

# Temporally diffeomorphic cardiac motion estimation from three-dimensional echocardiography by minimization of intensity consistency error

Zhijun Zhang<sup>a)</sup>

Department of Biomedical Engineering, Oregon Health and Science University (OHSU), 3181 Southwest Sam Jackson Park Road, Portland, Oregon 97239

Muhammad Ashraf

Department of Pediatric Cardiology, Oregon Health and Science University, 3181 Southwest Sam Jackson Park Road, Portland, Oregon 97239

David J. Sahn

Department of Biomedical Engineering, Oregon Health and Science University (OHSU), 3181 Southwest Sam Jackson Park Road, Portland, Oregon 97239 and Department of Pediatric Cardiology, Oregon Health and Science University, 3181 Southwest Sam Jackson Park Road, Portland, Oregon 97239

Xubo Song

Department of Biomedical Engineering, Oregon Health and Science University (OHSU), 3181 Southwest Sam Jackson Park Road, Portland, Oregon 97239

(Received 8 September 2013; revised 14 February 2014; accepted for publication 22 February 2014; published 15 April 2014)

**Purpose:** Quantitative analysis of cardiac motion is important for evaluation of heart function. Three dimensional (3D) echocardiography is among the most frequently used imaging modalities for motion estimation because it is convenient, real-time, low-cost, and nonionizing. However, motion estimation from 3D echocardiographic sequences is still a challenging problem due to low image quality and image corruption by noise and artifacts.

**Methods:** The authors have developed a temporally diffeomorphic motion estimation approach in which the velocity field instead of the displacement field was optimized. The optimal velocity field optimizes a novel similarity function, which we call the intensity consistency error, defined as multiple consecutive frames evolving to each time point. The optimization problem is solved by using the steepest descent method.

**Results:** Experiments with simulated datasets, images of an *ex vivo* rabbit phantom, images of *in vivo* open-chest pig hearts, and healthy human images were used to validate the authors' method. Simulated and real cardiac sequences tests showed that results in the authors' method are more accurate than other competing temporal diffeomorphic methods. Tests with sonomicrometry showed that the tracked crystal positions have good agreement with ground truth and the authors' method has higher accuracy than the temporal diffeomorphic free-form deformation (TDFFD) method. Validation with an open-access human cardiac dataset showed that the authors' method has smaller feature tracking errors than both TDFFD and frame-to-frame methods.

**Conclusions:** The authors proposed a diffeomorphic motion estimation method with temporal smoothness by constraining the velocity field to have maximum local intensity consistency within multiple consecutive frames. The estimated motion using the authors' method has good temporal consistency and is more accurate than other temporally diffeomorphic motion estimation methods.

© 2014 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4867864>]

Key words: cardiac motion estimation, diffeomorphic registration, echocardiography

## 1. INTRODUCTION

Heart disease is the leading cause of the death and morbidity in the U.S.<sup>1</sup> The study of its motion and deformation is important for the diagnosis and treatment of heart disease.<sup>2</sup> It helps in understanding the biomechanics of the heart, such as shortening, stretching, and twisting in a dynamic approach.<sup>3,4</sup> It can be used to quantitatively evaluate stress-strain variation in the myocardium.<sup>5-7</sup> Abnormality in cardiac motion can help doctors diagnose heart diseases such as myocardial infarction and/or ischemia.<sup>8</sup> It also helps physicians and surgeons val-

idate the effectiveness of cardiac surgery or other treatments such as cardiac resynchronization therapy.<sup>9</sup>

Cardiac imaging provides a noninvasive, three-dimensional, and dynamic way to study heart motion. Various image modalities can be used to acquire 3D dynamic images (3D+t) of the heart, which include MRI,<sup>10</sup> cardiac CT,<sup>11</sup> echocardiography,<sup>12,13</sup> and ultrasound tissue Doppler imaging.<sup>14</sup> Among these modalities, 3D echocardiography has the advantages of convenience, cost efficiency, real-time operation, and absence of radiation. However, motion estimation from 3D echocardiographic images is challenging

because of the low image quality and low spatiotemporal resolution.

Motion estimation algorithms for cardiac imaging generally fall into two categories: *model-based* and *intensity-based*. In model-based methods, models of the heart geometry and/or its deformation are used to constrain the estimation problem. The advantage of these methods is that these models lead to compact data representations. Various geometrical models have been proposed, such as surface models<sup>15</sup> and volume models.<sup>16,17</sup> A detailed survey is given in Ref. 18. However, model-based methods usually require the extraction of features such as boundary points or contours as a prerequisite, which is challenging in itself. Intensity-based methods estimate motion directly from voxel intensities, without any assumption about geometric shape or deformation.<sup>7</sup> They have the advantage that the motion estimation is not limited to a specific model. However, they tend to be more computationally intensive. Intensity-based registration methods such as the 3D B-spline based method,<sup>19</sup> Demon's method,<sup>20,21</sup> optical flow methods,<sup>22</sup> variational methods,<sup>23,24</sup> and speckle tracking<sup>12</sup> have been used to estimate cardiac motion from image sequences.<sup>6,25–31</sup>

Temporal smoothness is an important issue to consider for cardiac motion analysis because motion of heart wall is a smooth process. Introduction of temporal information can help to improve the motion estimation performance. Various temporal constraints have been integrated with the deformable models to regularize the estimated motion. Huang *et al.*<sup>32</sup> used a spatiotemporal (4D) freeform deformation (FFD) model to fit the corresponding points extracted from tagged MRI sequences. Lagrangian Dynamic motion equations have been used to govern deformation of surface models<sup>33</sup> such as superquadrics<sup>34</sup> and simplex surfaces.<sup>35</sup> Volumetric models such as finite elements (FE) with dynamic motion equations were also used in Refs. 17 and 36. Gerard *et al.*<sup>37</sup> used a statistical motion model learned from the tagged MRI to extend the 3D simplex surface model for motion tracking of echocardiographic sequences. Metz *et al.*<sup>38</sup> used a 4D statistical model estimated by principal component analysis to predict patient-specific cardiac motion. Wang *et al.*<sup>39</sup> tracked myocardial surface by maximizing the likelihood of a combined intensity model and a two-step motion prediction model, which was learned in advance. Comaniciu *et al.*<sup>40</sup> proposed a Kalman filter-based information fusion framework for shape tracking with a probabilistic subspace model constraint. McEachen *et al.*<sup>41</sup> used an adaptive filter to estimate the parameters of a harmonic series based temporal model by minimizing the error between the correspondence on image contours.

In intensity-based motion estimation methods, the transformation is usually represented by a continuous function. The temporal smoothness of the cardiac motion is usually taken into account in three ways. First, spatiotemporally smooth transformations<sup>38,42–46</sup> are used to represent the deformation functions. Second, the spatial transformations at different time points are regularized by constraints such as trajectory smoothness, inverse consistency, and transitivity.<sup>47–51</sup> Third, diffeomorphisms are used since they are intrinsically temporally smooth and topology preserving.<sup>52–55</sup> For the last

approach, the velocity field is estimated instead of the displacement field.<sup>56–58</sup> The transformation is temporally smooth and one-to-one if the velocity field is smooth.<sup>59</sup>

For both model-based and intensity-based methods, motion tracking is formulated as a registration problem in which an energy function consisting of a similarity term and a regularization term is optimized. The similarity measurement of the image sequences is usually adopted in two strategies according to how the reference frame is chosen. The first strategy considers the similarity between a fixed reference frame and all subsequent frames unwarped to it,<sup>42,44,47,49–51,60</sup> which we call the *reference-to-following* approach. The result is usually biased toward the reference frame, and is error-prone due to the decorrelation and dissimilarity between distant frames such as end-diastole (ED) and end-systole (ES) frames.<sup>61</sup> The second strategy evaluates the similarity between consecutive frames,<sup>6,25,27,28,62</sup> which we call the *frame-to-frame* approach. It has the advantage that consecutive frames have high correlation. Some modified methods based on these two strategies have also been proposed. In Metz *et al.*,<sup>38</sup> a global reference-to-following similarity measurement was used to adaptively update the reference image as the average of all unwarped subsequent frames. In Yigitsoy *et al.*,<sup>43</sup> a groupwise similarity measurement was used to minimize the summed SSDs between any two frames unwarped into the reference coordinate. In De Craene *et al.*,<sup>58</sup> a weighted sum of frame-to-frame and reference-to-following similarity measurements was used to take advantages of both methods. However, decorrelation between distant frames still exists in these improved measurements.

In our work, we proposed a temporally consistent diffeomorphic motion estimation method with similarity measurement evaluated at *multiple* consecutive frames. The transformation is spatiotemporally smooth because it is defined as the integral of a series of spatiotemporally smooth velocity fields. We used a compromise strategy between the reference-to-following and frame-to-frame methods for similarity measurement by using multiple (not just two) consecutive frames. As a result, there is less bias toward a fixed reference frame, and less decorrelation between distant frames. The use of multiple consecutive frames to estimate the velocity field also leads to improved temporal smoothness compared with the frame-to-frame method. In our experiments, we chose four consecutive frames for a balance of good temporal consistency, reduction of bias, and computation efficiency. Compared to the reference-to-following methods, our estimated transformation has higher fidelity with respect to the local image frames because it minimizes the local intensity consistency error (ICE), defined in Sec. 2.B, of several consecutive frames. Compared to the frame-to-frame method, our method has higher temporal consistency because it uses more consecutive frames. Experiments with simulated and real datasets showed improved performance.

Preliminary results of the method were presented in a conference paper.<sup>63</sup> In this paper, we extended our early work extensively by validating the proposed method thoroughly with simulated, *ex vivo*, *in vivo*, and human 3D image sequences. The paper is organized as follows. Section 2 describes our

method. In Sec. 3, the datasets and experiments are described in detail and in Sec. 4 the results are presented. Our discussion and conclusion are presented in Sec. 5.

## 2. METHODS

We first introduce the framework for diffeomorphic registration of two frames. Then we extend it to motion estimation of image sequences using our proposed intensity consistency error similarity measurement that involves multiple consecutive frames.

### 2.A. Parameterized diffeomorphic registration

We define a transformation  $\phi(\mathbf{x}, t)$ ,  $t \in [0, 1]$ ,  $\mathbf{x} \in \Omega \subset R^d$  ( $d=2, 3$ ) with smooth velocity field  $\mathbf{v}(\mathbf{x}, t)$  using the differential equation  $\frac{d\phi}{dt} = \mathbf{v}(\phi(\mathbf{x}), t)$ ,  $\phi(\mathbf{x}, 0) = \mathbf{x}$ . It has been proven that if  $\mathbf{v}(\mathbf{x}, t)$  is smooth with a differential operator  $L$  in a Sobolev space  $V$ , the transformation  $\phi(\mathbf{x}, t)$  defines a group of diffeomorphisms with  $t$  varying from 0 to 1.<sup>59</sup> The diffeomorphic image registration is stated as a variational problem, that given two images  $I_0$  and  $I_1$ , to find an optimal velocity field  $\hat{\mathbf{v}}$  which minimizes an energy functional consisting of a time integral of summed square difference (SSD) functions between the unwrapped images of  $I_0(\phi_{t,0}(\mathbf{x}))$  and  $I_1(\phi_{t,1}(\mathbf{x}))$  at each time point  $t$  and a distance metric between transformations  $\phi_{t,0}(\mathbf{x})$  and  $\phi_{t,1}(\mathbf{x})$ ,<sup>64</sup>

$$\hat{\mathbf{v}} = \arg \inf_{\mathbf{v} \in V} \lambda \int_0^1 \|\mathbf{v}(\mathbf{x}, t)\|_V^2 dt + \int_0^1 \int_{\Omega} (I_0(\phi_{t,0}(\mathbf{x})) - I_1(\phi_{t,1}(\mathbf{x})))^2 d\mathbf{x} dt, \quad (1)$$

with  $\lambda$  being the weight to balance these two energies,  $\phi_{t,0}(\mathbf{x})$  and  $\phi_{t,1}(\mathbf{x})$  being the transformations from time  $t$  to 0 and from  $t$  to 1. The optimal velocity field satisfies the Euler-Lagrange equation of the variational functional in Eq. (1). A direct solution for the partial differential equation (PDE) (the Euler-Lagrange equation) is computationally expensive.<sup>52</sup> Alternatively, a parameterized representation of the velocity field can be adopted,<sup>58,65</sup> where the velocity fields are represented with a series of 3D or a 4D B-spline function. The transformation is the integral of the velocity field and can be

approximated at discretized time points by using Euler forward integral with assumption that the velocity is piecewise constant within each time step.

In our work, we use a series of B-spline functions to parameterize the velocity field and the B-spline function at time point  $t_k$ , defined as  $\mathbf{v}(\mathbf{x}, t_k) = \sum_i \mathbf{c}_{i;k} \boldsymbol{\beta}(\mathbf{x} - \mathbf{x}_i)$ , with  $\mathbf{c}_{i;k}$  being the B-spline control vectors on a uniform grid of  $\mathbf{x}_i$  at  $t_k$  and  $\boldsymbol{\beta}(\mathbf{x} - \mathbf{x}_i)$  being the 3D cubic B-spline kernel function, which is the tensor product of the 1D B-spline functions. Denote  $\phi_k = \phi(\mathbf{x}, t_k)$  as the transformation from  $t_0$  to  $t_k$ , because the velocity is piecewise constant within each time step, we have

$$\begin{aligned} \phi_k &= \phi_{k-1} + \mathbf{v}(\phi_{k-1}, t_{k-1}) \Delta t \\ &= (\mathbf{I}d + \mathbf{v}_{k-1} \Delta t) \circ \phi_{k-1}, (k=1, 2, \dots, N_s, \Delta t=1/N_s), \end{aligned} \quad (2)$$

with  $\phi_0(\mathbf{x}) = \mathbf{x}$ ,  $N_s$ , and  $\Delta t$  being the total number of time steps of the discretized velocity field and the length of each time step, respectively, and  $\circ$  being the composition operator. The second term in Eq. (1) is discretized as  $\sum_{k=0}^{N_s} \int (I_0(\phi_{k,0}) - I_1(\phi_{k,N_s}))^2 d\mathbf{x}$ , which is a sum of SSD functions of the unwrapped images  $I_0(\phi_{k,0})$  and  $I_1(\phi_{k,N_s})$  at all time points  $t_k$  ( $k=0, 1, \dots, N_s$ ). The forward transformation  $\phi_{k,N_s}$  from  $t_k$  to  $t_{N_s}$  can be represented as  $\phi_{k,N_s} = (\mathbf{I}d + \mathbf{v}_{N_s-1} \Delta t) \circ \dots \circ (\mathbf{I}d + \mathbf{v}_k \Delta t)$ . By considering the reverse motion whose velocity is  $-\mathbf{v}(\mathbf{x}, t)$  at each time, the backward transformation from time  $t_k$  to  $t_0$  can be represented as  $\phi_{k,0} = (\mathbf{I}d - \mathbf{v}_1 \Delta t) \circ \dots \circ (\mathbf{I}d - \mathbf{v}_{k-1} \Delta t)$ . The registration energy functional Eq. (1) is then parameterized as a function of a group of parameters  $\{\mathbf{c}_{i;k}\}$  ( $k \in [0, N_s]$ ) which can be optimized with the gradient descent method.<sup>65</sup>

### 2.B. Diffeomorphic motion estimation with ICE for image sequences

Assume that we have an image sequence  $I_n$  ( $n=0, 1, 2, \dots, N_f$ ). We define a flow of diffeomorphisms  $\phi(\mathbf{x}, t)$ ,  $t \in [0, N_f]$ ,  $\mathbf{x} \in \Omega \subset R^d$  ( $d=2, 3$ ) with smooth velocity field  $\mathbf{v}(\mathbf{x}, t)$  using the differential equation  $\frac{d\phi}{dt} = \mathbf{v}(\phi(\mathbf{x}, t), t)$ ,  $\phi(\mathbf{x}, 0) = \mathbf{x}$ . We use  $N_s$  time steps for the velocity field between each two consecutive frames. Figure 1 shows the principle of intensity consistency along the point trajectories

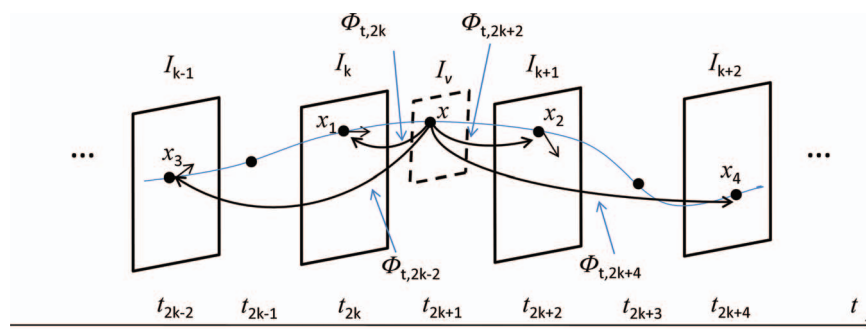


FIG. 1. The intensity of a point along the trajectory should be preserved. The optimal velocity field should minimize the difference of evolving image  $I_k(\phi_{t,t_{N_s,k}})$ ,  $I_{k+1}(\phi_{t,t_{N_s,k+1}})$ ,  $I_{k-1}(\phi_{t,t_{N_s,k-1}})$ ,  $I_{k+2}(\phi_{t,t_{N_s,k+2}})$  at time point  $t$  ( $t \in [k, k+1]$ ). We use two time steps ( $N_s = 2$ ) between each consecutive frames for illustration.

at a random time point (For convenience of illustration, we explain our method by using  $N_s = 2$  for an example, but the idea behind it is the same when other integers are used). Assume that in the image sequence a point  $\mathbf{x}$  in a virtual plane  $I_v$  at  $t$  moves through four consecutive frames  $I_{k-1}, I_k, I_{k+1}, I_{k+2}$  ( $t \in [k, k+1]$ ). We use the notation  $\phi_{t, t_{N_s k}}$  to indicate the transformation which maps  $\mathbf{x}$  from  $t$  to  $t_{N_s k}$ . Then the trajectories of point  $\mathbf{x}$  in the four consecutive frames are  $\phi_{t, t_{N_s k}}(\mathbf{x}), \phi_{t, t_{N_s(k+1)}}(\mathbf{x}), \phi_{t, t_{N_s(k-1)}}(\mathbf{x}), \phi_{t, t_{N_s(k+2)}}(\mathbf{x})$ . Their intensity values  $I_k(\phi_{t, t_{N_s k}}), I_{k+1}(\phi_{t, t_{N_s(k+1)}}), I_{k-1}(\phi_{t, t_{N_s(k-1)}}), I_{k+2}(\phi_{t, t_{N_s(k+2)}})$  in the four frames should be preserved since they are from the same physical point. Theoretically, the point  $\mathbf{x}$  at  $t$  can be replaced by any time point along the trajectory and the four frames can be extended to all image frames in the sequences.

In our implementation, we consider up to four frames as a compromise of rich intensity consistency context and computational efficiency (at the beginning and end of the sequence, three images are considered). On the one hand, considering consistency between two frames before and after a time point gives us reasonable temporal constraints on the point trajectories. On the other hand, consideration of more frames produces correspondence ambiguity due to speckle decorrelation. We define an energy term to measure the ICE as

$$E_{ICE}(t) = \int_{\Omega} ((I_k(\phi_{t, t_{N_s k}}) - I_{k+1}(\phi_{t, t_{N_s(k+1)}}))^2 + (I_{k-1}(\phi_{t, t_{N_s(k-1)}}) - I_{k+1}(\phi_{t, t_{N_s(k+1)}}))^2 + (I_k(\phi_{t, t_{N_s k}}) - I_{k+2}(\phi_{t, t_{N_s(k+2)}}))^2) d\mathbf{x}, \quad (3)$$

$\times (t \in [k, k+1]).$

The ICE function at time point  $t$  measures consistency of the intensity values of local frames under a velocity field. The optimal velocity field is estimated by minimizing a variational energy

$$\hat{\mathbf{v}} = \arg \inf_{\mathbf{v} \in V} \lambda \sum_{k=0}^{N_t} \|\mathbf{v}(\mathbf{x}, t_k)\|_V^2 + \sum_{k=0}^{N_t} E_{ICE}(t_k), \quad (4)$$

with  $N_t = N_s * N_f$  being the total number of time points used in the velocity field. The first term regularizes the velocity field to make it spatiotemporally smooth, and the second term assures that the optimal velocity field minimizes the sum of the local intensity consistency error at all time points  $t_k (k = 0, 1, 2, \dots, N_t)$ .

We use an adaptive scheme to choose the value of  $N_s$ . It is initialized as 2. The B-spline parameters are checked at each iteration to make sure that the transformation between each two time points, i.e.,  $\mathbf{I}d + \mathbf{v}_k \Delta t$ , is diffeomorphic.<sup>66</sup> If the

condition is broken due to large deformation between two frames, the number of  $N_s$  will be doubled to tolerate larger deformation, while the transformation between time points is diffeomorphic.

## 2.C. Spatiotemporal velocity field regularization

In order to assure that  $\phi(\mathbf{x}, t)$  is diffeomorphic, we need to define  $\mathbf{v}(\mathbf{x}, t)$  to be spatiotemporally smooth under a differential operator  $L$ . The linear operator we choose is:  $L = \nabla^2(\cdot) + w_t \frac{d}{dt}$ , with  $\nabla^2(\cdot)$  being a Laplacian operator,  $\frac{d}{dt}$  being the time derivative, and  $w_t$  being a constant weight. In the discrete time form of velocity field, the first term of Eq. (4) is approximated by

$$E_{reg} = \sum_{k=0}^{N_t} \sum_{\mathbf{x} \in \Omega} |\nabla^2 \mathbf{v}_k|^2 + w_t \sum_{k=1}^{N_t} \sum_{\mathbf{x} \in \Omega} |\mathbf{v}_k(\mathbf{x} + \mathbf{v}_{k-1} \Delta t) - \mathbf{v}_{k-1}|^2, \quad (5)$$

with  $\mathbf{v}_k = \mathbf{v}(\mathbf{x}, t_k)$ . The first term makes the velocity field spatially smooth, which is denoted as  $E_{sr}$ . The second term keeps the particle velocity smooth and it is indicated by  $E_{tr}$ . The overall effect is to make the velocity field spatiotemporally smooth.

## 2.D. Optimization

We use a steepest descent method to optimize the parameterized function. The derivative of the total registration energy with respect to the transformation parameters are calculated analytically.

Due to the fact that the ICE at each time point  $t_k$  is made up of SSD functions of the three closest pairs of frames across it, the derivative of the ICE energy term in Eq. (4) with respect to the B-spline parameters is equalized to the sum of the derivative of the three SSD functions. Suppose we have two images  $I_0$  and  $I_1$ , we want to estimate the derivative of SSD energy term  $E_{0,1}(t_j) = \int_{\Omega} (I_0(\phi_{j,0}(\mathbf{x})) - I_1(\phi_{j,N_s}(\mathbf{x})))^2 d\mathbf{x}$  at each  $t_j$  with respect to a series of discrete velocity field B-spline parameters  $\mathbf{v}_k (0 \leq k \leq N_1)$ , with  $N_s$  the total number of time steps. We know from Sec. 2.A that  $\phi_{j,N_s}$  is only affected by the velocity field  $\mathbf{v}_k$  with  $j \leq k \leq N_s - 1$ , and  $\phi_{j,0}$  is only related to  $-\mathbf{v}_k$  with  $0 \leq k \leq j - 1$ . If we denote  $c_{i,m;k}$  as the  $m$  component ( $m \in \{x, y, z\}$ ) of the  $i$ th B-spline parameter of  $\mathbf{v}_k$ , the derivative of  $E_{0,1}(t_j)$  with respect to  $c_{i,m;k}$  is

$$\frac{\partial E_{0,1}(t_j)}{\partial c_{i,m;k}} = \begin{cases} \sum_{\Omega} (I_1(\phi_{j,N_s}) - I_0(\phi_{j,0})) \nabla_m I_1(\phi_{j,N_s}) \frac{\partial \phi_{j,N_s}}{\partial c_{i,m;k}} & \text{if } (j \leq k \leq N_s - 1), \\ \sum_{\Omega} (I_0(\phi_{j,0}) - I_1(\phi_{j,N_s})) \nabla_m I_0(\phi_{j,0}) \frac{\partial \phi_{j,0}}{\partial c_{i,m;k}} & \text{if } (0 \leq k \leq j - 1) \end{cases} \quad (6)$$

with  $\Omega'$  being the local support of the B-spline kernel function,  $\nabla_m(\cdot)$  being the  $m$ th component of the image gradient, and  $\frac{\partial \phi_{i,N_f}}{\partial c_{i,m;k}}$  and  $\frac{\partial \phi_{i,0}}{\partial c_{i,m;k}}$  being the derivative of the concatenated B-spline function with respect to the B-spline parameters which are calculated using chain rule.<sup>57,65</sup> By replacing image pair of  $I_0$  and  $I_1$  with the three pairs of images used in intensity consistency error at each  $t_k$ , we get the derivative of the total similarity metric with respect to the velocity field parameters.

For the derivative of the spatial and temporal regularization energies with respect to  $c_{i,m;k}$ , we have

$$\frac{\partial E_{sr}}{\partial c_{i,m;k}} = \sum_{k=0}^{N_f} \sum_{\mathbf{x} \in \Omega'} (\nabla^2 v_k) \beta_m''(\mathbf{x} - \mathbf{x}_i), \quad (7)$$

with  $\beta_m''(\cdot)$  being the second derivative of the B-spline function with respect to  $m$  component. Considering that the displacement between two time steps is small, we have

$$\begin{aligned} \frac{\partial E_{tr}}{\partial c_{i,m;k}} &= w_t \sum_{k=0}^{N_f} \sum_{\mathbf{x} \in \Omega'} (2 * v_{i,m;k} - v_{i,m;k-1} - v_{i,m;k+1}) \\ &\quad \times \beta(\mathbf{x} - \mathbf{x}_i). \end{aligned} \quad (8)$$

The registration energy can be optimized by starting from initial position and descending along the negative gradient direction at each iteration until there is no significant decrease.

## 2.E. Implementation

In our experiments, we used a series of B-spline transformations with grid spacing of 10 voxels in each dimension to represent the velocity field. The values of  $\lambda$  and  $w_t$  were chosen to be 0.1 and 0.001 to allow the range of the ICE function and the regularization term to be comparable. The algorithm was implemented with Matlab<sup>67</sup> under a Windows XP 64-bit system on an eight-core 2.13 GHz Xeon CPU machine with 30 GB RAM. For 2D image sequence of 11 frames with a size of  $274 \times 192$ , it took about 10 min to estimate the motion with single core computing. For 3D sequences, the most frequently called functions such as the 3D B-spline interpolation of the displacement field and the trilinear interpolation of the image have been implemented by using OpenMP (Ref. 68) to effectively support the multicore processors. The typical execution time for a 3D sequence of 15 frames with frame size of  $137 \times 96 \times 124$  was about 2 h.

## 3. EXPERIMENTS AND DATASETS

We use simulated datasets and images of an *ex vivo* rabbit heart phantom, *in vivo* open-chest pig heart datasets, and human echocardiographic sequences to validate our method.

### 3.A. Simulated datasets

Both 2D and 3D sequences were generated in the simulated data experiments. In 2D test sequences, a long axis left

ventricle (LV) ED frame with a size of  $274 \times 192$  pixels was used as a reference frame. This frame was then deformed with a series of known continuous displacement field functions. The deformation was symmetric along the LV long axis to simulate the myocardial contraction along the radial and longitudinal directions. The displacement functions in  $x$  and  $y$  directions are defined as

$$f_x(i) = a_x \sin \frac{\pi(x - x_c)}{2r_d} \sin \left( \frac{i\pi}{N_f} \right), \quad (9)$$

$$\begin{aligned} f_y(i) &= a_y \sin \frac{\pi(y - y_{apex})}{2(y_{base} - y_{apex})} \\ &\quad \times \left( \sin \left( \frac{i\pi}{N_f} + \frac{\pi}{16} \right) - \sin \frac{\pi}{16} \right), \end{aligned} \quad (10)$$

with  $x_c$ ,  $r_d$  being the short axis center coordinate and the average LV short axis radius,  $y_{apex}$  and  $y_{base}$  being the height of base and apex,  $N_f$  and  $i$  being the number of frames and the frame index, and  $a_x$ ,  $a_y$  being the magnitudes of displacement field which are the largest shift in axial and longitudinal directions. An image sequence with  $N_f + 1$  frames (including the reference frame) was generated with  $i$  varying from 0 to  $N_f$  to simulate the motion in one cardiac cycle from ED to ES and then back to ED, where  $i = 0$  corresponds to the reference frame itself. Three sequences each with 11 frames were simulated with multiplicative noise added to each frames. The noise was sampled from a uniform distribution with zero mean and variances 0.06, 0.08, and 0.10, respectively, for each image sequence. The gray level of the original reference frame was within  $[0,1]$  and that of each simulated frame was rescaled to  $[0,1]$  after adding noise. The magnitude parameters  $a_x$  and  $a_y$  were set to be 5 and 15, respectively. The reference frame with geometrical parameters for simulation and the 6th simulated frame with noise variance 0.10 are shown together with the ground truth displacement field in Fig. 2.

In 3D test sequences, an ED volume frame was used as a reference. It was deformed with a spatiotemporal deformation, and a series of frames were generated. We assume that the LV short axis planes are parallel to  $xy$  coordinate plane and the long axis is along  $z$  direction. The displacement field functions  $f_{x,y}$  in  $xy$  plane and  $f_z$  in  $z$  direction are defined as

$$f_{x,y}(i) = a_{xy} \sin \frac{\pi |\mathbf{x}_p - \mathbf{x}_c|}{2r_d} \sin \left( \frac{i\pi}{N_f} \right) \cdot \vec{n}, \quad (11)$$

$$\begin{aligned} f_z(i) &= a_z \sin \frac{\pi(z - z_{apex})}{2(z_{base} - z_{apex})} \\ &\quad \times \left( \sin \left( \frac{i\pi}{N_f} + \frac{\pi}{16} \right) - \sin \frac{\pi}{16} \right), \end{aligned} \quad (12)$$

with  $(x, y, z)$  being the coordinate of each point,  $\mathbf{x}_p$  being the coordinate projected into  $xy$  plane,  $\mathbf{x}_c$  being the long axis projected in the  $xy$  plane,  $r_d$  being the average radius of the LV short axis plane,  $\vec{n}$  being the unit vector from the short axis plane center pointing to  $\mathbf{x}_p$ ,  $z_{apex}$  and  $z_{base}$  being the height of base and apex plane,  $N_f$  and  $i$  being the number of simulated frames and the frame index, and  $a_x$ ,  $a_y$  being the



FIG. 2. The reference frame and the 6th frame in the variance 0.1 sequence and the ground truth displacement field.

magnitudes of displacement fields which are the largest deformation in radial and longitudinal directions. Three sequences with noise variances of 0.06, 0.08, 0.10 were simulated using the same noise model as in 2D experiments. The reference volume size is  $137 \times 96 \times 124$  voxels and number of frames is 11. The magnitude parameters  $a_{xy}$  and  $a_z$  are set to 5 and 6, respectively. Figure 3 shows the reference frame, 6th simulated frame, and the corresponding displacement

fields in orthogonal slices of the sequence with noise variance of 0.1.

### 3.B. *Ex vivo* rabbit heart phantom datasets

We validated our method with real datasets from rabbit heart phantom with sonomicrometers<sup>69</sup> attached to the myocardial wall. Sonomicrometry, used as ground truth, is a

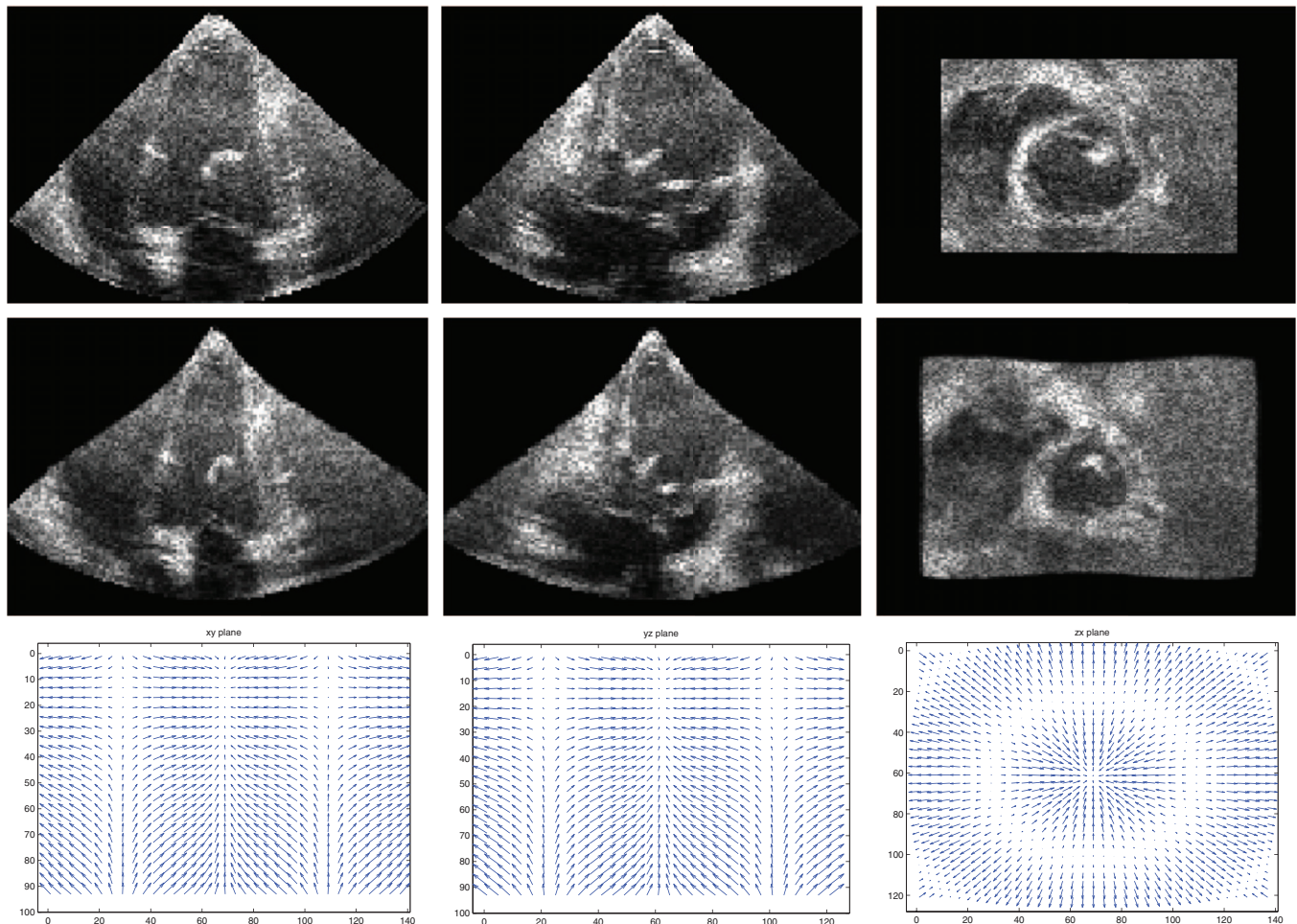


FIG. 3. From the top to bottom, the first row shows the three orthogonal view of the reference frame (ED). The second row shows those of the 6th frame (ES). The third row shows the displacement fields in the orthogonal planes between these two frames.

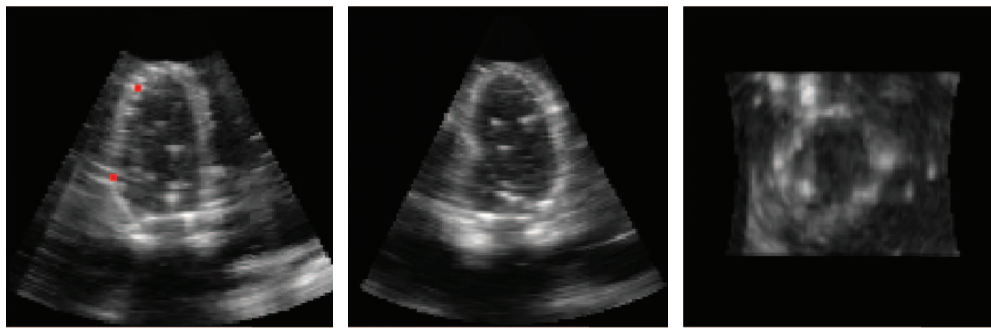


FIG. 4. The three center orthogonal views of the ED frame of the rabbit model. The two points in the long axis view shows the location of the two sonomicrometers on the phantom myocardium.

method to measure the distance between the sono (piezoelectric) crystals based on the speed of acoustic signals through the medium they are embedded in. In this test, two sono crystals were implanted on the apical and mid myocardium of the phantom (Fig. 4). Freshly harvested rabbit hearts were used. A balloon was inserted into the phantom and connected to a pump which produced systole and diastole phases periodically. The stroke volume can be varied to control the deformation of the myocardium. It varied uniformly from 1ml to 5ml with 1 ml increment. The image sequences were acquired using Philips IU22 system with an X6-1 transducer. Each sequence had ten frames with frame size  $109 \times 129 \times 129$  and voxel size  $0.57 \text{ mm}^3$ . The position of the crystals in the ED frame was manually located by a radiologist. Five sequences with varying stroke volume were acquired and the peak strain of the two crystals, which was defined as the largest relative displacement change to their initial distance within one heart cycle, were recorded with each sequence. We evaluated the two points peak strain by using our method and compared the estimated results with the sonomicrometry records.

### 3.C. *In vivo* open-chest pig heart datasets

We validated our method with image sequences of two open-chest adult pigs with sonomicrometry. The image sequences were acquired using a Philips IE33 system with an X7 cardiac probe placed directly on the cardiac apex, separated by a small piece of fresh liver as standoff (2–4 cm). Sector width and depth were adjusted to allow inclusion of LV walls within the image region. For each pig, three sequences were acquired to study the LV motion under different steady states. The first sequence was the baseline under normal conditions. The second sequence was acquired, while the inferior vena cava (IVC) was occluded by a pneumatic occluder until the preload was dropped to the ED pressure level. The third sequence was a myocardial ischemia case which was captured after the left anterior descending (LAD) coronary artery was ligated for 5–7 min. The number of frames for each sequence varied from 13 to 21 and the typical frame size was  $171 \times 155 \times 120$  with voxel size of  $0.836 \text{ mm}^3$ . The ED and ES frames of a baseline sequence are shown in Fig. 5 as an

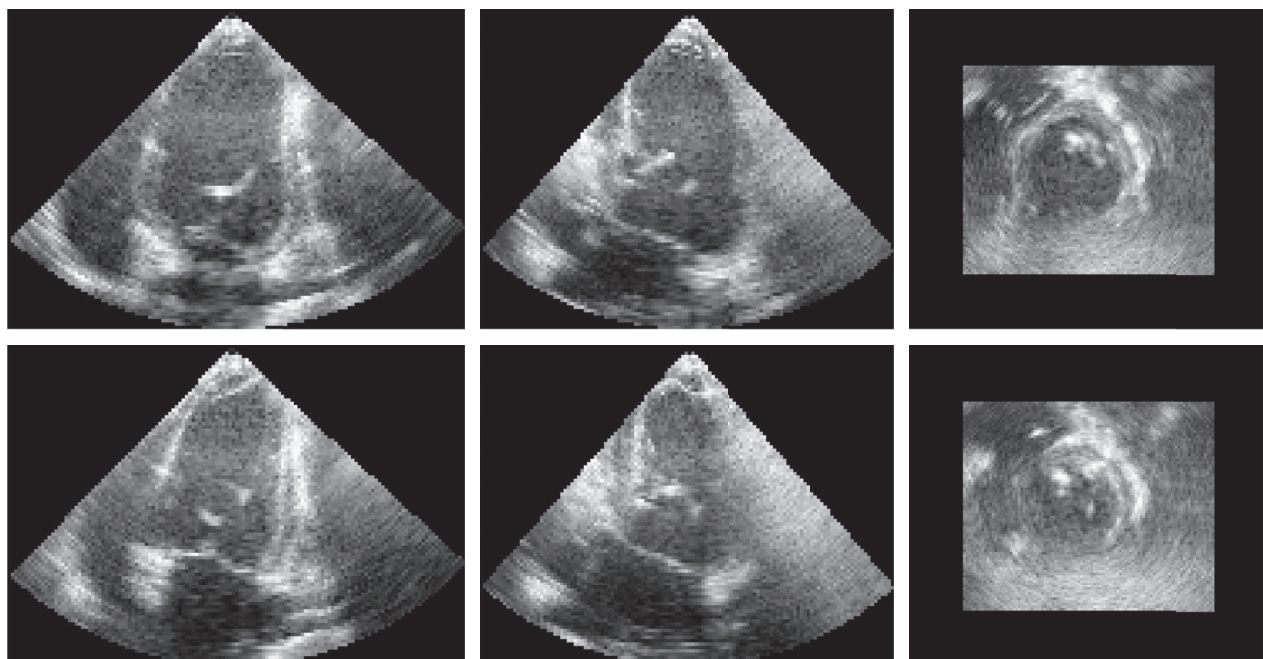


FIG. 5. The real pig baseline dataset. The top row shows the orthogonal views of the end-diastole frame. The bottom row shows those of the end-systole frame.

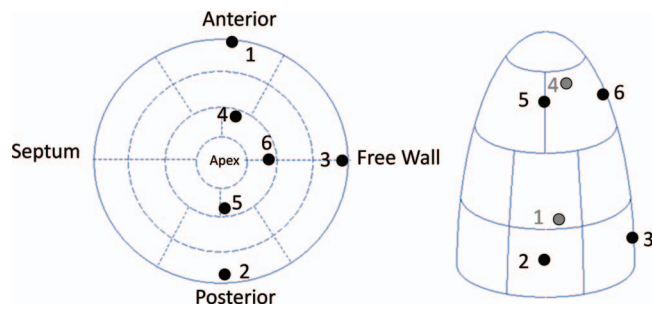


FIG. 6. The locations of six crystals on the myocardium.

example. For the first pig, we implanted six sonomicrometers in the epicardium for validation of the tracking accuracy. Atraumatic surgical and fine suturing techniques were used to fix the crystals in the epicardium with minimal damage. All the experiments were approved by the OHSU Institutional Animal Care and Use Committee. Three crystals (S1–S3) were implanted on the myocardium in the short axis plane with 20% long axis height from the base, and the other three crystals (S4–S6) were implanted in the short axis plane which was in 80% long axis height (Fig. 6). The crystal positions in the ED frames were manually located by a radiologist and their positions in other frames were propagated by the estimated transformation. The relative coordinates of the six crystals with one as the reference varied with time, and were recorded by sonomicrometry. The distance between crystals measured in the sonomicrometry was used as ground truth to compare with that estimated from the echocardiography.

### 3.D. Human datasets

We used the image sequences from an open access database for motion tracking.<sup>70</sup> The 3D echocardiography sequences of 15 healthy subjects were acquired by using an IE33 system with an X3-1 transducer. The typical frame size was  $208 \times 224 \times 208$  with voxel size  $0.82 \times 0.82 \times 0.72$  mm and the number of frames varied from 11 to 24. For each subject, a 3D tagged MRI sequence was also acquired by using a 3T Philips Achieva system. Twelve landmarks were tracked in the tagged MRI sequences by two expert observers. The landmarks in the tagged MRI were then transformed into the 3D echocardiography coordinate by using point based registration (Fig. 7). The transformed tagged MRI landmarks were

then used as ground truth to validate the estimated motion from echocardiography. Fifteen subjects were used for our validation (V1–V16, V3 was not included).

## 4. RESULTS

We compared our method with two recently developed diffeomorphic motion estimation methods. The first one is a diffeomorphic B-spline method with transformation transitivity as a temporal constraint. The constraint ensures that in each three consecutive frames, the composition of the transformations obtained from each two consecutive frame registrations is equal to the transformation from the first frame to the third one.<sup>51</sup> We refer to it as the transformation transitivity constraint (TTC) method. The second one is the temporal diffeomorphic free form deformation (TDFFD) method, which optimizes a spatiotemporal B-spline velocity field function by minimizing the SSD between the reference to each of the unwarped subsequent frames.<sup>58</sup> The same regularization parameters were used for our method and the TTC method. For the TDFFD method, identical spatial spacing of B-spline was used in the 4D velocity field, and the temporal spacing was equal to the time step between two consecutive frames. All three methods were compared in the simulated datasets. In the rabbit heart dataset, we compared all the three methods with sonomicrometry. In the open-chest pig and human volunteer experiments, we compared our method with the TDFFD method since they were found to have closer performance.

### 4.A. Simulated datasets

We compared our method with TTC and TDFFD methods by tracking the trajectories of the points on the myocardium during the motion process. For 2D sequences, we used a region of interest in the myocardium as an example to show the tracking results. The region of interest is displayed as a green rectangle in the left figure of Fig. 2. Figure 8 shows the 11-frame trajectories of nine tracked points in this region for the sequence with noise variance of 0.06. The ground truth trajectories are overlaid for comparison. We can see, generally, that coordinates of the points in each time step in our method are closer to the ground truth. Larger errors can be seen in the right part of the trajectories in the TTC method. In the TDFFD

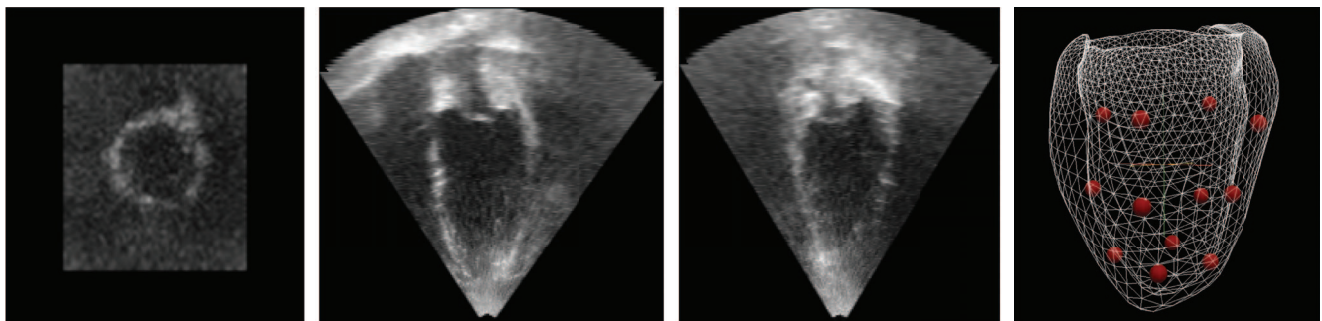


FIG. 7. The three orthogonal view of ED frame of one subject and the 12 landmark positions from one expert observer.



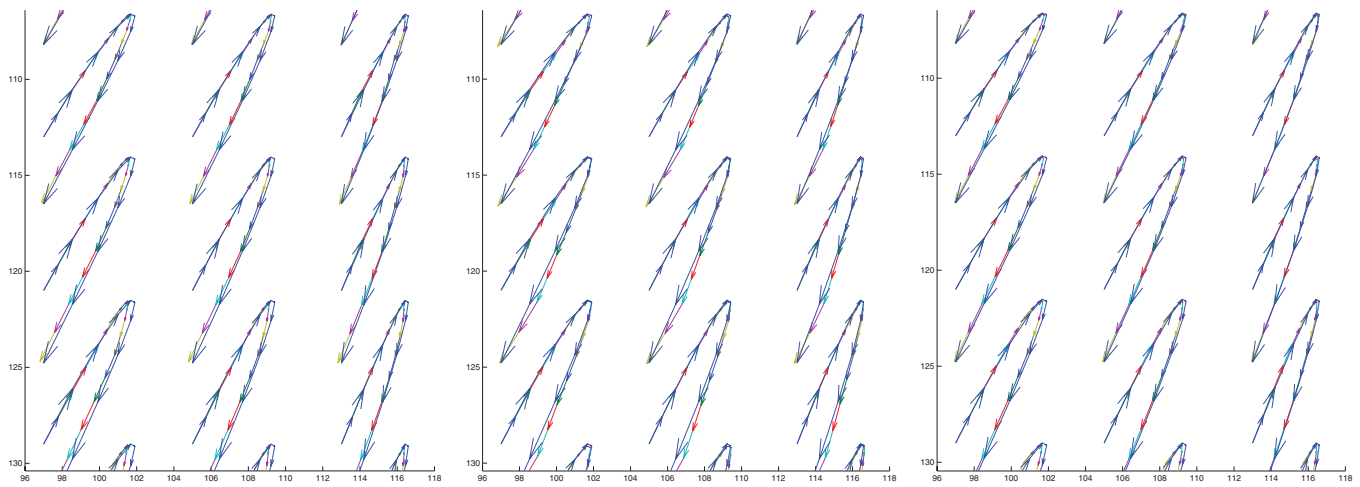


FIG. 8. Trajectories of nine points in 11 frames in TTC method (left), TDDFD method (middle), and our method (right) (A rectangle shows the bounding box of these nine points in Fig. 2.) The ground truth trajectories are overlaid with the estimated curves for comparison. The arrow shows the velocity at each time step.

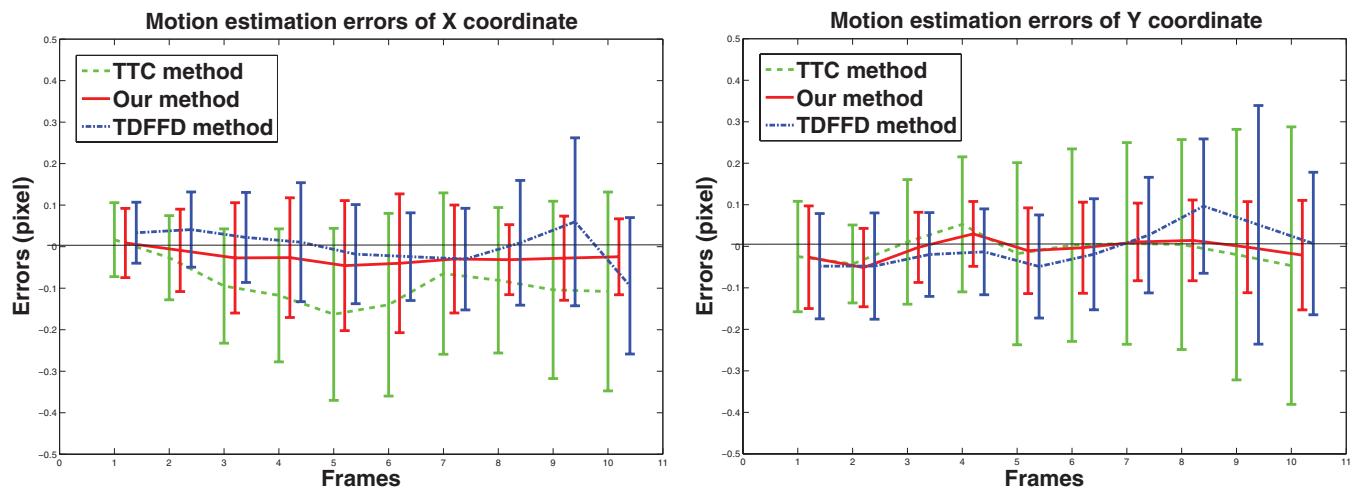


FIG. 9. The motion estimation errors (mean and std unit in pixel) in  $x$  (left) and  $y$  (right) coordinates in the TTC method (dashed line), TDDFD method (dash-dot line), and our method (solid line).

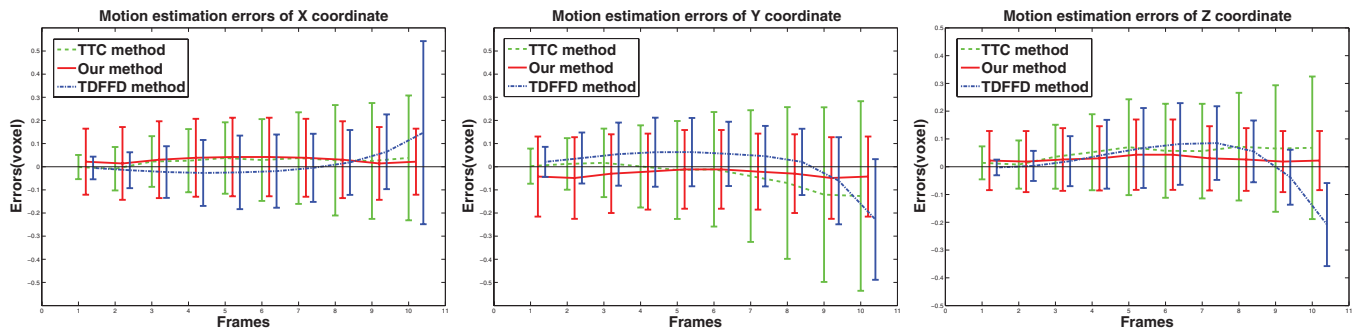


FIG. 10. The transformation estimation errors (mean and std unit in voxel) of  $x$  (left),  $y$  (middle), and  $z$  (right) coordinates in the TTC method (dashed line), TDDFD method (dash-dot line), and our method (solid line) in each frame.

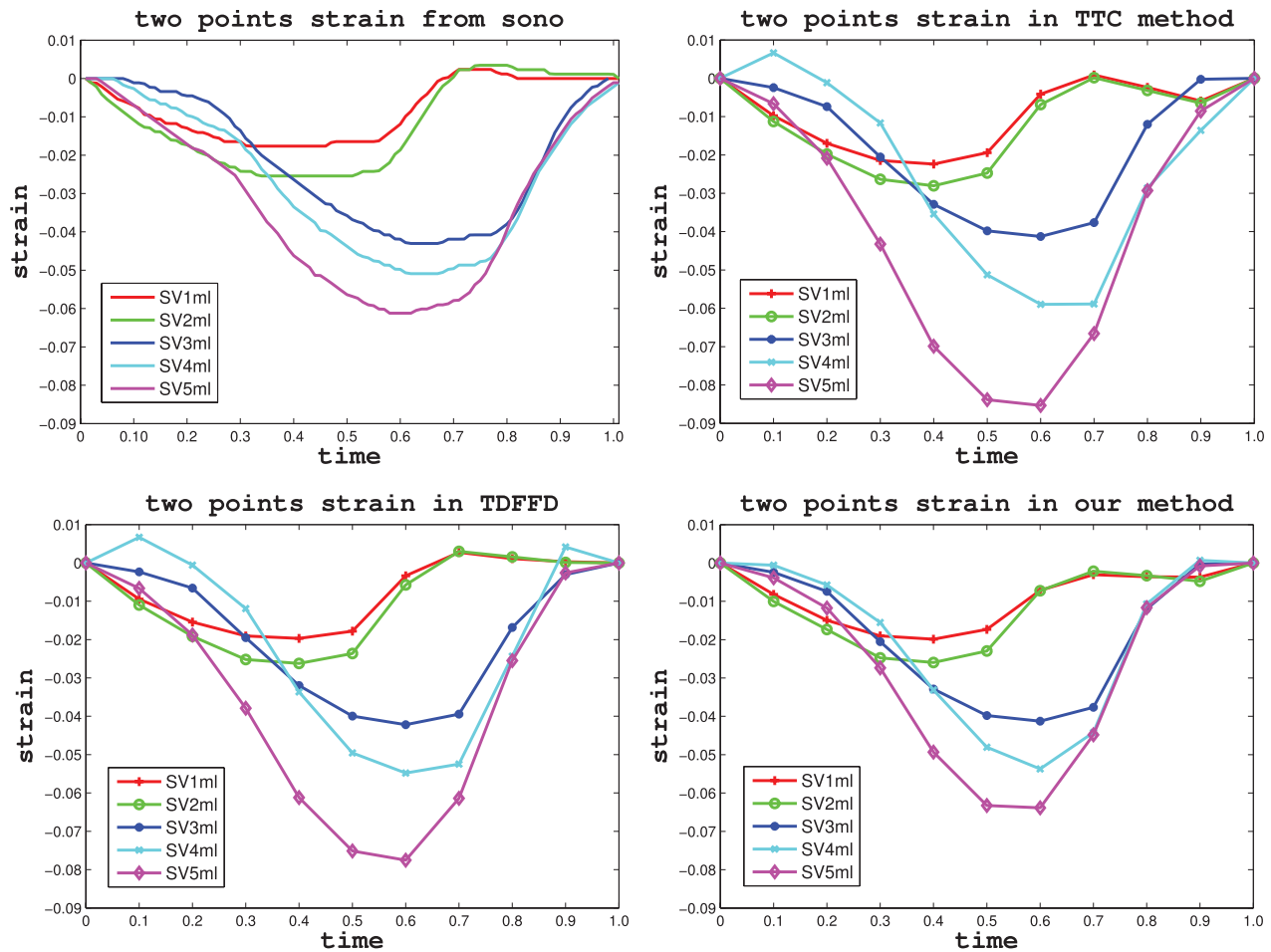


FIG. 11. Two point strain varying with time under different stroke volume from sonomicrometry (top left), TTC method (top right), TDFFD method (bottom left), and our method (bottom right).

method, larger errors appear in time steps near the end of the trajectories. We use the SSD between the estimated displacement fields and the ground truth in each frame to evaluate the transformation estimation errors. We show an example in Fig. 9 of the errors in both  $x$  and  $y$  coordinates in the sequence with noise variance of 0.08. We can see both the mean and

variance in  $x$  and  $y$  directions in our method are smaller than the other two methods.

We calculated the mean magnitude of transformation errors in the sequences of three noise variance using the three methods. In our method, the errors were 0.225, 0.267, and 0.312, while in the TTC method and the TDFFD method

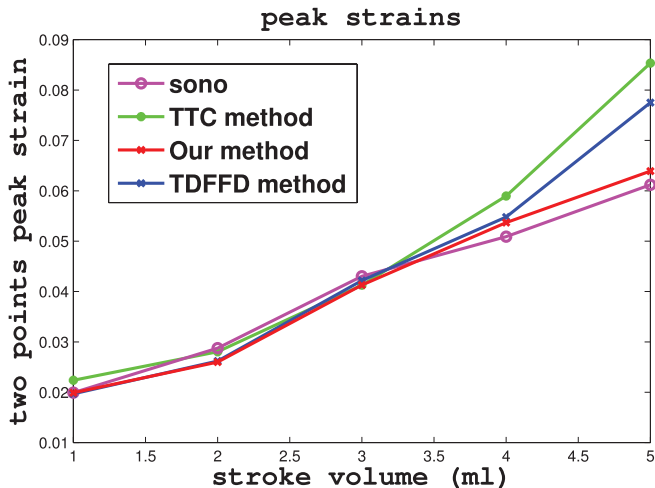


FIG. 12. The two point peak strains under different stroke volume estimated from sonomicrometry, TTC method, TDFFD method, and our method.

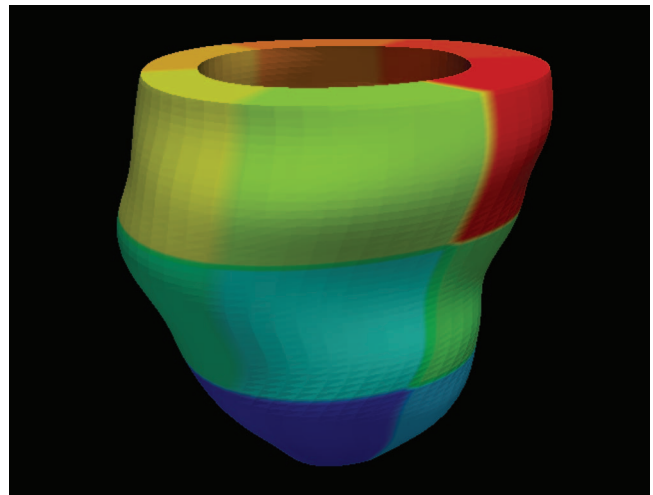


FIG. 13. An illustration of the 16 AHA segments.

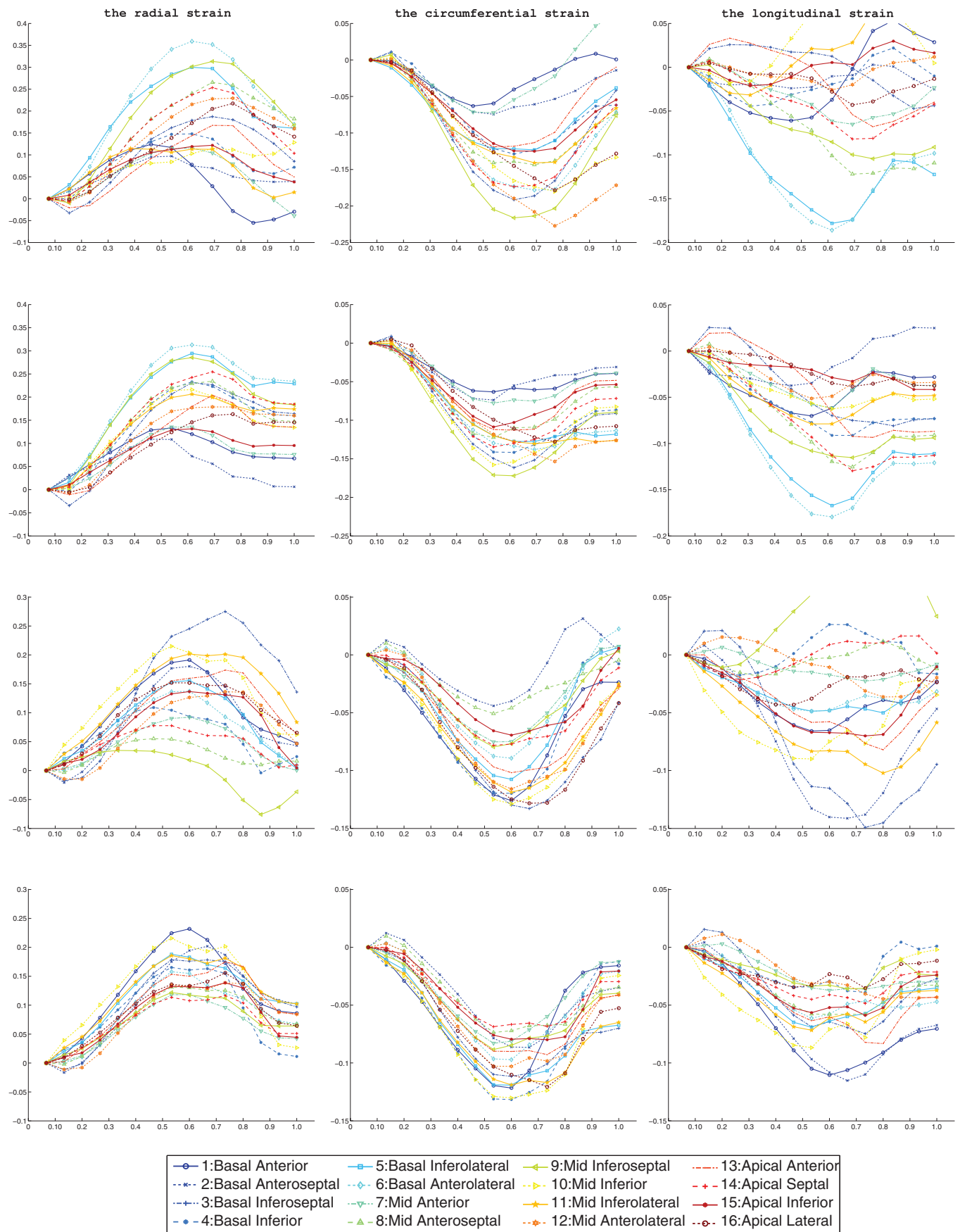


FIG. 14. The radial, circumferential, and longitudinal strains of the 16 AHA segments in the TDFFD method (first and third rows) and our method (second and fourth rows) estimated from two baseline sequences. The horizontal axis is the normalized cardiac time, the vertical axis is the strain. The markers used for the 16 AHA segments are shown in the bottom of the figure.

TABLE I. The correlations between the estimated pairwise distances and those from the sonomicrometry, with TDDFD method (numbers to the left), and our method (numbers to the right) in the baseline sequence. Numbers S1–S6 index the six sonomicrometry markers.

	S1	S2	S3	S4	S5	S6
S1	1.0/1.0	0.807/0.707	0.710/0.968	0.834/0.970	0.802/0.881	0.950/0.967
S2	0.807/0.707	1.0/1.0	0.780/0.726	0.946/0.979	0.823/0.929	0.907/0.828
S3	0.710/0.968	0.780/0.726	1.0/1.0	0.928/0.988	0.852/0.928	0.960/0.978
S4	0.834/0.970	0.946/0.979	0.928/0.988	1.0/1.0	0.989/0.981	0.855/0.945
S5	0.802/0.881	0.823/0.929	0.852/0.928	0.989/0.981	1.0/1.0	0.625/0.727
S6	0.950/0.967	0.907/0.828	0.960/0.978	0.855/0.945	0.625/0.727	1.0/1.0

they were 0.329, 0.365, 0.403, and 0.314, 0.347, 0.392, respectively.

For 3D datasets, we show in Fig. 10 the errors in  $x$ ,  $y$ , and  $z$  coordinates, for the sequence of noise variance 0.06. We can see that the variance of the error becomes larger toward the end of the sequence in the TTC method. In the TDDFD method, the means of the transformation estimation errors in three coordinates increase, which means the results were biased toward the reference frame. In our method, the mean and variance of the transformation estimation error change smoothly with the frame number. The average transformation estimation errors at three noise levels were 0.255, 0.274, 0.293 in our method and 0.322, 0.354, 0.391 in the TTC method and 0.279, 0.301, 0.325 in the TDDFD method.

#### 4.B. *Ex vivo* rabbit phantom datasets

In the rabbit phantom test, the two point strain of the crystals in five echocardiography sequences with stroke volume uniformly varying from 1 ml to 5 ml were estimated and compared with those of the sonomicrometry. Figure 11 shows the two point strains varying in a cardiac cycle under different stroke volumes by sonomicrometry, TTC, TDDFD, and our method. We can see that two point strains curve from our method were closer to the ground truth of sonomicrometry than were those from the TTC and TDDFD methods. Figure 12 shows the estimated two point peak strains varying with the stroke volume in sonomicrometry, TTC, TDDFD, and our method. We can also see that our method has closer two point peak strains with sonomicrometry, especially in the sequence with stroke volume 5 ml.

#### 4.C. *In vivo* open-chest pig heart datasets

For the *in vivo* open-chest pig heart datasets, we compared strain derived from our method with that from the TDDFD

method. We use the Green-Lagrangian strain<sup>71</sup> which measures how much the displacement function varies locally in a given direction. The myocardium strains were evaluated along the radial, circumferential, and longitudinal directions by using the transformation obtained.<sup>6</sup> We developed a program which can semiautomatically label the endocardium and epicardium contours in the short axis image planes between the mitral valve and the apex. For each contour, a closed curve parameterized by cubic B-spline was fitted to the manually picked boundary points around the myocardium. The contours were then resampled uniformly by a constant number of points along them. The endocardium and the epicardium surfaces were generated by connecting all their contours using triangles. Each vertex in both endocardium and epicardium was then assigned a segment attribute. We used the AHA myocardium segments convention<sup>72</sup> and the strains of 16 myocardium segments were evaluated as the average strain in each segment. The 17th (apex) segment was excluded because the directions were usually unstable. An example for the 16 AHA segments is shown in Fig. 13. The radial, circumferential, and longitudinal strains of the two baseline sequences are shown in Fig. 14. From the strain plots, we can see the motion consistency between segments is improved in our method. In the TDDFD method, the estimated strains after ES are more divergent. This is due to the larger intensity variations between the reference image with distant frames. As a result, the displacement fields obtained have larger estimation errors.

We compared the performance of the two algorithms by computing the correlations between the time varying functions of estimated pairwise distances from echo and those from sonomicrometry. The results for the baseline, IVC occlusion, and LAD ligation sequences are shown in Tables I, II, and III respectively. Numbers S1–S6 are the indices of the sono crystals. Each table entry shows the

TABLE II. The correlations between the estimated pairwise distances and those from the sonomicrometry, with TDDFD method (numbers to the left) and our method (numbers to the right) in the IVC occlusion sequence. Numbers S1–S6 index the six sonomicrometry markers.

	S1	S2	S3	S4	S5	S6
S1	1.0/1.0	0.857/0.955	0.874/0.841	0.853/0.928	0.759/0.891	0.906/0.880
S2	0.857/0.955	1.0/1.0	0.727/0.810	0.946/0.976	0.913/0.927	0.817/0.876
S3	0.874/0.841	0.727/0.810	1.0/1.0	0.932/0.933	0.856/0.889	0.890/0.893
S4	0.853/0.928	0.946/0.976	0.932/0.933	1.0/1.0	0.917/0.898	0.921/0.981
S5	0.759/0.891	0.913/0.927	0.856/0.889	0.917/0.898	1.0/1.0	0.760/0.732
S6	0.906/0.880	0.817/0.876	0.890/0.893	0.921/0.981	0.760/0.732	1.0/1.0

TABLE III. The correlations between the estimated pairwise distances and those from the sonomicrometry, with TDDFD method (numbers to the left), and our method (numbers to the right) in the LAD ligation sequence. Numbers S1–S6 index the six sonomicrometry markers.

	S1	S2	S3	S4	S5	S6
S1	1.0/1.0	0.972/0.975	0.967/0.977	0.731/0.870	0.987/0.986	0.979/0.960
S2	0.972/0.975	1.0/1.0	0.826/0.882	0.957/0.976	0.640/0.824	0.571/0.696
S3	0.967/0.977	0.826/0.882	1.0/1.0	0.963/0.981	0.829/0.935	0.630/0.901
S4	0.731/0.870	0.957/0.976	0.963/0.981	1.0/1.0	0.976/0.981	0.979/0.986
S5	0.987/0.986	0.640/0.824	0.829/0.935	0.976/0.981	1.0/1.0	0.926/0.919
S6	0.979/0.960	0.571/0.696	0.630/0.901	0.979/0.986	0.926/0.919	1.0/1.0

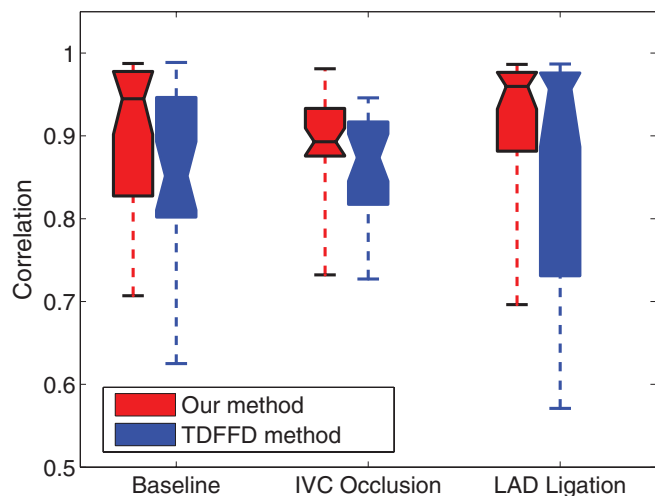


FIG. 15. The box plots of the correlation values in our method and the TDDFD method for the baseline, IVC occlusion, and LAD ligating cases. The bottom and top box boundary and the center line show the first quartile, third quartile, and median, respectively. The ends of the whiskers show the minimum and maximum of each case. For each dataset in each method 30 correlation values are averaged, except the six diagonal correlation value which is one.

correlation of the time function of pairwise crystal distance from sonomicrometry with that estimated from echo in the TDDFD (left side) and our method (right side). We can see that most of the correlation values are improved in our method in all three steady states. The box plots of the correlation values in the proposed and TDDFD methods for the three steady states are shown in Fig. 15. The increase of the average correlations in our method than TDDFD method is more than 5% for our datasets.

#### 4.D. Human datasets

We used the landmarks labeled by one observer as ground truth and the target registration error<sup>73</sup> was used for the validation of tracking errors. The target registration error of the 12 landmarks in the ES and the last frames were evaluated in our experiments.<sup>74,75</sup> We compared our method with the modified TDDFD method which used the weighted first-to-following and frame-to-frame similarity measurements. We also evaluated the results from a frame-to-frame method for comparison. The tracking errors in the three methods are shown in Fig. 16. Our method generally has smaller average tracking

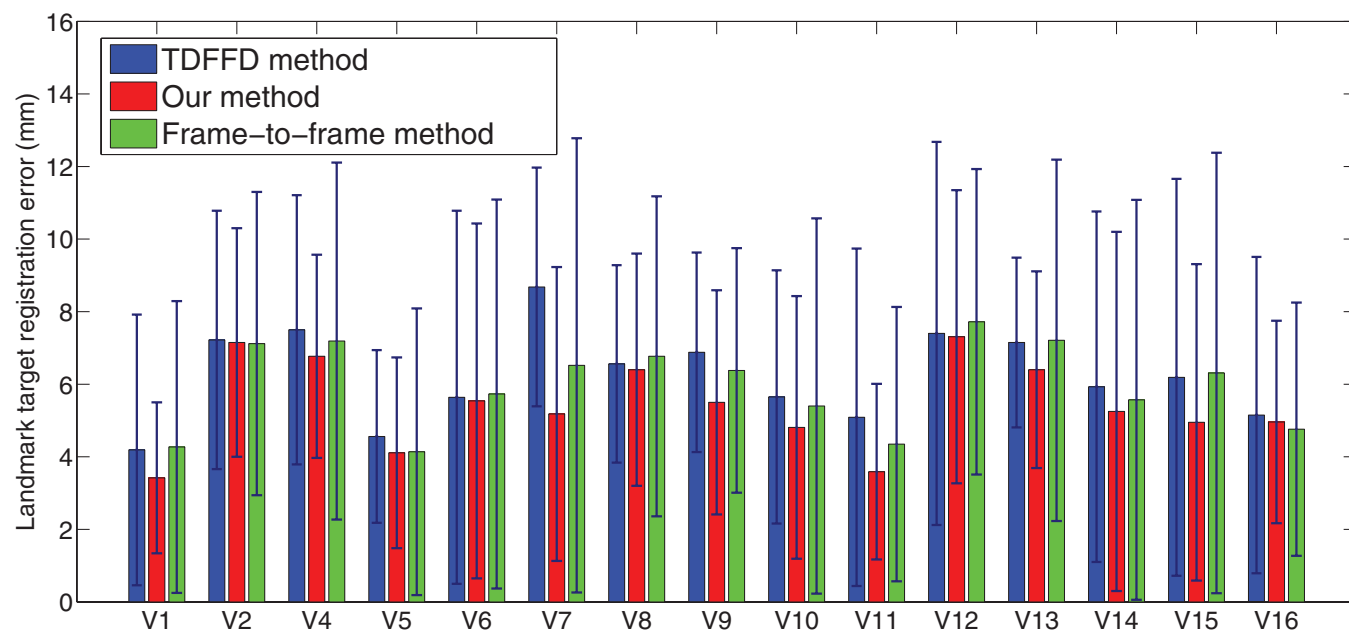


FIG. 16. The tracking error (mean and std) of the seven subjects in the TDDFD method, our method, and the frame-to-frame method.

error and variance than the TDFFD method and the frame-to-frame method.

## 5. DISCUSSION AND CONCLUSION

We proposed a diffeomorphic motion estimation method with temporal smoothness by constraining the velocity field to have maximum local intensity consistency within multiple image frames. Simulated and real cardiac sequences tests showed that results using our method are more accurate than other competing temporal diffeomorphic methods. Tests with sonomicrometry showed that the tracked crystal positions have good consistency with the ground truth and our method has higher accuracy than the TDFFD method. Validation with an open access human cardiac dataset showed that our method has smaller feature tracking errors than both TDFFD and frame-to-frame methods. From the human cardiac dataset test, we see that even when a weighted frame-to-frame and first-to-following similarity measurement is used, the dissimilarity between distant frames may cause larger feature tracking errors.

Our method extends the frame-to-frame motion estimation method to multiple frames. With more frames included into the similarity measurement, the similarity is more robust against noise than the frame-to-frame method. Furthermore, the ICE similarity measurement is defined on multiple time points between four consecutive frames, to allow the four frames unwarped to this time point to be as similar as possible. At each time point, the estimated spatiotemporal transformation maximizes the similarity of the four consecutive frames around it. The transformation is consistent with two frames before and after it while preserving good temporal smoothness. This introduces more intensity consistency for the spatiotemporal transformation than the frame-to-frame method and the TDFFD method.

One factor to consider is the number of consecutive frames to include in ICE. This will vary for different types of data, and will depend on several factors, such as the level of decorrelation from frame to frame and computation load. Inclusion of distant frames introduces uncorrelated intensity into similarity measurements. If only two frames are considered (the frame-to-frame case), the transformation errors tend to accumulate faster because the transformation parameters are only estimated by optimizing the similarity of two consecutive frames, instead of optimizing a groupwise similarity measurement. In this sense, more frames will be beneficial. In addition, more frames will lead to better temporal consistency. However, inclusion of more frames will increase the computation time. In our test, we use four frames as a compromise of these factors.

Our future research topics include the study of region-based intensity consistency between consecutive frames, and the design of improved groupwise similarity measurement.

## ACKNOWLEDGMENTS

This paper is supported by a NIH/NHLBI Grant No. 1R01HL102407-01 awarded to Xubo Song and David Sahn.

<sup>a)</sup> Author to whom correspondence should be addressed. Electronic mail: zjzhang@gmail.com

<sup>1</sup> L. M. Sherry, J. Q. Xu, and D. K. Kenneth, "Deaths: Final data for 2010," National Vital Statistics Reports No. 61, 2013.

<sup>2</sup> B. Bijnens, P. Claus, F. Weidemann, J. Strotmann, and G. R. Sutherland, "Investigating cardiac function using motion and deformation analysis in the setting of coronary artery disease," *Circulation* **116**, 2453–2464 (2007).

<sup>3</sup> G. Buckberg, J. I. E. Hoffman, A. Mahajan, S. Saleh, and C. Coghlan, "Cardiac mechanics revisited the relationship of cardiac architecture to ventricular function," *Circulation* **118**, 2571–2587 (2008).

<sup>4</sup> L. Axel, "Biomechanical dynamics of the heart with MRI," *Annu. Rev. Biomed. Eng.* **4**, 321–347 (2002).

<sup>5</sup> G. Kiss, D. Barbosa, K. Hristova, J. Crosby, F. Orderud, P. Claus, B. Amundsen, D. Loeckx, J. D'hooge, and H. G. Torp, "Assessment of regional myocardial function using 3D cardiac strain estimation: Comparison against conventional echocardiographic assessment," in *IEEE International Ultrasonics Symposium (IUS)* (IEEE, Rome, 2009), pp. 507–510.

<sup>6</sup> A. Elen, H. F. Choi, D. Loeckx, H. Gao, P. Claus, P. Suetens, F. Maes, and J. D'hooge, "Three-dimensional cardiac strain estimation using spatio-temporal elastic registration of ultrasound images: A feasibility study," *IEEE Trans. Med. Imaging* **27**(11), 1580–1591 (2008).

<sup>7</sup> T. Mäkelä, P. Clarysse, O. Sipilä, N. Pauna, Q. C. Pham, T. Katila, and I. E. Magnin, "A review of cardiac image registration methods," *IEEE Trans. Med. Imaging* **21**(9), 1011–1021 (2002).

<sup>8</sup> M. Bertini, P. P. Sengupta, G. Nucifora, V. Delgado, C. T. Ng, N. A. Marsan, M. Shanks, B. R. J. van, M. J. Schaliq, J. Narula, and J. J. Bax, "Role of left ventricular twist mechanics in the assessment of cardiac dyssynchrony in heart failure," *JACC* **2**(12), 1425–1435 (2009).

<sup>9</sup> R. M. Lang, S. A. Goldstein, I. Kronzon, and B. K. Khandheria, *Dynamic Echocardiography* (Elsevier, St. Louis, MO, 2010).

<sup>10</sup> L. Axel, A. Montillo, and D. Kim, "Tagged magnetic resonance imaging of the heart: A survey," *Med. Imag. Anal.* **9**(4), 376–393 (2005).

<sup>11</sup> B. M. Ohnesorge, C. R. Becker, T. G. Flohr, and M. F. Reiser, *Multi-Slice CT in Cardiac Imaging: Technical Principles, Clinical Applications and Future Developments* (Springer, New York, 2002).

<sup>12</sup> S. Mondillo, M. Galderisi, D. Mele, M. Cameli, V. S. Lomoriello, V. Zaca, P. Ballo, A. D. Andrea, D. Muraru, M. Losi, E. Agricola, A. D. 'Errico, S. Buralli, S. Sciomer, S. Nistri, and L. Badano, "Speckle-tracking echocardiography a new technique for assessing myocardial function," *J. Ultrasound Med.* **30**(1), 71–83 (2011).

<sup>13</sup> H. Geyer, G. Caracciolo, H. Abe, S. Wilansky, S. Carerj, F. Gentile, H. J. Nesser, B. Khandheria, J. Narula, and P. P. Sengupta, "Assessment of myocardial mechanics using speckle tracking echocardiography: Fundamentals and clinical applications," *J. Am. Soc. Echocardiogr.* **23**(4), 351–369 (2010).

<sup>14</sup> C. Y. Ho and S. D. Solomon, "A clinician's guide to tissue doppler imaging," *Circulation* **113**, e396–e398 (2006).

<sup>15</sup> J. Park, D. Metaxas, A. A. Young, and L. Axel, "Deformable models with parameter functions for cardiac motion analysis from tagged MRI data," *IEEE Trans. Med. Imaging* **15**(3), 278–289 (1996).

<sup>16</sup> J. Park, D. N. Metaxas, and L. Axel, "Analysis of left ventricular wall motion based on volumetric deformable models and MRI SPAMM," *Med. Image Anal.* **1**(1), 53–72 (1996).

<sup>17</sup> X. Papademetris, A. J. Sinusas, D. P. Dione, and J. S. Duncan, "Estimation of 3D left ventricular deformation from echocardiography," *Med. Imag. Anal.* **5**(1), 17–28 (2001).

<sup>18</sup> A. F. Frangi, W. J. Niessen, and M. A. Viergever, "Three-dimensional modeling for functional analysis of cardiac images: A review," *IEEE Trans. Med. Imaging* **20**(1), 2–25 (2001).

<sup>19</sup> D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Trans. Image Process.* **18**(8), 712–721 (1999).

<sup>20</sup> J. P. Thirion, "Image matching as a diffusion process: An analogy with Maxwell's demons," *Med. Imag. Anal.* **2**(3), 243–260 (1998).

<sup>21</sup> P. Cachier, E. Bardinet, D. Dormont, X. Pennec, and N. Ayache, "Iconic feature based nonrigid registration: The PASHA algorithm," *J. Comput. Vis. Image Understand.* **89**(2–3), 272–298 (2003).

<sup>22</sup> S. S. Beauchemin and J. L. Barron, "The computation of optical flow," *J. ACM Comput. Surv.* **27**(3), 433–466 (1995).

<sup>23</sup> G. E. Christensen, R. D. Rabbit, and M. I. Miller, "Deformable templates using large deformation kinematics," *IEEE Trans. Image Process.* **5**(10), 1435–1447 (1996).

- <sup>24</sup>G. Hermosillo, C. Chef'd'Hotel, and O. Faugeras, "Variational methods for multimodal image matching," *Int. J. Comput. Vis.* **50**(3), 329–343 (2002).
- <sup>25</sup>A. Myronenko, X. B. Song, and J. S. David, "LV motion tracking from 3D echocardiography using textural and structural information," in *Proceedings of MICCAI*, Lecture Notes in Computer Science Vol. 4792 (Springer, Heidelberg, 2007), pp. 428–435.
- <sup>26</sup>R. Chandrashekar, R. H. Mohiaddin, and D. Rueckert, "Analysis of 3-D myocardial motion in tagged MR images using nonrigid image registration," *IEEE Trans. Med. Imaging* **23**(10), 1245–1250 (2004).
- <sup>27</sup>T. Mansi, J. M. Peyrat, M. Sermesant, H. Delingette, J. Blanc, Y. Boujemline, and N. Ayache, "Physically-constrained diffeomorphic demons for the estimation of 3D myocardium strain from cine-MRI," in *Proceedings of Functional Imaging and Modeling of the Heart*, Lecture Notes in Computer Science Vol. 5528 (Springer, Heidelberg, 2009), pp. 201–210.
- <sup>28</sup>M. Sühling, M. Arigovindan, C. Jansen, P. Hunziker, and M. Unser, "Myocardial motion analysis from B-mode echocardiograms," *IEEE Trans. Image Process.* **14**(4), 525–536 (2005).
- <sup>29</sup>N. Rougon, C. Petitjean, F. Preteux, P. Cluzel, and P. Grenier, "A non-rigid registration approach for quantifying myocardial contraction in tagged MRI using generalized information measures," *Med. Image Anal.* **9**(4), 353–375 (2005).
- <sup>30</sup>V. Boldea, G. C. Sharp, S. B. Jiang, and D. Sarrut, "4D-CT lung motion estimation with deformable registration: Quantification of motion nonlinearity and hysteresis," *Med. Phys.* **35**(3), 1008–1018 (2008).
- <sup>31</sup>J. M. Reinhardt, K. Ding, K. L. Cao, G. E. Christensen, E. A. Hoffman, and S. V. Bodas, "Registration-based estimates of local lung tissue expansion compared to xenon CT measures of specific ventilation," *Med. Image Anal.* **12**(6), 752–763 (2008).
- <sup>32</sup>J. Huang, D. Abendschein, V. G. Dávila-Román, and A. A. Amini, "Spatio-temporal tracking of myocardial deformation with a 4D B-spline model from tagged MRI," *IEEE Trans. Med. Imaging* **18**(10), 957–972 (1999).
- <sup>33</sup>E. Bardinet, L. D. Cohen, and N. Ayache, "Dynamic 3D models with global and local deformations: Deformable superquadrics," *IEEE Trans. Pattern Anal. Mach. Intell.* **13**(9), 703–714 (1991).
- <sup>34</sup>E. Bardinet, L. D. Cohen, and N. Ayache, "Tracking and motion analysis of the left ventricle with deformable superquadrics," *Med. Imag. Anal.* **1**(2), 129–149 (1996).
- <sup>35</sup>J. Montagnat and H. Delingette, "4D deformable models with temporal constraints: Application to 4D cardiac image segmentation," *Med. Image Anal.* **9**(1), 87–100 (2005).
- <sup>36</sup>J. Schaerera, C. Casta, J. Pousina, and P. Claryssea, "A dynamic elastic model for segmentation and tracking of the heart in MR image sequences," *Med. Image Anal.* **14**(6), 738–749 (2010).
- <sup>37</sup>O. Gerard, A. C. Billon, J. M. Rouet, M. Jacob, M. Fradkin, and C. Al-louche, "Efficient model-based quantification of the left ventricular function in 3D echocardiography," *IEEE Trans. Med. Imaging* **21**(9), 1059–1068 (2002).
- <sup>38</sup>C. T. Metz, S. Klein, M. Schaap, T. Walsum, and W. J. Niessen, "Non-rigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach," *Med. Imag. Anal.* **15**(2), 238–249 (2011).
- <sup>39</sup>Y. Wang, B. Georgescu, H. Houle, and D. Comaniciu, "Volumetric myocardial mechanics from 3D+t ultrasound data with multi-model tracking," in *Proceedings of MICCAI Workshop of STACOM*, Lecture Notes in Computer Science Vol. 6364 (Springer, Heidelberg, 2010), pp. 184–193.
- <sup>40</sup>D. Comaniciu, X. Zhou, and S. Krishnan, "Robust real-time tracking of myocardial border: An information fusion approach," *IEEE Trans. Med. Imaging* **23**(7), 849–860 (2004).
- <sup>41</sup>J. C. McEachen, A. Nehorai, and J. S. Duncan, "Multiframe temporal estimation of cardiac nonrigid motion," *IEEE Trans. Image Process.* **9**(4), 651–665 (2000).
- <sup>42</sup>M. J. Ledesma-carbayo, K. Jan, D. Manuel, S. Andrés, M. Sühling, P. R. Hunziker, and M. Unser, "Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation," *IEEE Trans. Med. Imaging* **24**(9), 1113–1126 (2005).
- <sup>43</sup>M. Yigitsoy, C. Wachinger, and N. Navab, "Temporal groupwise registration for motion modeling," in *Proceedings of Information Processing in Medical Imaging*, Lecture Notes in Computer Science Vol. 6801 (Springer, Heidelberg, 2011), pp. 648–659.
- <sup>44</sup>B. Delhay, P. Clarysse, and I. E. Magnin, "Locally adapted spatio-temporal deformation model for dense motion estimation in periodic cardiac image sequences," in *Proceedings of Functional Imaging and Modeling of the Heart*, Lecture Notes in Computer Science Vol. 4466 (Springer, Heidelberg, 2007), pp. 393–402.
- <sup>45</sup>J. M. Peyat, H. Delingette, M. Sermesant, C. Y. Xu, and N. Ayache, "Registration of 4D cardiac CT sequences under trajectory constraints with multi-channel diffeomorphic demons," *IEEE Trans. Med. Imaging* **29**(7), 1351–1368 (2010).
- <sup>46</sup>D. Perperidis, R. H. Mohiaddin, and D. Rueckert, "Spatio-temporal free-form registration of cardiac MR image sequences," *Med. Image Anal.* **9**(5), 441–456 (2005).
- <sup>47</sup>E. Castillo, R. Castillo, J. Martinez, M. Shenoy, and T. Guerrero, "Four-dimensional deformable image registration using trajectory modeling," *Phys. Med. Biol.* **55**(1), 305–327 (2010).
- <sup>48</sup>J. Vandemeulebroucke, S. Rit, J. Kybic, P. Clarysse, and D. Sarrut, "Spatiotemporal motion estimation for respiratory-correlated imaging of the lungs," *Med. Phys.* **38**(1), 166–178 (2011).
- <sup>49</sup>H. Sundar, H. Littb, and D. G. Shen, "Estimating myocardial motion by 4D image warping," *Pattern Recog.* **42**(11), 2514–2526 (2009).
- <sup>50</sup>G. J. Klein and R. H. Huesman, "Four-dimensional processing of deformable cardiac PET data," *Med. Image Anal.* **6**(1), 29–46 (2002).
- <sup>51</sup>O. Skrinjar, A. Bistoquet, and H. Tagare, "Symmetric and transitive registration of image sequences," *Int. J. Biomed. Imaging* Hindawi, Cairo (2008).
- <sup>52</sup>M. F. Beg, M. I. Miller, A. Trouve, and L. Younes, "Computing large deformation metric mappings via geodesic flows of diffeomorphisms," *Int. J. Comput. Vis.* **61**(2), 139–157 (2005).
- <sup>53</sup>T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, "Diffeomorphic demons: Efficient non-parametric image registration," *NeuroImage* **45**(S1), S61–S72 (2009).
- <sup>54</sup>A. Q. Qiu, M. Albert, L. Younes, and M. I. Miller, "Time sequence diffeomorphic metric mapping and parallel transport track time-dependent shape changes," *NeuroImage* **45**(S1), S51–S60 (2009).
- <sup>55</sup>S. Durrleman, X. Pennec, A. Trouv, J. Braga, G. Gerig, and N. Ayache, "Toward a comprehensive framework for the spatiotemporal statistical analysis of longitudinal shape data," *Int. J. Comput. Vis.* **103**(1), 22–59 (2013).
- <sup>56</sup>A. R. Khan and M. F. Beg, "Representation of time-varying shapes in the large deformation diffeomorphic framework," in *Proceedings of the International Symposium on Biomedical Imaging* (IEEE, Paris, 2008), pp. 1521–1524.
- <sup>57</sup>M. De Craene, O. Camara, B. H. Bijnens, and A. F. Frangi, "Large diffeomorphic FFD registration for motion and strain quantification from 3D-US sequences," in *Proceedings of Functional Imaging and Modeling of the Heart*, Lecture Notes in Computer Science Vol. 5528, 437–446 (2009).
- <sup>58</sup>M. De Craene, G. Piella, O. Camara, N. Duchateau, E. Silvae, A. Doltrae, J. D'hooge, J. Brugadae, M. Sitgese, and A. F. Frangi, "Temporal diffeomorphic free-form deformation: Application to motion and strain estimation from 3D echocardiography," *Med. Imag. Anal.* **16**(1), 427–450 (2012).
- <sup>59</sup>P. Dupuis and U. Grenander, "Variational problems on flows of diffeomorphisms for image matching," *Q. Appl. Math.* **LVI**(3), 587–600 (1998).
- <sup>60</sup>A. Bistoquet, J. Oshinski, and O. Skrinjar, "Myocardial deformation recovery from cine MRI using a nearly incompressible biventricular model," *Med. Image Anal.* **12**(1), 69–85 (2008).
- <sup>61</sup>J. Meunier, "Tissue motion assessment from 3D echographic speckle tracking," *Phys. Med. Biol.* **43**(5), 1241–1254 (1998).
- <sup>62</sup>Z. J. Zhang, D. J. Sahn, and X. B. Song, "Frame to frame diffeomorphic motion analysis from echocardiographic sequences," in *Proceedings of the MICCAI Workshop of Third Mathematical Foundation of Computational Anatomy (MFCA)* (Inria, Paris, 2011), pp. 15–24.
- <sup>63</sup>Z. J. Zhang, D. J. Sahn, and X. B. Song, "Temporal diffeomorphic motion analysis from echocardiographic sequences by using intensity transitivity consistency," in *Proceedings of the Second International Conference on Statistical Atlases and Computational Models of the Heart: Imaging and modelling challenges (STACOM)*, Lecture Notes in Computer Science Vol. 7085 (Springer, Heidelberg, 2011), pp. 274–284.
- <sup>64</sup>M. F. Beg and A. Khan, "Symmetric data attachment terms for large deformation image registration," *IEEE Trans. Med. Imaging* **26**(9), 1179–1189 (2007).
- <sup>65</sup>J. Ashburner, "A fast diffeomorphic image registration algorithm," *NeuroImage* **38**(1), 95–113 (2007).
- <sup>66</sup>D. Rueckert, P. Aljabar, R. A. Heckemann, J. V. Hajnal, and A. Hammers, "Diffeomorphic registration using B-splines," in *Proceedings of MICCAI'06*, Lecture Notes in Computer Science Vol. 4191 (Springer, Heidelberg, 2006), pp. 702–709.

- <sup>67</sup>MATLAB version 7.11.0 (R2010b), The MathWorks Inc., Natick, MA, 2010.
- <sup>68</sup><http://www.openmp.com>.
- <sup>69</sup><http://www.sonometrics.com>.
- <sup>70</sup>C. Tobon-Gomez, M. D. Craene, K. McLeod, L. Tautz, W. Z. Shi, A. Hen-nemuth, A. Prakosa, H. Wang, G. Carr-White, S. Kapetanakis, A. Lutz, V. Rasche, T. Schaeffter, C. Butakoff, O. Friman, T. Mansi, M. Sermesant, X. Zhuang, S. Ourselin, H. O. Peitgen, X. Pennec, R. Razavi, D. Rueck-ert, A. F. Frangi, and K. S. Rhode, "Benchmarking framework for myocar-dial tracking and deformation algorithms: An open access database," *Med. Imag. Anal.* **17**(6), 632–648 (2013).
- <sup>71</sup>Y. C. Fung, *First Course in Continuum Mechanics* (Springer, Prentice Hall, NJ, 2010).
- <sup>72</sup>M. D. Cerqueira, N. J. Weissman, V. Dilsizian, A. K. Jacobs, S. Kaul, W. K. Laskey, D. J. Pennell, J. A. Rumberger, T. Ryan, and M. S. Ve-rani, "Standardized myocardial segmentation and nomenclature for tomo-graphic imaging of the heart: A statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the American Heart Association," *Circulation* **105**, 539–542 (2002).
- <sup>73</sup>J. M. Fitzpatrick and J. B. West, "The distribution of target registration error in rigid-body point-based registration," *IEEE Trans. Med. Imaging* **20**(9), 917–927 (2001).
- <sup>74</sup>A. A. Young, "Model tags: Direct three-dimensional tracking of heart wall motion from tagged magnetic resonance images," *Med. Image Anal.* **3**(4), 361–372 (1999).
- <sup>75</sup>C. Metz, N. Baka, H. Kirisli, M. Schaap, T. V. Walsum, S. Klein, L. Neef-jes, N. Mollet, B. Lelieveldt, M. D. Bruijne, and W. Niessen, "Condi-tional shape models for cardiac motion estimation," in *Proceedings of MIC-CAI*, Lecture Notes in Computer Science Vol. 6361 (Springer, Heidelberg, 2010), pp. 452–459.