Preprint version accepted to appear in IEEE Transactions on Medical Imaging.
Final version of this paper available at http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6837493

Improved myocardial motion estimation combining tissue Doppler and B-mode echocardiographic images

A. R. Porras, M. Alessandrini, M. De Craene, N. Duchateau, M. Sitges, B. H. Bijnens, H. Delingette, M. Sermesant, J. D'hooge, A. F. Frangi, Fellow, IEEE and G. Piella

***Abstract*—We propose a technique for myocardial motion estimation based on image registration using both B-mode echocardiographic images and tissue Doppler sequences acquired interleaved. The velocity field is modeled continuously using B-splines and the spatiotemporal transform is constrained to be diffeomorphic. Images before scan conversion are used to improve the accuracy of the estimation. The similarity measure includes a model of the speckle pattern distribution of B-mode images. It also penalizes the disagreement between tissue Doppler velocities and the estimated velocity field. Registration accuracy is evaluated and compared to other alternatives using a realistic synthetic dataset, obtaining mean displacement errors of about 1 mm. Finally, the method is demonstrated on data acquired from 6 volunteers, both at rest and during exercise. Robustness is tested against low image quality and fast heart rates during exercise. Results show that our method provides a robust motion estimate in these situations.**

***Index Terms*—ultrasound, image registration, tissue Doppler, echocardiography, motion estimation, multi-modal integration, TDFFD, FFD.**

I. Introduction

Quantification of cardiac motion and strain has proven to be helpful for cardiac function assessment, providing information on how a given pathology affects global and local mechanics of the myocardium [1], [2], [3]. Among the different imaging techniques available to quantify cardiac motion, ultrasound (US) imaging is one of the most used, since it captures a large range of information (i.e. valve flows, tissue velocities) dynamically and at a reasonable cost. Novel trends in the acquisition process, as the use of shear waves [4] and ultrafast imaging through planar waves [5], are believed to lead to significant improvements in spatiotemporal resolution. These advances will also enrich the spectrum of functional information that can be captured by this modality.

Many approaches for tracking anatomical structures in B-mode echocardiography were proposed during the last years. Speckle-tracking-based approaches [6] use block matching algorithms to track local speckle patterns along US sequences under the assumption that they are stable between consecutive frames. These algorithms do not make use of the temporal information in the whole image sequence and regularizations are performed in post-processing steps.

Several image registration approaches were explicitly designed as spatiotemporal registration schemes [7], [8], [9]. In these papers, temporal continuity of displacements is guaranteed. However, since displacement at a given time instant does not functionally depend on the displacement at previous time instants, the temporal continuity of the recovered velocity is not guaranteed. To overcome this limitation, a transform based on the velocity field was proposed in [10], calculating displacements by integrating velocities at all previous time instants. To preserve the topology and orientation of the anatomical structures, the transform was constrained to be diffeomorphic (smooth, invertible and with smooth inverse) [11]. An additional regularization term minimizing the compressibility of the myocardium was also added to the cost function. Other approaches to guide and constrain cardiac tracking including shape information were proposed in [12], [13], [14].

Complementary to B-mode echocardiographic images, tissue Doppler Imaging (TDI) is widely used in the clinical practice. It allows an objective quantification of true tissue velocities with higher temporal resolution than B-mode and better signal-to-noise ratio [2]. Estimating a displacement field from TDI data can be done by temporally integrating the measured velocities. However, only the projection of the velocity along the beam direction is available and, due to the low spatial resolution of TDI, local changes may not be captured. In addition, the noise present in the images is accumulated at each integration step, possibly leading to strong drift artifacts.

Some approaches have been proposed to overcome the problems of each technique by using both B-mode images and TDI together for a better quantification of heart motion. In [15], TDI was used to track heart motion along the beam direction, while B-mode sequences were used to track in the direction perpendicular to the beam. These two components

A. R. Porras, N. Duchateau, B. H. Bijnens and G. Piella are with the Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain. B. H. Bijnens is also with ICREA.

M. Alessandrini and J. D'hooge are with Department of Cardiovascular Sciences, Laboratory of Cardiovascular Imaging and Dynamics, KU Leuven, Leuven, Belgium.

M. De Craene is with Philips Research, Medisys, Suresnes, France.

M. Sitges is with the Cardiology Department, thorax Institute, Hospital Clínic de Barcelona - IDIBAPS, Barcelona, Spain.

H. Delingette and M. Sermesant are with the Inria-Asclepios Project, Sophia Antipolis, France

A. F. Frangi is with the Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB) - Department of Mechanical Engineering, University of Sheffield, Sheffield, UK.

were estimated separately and no additional regularization was performed. In [16], an optical flow based registration method was proposed, modeling the velocity using a spatial affine model. Registration was performed considering pairs of frames, so temporal coherence was not guaranteed. Furthermore, the evaluated B-mode and TDI frames had to coincide in time. Therefore, either temporal interpolation of the B-mode sequence was necessary or TDI frames had to be discarded, thus losing temporal resolution. Recently, some authors have introduced the possibility of estimating tissue velocities in two directions using transverse oscillation images [17], [18], although this technique is not yet ready for clinical use.

In this paper, we propose a registration framework that takes into account both B-mode images and tissue Doppler velocities, acquired interleaved with the same probe, to calculate one single continuous spatiotemporal diffeomorphic transform. In this case, TDI sequences and B-mode data are modeled by the same continuous velocity field, so both data representations coincide in space and no further spatial alignment is required. Preliminary results showing the interest of using both modalities together were presented in [19].

The method proposed in the current work extends the one from [19] by incorporating speckle statistics information to calculate similarities between consecutive B-mode frames. It also uses information before scan conversion, thus avoiding the use of spatially interpolated data. In addition, by using only recorded information, the number of samples in the images is lower than after scan-conversion, so less evaluations are necessary during the optimization process and the algorithm performs faster. Moreover, it is not necessary to mask the field of view, since all pixels in the non-scan converted images contain valuable information.

In [19], 2D B-mode synthetic images were created by slicing 3D synthetic images, while tissue Doppler data were generated by directly projecting velocities of ground truth volumetrical meshes on the beam direction and adding gaussian noise. In this paper, a more robust validation scheme is proposed, using a realistic US simulator as explained in Section III.A. The feasibility of using the current method for clinical cases is also shown. Stress echocardiography aims at understanding the relation between cardiac function and functional capacity during effort [20], [21]. In this protocol, image quality and temporal resolution may be low. We demonstrate the robustness of our method in this clinical setting. Results obtained with the proposed method were compared to results using only B-mode images. Our method allows estimating a realistic motion field in cases where tracking using one modality fails due to the quality of the data from this protocol.

II. Methodology

A basic scheme describing the proposed methodology is shown in Fig. 1. In the following sections, each component in the registration scheme will be described.

A. The transform

The proposed approach models the velocity field continuously in time and space using B-spline kernels, as proposed in [10]:

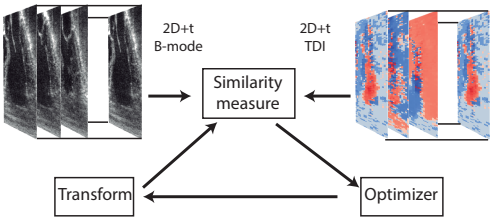

Fig. 1. Summary of the registration scheme.

$$\mathbf{v}(\mathbf{x},t;\mathbf{p}) = \sum_{i,j,k} \beta\left(\frac{x-q^i}{\Delta^i}\right)\beta\left(\frac{y-q^j}{\Delta^j}\right)\beta\left(\frac{t-q^k}{\Delta^k}\right)\mathbf{p}^{i,j,k} \tag{1}$$

where $\mathbf{x} = \{x, y\}$ are the spatial coordinates of the point whose velocity is evaluated, $\beta(\cdot)$ is a cubic B-spline kernel function, $\mathbf{Q} = \{q^i, q^j, q^k\}$ represents the grid of uniformly spaced control points, $\mathbf{\Delta} = \{\Delta^i, \Delta^j, \Delta^k\}$ are the spacings between control points and $\mathbf{p}$ is a vector containing the B-spline coefficients, which correspond to the velocities associated to each control point.

To map a point from $t = t_0$ to a time instant $t = T$, it is then necessary to integrate the velocity field from $t = t_0$ to $t = T$:

$$\varphi_{t_0}(\mathbf{x}, T; \mathbf{p}) = \mathbf{x} + \int_{t_0}^{T} \mathbf{v}\big(\varphi_{t_0}(\mathbf{x}, t; \mathbf{p}), t; \mathbf{p}\big)dt. \tag{2}$$

To compute the integral in Eq. 2, a forward Euler integration scheme was used. Thus, the transformation was approximated by:

$$\varphi_{t_0}(\mathbf{x}, t_n; \mathbf{p}) = \mathbf{x} + \sum_{k=0}^{n-1} \mathbf{v}\big(\varphi_{t_0}(\mathbf{x}, t_k; \mathbf{p}), t_k; \mathbf{p}\big)\Delta t_k \tag{3}$$

where $\Delta t_k = t_{k+1} - t_k$.

With this approach, the temporal sampling used to approximate the integration of the velocity field has to be small enough for a good estimation of the trajectory. In [10], a time interval of $1/2$ the spacing between B-mode frames was found to be small enough. In our implementation, it was reduced to $1/4$ the temporal spacing between consecutive TDI frames to increase the accuracy when estimating displacements. To ensure invertibility of the transformation, the determinant of its spatial Jacobian was constrained to be positive. To achieve this, if a negative value was detected, the temporal sampling was divided by a factor of 2 until no negative values were found [10].

In [19], a 2D+t grid of control points was set on Cartesian coordinates, as shown in Fig.2(a). In the current approach, we work with non-scan converted data in polar coordinates to avoid unnecessary interpolations [22]. Setting a grid in polar coordinates, such as in Fig.2(b), corresponds to a deformed grid in Cartesian coordinates, such as in Fig.2(c). As it can be observed from Fig. 2, the Cartesian spacing between control points increases with their distance to the transducer placed at the origin of the field of view.

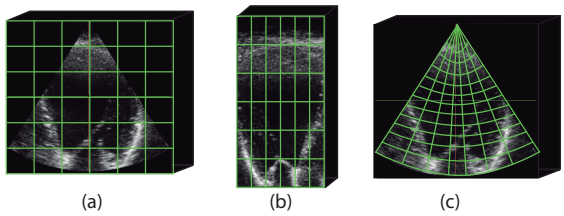

Fig. 2. (a) Grid of control points in Cartesian coordinates as in [19]. (b) Grid of control points in polar coordinates. (c) Representation of the image in (b) in Cartesian coordinates.

B. The similarity measure

The similarity measure to be minimized is the following:

$$M(\mathbf{p}) = (1-\lambda)U(\mathbf{p}) + \lambda D(\mathbf{p}) \tag{4}$$

where $U(\mathbf{p})$ represents the matching between the registered B-mode frames, $D(\mathbf{p})$ measures the agreement between the estimated velocity field and the velocity values provided by TDI, and λ is a term balancing the contribution of $U(\mathbf{p})$ and $D(\mathbf{p})$. Note that $U(\mathbf{p})$ and $D(\mathbf{p})$ contribute separately to estimate one single and continuous transform. Therefore, B-mode and tissue Doppler frames do not need to coincide at specific discrete locations, so data interpolation is not required to compensate for different resolutions.

In [19], a mean squared error approach was used for $U(\mathbf{p})$, where each frame was compared to a reference frame. Assuming that speckle patterns can be represented as a multiplicative Rayleigh distributed noise, and that they are preserved among consecutive frames, a similarity term based on [23] was used in the current approach, similar to [24]:

$$U(\mathbf{p}) = \sum_{n=1}^{N} \sum_{\mathbf{x}\in\Omega_{I_{n-1}}} \ln(e^{2\Delta_{n-1}^{n}(\mathbf{x};\mathbf{p})}+1) - \Delta_{n-1}^{n}(\mathbf{x};\mathbf{p}) \tag{5}$$

where $N+1$ is the number of B-mode frames, $\Omega_{I_{n-1}}$ is the spatial domain of I_{n-1} and $\Delta_{n-1}^{n}(\mathbf{x};\mathbf{p}) = I_{n-1}(\mathbf{x}) - I_n\big(\varphi_{t_{n-1}}(\mathbf{x},t_n;\mathbf{p})\big)$.

When using a traditional B-spline based registration approach with Cartesian images (as shown in Fig.2(a)), all samples contribute equally to find the optimal velocity value at each control point. In this case, the magnitude of the displacement error at one pixel is independent from its location. When working with images in polar coordinates as shown in Fig.2(b), the Cartesian space represented by each sample is not homogeneous and increases with its distance to the transducer, as seen in Fig.2(c). An error of one pixel far from the transducer is larger in Cartesian units (in the physical space) than an error of one pixel close to it. Thus, it is necessary to compensate for this difference.

The relation between a displacement error of a sample at two different locations is proportional to the relation between the Cartesian space between samples at these locations in the direction perpendicular to the beam. Since this space is a function of the distance between the samples and the transducer, an approach weighting each sample contribution depending on its distance to the transducer is proposed:

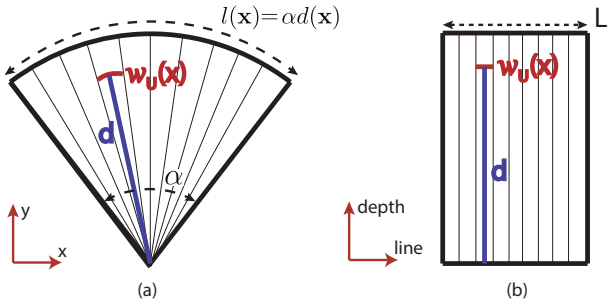

Fig. 3. Representation of $w_U(\mathbf{x})$ in scan converted (a) and non-scan converted format (b)

$$U(\mathbf{p}) = \sum_{n=1}^{N} \sum_{\mathbf{x}\in\Omega_{I_{n-1}}} w_U(\mathbf{x})\big(\ln(e^{2\Delta_{n-1}^{n}(\mathbf{x};\mathbf{p})}+1) - \Delta_{n-1}^{n}(\mathbf{x};\mathbf{p})\big) \tag{6}$$

where $w_U(\mathbf{x})$ represents the weight applied to each sample in the B-mode images. The weighting function we proposed was:

$$w_U(\mathbf{x}) = \frac{l(\mathbf{x})}{L} = \frac{\alpha d(\mathbf{x})}{L} \tag{7}$$

where $l(\mathbf{x})$ is the length of the arch perpendicular to the beam at position $\mathbf{x}$, L is the number of US beams, α is the angle of the whole field of view in radians and $d(\mathbf{x})$ is the distance between $\mathbf{x}$ and the origin of the field of view, as represented in Fig. 3. In the implementation, $w_U(\mathbf{x})$ was normalized so that $\sum w_U(\mathbf{x}) = 1$.

Since TDI only measures tissue velocity along the beam direction, the estimated velocity field was projected on this direction and compared to the velocity values provided by TDI:

$$D(\mathbf{p}) = \sum_{m=0}^{M} \sum_{\mathbf{x}\in\Omega_{V_0}} w_D(\mathbf{x})\Big(\mathbf{b}\big(\varphi_{t_0}(\mathbf{x},t_m;\mathbf{p})\big)\cdot \mathbf{v}\big(\varphi_{t_0}(\mathbf{x},t_m;\mathbf{p}),t_m;\mathbf{p}\big) - V_m(\mathbf{x})\Big)^2 \tag{8}$$

where $w_D(\mathbf{x})$ is the weight function applied to each sample in the TDI data similar to $w_U(\mathbf{x})$ in Eq.7, $\mathbf{b}(\varphi_{t_0}(\mathbf{x},t_m;\mathbf{p}))$ is a unitary vector in the beam direction, $V_m(\mathbf{x})$ is the velocity value provided by TDI at location $\mathbf{x}$ and frame at time $t = t_m$, $M+1$ is the number of tissue Doppler frames, and Ω_{V_0} corresponds to the spatial domain of the first TDI frame.

Finally, to find the minimum of the proposed similarity metric, the L-BFGS-B (Limited memory Broyden-Fletcher-Goldfarb-Shannon with simple Bounds) [25] optimizer was chosen, which is a limited memory quasi-Newton algorithm for solving large nonlinear optimization problems with simple bounds on the variables.

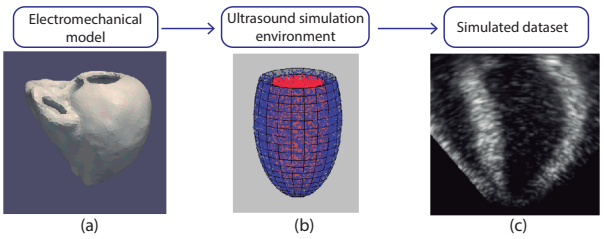

Fig. 4. Schematic of the pipeline for the simulation of cardiac ultrasound sequences in [26].

III. Experiments

For the validation of the proposed method, a realistic synthetic dataset including simulations of healthy and pathologic hearts was used. First, the method was evaluated and compared to other alternatives. Afterwards, the method was applied to 6 real cases to show that it is able to estimate a realistic motion field in situations where the use of one single modality is challenging.

For all the experiments a grid of 6 B-spline control points in each spatial dimension was used. In [10], it was demonstrated that increasing the number of temporal control points resulted in higher accuracy, and that taking more temporal control points than the number of frames would oversample the velocity field. In the current work, we set the number of temporal control points equal to the number of B-mode frames. In our dataset, the patient with the least number of B-mode frames at rest had 25 frames. Therefore, we constrained the maximum number of temporal control points to 25 for a fair analysis between patients.

A. Generation of synthetic data

Synthetic B-mode and TDI data were generated using a modification of the pipeline described in [26] and summarized in Fig. 4. A realistic 3D volumetric tetrahedral mesh (as shown in Fig. 5(a)) was built from the segmentation of a Magnetic Resonance Image (MRI) acquisition on an healthy patient. The SOFA (Simulation Open Framework Architecture) [27] simulation framework, which is targeted at real-time simulation, was then used to apply the Bestel-Clement-Sorine electromechanical model [28] to the 3D heart geometry to simulate one cardiac cycle, starting from end of diastole at a frame rate of 90 Hz.

By modifying the value of the mechanical parameters of the model at a segmental level, several degrees of ischemia can be simulated. In particular, the same 9 cases considered in [26] were included in the present study: one normal case; two cases with ischemia (one mild and one severe) in the proximal region of the left anterior descending artery (LADprox); two cases with ischemia (one mild and one severe) in the distal region of the left anterior descending artery (LADdist); two cases with ischemia (one mild and one severe) in the region of the right coronary artery (RCA); two cases with ischemia (one mild and one severe) in the region of the left circumflex coronary artery (LCX).

The output meshes from the electromechanical simulation were used to displace a cloud of point scatterers mimicking the acoustic tissue response of the myocardium. From each scatter map, US radiofrequency lines were generated by convolving the cloud of scatter points with the point spread function (PSF) of the imaging system. Note that scatterers do not move during the simulation of a single frame, so the acquisition is supposed to be instantaneous. As in [26], COLE [29] was adopted as a fast US simulation environment due to its high computational efficiency. From each time variant scatter map, a cine B-mode sequence and a TDI sequence were generated. In particular, 2D apical four chamber views were considered in this study. For both modalities, the simulated system implemented a 1D phased array with 64 elements, each element of width $\lambda/2$, height 14 mm and a kerf of $\lambda/10$. The simulated probe had a center frequency (f_0) of 4 MHz, a -6 dB relative bandwidth of 65% and a sampling frequency of 50 MHz. The scan angle was 75 degrees, scan depth was 14 cm and the focus was positioned at 7 cm. Note that all the settings of the synthetic system were chosen to be the closest to the real setup used in the in-vivo evaluation (see Table II and Table III).

1) Simulation of B-mode images: One mesh out of three from the electromechanical simulation was considered leading to an imaging frame rate of 30 Hz. From the associated scatter map, 100 radiofrequency lines were acquired by uniformly sweeping through the scan angle. As such, after envelope detection, log-compression and scan conversion, a set of B-mode images were obtained (pixel size = $0.34\text{x}0.34$ mm^2). Fig. 4(b) and Fig. 4(c) show an example of point scatterers and simulated B-mode image.

2) Simulation of TDI: All meshes from the electromechanical simulation were considered, leading to a frame rate of 90 Hz. For each frame, 20 scan lines were considered by uniformly sweeping through the scan angle. For each scan line, a set of four radiofrequency lines were simulated with a pulse repetition frequency (PRF) of 2 KHz. This led to a Nyquist velocity of $\sim$19 cm/s, given:

$$v_{Nyq} = \frac{c \cdot PRF}{4 f_0} \tag{9}$$

where c is the sound speed in the tissue ($\sim$1540 m/s) and f_0 is the center frequency. Tissue motion between two successive firings in the same direction was simulated by linearly interpolating the position of the scattering centers in the considered frame and the following one. For each scan line and each depth, the tissue velocity was computed from the corresponding package of four signals by means of a standard phase shift based estimator [30]. After color rendering of the computed velocity and scan conversion, TDI images, as the ones of Fig. 5 (b)(c), were obtained.

B. Selection of the weighting parameter λ

Before registering an echocardiographic image sequence using the proposed method, it is first necessary to set a value for the parameter λ that balances the contribution of each term in the similarity measure. A high value of λ gives more importance to $D(\mathbf{p})$, while a low value gives more importance to $U(\mathbf{p})$.

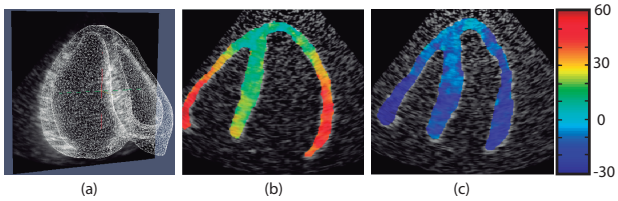

Fig. 5. A simulated US frame together with the wireframe of the mesh from the electromechanical model is represented in (a). Example of TDI images for early systole and early diastole are shown in (b) and (c), respectively. Colors represent the velocity in mm/s, which is considered positive when directed towards the transducer.

TABLE I
SUMMARY OF THE METHODS COMPARED WITH THE SYNTHETIC DATA.

Method	Image format	B-mode similarity
M1	Non-scan converted (Fig. 2(b))	US-specific (Eq. 6)
M2	Scan converted (Fig. 2(a))	US-specific (Eq. 6)
M3	Non-scan converted (Fig. 2(b))	Mean squared error [19]
M4	Scan converted (Fig. 2(a))	Mean squared error [19]

To choose an adequate value of λ, two main aspects have been considered: (i) the confidence on the value provided by each term, and (ii) the way each term converges to its minimum. The former was considered to be equivalent to the confidence on the data provided by each modality. The latter was related to the slope of the function that is minimized during the optimization. Thus, an adequate value of λ would balance the magnitude of the slope of both terms. Considering these two aspects, we set λ as:

$$\lambda = \gamma r \tag{10}$$

where γ is the ratio between the confidence on the tissue Doppler data and the B-mode images, and r is the ratio between the slope of the two terms in the similarity measure. In our experiments, the confidence ratio was set to 1, thus assuming the same quality for TDI and B-mode images. r was approximated as the magnitude of the initial derivative (with respect to the transformation parameters) of $U(\mathbf{p})$ divided by the derivative of $D(\mathbf{p})$.

C. Validation

The proposed method was applied to the previously described synthetic dataset to evaluate its accuracy in terms of displacement. The possible combinations between the similarity measure in [19] and the one from Eq. 6, and between using Cartesian and polar data were tested, as summarized in Table 1. Fig. 6 shows the mean error and the standard deviation of the displacements calculated from the first frame for all the synthetic patients during one cardiac cycle implementing these different alternatives.

Results from Fig. 6 show that using both non-scan converted images and a similarity measure including speckle distribution statistics gives better performance in average than the other three alternatives evaluated, reducing the mean error of the displacements estimated with methods M2, M3 and M4 by

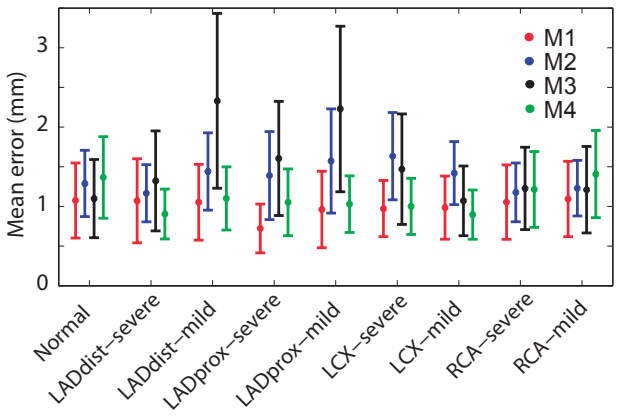

Fig. 6. Mean errors (circles) and standard deviations (lines) calculated for the 9 synthetic datasets during one cardiac cycle with the methods summarized in Table 1.

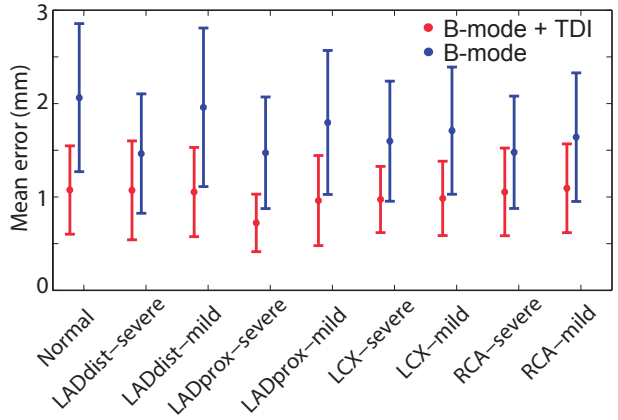

Fig. 7. Mean errors (circles) and standard deviations (lines) calculated for the 9 synthetic patients during one cardiac cycle with method M1 (red) and using only B-mode images (blue).

35%, 48% and 10%, respectively. Mean differences between the proposed method (M1) and methods M2 and M3 were found to be statistically significant with $p < 0.05$ using a Student's t-test. Mean differences between methods M1 and M4 were not found to be statistically significant, obtaining a p-value of 0.15.

Fig. 7 compares displacement error results obtained with M1 with respect to the results obtained taking into account only the B-mode images, using Eq. 6. The results show that integrating B-mode images and tissue Doppler velocities results in a more accurate estimation of the displacement field, thus agreeing with the results from [19]. Differences between the mean displacement error obtained by integrating B-mode and TDI data, and the error obtained using only B-mode were found to be statistically significant for $p < 0.05$.

D. Clinical data

The method presented was also applied to a set of US images acquired from 6 volunteers (age 30 $\pm$ 5.5, males) with a Vivid-Q system (GE Healthcare Milwaukee, WI)

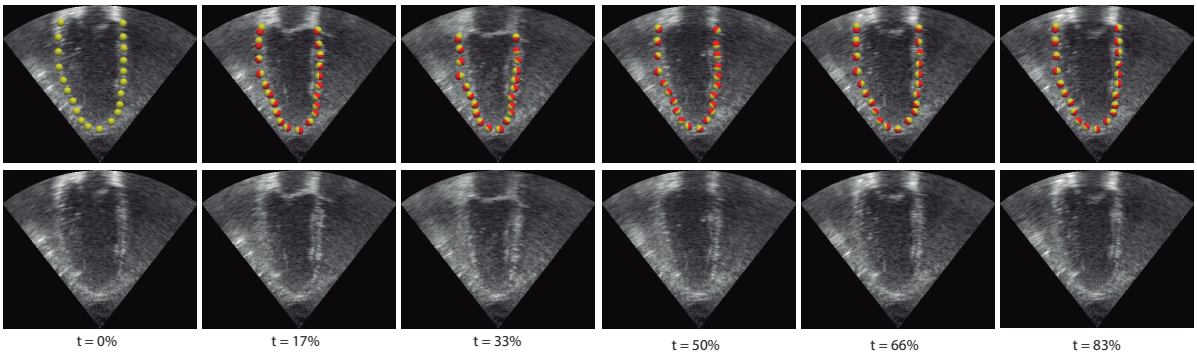

Fig. 8. The top row shows a set of landmarks placed on the left ventricle and displaced according to the transformation calculated using only B-mode (yellow), and using both B-mode and TDI together (red) at different times of the cardiac cycle for one example patient at rest. The bottom row shows the corresponding B-mode images.

TABLE II
COMPARISON BETWEEN B-MODE IMAGE INFORMATION FOR REAL CASES AT REST, DURING EXERCISE AND FOR SYNTHETIC DATA.

	Rest	Exercise	Synthetic
Frames/cycle	39 ± 9	14 ± 2	27
Frame rate (Hz)	35 ± 9	34 ± 7	30
Beams	92.3 ± 9.4	85.7 ± 18.9	100
Samples/beam	577 ± 47.9	584.2 ± 64.9	248
Angle (°)	70 ± 7	65 ± 14.1	75
Depth (mm)	133 ± 11	135 ± 15	140

TABLE III
COMPARISON BETWEEN TISSUE DOPPLER IMAGE INFORMATION FOR REAL CASES AT REST, DURING EXERCISE AND FOR SYNTHETIC DATA.

	Rest	Exercise	Synthetic
Frames/cycle	105 ± 28	45 ± 7	79
Frame rate (Hz)	116 ± 28	107 ± 22	90
Beams	14 ± 3.1	15.3 ± 2.7	20
Samples/beam	190.5 ± 180.4	194.7 ± 21.4	248
Angle (°)	69 ± 8.4	60.5 ± 12.4	72
Depth (mm)	133 ± 11	135 ± 15	140
Nyquist vel. (cm/s)	18.1 ± 3	31.6 ± 1.5	19
PRF (Hz)	1125 ± 191	1958.3 ± 93.2	2000

and a GE Healthcare M4S probe (1.7/3.4 MHz frequency). For each case, both B-mode and tissue Doppler images of the left ventricle were acquired at rest. Afterwards, stress echocardiography was performed with an ergometric bicycle and images were acquired at maximum effort. Table II and Table III show the mean values and standard deviation for the different parameters of the acquired B-mode and tissue Doppler images, respectively. Finally, our method integrating both B-mode and TDI information for motion estimation was applied to the acquired images and results were compared to the ones obtained by using only B-mode images.

Fig. 8 shows the result of tracking a set of landmarks placed on the left ventricle along one cardiac cycle using only B-mode images (yellow), and integrating both B-mode and TDI (red) with the proposed method, for one patient at rest. As it can be observed, most of the differences between both methods

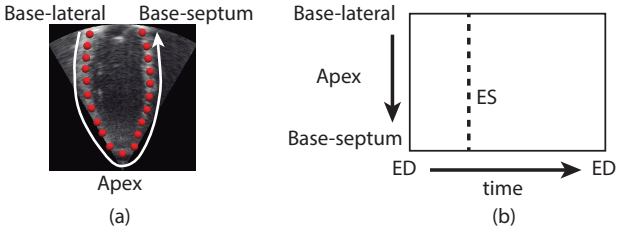

Fig. 9. Schematic describing a spatiotemporal map. (a) shows a curved line placed along the myocardium in a B-mode image and (b) shows the corresponding spatiotemporal map. The vertical axis represents spatial locations along the myocardium and the horizontal axis represents time in the cardiac cycle extracted between consecutive R-peaks in the ECG. The dotted vertical line indicates end systole. ED and ES stand for end diastole and end systole, respectively.

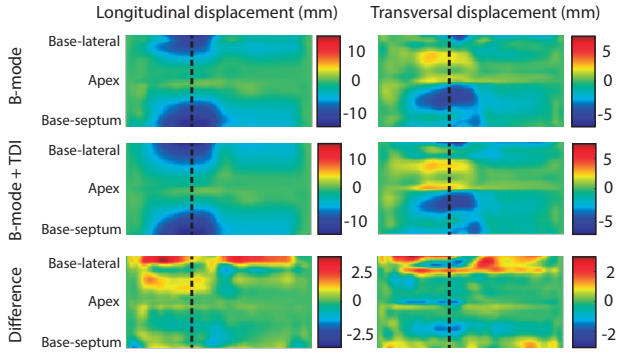

Fig. 10. Displacements estimated using only B-mode images (top row), using B-mode and TDI together (middle row) and the difference between both approaches (bottom row) for one patient at rest.

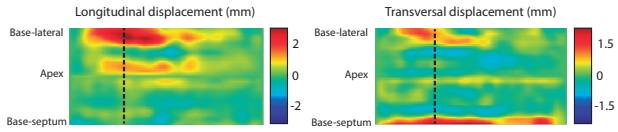

Fig. 11. Mean difference between displacements estimated using only B-mode and using both B-mode and TDI at rest.

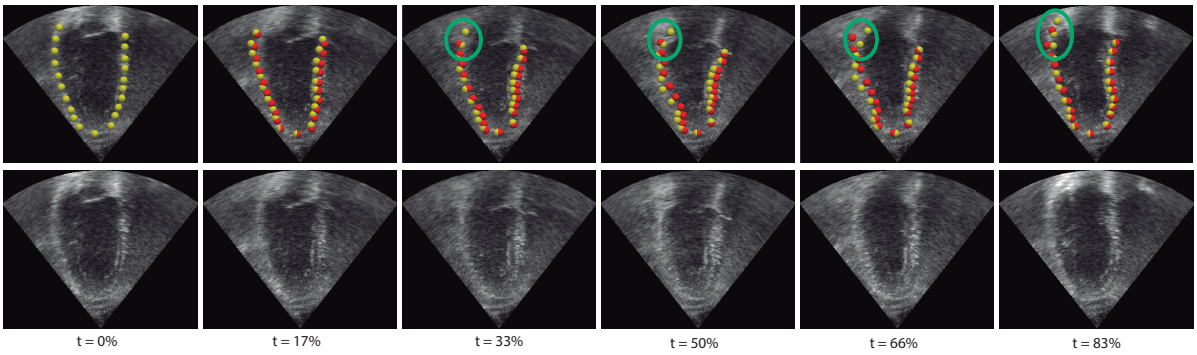

Fig. 12. The top row shows a set of landmarks placed on the left ventricle and displaced according to the transformation calculated using only B-mode (yellow), and using both B-mode and TDI together (red) at different times of the cardiac cycle during exercise for the patient in Fig. 8. Green circles indicate where tracking using B-mode only has problems due to image quality. The bottom row shows the corresponding B-mode images.

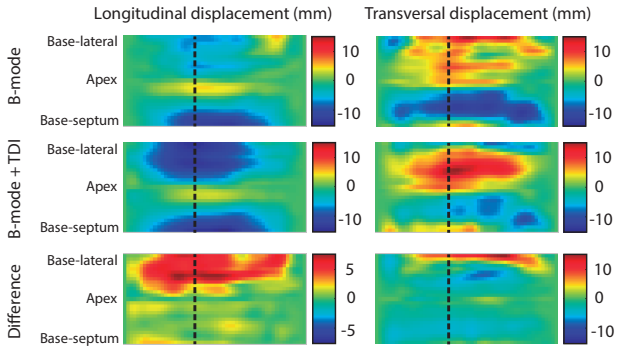

Fig. 13. Displacements estimated using only B-mode images (top row), using B-mode and TDI together (middle row) and the difference between both approaches (bottom row) for one patient during exercise.

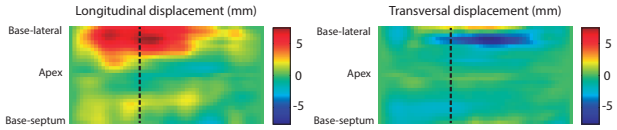

Fig. 14. Mean difference between displacements estimated using only B-mode and using both B-mode and TDI during exercise.

lay on the lateral wall. Fig. 10 shows the displacements calculated using the two methods in both longitudinal and transversal directions. Spatiotemporal color maps were used to visualize the results. An illustration to understand this kind of map is shown in Fig. 9. The vertical axis represents spatial locations along the myocardium (which has been unfolded inspired from anatomical M-mode echocardiographic images) and the horizontal axis represents time in the cardiac cycle. Positive longitudinal displacement was defined from apex to base, while positive transversal displacement was defined from lateral wall to septum. One can see that the difference between the displacements estimated with the two methods is larger at the base of the lateral wall.

Fig. 11 shows the mean displacement differences between the two methods for the 6 patients analyzed at rest. Results show that the displacement estimated at end systole using only B-mode images is lower, in average, than the displacement estimated with the proposed method, and with maximum difference located at the base of the ventricle. These results are in line with the results obtained for the patient shown in Fig. 10.

Fig. 12 shows the result of tracking a set of landmarks placed on the left ventricle with the two methods for the patient presented in Fig. 8, using the data acquired during exercise. Fig. 13 also shows the displacements calculated during exercise for this patient. As it can be observed, displacement differences between the two approaches are increased with respect to the results obtained at rest. Moreover, the displacement pattern estimated using only B-mode images in Fig. 13 does not look physiologically realistic.

The mean displacement differences between the two approaches were finally calculated using the data acquired during exercise for the 6 patients analyzed, as shown in Fig. 14. It is possible to see that the magnitude of the mean difference is increased when compared to the differences from the data acquired at rest. In addition, most of these differences are located at the base of the lateral wall.

IV. Discussion

The proposed method was validated using a synthetic dataset and compared with the alternatives shown in Table 1. Results in Fig. 6 show that the proposed method (M1) performs better, in average, than the other tested alternatives. Furthermore, M1 uses non-scan converted images, so the registration process is computationally less expensive because the number of samples in the images before scan conversion is lower. In addition, it is not necessary to mask the field of view, since all samples in the images provide valuable information.

From the results in Fig. 6, one can conclude that, among the generated synthetic cases, M4 performs better in those cases where ischemia is simulated close to the Left Anterior Descending (LAD) artery and the Left Circumflex (LCX). The 2D simulated images include the areas in the septum and lateral wall that are more affected by the ischemia in the LAD and LCX regions, respectively. Since local motion in ischemic

regions is reduced, local changes between consecutive frames are also reduced. Method M4 uses a similarity measure that compares all frames with a reference frame, so its performance varies depending on how similar each frame is to the reference. Therefore, in those cases where motion is reduced, the performance of M4 may improve, producing slightly better results than M1 in "LADdist-severe" and "LCX-mild" cases. Method M2 likely fails because it applies a speckle statistics-based measure to Cartesian B-mode images. This kind of metric uses the Rayleigh distribution to model the speckle. However, when working with Cartesian images, many pixels in the image contain interpolated information, resulting in a speckle pattern whose properties may differ from the Rayleigh statistics. Finally, M3 performs worse than the others. The similarity measure used by this method compares each frame, in non-scan converted format, to a reference frame using a mean squared error-based approach. In this case, if the tissue structures imaged in the reference frame fall completely or partially out of the US lines, since there is no interpolation, this information will be completely lost. Thus, M3 could fall into a local minimum during the optimization process, not being able to recover an accurate displacement field.

The proposed method was also applied to 6 healthy volunteers, acquiring images both at rest and during exercise. Since the quantitative parameters extracted from echocardiographic sequences [3] are highly conditioned by the accuracy of the tracking, clinical studies targetting a better understanding of cardiac function during exercise may benefit from the improvements of the proposed method. These protocols are currently receiving more attention to understand the cardiac function contribution to the functional capacity during effort and to prevent from any potentially high-risk impairment [20], [21].

Results from Fig. 11 show the differences between displacements estimated using only B-mode and integrating both B-mode and TDI at rest. As it can be seen, the magnitude of the difference in longitudinal displacements is low in average, being higher in the lateral wall with a maximum of 2.5 mm at end systole. Transversal displacement differences are lower than in the case of longitudinal displacements as expected, since TDI only gives information in the longitudinal direction. Most of the differences in the transversal direction are accumulated at the base and differences may occur as a consequence of different estimations in the longitudinal direction. Fig. 8 and Fig. 10 show the displacements estimated for one example case using only B-mode and using both B-mode and TDI. These results are in line with the average results previously discussed.

Displacements were also estimated for the same 6 volunteers during exercise. In this situation, the quality of the B-mode images acquired is worse due to the difficulty of the acquisition, as it can be observed in Fig. 12. This, added to the reduced number of frames because of higher heart rates, makes displacement estimation using only B-mode images challenging. In the present study, the number of frames during exercise is reduced to 41% of the number of frames acquired at rest. Results from Fig. 14 show larger differences between the displacements estimated using only B-mode and using both B-mode and TDI together when compared with the results obtained at rest. These larger differences are more visible at the base of the lateral wall, with a maximum error of about 6 mm at end systole. The basal level (mitral annulus) and the lateral wall are subject to lower image quality (border of the echocardiographic window, reflections, and position of the valve with respect to the myocardium). Green circles in Fig. 12 indicate a region where tracking using B-mode only is prone to large errors. Our method might also lead to inaccuracies in such challenging cases (radial motion tends to be overestimated at the apical lateral level). However, due to the use of TDI data, these artifacts are more localized and along the radial direction only. The example case displayed in Fig. 13 also shows large differences in the estimation of displacements in the lateral wall, in line with the average results.

One limitation of the current work may be the number of real cases included. However, our primary objective was to show the feasibility and the added value of our approach as compared to the ones considering one single modality. Application to a larger number of subjects within a specifically-designed clinical study is left for further works.

V. Conclusions

A method to integrate B-mode images and TDI velocities in one single registration framework has been proposed. Unlike other methods to integrate these two modalities [15], [16], [31], B-mode and TDI samples are evaluated separately and they do not have to coincide in time, so temporal/spatial interpolation is not needed to calculate one single and continuous transform.

The improvements of using both modalities together instead of a singe modality were demonstrated. Results from validation with synthetic data showed that using speckle statistics information and data before scan conversion outperforms the results obtained by using a more classical pixel intensity-based measure to compare scan-converted frames. Experiments with real cases also showed that integrating both modalities gives a realistic motion estimation in cases where using only B-mode images is challenging.

Future work includes the extension to clinical protocols including fast motion, such as LBBB-related patterns and athletes undergoing exercise.

Acknowledgements

This research was partially funded by the Spanish Ministry of Economy and Competitiveness (TIN2012-35874) and the Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 611823. The work of A. R. Porras was supported by the Spanish Government with a FPU grant. The work of M. Alessandrini was supported by the Research Foundation Flanders (FWO).

References

[1] C. Petitjean, N. Rougon, and P. Cluzel, "Assessment of myocardial function: a review of quantification methods and results using tagged MRI.," *J Cardiovasc Magn Reson*, vol. 7, no. 2, pp. 501–516, 2005.

[2] G. Sutherland, L. Hatle, P. Claus, J. D'hooge, and B. Bijnens, *Doppler myocardial imaging-a textbook*. BSWK Bvba, 2006.

[3] B. Bijnens, M. Cikes, C. Butakoff, M. Sitges, and F. Crispi, "Myocardial motion and deformation: What does it tell us and how does it relate to function?," *Fetal Diagn Ther*, vol. 32, no. 1-2, pp. 5–16, 2012.

[4] A. P. Sarvazyan, O. V. Rudenko, S. D. Swanson, J. B. Fowlkes, and S. Y. Emelianov, "Shear wave elasticity imaging: a new ultrasonic technology of medical diagnostics.," *Ultrasound Med Biol*, vol. 24, pp. 1419–1435, Nov 1998.

[5] G. Montaldo, M. Tanter, J. Bercoff, N. Benech, and M. Fink, "Coherent plane-wave compounding for very high frame rate ultrasonography and transient elastography.," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 56, pp. 489–506, Mar 2009.

[6] R. Jasaityte, B. Heyde, and J. D'hooge, "Current state of three-dimensional myocardial strain estimation using echocardiography.," *J Am Soc Echocardiogr*, vol. 26, pp. 15–28, Jan 2013.

[7] M. J. Ledesma-Carbayo, J. Kybic, M. Desco, A. Santos, M. Shling, P. Hunziker, and M. Unser, "Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation.," *IEEE Trans Med Imaging*, vol. 24, pp. 1113–1126, Sep 2005.

[8] C. T. Metz, S. Klein, M. Schaap, T. van Walsum, and W. J. Niessen, "Nonrigid registration of dynamic medical imaging data using nd + t B-splines and a groupwise optimization approach.," *Med Image Anal*, vol. 15, pp. 238–249, Apr 2011.

[9] M. Yigitsoy, C. Wachinger, and N. Navab, "Temporal groupwise registration for motion modeling," in *Inf Process Med Imaging*, vol. 6801 of *LNCS*, pp. 648–659, Springer Berlin Heidelberg, 2011.

[10] M. De Craene, G. Piella, O. Camara, N. Duchateau, E. Silva, A. Doltra, J. D'hooge, J. Brugada, M. Sitges, and A. F. Frangi, "Temporal diffeomorphic free-form deformation: Application to motion and strain estimation from 3D echocardiography," *Med Image Anal*, vol. 16, no. 2, pp. 427 – 450, 2012.

[11] A. Trouvé, "Diffeomorphisms groups and pattern matching in image analysis," *Int J Comput Vis*, vol. 28, pp. 213–221, 1998.

[12] N. Paragios, "A level set approach for shape-driven segmentation and tracking of the left ventricle.," *IEEE Trans Med Imaging*, vol. 22, pp. 773–776, Jun 2003.

[13] I. Dydenko, F. Jamal, O. Bernard, J. D'hooge, I. E. Magnin, and D. Friboulet, "A level set framework with a shape and motion prior for segmentation and region tracking in echocardiography," *Med Image Anal*, vol. 10, no. 2, pp. 162 – 177, 2006.

[14] C. Compas, E. Wong, X. Huang, S. Sampath, B. Lin, P. Pal, X. Papademetris, K. Thiele, D. Dione, M. Stacy, L. Staib, A. Sinusas, M. O'Donnell, and J. Duncan, "Radial basis functions for combining shape and speckle tracking in 4d echocardiography," *IEEE Trans. Med. Imag.*, vol. 33, no. 6, pp. 1275–1289, 2014.

[15] B. H. Amundsen, J. Crosby, P. A. Steen, H. Torp, S. A. Slrdahl, and A. Stylen, "Regional myocardial long-axis strain and strain rate measured by different tissue doppler and speckle tracking echocardiography methods: a comparison with tagged magnetic resonance imaging," *Eur J Echocardiogr*, vol. 10, no. 2, pp. 229–237, 2009.

[16] M. Suhling, M. Arigovindan, C. Jansen, P. Hunziker, and M. Unser, "Bimodal myocardial motion analysis from B-mode and tissue doppler ultrasound," in *Biomedical Imaging: From Nano to Macro. IEEE ISBI*, pp. 308 – 311 Vol. 1, april 2004.

[17] H. Liebgott, A. B. Salem, A. Basarab, H. Gao, P. Claus, J. D'hooge, P. Delachartre, and D. Friboulet, "Tangential sound field oscillations for 2D motion estimation in echocardiography," in *IEEE Int. Ultrasonics Symp.*, pp. 498–501, 2009.

[18] M. Alessandrini, A. Basarab, L. Boussel, X. Guo, A. Serusclat, D. Friboulet, D. Kouam, O. Bernard, and H. Liebgott, "A new technique for the estimation of cardiac motion in echocardiography based on transverse oscillations: a preliminary evaluation in silico and a feasibility demonstration in vivo," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1148–1162, 2014.

[19] A. R. Porras, M. De Craene, N. Duchateau, M. Sitges, B. H. Bijnens, A. F. Frangi, and G. Piella, "Myocardial motion estimation combining tissue doppler and b-mode echocardiographic images," in *MICCAI*, vol. 8150 of *LNCS*, pp. 484–491, 2013.

[20] M. Wilhelm, L. Roten, H. Tanner, J. P. Schmid, I. Wilhelm, and H. Saner, "Long-term cardiac remodeling and arrhythmias in nonelite marathon runners," *Am J Cardiol*, vol. 110, pp. 1060–1065, 2012.

[21] B. M. Pluim, A. H. Zwinderman, A. van der Laarse, and E. E. van der Wall, "The athlete's heart: a meta-analysis of cardiac structure and function," *Circulation*, vol. 101, pp. 336–344, 2000.

[22] A. Myronenko, X. Song, and D. J. Sahn, "Maximum likelihood motion estimation in 3D echocardiography through non-rigid registration in spherical coordinates," in *FIMH* (N. Ayache, H. Delingette, and M. Sermesant, eds.), vol. 5528 of *LNCS*, pp. 427–436, Springer Berlin Heidelberg, 2009.

[23] D. Cohen and I. Dinstein, "New maximum likelihood motion estimation schemes for noisy ultrasound images," *Pattern Recogn.*, pp. 455–463, 2002.

[24] G. Piella, M. D. Craene, C. Butakoff, V. Grau, C. Yao, S. Nedjati-Gilani, G. P. Penney, and A. F. Frangi, "Multiview diffeomorphic registration: Application to motion and strain estimation from 3D echocardiography," *Med Image Anal*, vol. 17, no. 3, pp. 348 – 364, 2013.

[25] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J. Sci. Comput.*, vol. 16, pp. 1190–1208, Sept. 1995.

[26] M. De Craene, S. Marchesseau, B. Heyde, H. Gao, M. Alessandrini, O. Bernard, G. Piella, A. R. Porras, E. Saloux, L. Tautz, A. Hennemuth, A. Prakosa, H. Liebgott, O. Somphone, P. Allain, S. M. Ebeid, H. Delingette, M. Sermesant, and J. D'hooge, "3D strain assessment in ultrasound (straus): A synthetic comparison of five tracking methodologies.," *IEEE Trans Med Imaging*, vol. 32, no. 9, pp. 1632–1646, 2013.

[27] J. Allard, S. Cotin, F. Faure, P.-J. Bensoussan, F. Poyer, C. Duriez, H. Delingette, and L. Grisoni, "SOFA - an open source framework for medical simulation," in *MMVR 15*, pp. 1–6, 2007.

[28] S. Marchesseau, H. Delingette, M. Sermesant, and N. Ayache, "Fast parameter calibration of a cardiac electromechanical model from medical images based on the unscented transform.," *Biomech Model Mechanobiol*, vol. 12, pp. 815–831, Aug 2013.

[29] H. Gao, H. F. Choi, P. Claus, S. Boonen, S. Jaecques, G. Van Lenthe, G. Van der Perre, W. Lauriks, and J. D'hooge, "A fast convolution-based methodology to simulate 2-D/3-D cardiac ultrasound images," *IEEE Trans Ultrason Ferroelectr Freq Control*, vol. 56, no. 2, pp. 404–409, 2009.

[30] J. A. Jensen, *Estimation of Blood Velocities Using Ultrasound, A Signal Processing Approach*. Cambridge University Press, 1996.

[31] V. Tavakoli, M. F. Stoddard, and A. A. Amini, "Improved cardiac motion detection from ultrasound images using TDIOF: a combined B-mode/ tissue Doppler approach," in *Proc. SPIE*, vol. 8672, pp. 86720C–86720C–7, 2013.