J. Vis. Commun. Image R. 20 (2009) 478–490

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Surface Function Actives

Qi Duan [a,c,*], Elsa D. Angelini [b], Andrew F. Laine [a]

[a] Department of Biomedical Engineering, Columbia University, ET-351, 1210 Amsterdam Avenue, New York, NY 10027, USA
[b] Department of Image and Signal Processing, Ecole Nationale Supérieure des Télécommunications, LTCI, CNRS-UMR 5141, France
[c] Center for Biomedical Imaging, NYU School of Medicine, 660 1st Ave., FL1, New York, NY 10016, USA

ARTICLE INFO

Article history:
Received 12 October 2007
Accepted 8 June 2009
Available online 13 June 2009

Keywords:
Surface Function Actives
Image segmentation
Deformable model
Real-time segmentation
Variational approach
Interface representation

ABSTRACT

Deformable models have been widely used in image segmentation since the introduction of the snakes. Later the introduction of level set frameworks to solve the energy minimization problem associated with the deformable model overcame some limitations of the parametric active contours with respect to topological changes by embedding surface representations into higher dimensional functions. However, this may also bring in more computational load so that recent advances in spatio-temporal resolutions of 3D/4D imaging raised some challenges for real-time segmentation, especially for interventional imaging. In this context, a novel segmentation framework, Surface Function Actives (SFA), is proposed for real-time segmentation purpose. SFA has great advantages in terms of potential efficiency, based on its dimensionality reduction for the surface representation. Utilizing implicit representations with variational framework also provides flexibility and benefits currently shared by level set frameworks. An application for minimally-invasive intervention is shown to illustrate the potential applications of this framework.

1. Introduction

Image segmentation is a critical step for quantitative image analysis. In medical imaging, image segmentation is the prerequisite for quantitative evaluation of morphologies and pathologies. For example, in cardiac imaging, delineating borders of chambers of the heart and valves are of clinical importance to quantify cardiac function. Segmentation of the left ventricular endocardium is required for quantitative evaluation of the LV function, such as ejection fraction or 3D fractional shortening [1]. With recent advances in 3D and 4D imaging techniques towards real-time imaging, the amount of data is becoming prohibitively overwhelming. Manual tracing of these large data sets is tedious and impractical in clinical setting.

In this context, automated or semi-automated segmentation methods have been proposed and applied to medical image analysis to leverage the human efforts involved in the segmentation task. Based on the mathematical foundation of each method, segmentation approaches can be roughly divided into several classes: classification (e.g. thresholding, k-means), region growing (such as fuzzy connectedness [2]), deformable models (e.g. snake [3], level set [4–7]), active shape [8] and active appearance models [9], and stochastic methods (Markov random field [10], graph cut [11]). Hybrid methods [12] combining different existing methods were also proposed. Among segmentation methods, deformable models are still widely used in medical image analysis, especially for cardiac imaging.

The first deformable model parametric formulation was proposed by Kass et al. in 1987 [3]. The idea was to digitize the object boundary into a set of points with predefined connectivity among the node; the contour then deforms under the combination of internal forces (such as elasticity and inertia) and external forces (such as image gradient force) to align with the desired object's boundaries. Parametric deformable models, also called snakes in 2D, have been widely applied in various segmentation tasks [13]. In 1998, Xu et al. [14] proposed the Gradient Vector Flow or GVF as a novel driving force for the snakes. This newly designed force overcame several drawbacks from the original snake framework and increased the performance of the active contour. However, there were still some limitations related to the parametric formulation of active contours, such as difficulties to adapt the contour to topological changes, especially in 3D.

In the late 1990s, Sethian et al. [15] proposed a new framework called level set to overcome these limitations. The basic idea was to embed the contour evolution into iso-value curves of a function with higher dimensionality. Such functions were called level set functions. Topological changes could be naturally handled without additional efforts. Moreover, highly convoluted surfaces, which were very hard to handle for parametric deformable models, could also be easily represented. For this reason, level set formulations have become a research focus in image segmentation in recent years. In 2001, Chan and Vese [16] introduced their "active contour without edges" approach. In their framework, no image gradient information was needed as with traditional deformable models.

* Corresponding author. Address: Center for Biomedical Imaging, NYU School of Medicine, 660 1st Ave., FL1, New York, NY 10016, USA. Fax: +1 212 854 5995.
E-mail address: Qi.Duan@nyumc.org (Q. Duan).

1047-3203/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2009.06.002

Instead, driving forces were derived via energy minimization of the Mumford–Shah segmentation functional [17] for piecewise-constant regions. Their method could easily deal with noisy images. And as a result, this framework has been widely used in ultrasound segmentation [18], brain segmentation [19], and many other applications. However, the introduction of level set functions implicitly increased the number of parameters of the surface model, which increases the demand for computational power. Although many optimization modifications such as narrowbanding [20] or fast marching schemes were proposed, generally speaking, level set framework is still a relatively "slow" approach especially for 3D or 4D data.

As imaging technology evolves, demands for real-time feedback also increases, mostly for interventional imaging and minimum-invasive surgery. Latest 3D and 4D imaging techniques and real-time imaging techniques not only provide better appreciations of the anatomy and function of the body, but also raise a challenge for image segmentation in terms of computational efficiency. In this context, a new framework called Surface Function Actives (or SFA) is proposed in this paper to push the limits of real-time segmentation.

2. Mathematical models

2.1. Interface representation

"Deformable model" is composed of two critical parts: the "model" or the surface representation, which represents the interface, and the "deformation scheme" driven by applied forces, which fits the model to the image for desired segmentation results. In all deformable model methods, interface representation is fundamental since the interface or boundary is the target object that needs to be fitted to the image information to find the desired boundaries. Mathematically, there are two ways to represent the interface:

1. *Explicit representation:* that is representing the surface by explicitly listing the coordinates of the boundary points (i.e. a parametric representation). This is the representation that original snakes [21] used. The coordinate system can use either natural basis or other basis as well, depending on applications;
2. *Implicit representation:* that is representing the surface by embedding the boundary as the iso-value curves of some function f called the *representation function*. Level set functions [4–7] are a good example to embed the interface as the zero level set of a distance function.

2.2. Surface functions

Most of the recent efforts in segmentation based on implicit interface representation were focused on the level set framework, given its advantages for topological changes and feasibility to represent convoluted surfaces. As mentioned above, level set functions add one extra dimension beyond the dimensionality of the image data. For example, to represent a surface in 3D space, the level set function corresponding to the surface will be 4D. For comparison, original parametric deformable models only required a list of point coordinates in 3D. For level set, this extra dimension brings various benefits as well as additional computation load, which may degrade computational efficiency.

By looking the *opposite* way of level set frameworks, it is very natural to think of dimensionality reduction in surface representation to reduce the computational complexity. Using terminology of interface representation, we are looking for a representation function which has fewer dimensions than the image data, i.e. using a 2D function to represent a 3D surface in space. We call such function a *surface function*.

Mathematically, in N dimensional space, we can define a surface function $g : \mathbb{R}^{N-1} \rightarrow \mathbb{R}$ as a special set of functions representing one of the coordinates constrained by the others. Without losing generality, we can assume that this special coordinate is x_0 and the other coordinates are x_1 to x_{N-1}. That is:

$$x_0 = g(x_1, \ldots x_{N-1}) \tag{1}$$

The corresponding representation function f is defined as:

$$f = x_0 - g(x_1, \ldots x_{N-1}). \tag{2}$$

So that the corresponding boundary is the zero-value curve of the function f, i.e. $f(\vec{X}) = 0, \quad \vec{X} = (x_0, x_1, \ldots x_{N-1})$.

For example, in 2D to represent a straight line with slope = 1 and passing through the origin $(0, 0)$, there can be three different representations:

1. *Explicit representation:* such as the list of points $\{\ldots(-1,-1), (0,0), (1,1), \ldots\}$;
2. *Level set:* such as the zero level set of $\Phi(x, y, t) = \frac{1}{\sqrt{2}}(x - y)$, note that Φ is a signed distance function, and it is defined on the whole x–y plane instead of just on the boundary;
3. *Surface function:* such as $y = x$.

Fig. 1 illustrates each representation. Note that since these three surface representations encode the same surface, there are some similarities between them. The corresponding representation function for the level set framework is $f = \frac{1}{\sqrt{2}}(x - y) = 0$; the corresponding representation function for surface function representation is

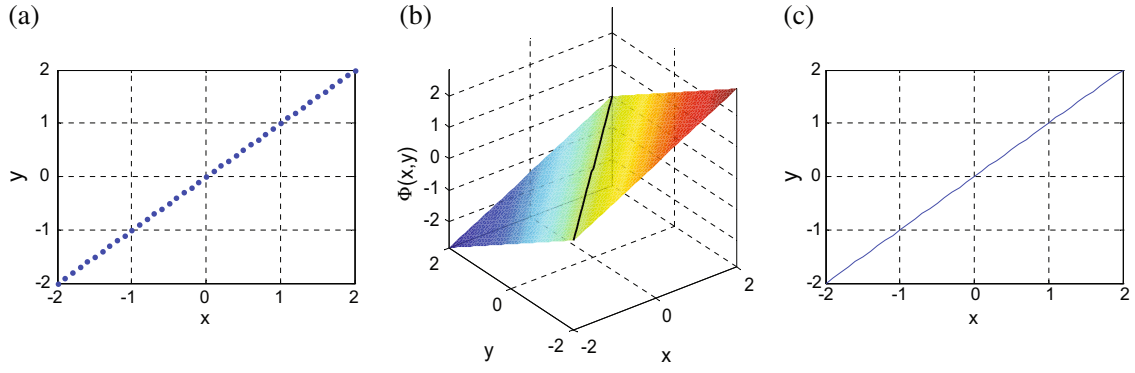

Fig. 1. Illustration of interface representation by three difference frameworks: (a) explicit representation; (b) level set framework with the black line showing the zero level set; and (c) surface function.

$f = y - x = 0$. It is obvious that both functions have the same roots, with the fact that the coordinates used in the explicit representation are digitized version of these roots. And geometrically, these roots form the surface we are trying to represent: a straight line with slope = 1 and passing through (0, 0). In other words, explicit representations, level set distance functions, and surface functions are just three different equivalent forms of the actual interface function f for the targeting surface.

In this example, however, besides similarity and equivalence, it is more interesting to notice their differences: the explicit representation is a set of coordinates defined on the 2D x–y plane; the level set function is an $\mathbb{R}^2 \to \mathbb{R}$ distance function defined on the whole 2D x–y plane; the surface function is an $\mathbb{R} \to \mathbb{R}$ function defined only on a 1D x-axis. As a 1D function, surface function representation has the advantage in efficiency compared with the other two common representations. From the example, we can see that interface representation based on surface function has less dimensionality than both other methods; compared with explicit representation, surface function can utilize function expression or function basis to efficiently represent the interface. Given the fact that most anatomical surfaces are smooth [22], anatomical surfaces can be efficiently and accurately represented by the surface function framework, which simplifies the downstream mathematical computation such as energy minimizations during image segmentation.

2.3. Driving forces

Similar to other deformable models, we adopted a variational framework in deriving the driving forces. For example, we can use the Mumford–Shah segmentation energy functional:

$$E(f,\vec{C}) = \beta \int_{\Omega} (f-g)^2 dV + \alpha \int_{\Omega\setminus\vec{C}} |\nabla f|^2 dV + \gamma \oint_{\vec{C}} ds, \tag{3}$$

in which $\vec{C}$ denotes the smoothed and closed segmented interface, g represents the observed image data, f is a piecewise smoothed approximation of g with discontinuities only along $\vec{C}$, and Ω denotes the image domain. The first integral enforces similarity between f and g, which is equivalent to homogeneity constraint if f is piecewise smoothed; the second integral controls the smoothness of f; and the last integral is limits the length of the segmented boundary, which, acts as internal elasticity constraint to prevent leaking at attachment to weak boundaries.

Given the flexibility of variational frameworks, other segmentation energy functionals can be also easily adopted.

In general, deformable models usually utilize iterative methods to find the optimal solution for the associated energy minimization framework via curve evolution, which requires an additional variable as an artificial time step added into the functions. In this case, curve evolution with explicit representation with K node points becomes an $N \times K$ variable minimization problem since the evolving curve is represented by

$$\begin{bmatrix} \vec{X}^0(t) \\ \vec{X}^1(t) \\ \vdots \\ \vec{X}^{K-1}(t) \end{bmatrix} = \begin{bmatrix} x_0^0(t) & x_1^0(t) & \cdots & x_{N-1}^0(t) \\ x_0^1(t) & x_1^1(t) & \cdots & x_{N-1}^1(t) \\ \vdots & \vdots & \vdots & \vdots \\ x_0^{K-1}(t) & x_1^{K-1}(t) & \cdots & x_{N-1}^{K-1}(t) \end{bmatrix}, \tag{4}$$

with $N \times K$ evolving variables.

Curve evolution with level set becomes an $(N+1)$-variate functional minimization problem since the evolving curve is represented by

$$\phi(\vec{X},t) = \phi(x_0, x_1, \ldots x_{N-1}, t), \tag{5}$$

which has to be solved for every point on the entire image domain or within the narrowband.

Curve evolution with Surface Function Actives becomes an N-variate functional minimization problem since the evolving curve can be represented by

$$x_0(t) = f(x_1, x_2, \ldots x_{N-1}, t). \tag{6}$$

The advantage in dimensionality reduction for Surface Function Actives over level set framework is evident.

The advantage of SFA over explicit expression is in two aspects. First, in explicit representation, for each node point, there are N evolving variables, whereas in surface function representation, there is only one variable for each corresponding points. This will become more evident if we digitize Eq. (6) and reformulate in a similar form as in Eq. (4):

$$\begin{bmatrix} \vec{X}^0(t) \\ \vec{X}^1(t) \\ \vdots \\ \vec{X}^{K-1}(t) \end{bmatrix} = \begin{bmatrix} x_0^0(t) & x_1^0 & \cdots & x_{N-1}^0 \\ x_0^1(t) & x_1^1 & \cdots & x_{N-1}^1 \\ \vdots & \vdots & \vdots & \vdots \\ x_0^{K-1}(t) & x_1^{K-1} & \cdots & x_{N-1}^{K-1} \end{bmatrix}. \tag{7}$$

Although the memory usage of Eq. (7) is the same as Eq. (4), the curve evolution of Eq. (7) has $N-1$ less dimensionality than Eq. (4), which usually leads to faster and more stable convergence. Generally speaking, the more parameters to be optimized, the larger possibility that local minimums and saddle points exist, especially with presence of noise. Of course it is not necessarily true for every case that 1D optimization is more stable than N-D; they could be equivalent. But even for that, the searching space for 1D case is much smaller than the N-D one, which leads to faster convergence.

Another aspect is that Eq. (6) can be represented via function basis, such as cubic Hermite functions, in which case only a few weighting parameters rather than a lot of digitized node points have to be stored and iterated on. This can further improve the accuracy, efficiency, and numerical stability.

2.4. Comparison with other deformable models

Although as mentioned above, the interface functions for the three deformable models are equivalent in terms of surface representation, different ways to approach interface formulation provide different benefits and limitations.

Parametric active contours with explicit representations provide relative simple representations through interface point coordinates and do not add additional dimensionality to the optimization problem. However, it cannot easily handle topological changes, and usually requires some prior knowledge about the target topology for proper initialization. It is also not trivial to determine whether an arbitrary pixel is inside or outside the segmented objects. Moreover, in order to compare to other segmentation results such as manual tracing, it is usually not very easy to directly compute quantitative metrics such as surface distances since it requires pairing of closest points.

The level set framework based on implicit representations via distance functions can automatically deal with topological changes and allows easy determination of whether a point is inside the object or not by simply looking at the sign of the level set function at the point location. However, the level set formulation implicitly introduces a new dimension, i.e. the value of the level set function, for each voxel in the whole image data space, whereas the other two models only focus on the interface itself. This type of formulation implicitly increases the dimensionality of the variational problem and thus increases the computational cost of the optimization process. Even though a narrowband approach can improve the efficiency by focusing only around the interface, it still requires more

voxel information than the other two formulations. In terms of segmentation comparison, if the level set function is the signed distance function, it is very easy to compute the distance between surfaces, although in most of implementations, level set functions after few iterations do not necessarily remain as signed distance functions, especially for those using narrowband approaches.

Surface Function Actives is a kind of marriage of the previously discussed models: it focuses only on the interface as the explicit representation, while being formulated as an implicit representation like the level set framework. It has advantage on dimensionality reduction in surface representation compared with level set. It can utilize function basis to avoid memory-inefficient boundary point digitizing. Even if a digitized form has to be used and the surface representation is similar to explicit expression, Surface Function Actives still has faster convergence when compared with parametric deformable model. This dimensionality reduction gives SFA advantages in efficiency in both aspects of the deformable model (i.e. surface modeling and deformation scheme). Further more, with an implicit representation, it is straightforward to determine whether a point is inside the contour by simply comparing the value of the surface function for that point with the value of the surface function on the boundary. In addition, surface functions enable immediate quantitative evaluation of the segmentation results via surface comparisons and differences in surface function values. However, similar to parametric active contours, it is not trivial to deal with topological changes.

2.5. Further extension in flexibility

Beside the advantage brought by dimensionality reduction, SFA framework is also benefited from basis representation. By utilizing coordinate basis other than Cartesian coordinates, SFA can not only easily dealt with enclosed shape as heart, liver and various tumors, but also easily incorporate shape prior information. By utilizing function basis other than natural basis, SFA can not only efficiently represent convoluted surfaces, but also naturally enforce prior knowledge on surface smoothness.

By using the concept of piecewise function, SFA can be extended with combination of finite element patches to capture much more complex shape, like left ventricle. By incorporating repositioning and reorientation, the capture range of SFA can be largely increased, giving less dependence on initialization.

A sample application illustrating all these flexibility extensions in the context of real-time cardiac segmentation was provided later in the paper. But first, some basic idea of SFA was illustrated on synthetic images using Cartesian coordinate system.

3. Experimental results on synthetic image

To illustrate the performance and some advantages of the proposed Surface Function Actives (SFA), several segmentation examples are presented in this section on a synthetic image. This section specifically focuses on two implicit representation methods: the proposed SFA method and the level set representation. Both segmentation frameworks use variational formulae and interface functions. A fair head-to-head comparison is possible by setting identical segmentation energy functional and numerical schemes for both methods.

3.1. Synthetic image

To illustrate the flexibility of the proposed SFA framework, instead of using an example on common piecewise smooth images, in this section, both SFA and level set approach were challenged with textured regions segmentation.

The synthetic image, as shown in Fig. 2a, was composed of two parts. Pixel intensities for each part were randomly sampled from normal distributions with identical mean values and different standard deviations. The corresponding ground-true binary image is shown in Fig. 2b. The blue region had a standard deviation of 5 and the red region had a value of 10. The interface between the regions was a sine function. The dimension of the image was 65 by 65 pixels.

3.2. SFA using numerical solution

Usually in image segmentation, especially for the level set framework, it is not easy to find a closed form interface function. Instead, a numerical solution or approximation of the interface function is computed via iterative numerical energy minimization. With the level set functions, this requires computation values of the level set function at each pixel (or on a narrowband near the interface). With SFA, we only need to compute the surface function values at each pixel on the interface.

Given the texture-based segmentation problem presented in Fig. 2, the following energy functional was selected:

$$E = \int_{\Omega} (\sigma(x,y) - \delta_1)^2 H(x,y) dxdy + \int_{\Omega} (\sigma(x,y) - \delta_2)^2 (1 - H(x,y)) dxdy \tag{8}$$

where Ω is the image domain, $\sigma(x, y)$ is a standard deviation estimator for pixel (x, y) within a small neighborhood, and H is the Heaviside function, which equals 1 inside the current interface and 0 outside. The parameters $\delta 1$ and $\delta 2$ are computed as the average standard deviations inside and outside the current interface, respectively. The optimal segmentation will partition the image into two regions, with relative homogeneous distributions of the standard deviations within each region. This approach is equivalent to segmenting a representation of local standard deviations $\sigma(x, y)$ values of the image, knowing that for normal distributions $N(\mu, \sigma)$, average standard deviations converge to the scale parameter σ. The Chan-Vese level set numerical schemes described in [6] were used for the level set implementation. For simplification, no curvature constraints were used.

Both methods were initialized as a straight line at the center of the image, as shown in Fig. 3a. The ground truth boundary is shown in green on the same figure. Corresponding surface functions for SFA was just a 1D constant function as $y(x) = 0$, $-32 \leqslant x \leqslant 32$, whereas the corresponding level set function was a plane with slope 1 as shown in Fig. 3b.

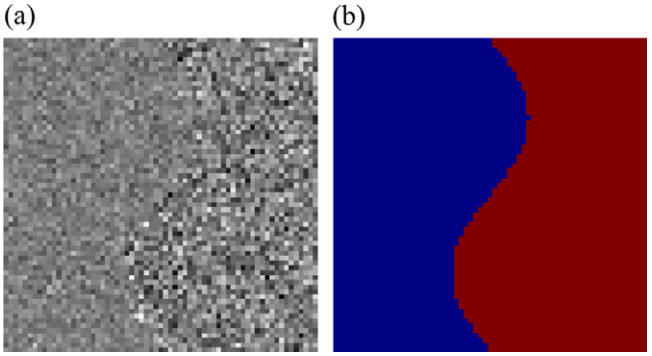

Fig. 2. (a) Synthetic image composed of two regions with normal distributions with the same mean values but different standard deviations; (b) corresponding binary images indicating the ground truth segmentation. The blue region has a standard deviation of 5 and the red one has a value of 10 in the original image. The interface is a sine function.

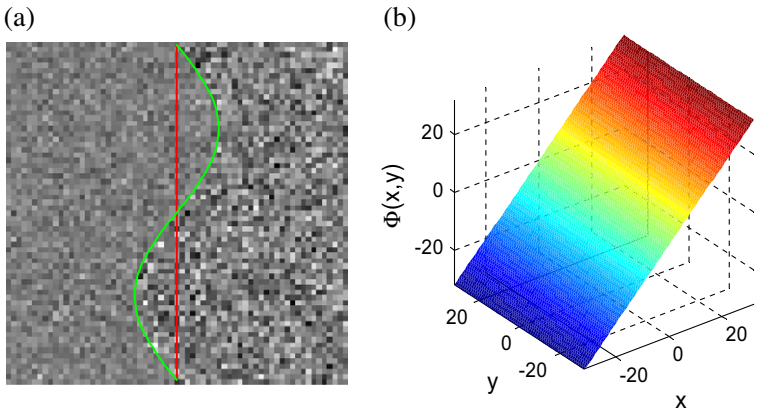

Fig. 3. (a) Initialization of the deformable model as in red and the ground truth interface in green; (b) corresponding signed distance function for the level set initialization.

Both methods were implemented in Matlab©. All computations were executed on a 2.4 GHz 64-bit AMD server, running Red Hat Linux Enterprise AS. For each pixel, a 7×7 neighborhood was used to compute the local standard deviations. An artificial time step was set to 5 for both methods. Stability of the surface was considered as the convergence criterion.

In order to quantitatively evaluate the segmentation result for each method, true positive (TP) fraction ratio and false positive (FP) fraction ratio were computed. In addition, root-mean squared error in distances from the interfaces to the ground truth was estimated.

It took eight iterations (0.38 s) for the SFA to converge, each iteration taking about 47.5 ms. Final segmentation result is shown in Fig. 4a, with red line indicating the automatic segmentation and green line indicating the ground truth. The TP fraction ration was 99.7%, whereas the FP ration was 3.8%. RMS error of the distance to the ground truth was 1.56 pixels.

For comparison, it took 36 iterations (3.43 s) for the level set approach to converge, each iteration taking about 95.3 ms. Final segmentation result is shown in Fig. 4b, with the red line indicating the automatic segmentation and green line indicating the ground truth. The TP fraction ration was 99.6%, whereas the false positive fraction ration was 4.0%. The RMS error of the distance to the ground truth was 2.15 pixels.

From this experiment, the SFA framework clearly demonstrates advantages in computational efficiency when compared to the level set framework, with not only a shorter time per iteration, but also fewer iterations. This is mainly due to the fact that the level set function has to be updated over the entire image domain whereas the SFA is only updated at the interface. Quantitative segmentation comparison yielded comparable segmentation performance for both methods, SFA having slightly better performance.

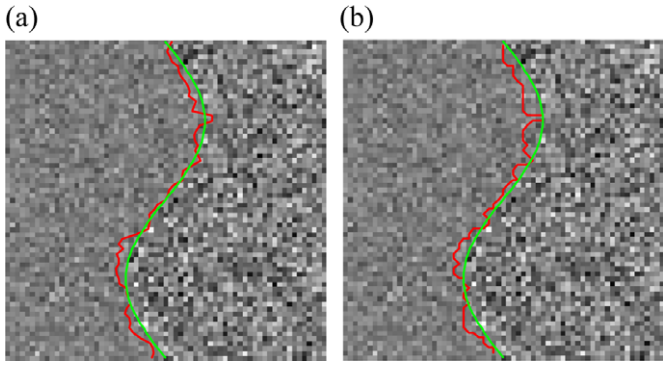

Fig. 4. Final segmentation (red line) compared to the ground truth (green line) for (a) Surface Function Actives and (b) Chan-Vese level set with identical segmentation energy functionals and numerical schemes.

To test the ability of the proposed SFA framework under different and more challenging initialization setups, another initialization on the left-most boundary was tested as well for both methods. The initialization and final results from both methods are shown in Fig. 5.

Both methods spent much more time to reach the final results. It took 25 iterations for the SFA and 75 iterations for the level set method to converge. Segmentation by SFA yielded a TP of 99.81% and a FP of 4.03%. Segmentation by the level set framework generated a TP of 99.91% and a FP of 6.58%. RMS errors were 1.61 pixels for SFA and 2.80 pixels for the level set. Both methods had slightly poorer performance compared with the results using closer initialization. However, the SFA framework still exhibited advantages in efficiency and slightly better performance.

It has to be noted that in numerical solution, SFA was operating in a "downgraded" form which is similar to the parametric deformable model. But it still has the advantage in efficiency over parametric deformable model since only one variable per node is evolving, rather than N variables per node for the explicit representation.

3.3. SFA using analytical solution or function basis

Another advantage of the SFA framework is to provide closed form solution or approximation for the interface, which, as indicated in the method section, can bring additional gain in efficiency. Since the surface function is a $(N-1)$-D function for N-D image data, we can choose arbitrary bases (natural bases or other bases) in the function space to express the surface function. Especially when some prior knowledge about the interface properties is known, proper choice of basis functions can not only seamlessly incorporate such prior knowledge, but also further increase the efficiency of the SFA framework. Moreover, having an analytical form of the interface can provide additional benefits like analytical differentiation in a downstream analysis.

To illustrate this point, assuming that by looking at the original image data, we decided to use the following sine function to describe the interface.

$$y(x) = a\sin(x) + b \tag{9}$$

the problem was then converted to the identification of the optimal parameters (a, b) that minimize the segmentation functional. Since there were only two parameters to be optimized, only a few pixels

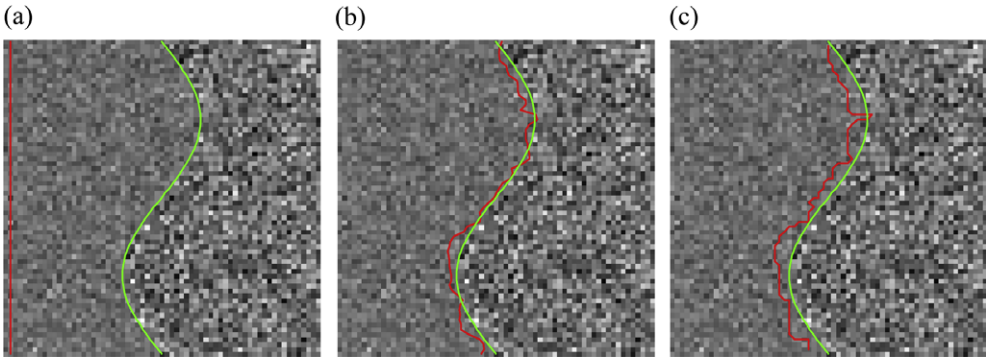

Fig. 5. (a) Initialization of the deformable model, in red, at left-most of the image with the ground-true interface in green; Corresponding final segmentation (red line) compared to the ground truth (green line) for (b) Surface Function Actives and (c) Chan-Vese level set under identical segmentation energy functional and numerical schemes.

were needed during each iteration. In this example, we worked with 9 pixels on the boundary to ensure the overcompleteness of the estimation, compared to all pixels on the boundary as in the numerical solution scenario. This formulation reduces computational cost by utilizing prior knowledge on the shape of the interface. Close and far-away initializations were tested. Final segmentation results are shown in Fig. 6.

It took the SFA 10 iterations to converge under close initialization. Each iteration took 21 ms. The TP ratio was 100%; the FP ratio was 3.79%. The RMS error was 1.26 pixels. Under far-away initialization, it took the SFA 25 iterations to converge. Each iteration took 17 ms. The TP ratio was 100%; the FP ratio was 4.50%. The RMS error was 1.40 pixels.

The presented results showed that by utilizing prior knowledge, both efficiency and accuracy could be improved under the same initialization setups and numerical schemes. Besides, the use of a closed form function provides a continuous expression of the interface which may be of great benefit for downstream analysis.

In real applications, it can be difficult to identify an accurate surface function. In this context, an approximation function or approximation basis functions can be used. In the same example, when we used cubic polynomials to approximate the interface, it took longer time, 3.59 s, to converge, with corresponding RMS errors of 1.66 pixels. The final segmentation result is shown in Fig. 7.

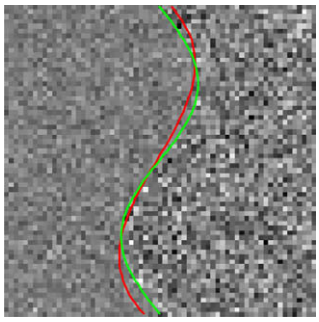

Fig. 7. Final segmentation (red line) compared to the ground truth (green line) for the SFA using cubic polynomials as the surface approximation.

4. Sample applications

Static X-ray computed tomography (CT) volumes are often used as anatomic roadmaps during catheter-based cardiac interventions performed under X-ray fluoroscopy guidance. These CT volumes provide a high-resolution depiction of soft-tissue structures, but at only a single time within the cardiac and respiratory cycles. Combining these static CT roadmaps with segmented myocardial borders extracted from 4D ultrasound (US) data provides intra-operative access to real-time dynamic information about the cardiac anatomy. In this work, using a customized segmentation method based on a 3D Surface Function Actives, endocardial borders of the left ventricle were extracted from US image volumes (4D data sets) at a frame rate of approximately five frames per second. Coordinate systems for the CT and US modalities were registered using rigid body registration at end diastole based on manually selected landmark points, and the segmented endocardial surfaces were overlaid onto the CT volume. The RMS fiducial registration error was 3.80 mm. The accuracy of the segmentation was quantitatively evaluated on phantom and human volunteer studies via comparison with manual tracings on nine randomly selected frames using a finite element model (US image resolutions of the phantom and volunteer data were $1.3 \times 1.1 \times 1.3$ and $0.70 \times 0.82 \times 0.77$ mm, respectively). This comparison yielded a RMS error of 3.70 ± 2.5 mm (approximately 3 pixels) in the phantom study and 2.58 ± 1.58 mm (approximately 3 pixels) in the clinical study. The combination of static anatomical roadmap volumes and dynamic intra-operative anatomic information can enable better guidance and feedback for image-guided minimally invasive cardiac interventions, given the fact that only 75 ms was needed for each clinical 3D image.

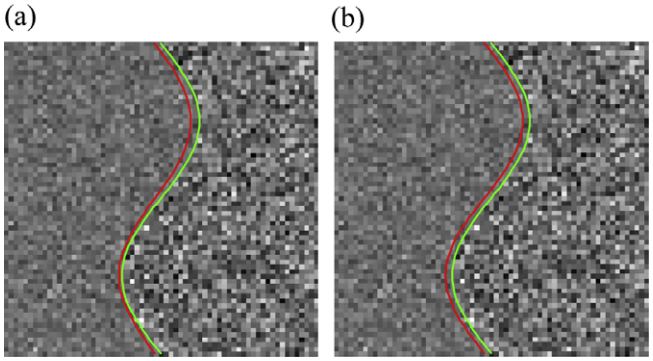

Fig. 6. Final segmentation (red line) compared to the ground truth (green line) for SFA under (a) close initialization and (b) far-away initialization.

4.1. Introduction

In cardiovascular minimally invasive interventional procedures such as catheter-based radiofrequency ablation [23], pre-procedural roadmaps can be valuable for surgical planning and intra-procedural guidance by complementing the intra-procedural imaging modality. Though multiple imaging modalities have been

proposed for acquiring such roadmaps, cardiovascular X-ray computed tomography (CT) is widely used since it can provide 3D high-resolution depiction of soft-tissues [24] not usually attainable with intra-procedural imaging modalities. The utility of pre-procedural roadmaps is usually limited in that they are static volumes acquired at a single time point within cardiac and respiratory cycles, whereas during the interventional procedure, the geometry and position of the heart changes over time due to cardiac motion and breathing. Several modalities, such as X-ray fluoroscopy [25], interventional magnetic resonance imaging (iMRI) [26], and 2D/3D ultrasound (US) have been used for intra-procedural imaging to acquire dynamic information about the targeted tissues, though each modality has its own limitation. For example, X-ray based methods do not depict soft-tissue structures well, are limited to 2D projection perspectives, and introduce additional ionizing radiation to the patient and the operators; MR-based methods require an MR-compatible operating room and instruments. On the other hand, real-time 3D (RT3D) US is used for intra-procedural live monitoring since it is portable, does not generate ionizing radiation, and does not require any special operating environments.

Development of RT3D echocardiography started in the late 1990s by Volumetrics© [27] based on matrix phased arrays transducers. Current systems can acquire roughly one quadrant of the heart in a single 3D field-of-view in real-time, or can stitch together acquisitions over four cardiac cycles for a fully sampled cardiac volume. This quadrant-wise acquisition design enables a dramatic increase in spatial resolution and image quality compared to those reconstructed 3D ultrasound from 2D B-mode slices, which makes such 3D ultrasound techniques increasingly attractive for daily cardiac clinical diagnosis. Since RT3D ultrasound acquires volumetric ultrasound sequences with fairly high temporal resolution (about 20 frames per seconds per quadrant) and a stationary transducer, complex 3D cardiac motion can be captured with high fidelity. RT3D US data from one quadrant can roughly cover the left ventricle and can be recorded in true real-time fashion, with a temporal resolution of about 20 frames per second. This dynamic 3D imaging modality provides new opportunities for non-invasive monitoring of intra-operative tissue dynamics and functional information.

The purpose of this work was to explore the feasibility of augmenting static cardiac CT roadmaps with segmented myocardial borders extracted from RT3D US streams. Emphasis was placed on the integration of information based on the inherent strengths of each imaging modality. CT provides a static high-resolution depiction of cardiac soft-tissue structures; RT3D US provides intra-operative access to real-time dynamic information about the myocardium without exposing the patient to additional ionizing radiations. Integration and fusion of image data and information from multiple modalities is a critical enabling technology that will improve the quality of intra-operative guidance for many minimally invasive cardiac interventions.

4.2. Methods

There are different ways of incorporating information from streamed US data into a CT roadmap. One direct way is to overlay ultrasound images onto the CT volume. However, even though the multi-modality registration step and the volume rendering step can be implemented with real-time performance, overlaying all of the image data from the high speed volumetric streaming ultrasound on top of the CT volume can be overwhelming to the operators. In this work, we extracted the relevant dynamic cardiac information by online segmentation of the endocardial borders from RT3D US images; the extracted endocardium was overlaid on the static CT roadmap in real-time to augment the intra-procedural display. This work involves multi-modality registration and online segmentation of RT3D US image data. In the following sections, the hardware and imaging systems will be introduced; the registration framework will be detailed; and finally, the design and performance of a real-time segmentation method will be analyzed.

5. Hardware and Imaging

A multi-modality imaging phantom was constructed by Philips Research North America to facilitate technology development and validation. The custom-built phantom was tuned for realistic soft-tissue contrast under both CT and US imaging and mimics the geometry of a human heart (Fig. 8a).

X-ray CT imaging was performed on a Philips Brilliance 16 slice CT scanner. The CT image data was acquired with a spatial resolution of $0.59 \times 0.59 \times 1$ mm and field-of-view of $300 \times 300 \times 200$ mm. This CT volume was used as the static pre-procedural roadmap (Fig. 8b). Ultrasound imaging was performed on a Philips iE33 system, using the X3-1 3D imaging probe (Fig. 8c). For development purposes, RT3D US data sets were saved to disk for repeated use. In our segmentation experiments, streaming of the data was simulated by loading a single 3D frame at a specific time from the disk and applying our processing steps before loading the next 3D frame. During US scanning, the cardiac phantom was manually deformed to simulate cardiac motion. The spatial resolution of the RT3D US image data of the phantom was $1.3 \times 1.1 \times 1.3$ mm. The temporal resolution was 18 fps.

Cardiac US image data was collected on two human volunteers. The clinical protocols were approved by the institutional review board and informed consent was secured. The RT3D US images were acquired with the same hardware that was used in the phantom experiments and with a spatial resolution of $0.70 \times$

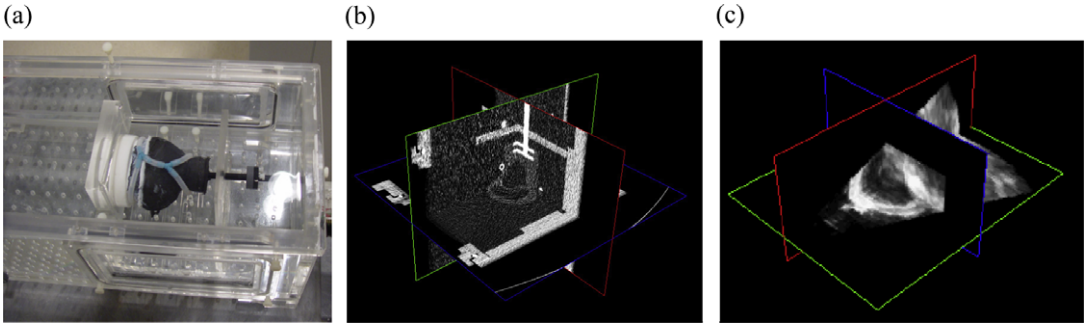

Fig. 8. (a) A photograph of the multi-modality phantom under experimental settings; three orthogonal planes through (b) the corresponding CT roadmap and (c) one sample volume taken from the 4D ultrasound stream.

0.82 × 0.77 mm, and a temporal resolution of 23 fps. CT imaging was not performed on the human volunteers.

6. Registration

The CT coordinate system was chosen as the reference coordinate system, and the ultrasound data was registered to it. In this initial work, the ultrasound probe was left in a fixed position relative to the cardiac anatomy, so that a one-time coordinate registration was sufficient.

$$X_{CT} = R_{US}^{CT} X_{US} \tag{10}$$

Eq. (10) describes the relationship between the CT and US coordinate systems. X_{US} refers to image coordinates in the US data space. When multiplied by the registration matrix, R_{US}^{CT}, US coordinates are transformed to corresponding image coordinates in CT data space, X_{CT}. The registration matrix, R_{US}^{CT}, describes a rigid body transform between the two coordinate systems. Anatomical features such as the mitral valve annulus, tip of the apex, and cross-sectional points of the artificial coronary artery, which were visible in both imaging modalities, were used to define the corresponding points in the CT image and in the first frame of the ultrasound image data. The registration matrix, R_{US}^{CT}, was calculated analytically using a least squares technique for eight manually selected landmark points in the two image datasets.

7. Segmentation

In ultrasound image analysis, including segmentation, a preprocessing step is usually applied to suppress the speckle noise in the ultrasound images [28]. However, such processing usually needs full volume linear or non-linear processing, which may require substantial computational power. This is acceptable for offline processes but this requirement is incompatible with this online or "real-time" segmentation application.

In cardiac ultrasound segmentation, besides the classical methods based on thresholding [29,30] and morphological operations [31,32], parametric active surface models [33–36], level set frameworks [18,37], and active shape/appearance models (ASM/AAM) [38–40] have been proposed as alternative methods. Since the general level set implementation is relatively slow, and the requirement of a large training database for ASM/AAM is non-realistic in the cardiac intervention context, we chose to develop a segmentation algorithm using a customized 3D Surface Function Actives [41].

7.1. Surface representation

In order to efficiently capture the shape of endocardium, instead of using a traditional triangulated mesh with a linear surface within each patch, 3rd order Hermite polynomials *in spherical coordinate system* were used as surface descriptors [33,42].

In 1D, there are four cubic Hermite basis functions $H_i^j(\xi)$:

$$\begin{aligned} H_0^0(\xi) &= 1 - 3\xi^2 + 2\xi^3 \\ H_0^1(\xi) &= \xi(\xi - 1)^2 \\ H_1^0(\xi) &= \xi^2(3 - 2\xi) \\ H_1^1(\xi) &= \xi^2(\xi - 1) \end{aligned} \tag{11}$$

On a 2D finite element patch, as shown in Fig. 9, there will be four basis functions associated with each local coordinate direction (ξ_1 or ξ_2), which generates a total of 16 2D basis functions. Fig. 9b shows an example of using 3rd order Hermite polynomials to represent a convoluted 3D surface in 3D space using a single 2D finite element patch. For comparison, if linear quadrilateral patches were used, about 100 linear patches would be required to represent the same surface with comparable accuracy. For this reason, Hermite polynomials are widely used in cardiac biomechanics studies for surface representation [1,42,43]. A simple 8 × 8 finite element model (FEM) with intrinsic C^1 continuity can sufficiently represent the geometry of the endocardium [1,43]. In our implementation, this 8 × 8 convention was followed, i.e. an 8 × 8 FEM using cubic Hermite polynomials as surface descriptors was used to represent the endocardium. The model was driven to segment the endocardium with forces derived from an energy functional described in the following section.

7.2. Segmentation energy functional

Since it would be too computationally expensive to pre-process or smooth the US images in our real-time application, traditional segmentation approaches that use an image gradient based energy minimization function are not robust enough due to the high noise level of ultrasound images. In order to overcome this problem, we borrowed an idea from the "active contour without edges" concept, which was originally proposed by Chan and Vese [16] for a level set framework.

The 3D active mesh model is deformed under external forces derived from a minimization of the Mumford–Shah energy functional [17]:

$$E(f, \vec{C}) = \beta \int_{\Omega} (f - g)^2 dV + \alpha \int_{\Omega \setminus \vec{C}} |\nabla f|^2 dV + \gamma \oint_{\vec{C}} ds, \tag{12}$$

in which $\vec{C}$ denotes the smoothed and closed segmentation, g represents the observed US data, f is a piecewise smoothed approximation to g with discontinuities only along $\vec{C}$, and Ω denotes the image domain. The first integral enforces the similarity between f and g, which is equivalent to homogeneity constraint since f is piecewise smoothed; the second integral controls the smoothness

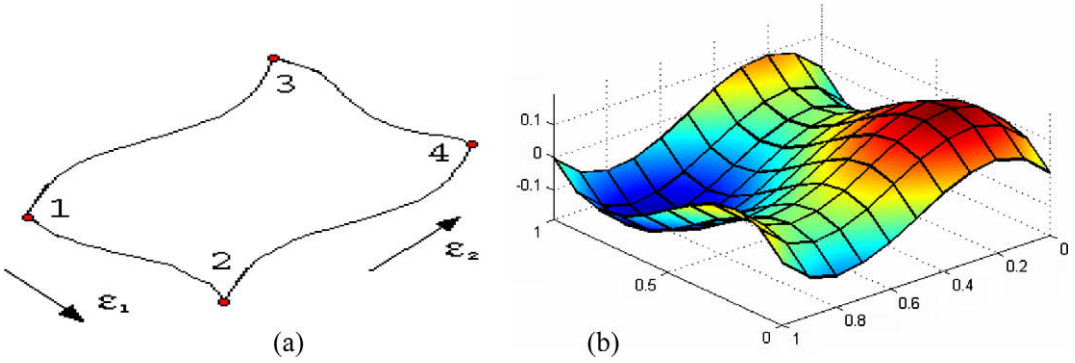

Fig. 9. Illustration of cubic Hermite surface representation: (a) a 2D surface with finite element patches used in our model; (b) a single patch with cubic Hermite representation that can efficiently characterize a convoluted surface as shown in color, whereas it takes about 100 linear quadrilateral patches to achieve similar accuracy.

of f; and the last integral is actually the length of the segmented boundary, which, acts as internal elasticity to prevent leaking at weak boundaries. The external forces driving the 3D active mesh were formulated using the same image homogeneity-based rationale proposed by Chan and Vese [16], using homogeneity measures inside and outside regions based on the current segmentation. Specifically, the optimum segmentation, corresponding to endocardium, divides the image into two relatively homogeneous regions. In this application, these regions correspond to the blood pool and the myocardium.

The Mumford–Shah equation (Eq. (12)) was only evaluated at discrete sampled sub-node points. To ensure that the system was well-defined, each surface patch was super-sampled by a 4×4 sub-node grid. In this way, the whole surface model was over-constrained with continuity constraints between each adjacent surface patch since the surface basis functions were cubic. The Mumford–Shah equation (Eq. (12)) was minimized using a Newton Downhill method, chosen for its computational efficiency

$$H^{i,t+dt} = H^{i,t} - dt\frac{\partial E}{\partial H^i}, \tag{13}$$

with dt representing the artificial time step in numerical iterations.

7.3. Repositioning and reorientation of the surface

One common drawback of segmentation using parametric active surface models is that the capture range of the method is usually small compared to other methods. For ultrasound segmentation applications, the initial contour is usually required to be positioned, usually manually, fairly close to the actual boundary [13]. Automated and semi-automated methods have been proposed to avoid manual intervention, using optical flow tracking [13], the Hough transform [44], or multi-scale approaches [45]. Since temporal performance is critically important for our real-time interventional application, a computationally more efficient approach was needed.

In order to reduce the dependence of the segmentation result on the initial position, after segmentation convergence for the current frame, the interface model used for the initialization of the next frame repositions itself so align its center and long-axis with the centroid and the long-axis of the segmented surface. This extra step in the initialization procedure speeds up the convergence, and keeps the Hermite coefficients at each node as small as possible, which increases the numerical stability of the optimization process.

7.4. Implementation

The segmentation software was implemented in C++ using an ITK [46]/VTK [47] compatible framework. These open source libraries were used for rendering, visualization, and interaction with the image data.

8. Validation

To quantitatively evaluate the segmentation's performance, nine 3D image volumes from a 4D phantom data set, (which contained a total of 500 3D frames), and four image volumes from a 4D clinical data set, (which contained a total of 87 3D fames), were randomly selected. The endocardial surfaces were manually traced using a customized tracing interface implemented using Matlab© (The MathWorks, Inc, Natick, Massachusetts); between 200 and 300 boundary points were selected for each volume. To quantitatively compare manual and automated segmentation results, the output of the automatic segmentation for the selected frames, i.e. after convergence of the FEM surface, was fitted to the corresponding manual tracing point cloud via updating of the Hermite parameters at each node. The endocardial surfaces, identified by manual tracing and automated segmentation were therefore efficiently represented by two sets of Hermite parameters at each node of a single finite element surface model, which enabled point-wise quantitative surface comparison. In this study, surface distance at each corresponding node point from the FEM for each surface was used as the criterion for surface discrepancy. A similar approach for segmentation comparison can be found in previous studies [48].

9. Results

9.1. Results on phantom data

The CT and US coordinate systems were rigidly registered using eight manually selected landmarks, and the RMS fiducial registration error was 3.80 mm.

The segmentation algorithm was applied to each ultrasound frame (500 in total) in the sequence, requiring 4.12 ± 0.73 (mean ± stdev) iterations per frame to converge on the endocardial border. The processing for each image frame took 150–200 ms on a Pentium 4 (2.80GHz desktop with 1.00 GB RAM running Microsoft Windows XP), enabling a 5 frame/second application update. The actual time required for the segmentation was 50 ± 8 ms; the remaining roughly 130 ms were required for data loading from the hard disk and rendering.

Fig. 10(a) shows a snapshot of the endocardial surface segmentation (red surface) overlaid with the ultrasound image data. After applying the rigid body transformation, the segmented endocardial surface can be mapped onto the pre-procedural CT dataset shown in (b) using the transform R_{US}^{CT}. As the ultrasound images capture

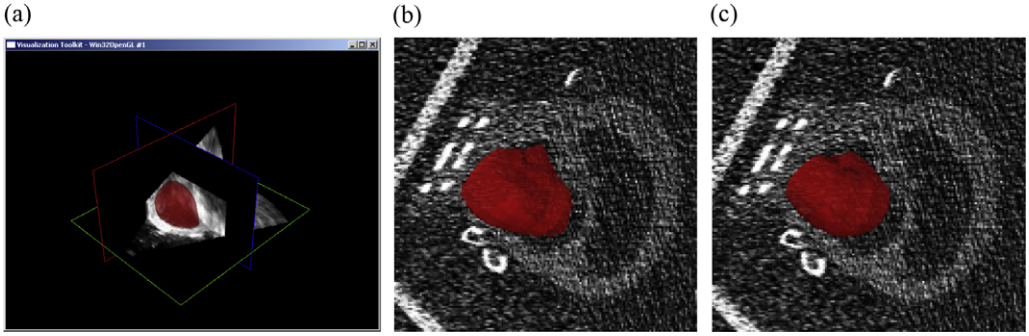

Fig. 10. (a) A representative ultrasound frame from 4D phantom ultrasound data set, with the endocardial surface (red) being segmented at 5 frames per second. (b) after applying the rigid body transformation, the segmented endocardial surface can be mapped onto the pre-procedural CT dataset. (c) As the ultrasound images capture dynamic changes in the shape of the phantom, the segmentation is updated, and continuously projected into the static CT dataset.

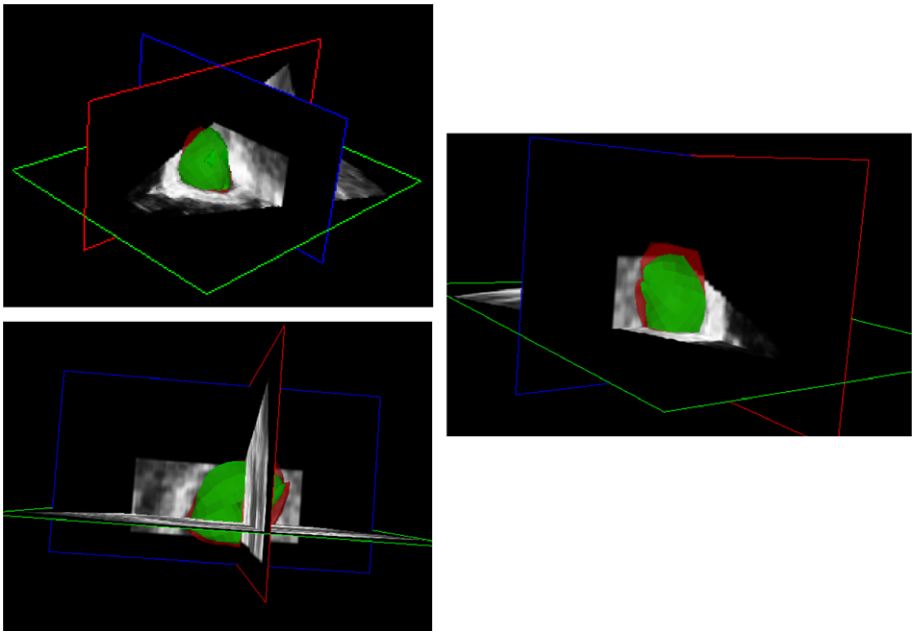

Fig. 11. Comparison of the segmentation of the phantom endocardial surface: automated method (red) and manual tracing (green) for one US frame at three different view angles.

dynamic changes in the shape of the phantom, the segmentation is updated, and continuously projected into the static CT dataset (Fig. 10c). Quantitative evaluation of the automatic segmentation compared to manual tracings, as described in Section 2.4, yielded an RMS error of 3.70 ± 2.5 mm which represents less than 3 pixels. Sample comparison frames are shown in Fig. 11, where the red surface displays the endocardial surface segmented by our automatic segmentation method and the green surface displays the endocardial surface from manual tracing.

9.2. Results on clinical data

The segmentation algorithm was also tested on each frame of a clinical patient data set (87 frames in total). The clinical data set had better image resolution than the phantom data. Quantitative evaluation of the segmentation via comparison with manual tracings, as described above, yielded a RMS error of 2.58 ± 1.58 mm which corresponds to approximately 3 pixels. The actual time required for segmentation of each frame of the clinical US data was 75 ± 12 ms (this time does not include loading the data from the disk or rendering). Sample comparison frames are shown in Fig. 12.

10. Discussions

This work presented a near real-time method for extracting endocardial surfaces in RT3D cardiac US streamed data and using this intra-procedural information to augment static CT roadmaps. This work demonstrated the feasibility of using streaming RT3D cardiac ultrasound as a real-time non-invasive tool for qualitative assessment of dynamic cardiac tissue morphology and proposed a framework for incorporating this information into an intra-procedural display for image guidance.

10.1. Registration and segmentation accuracy

The RMS segmentation error of the phantom data was 3.70 ± 2.5 mm. This is comparable to the findings in other studies using active contours for 2D cardiac ultrasound segmentation [13]. The segmentation errors in the clinical data set were smaller (2.58 ± 1.58 mm RMS error). The RMS error values reported in this study were calculated via a comparison between two finite element surface models (FEM), one fit by the automatic segmentation algorithm to the endocardial surface within the image data, and one fit to a manually traced surface. Inaccurate manual tracings

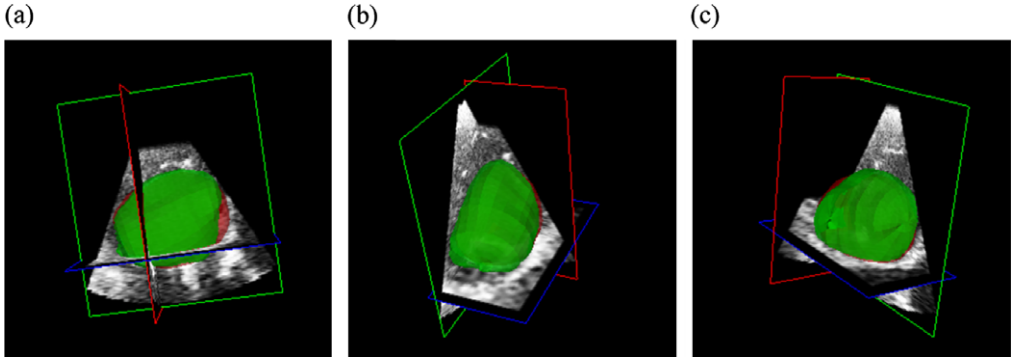

Fig. 12. Comparison of clinical data between the automated method (red) and manual tracing (green) on one frame at three different view angles.

and/or a non-perfect fit of the FEM surface to the manual tracing also contributed to the RMSE values. The improved segmentation accuracy with the clinical cardiac images was likely due to the data set's higher spatial resolution. The increased image resolution provided more support points along the endocardial surface. Additionally, higher image resolution might have facilitated more accurate manual tracings.

The multi-modal RMS registration error was 3.80 mm. This registration error was possibly due to the variation in manually picked landmark points, the limited resolution of the ultrasound image data, and deformation of the phantom between the CT and US scans. This error was comparable to the RMS error in other similar studies [49] (reporting 5 mm RMS surface distance between CT and MR co-registration).

In the multi-modality intra-operative display, the segmented endocardial surface from RT3D US was registered to the CT volume. Registration and segmentation errors both contributed to the overall error in the augmented roadmap display. The results of the phantom registration and segmentation experiments in this study indicated that this overall error could be as large as 7.5 mm. It is important to note that in the context of this application, i.e. augmentation of a static CT roadmap, segmentation was not intended to provide precise quantitative functional measurements, such as ejection fraction. Therefore the implementation emphasized temporal performance at the cost of optimal segmentation accuracy.

10.2. Sensitivity to initialization

In order to minimize the amount of manual interaction required to initialize the segmentation model, the algorithm was simply initialized as a small sphere in the center of the ultrasound volume. A sample screenshot of the initial state is shown in Fig. 13a. Only a small portion of the surface was inside the left ventricle at initialization. As shown in Fig. 13a, the surface was initialized with a major part of the model residing within the myocardium; only a small portion of the model was actually inside the ventricle. Several frames later, as shown in Fig. 13b, the model began to grow inside the ventricle. The segmentation depicted in Fig. 13b is the result after the model had converged. The incorrect segmentation in Fig. 13b is a transient result due to the poor initialization. Repositioning and reorientation of the model, based on the current segmentation, were key algorithmic steps that enabled the model to adjust itself and overcome poor initialization. In our experiments, the model only required about 1–2 s (about 25 frames) to properly align itself with the left ventricle as long as some of the *surface function actives* were within the ventricle at initialization. This 1–2 s overhead at the start of US imaging is acceptable in the context of most interventional applications. This analysis, regarding the segmentation method's relative insensitivity to initialization conditions, was based on qualitative observations in a small number of experiments. A rigorous investigation of capture range will be needed in future studies.

In the current implementation, repositioning and reorientation accounted for an average of 66% of the overall processing time for each frame. This could be a bottleneck of the algorithm and a candidate for further optimization to achieve higher frame rates. Fortunately, it has been found in cardiac biomechanics [1] that the long-axis of the left ventricle is relatively stable through the cardiac cycle. Therefore, it might be possible to turn off the reorientation/reposition step after the initial several seconds which were required for the model to adapt to actual orientation of the image data. This modification to the algorithm could provide a threefold gain in the update rate.

10.3. Segmentation speed

The update rate was about five frames per second, which was well below the maximum frame rate at which RT3D US image data can be acquired. The temporal performance of the segmentation could be increased by several means. In the current implementation, the 4D data stream was stored on the system hard disk. Therefore, the algorithm needed to load the data from the hard disk frame by frame, an effort that takes about 50 ms per frame. In a real clinical setting, the US data would be streaming into the system memory and would be processed in place; loading from the hard disk would not be required. Additionally, the time spent on visualization could also be decreased by using high performance graphics cards and rendering algorithms.

As pointed out in Section 4.2, since repositioning and reorientation accounts for an average of 66% of the overall processing time for each frame, the computational load could potentially be cut by a factor of 3 by reducing the frequency of these steps once the model is roughly aligned with the anatomy. Also, since most of the operations in the current implementation were node-wise, the algorithm could be parallelized in order to utilize the efficiency of multi-threaded and parallel computing. Finally, running the whole framework on a more powerful hardware platform would reduce the overall processing time.

By taking all these factors into account, the update rate of the overall framework might be able to reach 20 frames per second or higher rate for true real-time performance.

10.4. Future Developments

This work represented preliminary implementation towards the extraction of dynamic information from a 3D intra-procedural US stream. In addition to the endocardium's position and morphology, further valuable information might also be extracted using the current framework. Since the segmentation model is a parametric surface model based on finite elements, the segmentation process was actually a LV parameterization process. Once the segmentation

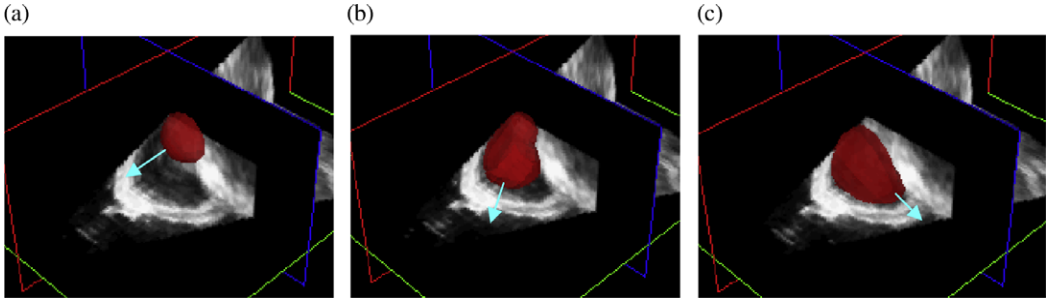

Fig. 13. Segmented model at (a) initial state, (b) 0.5 s later, and (c) 2 s later. The arrow indicates the orientation of the model's long-axis in each frame.

converged, a patient specific finite element LV surface model was automatically built. Based on this surface, model-based LV motion analysis [1] could be easily adapted into this framework. More detailed cardiac functional metrics, such as mechanical dyssynchrony, ejection fraction, local 3D fractional shortening, etc, might be directly computed or estimated from the model.

This framework could be also applied to other applications, such as tumor tracking for liver ablation guidance or tracking of the atria or right ventricle in minimally invasive cardiac interventions. Ultimately, this segmentation method could be applied in the context of a multi-modality guidance system using electromagnetic or optical tracking to localize the ultrasound probe for automatically and dynamically computing the registration with datasets from other imaging modalities.

11. Conclusions

Surface Function Actives (SFA) was presented as a new framework for deformable model. It offers great advantages in computational efficiency, benefited from dimensionality reduction in the interface representation. It utilizes implicit surface representation, enabling easy determination of inside from outside surface areas as well as straightforward quantitative segmentation comparison. Moreover, besides providing numerical solutions to the desired interface like a level set framework, SFA can use closed form expressions as well to achieve even better efficiency and accuracy. The continuous form of interface function can also benefit downstream analysis based on shape or other information from the interface. Although illustrated in a minimum-invasive application, SFA is suitable for any applications where real-time feedback is desired. The variational framework that SFA adopts provides flexibility of further expansion in terms of additional constraints or the ability to deal with multi-channel multi-phase segmentation problems.

Acknowledgments

The authors would like to thank Dr. Daniel R. Elgort, Dr. Guy Shechter, and Dr. Luis F. Gutiérrez from Philips Research North America to provide the streaming ultrasound data in this paper.

References

[1] S. Herz, C. Ingrassia, S. Homma, et al., Parameterization of left ventricular wall motion for detection of regional ischemia, Annals of Biomedical Engineering 33 (7) (2005) 912–919.
[2] J.K. Udupa, W.L.S. Samarasekera, et al., Multiple sclerosis lesion quantification using fuzzy Connectedness principles, IEEE Transactions in Medical Imaging 16 (1997) 598–609.
[3] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, International Journal of Computer Vision 1 (1987) 321–331.
[4] J. Sethian, Level Set Methods and Fast Marching Methods, second ed., vol. 3, Cambridge University Press, Cambridge, 1999.
[5] S. Osher, R. Fedkiw, Level Set Methods and Dynamic Implicit Surfaces, vol. 153, Springer, New York, 2003.
[6] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Transactions on Image Processing 10 (2) (2001) 266–277.
[7] E. Angelini, S. Homma, G. Pearson, et al., Segmentation of real-time three-dimensional ultrasound for quantification of ventricular function: a clinical study on right and left ventricles, Ultrasound in Medicine and Biology 31 (9) (2005) 1143–1158.
[8] T.F. Cootes, C.J. Taylor, D.H. Cooper, et al., Active shape models–their training and application, Computer Vision and Image Understanding 61 (1) (1995) 38–59.
[9] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, Lecture Notes in Computer Science, vol. 1407, Springer, Berlin/Heidelberg, 1998, pp. 484–498.
[10] K. Held, E.R. Kops, B.J. Krause, et al., Markov random field segmentation of brain MR images, IEEE Transactions on Medical Imaging 16 (6) (1997) 878–886.
[11] Y.Y. Boykov, M.-P. Jolly, Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images, in: Eighth International Conference on Computer Vision (ICCV'01), 2001, pp. 105–112.
[12] Y. Jin, C. Imielinska, A. Laine, et al., Segmentation and evaluation of adipose tissue from whole body MRI scans, in: Proceedings of the Sixth International Conference on Medical Image Computing and Computer Assisted Interventions (MICCAI 2003), Montreal, Canada, 2003, pp. 635–642.
[13] I. Mikic, S. Krucinski, J.D. Thomas, Segmentation and tracking in echocardiographic sequences: active contours guided by optical flow estimates, IEEE Transactions on Medical Imaging 17 (2) (1998) 274–284.
[14] C. Xu, J.L. Prince, Snakes, Shapes and Gradient Vector Flow, IEEE Transaction on Image Processing 7 (3) (1998) 359–369.
[15] J.A. Sethian, Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science, Cambridge University Press, Cambridge, UK, 1999.
[16] T.F. Chan, L.A. Vese, Active contours without edges, IEEE Transactions on Image Processing 10 (2) (2001) 266–277.
[17] D. Mumford, J. Shah, Boundary detection by minimizing functional, in: International Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 1985, pp. 22–26.
[18] E. Angelini, A. Laine, S. Takuma, et al., LV volume quantification via spatio-temporal analysis of real-time 3D echocardiography, IEEE Transactions on Medical Imaging 20 (6) (2001) 457–469.
[19] T. Song, E.D. Angelini, B.D. Mensh, et al., Comparison study of clinical 3D MRI brain segmentation evaluation, in: Annual International Conference IEEE Engineering in Medicine and Biology Society (EMBS), San Francisco, CA, USA, 2004, pp. 1671–1674.
[20] A. Yezzi, A. Tsai, A. Willsky, A fully global approach to image segmentation via coupled curve evolution equations, Journal of Visual Communication and Image Representation 13 (2002) 195–216.
[21] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, in: Proceedings of 1st International Conferences on Computer Vision, 1987, pp. 259–268.
[22] X. Han, C. Xu, J. Prince, A topology preserving deformable model using level sets, Computer Vision and Pattern Recognition (2001) 765–770.
[23] M.R.M. Jongbloed, M.S. Dirksen, J.J. Bax, et al., Atrial Fibrillation: Multi-Detector Row CT of Pulmonary Vein Anatomy prior to Radiofrequency Catheter Ablation–Initial Experience, Radiology 234 (2005) 702–709.
[24] P. Schoenhagen, Three-dimensional imaging for the guidance of coronary interventional procedures: impact on clinical decision making?, The International Journal of Cardiovascular Imaging 20 (2004) 531–532
[25] A.M. Wink, J.B.T.M. Roerdink, Denoising functional MR images: a comparison of wavelet denoising and Gaussian smoothing, IEEE Transactions on Medical Imaging 23 (3) (2004) 374–387.
[26] C.M. Hillenbrand, D.R. Elgort, E.Y. Wong, et al., Active device tracking and high-resolution intravascular MRI using a novel catheter-based, opposed-solenoid phased array coil, Magnetic Resonance in Medicine 51 (4) (2004) 668–675.
[27] O.T.V. Ramm, S.W. Smith, Real time volumetric ultrasound imaging system, Journal of Digital Imaging 3 (4) (1990) 261–266.
[28] Q. Duan, E.D. Angelini, A. Laine, Assessment of visual quality and spatial accuracy of fast anisotropic diffusion and scan conversion algorithms for real-time three-dimensional spherical ultrasound, in: SPIE International Symposium Medical Imaging, San Diego, CA, USA, 2004, pp. 331–342.
[29] M. Mulet-Parada, J.A. Noble, 2D+T acoustic boundary detection in echocardiography, in: Medical Image Computing and Computer-Assisted Intervention—MICCAI'98, Cambridge, MA, 1998, pp. 806–813.
[30] D. Boukerroui, O. Basset, A. Baskurt, et al., Segmentation of echocardiographic data. Multiresolution 2D and 3D Algorithm Based on Grey Level Statistics, in: MICCAI, Cambridge, UK, 1999, pp. 516–523.
[31] J.G. Thomas, R.A. Peters II, P. Jeanty, Automatic segmentation of ultrasound images using morphological operators, IEEE Transactions on Medical Imaging 10 (2) (1991) 180–186.
[32] H. Zhang, Z. Bian, D. Jiang, et al., Level set method for pulmonary vessels extraction, in: International Conference on Image Processing, 2003, pp. II-1105–II-1108.
[33] B. Vallet, E.D. Angelini, A.F. Laine, Variational segmentation framework in prolate spheroidal coordinates for 3D real-time echocardiography, in: Medical Imaging 2006: Image Processing, San Diego, CA, USA, 2006, pp. 1370–1380.
[34] V. Chalana, D.T. Linker, D.R. Haynor, et al., A multiple active contour model for cardiac boundary detection on echocardiographic sequences, IEEE Transactions on Medical Imaging 15 (3) (1996) 290–298.
[35] R. Drezek, G.D. Stetten, T. Ota, et al., Active contour based on the elliptical Fourier series, applied to matrix-array ultrasound of the heart, in: 25th AIPR Workshop: Emerging Applications of Computer Vision, 1997, pp. 26–34.
[36] Z. Tao, H.D. Tagare, Stopping rules for active contour segmentation of ultrasound cardiac images, in: Medical Imaging 2005: Image Processing, San Diego, CA, USA, 2005, pp. 475–484.
[37] N. Paragios, A level set approach for shape-driven segmentation and tracking of the left ventricle, IEEE Transactions on Medical Imaging 22 (6) (2003) 773–776.
[38] R. Valdes-Cristerna, V. Medina-Banuelos, O. Yanez-Suarez, Coupling of radial-basis network and active contour model for multispectral brain MRI segmentation, IEEE Transactions on Biomedical Engineering 51 (3) (2004) 459–470.
[39] J.G. Bosch, S.C. Mitchell, B.P.F. Lelieveldt, et al., Fully automated endocardial contour detection in time sequences of echocardiograms by three-dimensional active appearance models, in: Medical Imaging 2002: Image Processing, San Diego, CA, USA, 2002, pp. 452–462.

[40] S.C. Mitchell, J.G. Bosch, B.P.F. Lelieveldt, et al., 3-D active appearance models: segmentation of cardiac MR and ultrasound images, IEEE Transactions on Medical Imaging 21 (9) (2002) 1167–1178.
[41] B. Vallet, E. Angelini, A. Laine, Variational segmentation framework in prolate spheroidal coordinates for 3D real-time echocardiography, in: SPIE Medical Imaging Conference, San Diego, CA, USA, 2006, pp. 61444A.1–61444A.11.
[42] G.R. Christie, D.P. Bullivant, S.A. Blackett, et al., Modelling and visualising the heart, Computing and Visualization in Science 4 (4) (2004) 227–235.
[43] P.M. Nielsen, I.J. LeGrice, B.H. Smaill, et al., Mathematical model of geometry and fibrous structure of the heart, American Journal of Physiology—Heart and Circulation Physiology 260 (1991) H1365–H1378.
[44] T. Wang, I.Y.H. Gu, M. Viberg, et al., Tracking moving objects in video using enhanced mean shift and region-based motion field, in: EUSIPCO, 2007, pp. 307–311.
[45] J. Chung, P. Abraszewski, X. Yu, et al., Paradoxical increase in ventricular torsion and systolic torsion rate in type I diabetic patients under tight glycemic control, Journal of the American College of Cardiology 47 (2) (2006) 384–390.
[46] P. Hellier, C. Barillot, I. Corouge, et al., Retrospective evaluation of intersubject brain registration, IEEE Transaction on Medical Imaging 22 (9) (2003) 1120–1130.
[47] E.D. Angelini, T. Song, B.D. Mensh, et al., Segmentation and quantitative evaluation of brain MRI data with a multi-phase three-dimensional implicit deformable mode, in: SPIE International Symposium, Medical Imaging 2004, San Diego, CA, USA, 2004, pp. 526–537.
[48] Q. Duan, E. D. Angelini, S. L. Herz, et al., Tracking of LV endocardial surface on real-time three-dimensional ultrasound with optical flow, in: Third International Conference on Functional Imaging and Modeling of the Heart 2005, Barcelona, Spain, 2005, pp. 434–445.
[49] T. O'Donnell, S. Aharon, S.S. Halliburton, et al., Multi-modality model-based registration in the cardiac domain, in: IEEE Conference on Computer Vision and Pattern Recognition, 2000, pp. 790–791.