: comparison with magnetic resonance imaging," *J. Amer. Soc. Echocardiogr.*, vol. 10, p. 853, 1997.
- [2] —, "Freehand 3-dimensional echocardiography for measurement of left ventricular mass: *in vivo* anatomic validation using explanted human hearts," *J. Amer. Coll. Cardiol.*, vol. 30, p. 802, 1997.
- [3] M. D. Handschmaker, J. P. Lethor, S. C. Siu, D. Mele, J. M. Rivera, M. H. Picard, A. E. Weyman, and R. A. Levine, "A new integrated system for three-dimensional echocardiographic reconstruction: development and validation for ventricular volume with application in human subjects," *J. Amer. Coll. Cardiol.*, vol. 21, pp. 743–753, 1992.

- [4] P. Sapin, K. Schroeder, M. Smith, A. DeMaria, and D. King, "Three-dimensional echocardiographic measurement of left ventricular volume *in vitro*: Comparison with two-dimensional echocardiography and cinerentriology," *J. Amer. Coll. Cardiol.*, vol. 22, 1993.
- [5] T. Menzel, S. Mohr-Kahaly, and J. Meyer *et al.*, "Quantitative assessment of aortic stenosis by 3-dimensional echocardiography," *J. Amer. Soc. Echocardiogr.*, vol. 10, p. 215, 1997.
- [6] G. Stetten, T. Ota, C. Ohazama, C. Fleishman, J. Castelucci, J. Oxaal, T. Ryan, J. Kisslo, and O. V. Ramm, "Real-time 3D ultrasound: A new look at the heart," *J. Cardiovasc. Diagnosis Procedures*, vol. 15, pp. 73–84, 1998.
- [7] S. W. Smith and O. T. Ramm, "High-speed ultrasound volumetric imaging system-Part I: Transducer design and beam steering," *IEEE Trans. Ultrason., Ferroelect. Freq. Contr.*, vol. 38, pp. 100–108, Mar. 1991.
- [8] O. T. V. Ramm, S. W. Smith, and J. H. G. Pavy, "High-speed ultrasound volumetric imaging system-Part II: Parallel processing and image display," *IEEE Trans. Ultrason., Ferroelect. Freq. Contr.*, vol. 38, pp. 109–115, Mar. 1991.
- [9] O. T. V. Ramm and S. W. Smith, "Real time volumetric ultrasound imaging system," *J. Digital Imag.*, vol. 43, pp. 261–266, May 1990.
- [10] T. Ota, C. E. Fleishman, C. J. Ohazama, G. Stetten, C. W. Lewis, D. D. Glower, J. Li, T. Ryan, J. Kisslo, and O. T. V. Ramm, "Measurement of left ventricular volume by real-time, three-dimensional echocardiography in dogs," *Circulation*, vol. 94, p. 379, 1990.
- [11] S. W. Smith and O. T. Ramm, "High-speed ultrasound volumetric imaging system-Part II: Parallel processing and image display," *IEEE Trans. Ultrason., Ferroelect. Freq. Contr.*, vol. 38, pp. 109–115, Mar. 1991.
- [12] I. Herlin, D. Berezziat, G. Giraudon, C. Nguyen, and C. Graffigne, "Segmentation of echocardiographic images with Markov random fields," in *Proc. Eur. Conf. Computer Vision*, 1994, pp. 201–206.
- [13] J. M. B. Dias and J. M. N. Leitao, "Wall position and thickness estimation from sequences of echocardiographic images," *IEEE Trans. Med. Imag.*, vol. 15, pp. 25–38, Feb 1996.
- [14] D. Boukerroui, O. Basset, A. Baskurt, and A. Noble, "Segmentation of echocardiographic data. Multiresolution 2-D and 3-D algorithm based on grey level statistics," in *Proc. MICCAI*, Cambridge, UK, 1999, pp. 516–523.
- [15] E. A. Ashton and K. J. Parker, "Multiple resolution Bayesian segmentation of ultrasound images," *Ultrason. Imag.*, vol. 17, pp. 291–304, 1995.
- [16] S. K. Setarehdan and J. J. Soraghan, "Automatic cardiac LV boundary detection and tracking using hybrid fuzzy temporal and fuzzy multiscale edge detection," *IEEE Trans. Biomed. Eng.*, vol. 46, pp. 1364–1378, Nov. 1999.
- [17] J. Feng, W. Lin, and C. Chen, "Epicardial boundary detection using fuzzy reasoning," *IEEE Trans. Med. Imag.*, vol. 10, pp. 187–199, June 1991.
- [18] C. Kotropoulos, X. Magnisalis, I. Pitas, and M. G. Strintzis, "Nonlinear ultrasonic image processing based on signal-adaptive filters and self-organizing neural networks," *IEEE Trans. Image Processing*, vol. 3, pp. 65–77, Jan 1994.
- [19] G. Coppini, R. Poli, and G. Valli, "Recovery of the 3-D shape of the left ventricle from echocardiographic images," *IEEE Trans. Med. Imag.*, vol. 14, pp. 301–317, June 1995.
- [20] P. R. Detmer, G. Bashein, and R. W. Martin, "Matched filter identification of left-ventricular endocardial borders in transoesophageal echocardiograms," *IEEE Trans. Med. Imag.*, vol. 9, pp. 396–404, Dec 1990.
- [21] A. Sarti, K. Mikula, and F. Sgallari, "Nonlinear multiscale analysis of three-dimensional echocardiographic sequences," *IEEE Trans. Med. Imag.*, vol. 18, pp. 453–466, June 1999.
- [22] V. Chalana, D. T. Linker, D. R. Haynor, and Y. Kim, "A multiple active contour model for cardiac boundary detection on echocardiographic sequences," *IEEE Trans. Med. Imag.*, vol. 15, pp. 290–298, June 1996.
- [23] I. Mikic, S. Kruncinski, and J. D. Thomas, "Segmentation and tracking in echocardiographic sequences: Active contour guided by optical flow estimates," *IEEE Trans. Med. Imag.*, vol. 17, pp. 274–284, Apr. 1998.
- [24] S. Malassiotis and M. G. Strintzis, "Tracking the left ventricle in echocardiographic images by learning heart dynamics," *IEEE Trans. Med. Imag.*, vol. 18, pp. 282–290, Mar. 1999.
- [25] R. Drezek, G. D. Stetten, T. Ota, C. Fleishman, E. Lily, C. Lewis, C. J. Ohazama, T. Ryan, D. Glower, J. Kisslo, and O. T. v. Ramm, "Active contour based on the elliptical Fourier series, applied to matrix-array ultrasound of the heart," in *Proc. 25th AIPR Workshop: Emerging Applications of Computer Vision*, 1997, pp. 26–34.
- [26] A. Salustri and J. R. T. C. Roelandt, "Ultrasonic three-dimensional reconstruction of the heart," *Ultrasound Med. Biol.*, vol. 21, pp. 281–293, 1995.

- [27] R. N. Rohling, A. H. Gee, and L. Berman, "3-D spatial compounding of ultrasound images," *Med. Image Anal.*, vol. 1, pp. 177–193, 1997.
- [28] G. M. Treece, R. W. Prager, A. H. Gee, and L. Berman, "Fast surface and volume estimation from nonparallel cross-sections for freehand three-dimensional ultrasound," *Med. Image Anal.*, vol. 3, pp. 141–174, 1999.
- [29] E. O. Ofili and N. C. Nanda, "Three-dimensional and four-dimensional echocardiography," *Ultrasound Med. Biol.*, vol. 20, 1994.
- [30] M. Mulet-Parada and J. A. Noble, " $2D + T$  acoustic boundary detection in echocardiography," in *Proc. Medical Image Computing and Computer-Assisted Intervention-MICCAI'98*, Cambridge, MA, 1998, pp. 806–813.
- [31] G. D. Stetten and S. M. Pizer, "Medial-node models to identify and measure objects in real-time 3-D echocardiography," *IEEE Trans. Med. Imag.*, vol. 18, pp. 1025–1034, Oct. 1999.
- [32] G. Jacob, J. A. Noble, M. Mulet-Parada, and A. Blake, "Evaluating a robust contour tracker on echocardiographic sequences," *Med. Image Anal.*, vol. 3, pp. 63–75, 1999.
- [33] F. Meyer and R. R. Coifman, "Brushlets: A tool for directional image analysis and image compression," *Appl. Computational Harmonic Anal.*, vol. 4, pp. 147–187, 1997.
- [34] T. McInerney and D. Terzopoulos, "Deformable models in medical image analysis: A survey," *Med. Image Anal.*, vol. 1, pp. 91–108, 1996.
- [35] X. Papademetris, A. J. Sinusas, D. P. Dione, and J. S. Duncan, "3-D cardiac deformation from ultrasound images," in *Proc. MICCAI'99*, Cambridge, UK, 1999, pp. 420–429.
- [36] J. Montagnat, H. Delingette, and G. Malandain, "Cylindrical echocardiographic image segmentation based on 3-D deformable models," in *Proc. Medical Imaging and Computer-Assisted Intervention—MICCAI'99*, Cambridge, U.K., 1999, pp. 168–175.
- [37] R. Coifman and D. Donoho, "Translation invariant denoising," in *Wavelet and Statistics*. Berlin, Germany: Springer-Verlag, 1995, Lecture Notes in Computer Science.
- [38] K. Berkner and R. O. W. Jr, "Smoothness estimates for soft-threshold denoising via translation invariant wavelet transforms," Computational Math. Lab., Rice Univ., Houston, TX, CML TR 98-01, 1998.
- [39] S. S. B. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Scientif. Computing*, vol. 20, pp. 33–61, 1999.
- [40] S. G. Chang, B. Yu, and M. Veterlli, "Spatially adaptive wavelet thresholding with context modeling for image denoising," presented at the Proc. Int. Conf. Image Processing, Chicago, IL, 1998.
- [41] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, "Noise reduction using an undecimated discrete wavelet transform," *IEEE Signal Processing Lett.*, vol. 31, pp. 532–546, Jan. 1996.
- [42] R. R. Coifman and Y. Meyer, "Remarques sur l'analyse de Fourier a fenetre," in *Comptes Rendus de l'Academie des Sciences*, 1991, pp. 259–261.
- [43] H. Malvar, "Lapped transforms for efficient transform/subband coding," *IEEE Trans. Acoust. Speech Signal Processing*, vol. 38, pp. 969–978, June 1990.
- [44] P. Ausher, G. Weiss, and M. V. Wickerhauser, "Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets," in *Wavelets—A tutorial in Theory and Applications*, C. K. Chui, Ed. San Diego, CA: Academic, 1992, vol. 2, Wavelet Analysis and its Applications, pp. 237–256.
- [45] F. G. Meyer and R. R. Coifman, "Adaptive directional image compression with oriented wavelets," in *Proc. Int. Conf. Image Processing*, Lausanne, Switzerland, 1996, pp. 601–603.
- [46] F. G. Meyer, A. Z. Averbuch, J. O. Stromberg, and R. R. Coifman, "Multi-layered image representation: application to image compression," in *Proc. Wavelet Applications in Signal and Image Processing VI*, vol. 3458, San Diego, CA, 1998, pp. 128–136.
- [47] S. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 11, pp. 674–693, July 1989.
- [48] X. Zong, A. F. Laine, and E. A. Geiser, "Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing," *IEEE Trans. Med. Imag.*, vol. 17, pp. 532–540, Aug. 1998.
- [49] I. Koren, A. F. Laine, J. Fan, and F. J. Taylor, "Edge detection in echocardiographic image sequences by 3-D multiscale analysis," in *Proc. IEEE Int. Conf. Image Processing*, 1994, pp. 288–292.
- [50] J. W. Lin, A. F. Laine, and S. R. Bergmann, "Improving PET-based physiological quantification through wavelet denoising," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 202–212, Feb 2001.
- [51] T. McInerney and D. Terzopoulos, "A finite element model for 3-D shape reconstruction and nonrigid motion tracking," in *Proc. 4th Int. Conf. Computer Vision*, Berlin, Germany, 1993, pp. 518–523.
- [52] P. Shi, A. Amini, G. Robinson, A. Sinusas, C. T. Constable, and J. Duncan, "Shape-based 4-D left ventricular myocardial function analysis," in *Proc. IEEE Workshop Biomedical Image Analysis*, Seattle, WA, 1994, pp. 88–97.
- [53] D. N. Metaxas, *Physics-based deformable models. Applications to computer vision, graphics and medical imaging*, 1997.
- [54] L. D. Cohen, "On active contour models and balloons," in *Proc. Computer Vision, Graphics, and Image Processing: Image Understanding*, 1991, pp. 211–218.
- [55] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, vol. 1, pp. 321–331, 1987.
- [56] D. Terzopoulos, A. Witkin, and M. Kass, "Constraints on deformable models: Recovering 3-D shape and nonrigid motion," *Artif. Intell.*, vol. 36, pp. 91–123, 1988.
- [57] L. D. Cohen and I. Cohen, "Finite-element methods for active contour models and balloons for 2-D and 3-D Images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 1131–1147, Nov. 1993.
- [58] F. Leymarie and M. D. Levine, "Tracking deformable objects in the plane using an active contour model," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 15, pp. 617–634, June 1993.
- [59] C. Xu and J. L. Prince, "Snakes, shapes and gradient vector Flow," *IEEE Trans. Image Processing*, pp. 359–369, Mar. 1998.
- [60] V. Dutt, "Statistical Analysis of Ultrasound Echo Envelope," in *Ultrasound Research Laboratory*. Rochester, MN: Mayo Foundation, 1995, p. 181.
- [61] E. Angelini, S. Takuma, A. Laine, and S. Homma, "Spatio-temporal directional analysis of 4-D echocardiography," in *Proc. SPIE-45th Annu. Meeting*, San Diego, CA, 2000, pp. 605–614.
- [62] G. Stetten and R. Morris, "Shape detection with the flow integration transform," *Inform. Sci.*, vol. 85, pp. 203–221, 1995.
- [63] G. Stetten, M. Caines, C. Ohazama, and O. T. V. Ramm, "Toward the volumetricardiogram (VCG): Volume determination of cardiac chambers using 3-D matrix-array ultrasound," in *Proc. SPIE Symp. Medical Imaging*, 1995, pp. 185–196.
- [64] E. Angelini, A. Laine, S. Takuma, and S. Homma, "Directional representations of 4-D echocardiography for temporal quantification of LV volumes," in *Proc. Medical Imaging and Computer-Assisted Intervention—MICCAI'99*, Cambridge, U.K., 1999, pp. 430–440.
- [65] R. R. Coifman and V. M. Wickerhauser, "Entropy based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, pp. 713–718, Mar. 1992.
- [66] S. Takuma, T. Ota, T. Muro, I. Oropesa, R. Sciacca, L. Mendoza, M. R. D. Tullio, D. K. Blood, J. Yoshikawa, and S. Homma, "Assessment of left ventricular function by real-time 3-D echocardiography compared with conventional noninvasive methods," *Amer. J. Cardiol.*, 1999.
- [67] T. N. Jones and D. N. Metaxas, "Segmentation using deformable models with affinity-based localization," in *Proc. CVRMD'97*, 1997, pp. 53–62.
- [68] J. Montagnat, H. Delingette, N. Scapel, and N. Ayache, "Representation, shape, topology and evolution of deformable surfaces. Application to 3-D medical image segmentation," in *INRIA*, 2000, Sophia Antipolis RR-3954.