Registration of Multiview Echocardiography Sequences Using a Subspace Error Metric

Devis Peressutti*, Alberto Gomez, Graeme P. Penney, and Andrew P. King

Abstract—Objective: 3-D+t echocardiography (3DtE) is widely employed for the assessment of left ventricular anatomy and function. However, the information derived from 3DtE images can be affected by the poor image quality and the limited field of view. Registration of multiview 3DtE sequences has been proposed to compound images from different acoustic windows, therefore improving both image quality and coverage. We propose a novel subspace error metric for an automatic and robust registration of multiview intrasubject 3DtE sequences. *Methods:* The proposed metric employs linear dimensionality reduction to exploit the similarity in the temporal variation of multiview 3DtE sequences. The use of a low-dimensional subspace for the computation of the error metric reduces the influence of image artefacts and noise on the registration optimization, resulting in fast and robust registrations that do not require a starting estimate. *Results:* The accuracy, robustness, and execution time of the proposed registration were thoroughly validated. Results on 48 pairwise multiview 3DtE registrations show the proposed error metric to outperform a state-of-the-art phase-based error metric, with improvements in median/75th percentile of the target registration error of 21%/31% and an improvement in mean execution time of 45%. *Conclusion:* The proposed subspace error metric outperforms sum-of-squared differences and phase-based error metrics for the registration of multiview 3DtE sequences in terms of accuracy, robustness, and execution time. *Significance:* The use of the proposed subspace error metric has the potential to replace standard image error metrics for a robust and automatic registration of multiview 3DtE sequences.

Index Terms—Dimensionality reduction, echocardiography, multiview registration, principal component analysis (PCA) error metric.

I. INTRODUCTION

ECHOCARDIOGRAPHY imaging is routinely employed for the assessment of left ventricular (LV) anatomy and

function due to its high temporal and spatial resolution, nonionizing nature, low cost, and portability. In the last decade, particular interest has been focused on 3-D+t echocardiography (3DtE), which allows volumetric imaging of the heart for accurate estimation of LV indices, such as mass, ejection fraction, and volume [1]. However, 3DtE suffers from characteristic image artefacts, such as tissue inhomogeneities, multiple and off-axis reflections, and shadowing that degrade image quality [2], [3]. Furthermore, the imaging field of view is typically limited and often does not cover the whole LV anatomy. Moreover, LV structures appear significantly different depending on the 3DtE beam incidence angle and acoustic window, therefore impairing the assessment of LV anatomy and function.

Compounding or fusion of multiview 3DtE sequences has been proposed to overcome these issues and increase both 3DtE image quality and coverage. Several studies [4]–[8] have shown the advantages of combining 3DtE images from multiple acoustic windows for a range of applications, from LV segmentation to motion and strain estimation. Fusion of multiview 3DtE sequences requires the correct registration of the LV geometry from different acoustic windows. The accuracy and robustness of this registration are affected by the 3DtE angle-dependent image quality and image artefacts, which make standard image error metrics, such as sum-of-squared differences (SSD) or normalized cross correlation, unsuitable for this task [9].

A phase-based error metric based on the monogenic signal [10] has been proposed to overcome the limitations of intensity-based error metrics. This technique combines phase and orientation images to derive an error metric which, unlike intensity, is invariant to changes in brightness and contrast. As a result, the phase-based error metric has been successfully employed in several 3DtE compounding techniques [4]–[8]. However, as highlighted in [10], a reasonably good starting initialization is required to avoid local minima during optimization. Furthermore, the computation and optimization of the phase-based error metric can be computationally expensive when considering multiple cardiac phases.

Due to advances in medical imaging technology, 3-D+t medical images have become increasingly available for spatiotemporal analysis of organs of interest. Registration techniques have been accordingly adapted to exploit the additional information provided by the temporal dimension for the performance of 3-D+t to 3-D+t image registration [11], [12]. For instance, Schreibmann *et al.* [11] proposed registration of 3-D+t computed tomography (CT) datasets for radiotherapy applications, while Peyrat *et al.* [12] proposed to use trajectory constraints, also for the registration of 3-D+t CT

cardiac sequences. By exploiting the temporal dimension, these 3-D+t registrations proved to be more robust than 3-D-to-3-D registrations. However, methods relying on intensity-based error metrics are limited by the accuracy and robustness of these metrics when applied to echocardiography images.

In this paper, we propose a novel error metric for the intrasubject registration of multiview 3DtE sequences. The proposed metric exploits the similarity in the temporal variation of multiview 3DtE sequences. In particular, 3DtE sequences are projected from the image domain onto a lower dimensional subspace defined by a principal component analysis (PCA) of the 3DtE temporal sequences. This PCA projection reduces the influence of image artefacts and noise on the registration optimization, resulting in a fast and robust registration that has a large capture range.

Dimensionality reduction techniques have been previously employed for the extraction of spatial and temporal features for 3-D multimodal registration [13], [14] and 3-D+t image reconstruction [15]. In [14], manifold learning was employed to derive a structural representation of multimodality 3-D images, such that intensity-based error metrics could be used on the structural 3-D images rather than the original multimodality images. Similarly, in [15], manifold learning was employed to derive a low-dimensional representation of the image sequences for the reconstruction of 3-D+t images from 2-D sequences. Although the concept of using dimensionality reduction techniques for the extraction of the underlying data structure is similar, our proposed registration differs significantly from previous works in the way the spatial and temporal similarity in variation is simultaneously exploited for the derivation of an accurate and robust error metric. A different approach was proposed in [16], where the physical properties of echo imaging were exploited by computing an error metric in a spherical coordinate system. Our proposed approach exploits the spatiotemporal covariances of the echocardiographic sequences.

This paper is structured as follows. The proposed method is described in Section II, while the materials and experiments used to validate the proposed technique are presented in Section III and IV, respectively. Results of the validation experiments are reported in Section V, while a discussion of the proposed technique and the key findings is reported in Section VI.

II. METHOD

The novelty of the proposed multiview 3DtE sequences registration lies in the use of linear dimensionality reduction for the projection of the input sequences onto a lower dimensional subspace for the computation of the error metric. In the following, we detail the method in the context of a pairwise registration of two 3DtE sequences. An overview of the proposed method is shown in Fig. 1.

We denote by $I_t = \{I_{t_i}\}_{i=1}^{N_t}$ the target 3DtE sequence defined over the domain Ω_{I_t} and by $I_s = \{I_{s_i}\}_{i=1}^{N_s}$ the source 3DtE sequence defined over the domain Ω_{I_s}. N_t and N_s denote the number of cardiac phases in the target and source image sequences, respectively. The aim of the registration is to find the optimal spatial mapping $\mathcal{T}$ from the source to the target image sequence [17]. In the context of intrasubject mul-

tiview 3DtE registration, we assume that the domains Ω_{I_t} and Ω_{I_s} remain constant during the sequence acquisition. In other words, we assume that the probe remained stationary during the acquisition of both sequences. Furthermore, we assume that the 3DtE sequences were acquired during consecutive breath-holds at the same respiratory position, so a 6-degree-of-freedom (DOF) rigid-body transformation suffices to compensate for the misalignments due to the different probe location and orientation. Therefore, the transformation $\mathcal{T}$ is parameterized by three translations and three rotations $\theta = [t_x, t_y, t_z, r_x, r_y, r_z]$. In the following, we also assume that N_t and N_s evenly cover the entire cardiac cycle.

PCA [18] has been widely employed for linear dimensionality reduction and feature extraction of high-dimensional data in many scientific applications, ranging from chemistry to medical imaging. In particular, PCA seeks to compute a low-dimensional linear subspace that preserves the variance of the input data. An eigendecomposition of the covariance matrix is employed to compute the directions of maximum variation of the data, also known as principal components (PCs). Compared to the input high-dimensional space, PCs better describe the underlying phenomenon that causes variation in the observed data, i.e., LV motion. For this reason, PCA generates a subspace suitable for the comparison of multiview 3DtE sequences. The proposed error metric is computed as follows.

If $N_t \neq N_s$, temporal interpolation is employed to resample both target and source sequences to the same number N of cardiac phases. After smoothing and resampling (see Section IV for details), the target sequence I_t and the transformed source sequence $\mathcal{T}(I_s, \theta)$ are represented as $\mathbf{X}_{I_t} \in \mathbb{R}^{D \times N}$ and $\mathbf{X}_{\mathcal{T}(I_s,\theta)} \in \mathbb{R}^{D \times N}$, where D is the number of voxels in the overlapping domain $\Omega = \Omega_{I_t} \cap \Omega_{\mathcal{T}(I_s,\theta)}$. That is, the column vectors of $\mathbf{X}_{I_t}$ and $\mathbf{X}_{\mathcal{T}(I_s,\theta)}$ represent single vectorized 3DtE echo images, and the number of columns is the number N of cardiac phases. Since the dimensionality of $\mathbf{X}_{I_t}$ and $\mathbf{X}_{\mathcal{T}(I_s,\theta)}$ is much larger than the number of observations, i.e., $N \ll D$, a dual formulation [18] of PCA is employed to avoid the computation of a $D \times D$ covariance matrix. After removal of the mean value $\bar{\mathbf{X}}_{I_t} \in \mathbb{R}^D$, the eigendecomposition of the Gram matrix

$$\mathbf{Q}_{I_t} = \mathbf{X}_{I_t}^T \mathbf{X}_{I_t} \in \mathbb{R}^{N \times N} \qquad (1)$$

provides the diagonal eigenvalue matrix $\mathbf{E} \in \mathbb{R}^{N \times N}$ and eigenvector matrix $\mathbf{V} \in \mathbb{R}^{N \times N}$. By retaining only the largest $d \leq N \ll D$ eigenvalues $\hat{\mathbf{E}} \in \mathbb{R}^{d \times d}$ and associated eigenvectors $\hat{\mathbf{V}} \in \mathbb{R}^{N \times d}$, the projection matrix is computed as

$$\mathbf{U}_{I_t} = \mathbf{X}_{I_t} \hat{\mathbf{V}} \hat{\mathbf{E}}^{-\frac{1}{2}} \in \mathbb{R}^{D \times d}. \qquad (2)$$

This way, the target 3DtE sequence is represented in the lower dimensional subspace by

$$\mathbf{Z}_{I_t} = \mathbf{U}_{I_t}^T \mathbf{X}_{I_t} \in \mathbb{R}^{d \times N}. \qquad (3)$$

Since I_t and I_s are acquired using the same imaging modality, we can project $\mathbf{X}_{\mathcal{T}(I_s,\theta)}$ (after mean value $\bar{\mathbf{X}}_{\mathcal{T}(I_s,\theta)} \in \mathbb{R}^D$ removal) onto the subspace of the target sequence, obtaining the representation of the transformed source sequence in the target sequence subspace

$$\mathbf{Z}_{\mathcal{T}(I_s,\theta) \to I_t} = \mathbf{U}_{I_t}^T \mathbf{X}_{\mathcal{T}(I_s,\theta)} \in \mathbb{R}^{d \times N}. \qquad (4)$$
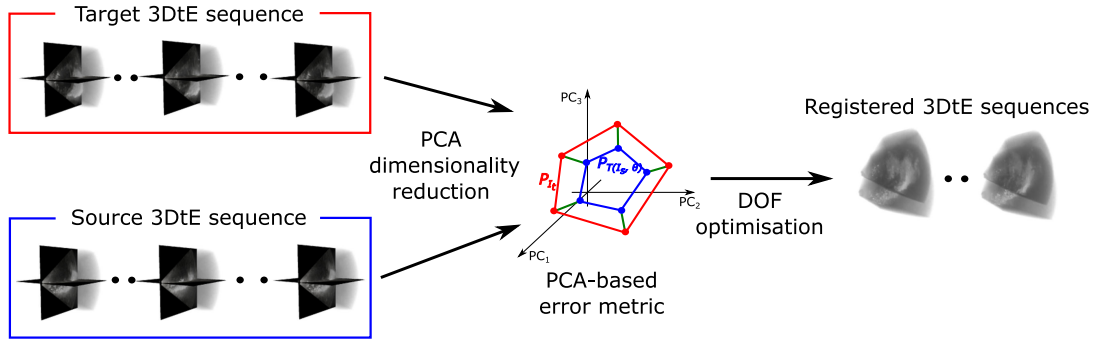
Fig. 1. Overview of the proposed method. PCA is applied to the input 3DtE sequences to compute a low-dimensional subspace that captures the variation due to the LV motion. The proposed error metric is computed in the low-dimensional subspace during the optimization of the DOFs of the transformation.
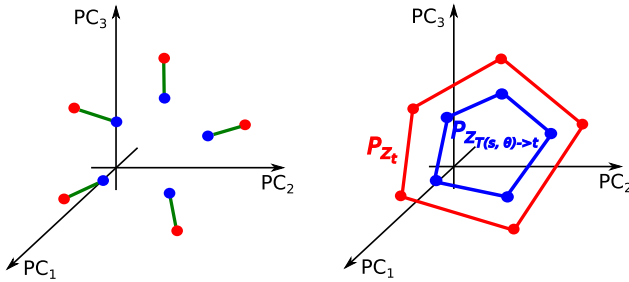

Fig. 2. Toy example of computation of the proposed error metric. Each red dot represents a 3DtE target image in the subspace $\{z_{I_{t_n}}\}_{n=1}^5$, while blue dots represent a transformed 3DtE source image projected into the same subspace $\{z_{\mathcal{T}(I_{s_n},\theta)\to I_t}\}_{n=1}^5$. In the left figure, the sum of the green segments represents the first term of (5), while the second term in (5) seeks to minimize the difference between the perimeters of the polygons formed by the red dots and the blue dots (see the right figure).

The proposed error metric is then formulated as

$$C(\theta) = \gamma \cdot \sum_{n=1}^{N} \| z_{I_{t_n}} - z_{\mathcal{T}(I_{s_n},\theta)\to I_t} \|_2^2$$
$$+ (1-\gamma) \cdot \left(\mathcal{P}_{\mathbf{Z}_{I_t}} - \mathcal{P}_{\mathbf{Z}_{\mathcal{T}(I_s,\theta)\to I_t}} \right) \quad (5)$$

where $z_{I_{t_n}} \in \mathbb{R}^d$ and $z_{\mathcal{T}(I_{s_n},\theta)\to I_t} \in \mathbb{R}^d$ are the PC weights of the nth target and transformed source image, respectively, and $\mathcal{P}$ denotes the perimeter of the closed d-dimensional polygons having $\{z_{I_{t_n}}\}_{n=1}^N$ and $\{z_{\mathcal{T}(I_{s_n},\theta)\to I_t}\}_{n=1}^N$ as vertices. The first term of (5) represents a measure of the distance between the target sequence and the transformed source sequence in the subspace, while the second term ensures that the overall variation captured by the transformed source sequence matches the original target sequence variation. Effectively, the second term prevents incorrect transformations that would overlap regions of noise only, i.e., with no cardiac motion variation. We employed this relatively simple perimeter-based measure to quantify the overall variation, but more complex indices could be employed instead. The parameter γ is the only free parameter of the proposed method and weights the contribution of the two terms. A toy example of the contributions of the two terms to the error metric is shown in Fig. 2. It is worth noting that the standard intensity-based SSD error metric corresponds to the first term of

(5) when an identity projection $\mathbf{U}_{I_t} = \mathbb{I} \in \mathbb{R}^{D \times D}$ is employed in (3).

The optimal rigid-body transformation parameters $\hat{\theta}$ are found by minimizing the proposed error metric

$$\hat{\theta} = \mathrm{argmin}_\theta C(\theta). \quad (6)$$

In this study, a simple hill-climbing algorithm is employed for the optimization of (6). At each iteration, a step is taken in the direction of the DOF providing the largest gain in the minimization of (5).

III. MATERIALS

For the validation of the proposed multiview 3DtE sequence registration, four healthy subjects were imaged. A iE33 3-D real-time echocardiography system with a X31 3 to 1-MHz broadband matrix array transducer (Philips Healthcare) was employed to acquire the 3DtE sequences. Infrared light-emitting diodes were rigidly attached to the echo probe to enable the tracking of its spatial position using an optical tracking system (Optotrak 3020, Northern Digital Inc.) [19], [20]. Before image acquisition, the probe was calibrated using the method described in [21]. For each subject, four to five 3DtE sequences were acquired from apical and modified parasternal acoustic windows during consecutive breath-holds at end-exhale. Images were acquired with the subject lying in a supine position. The 3DtE sequences were cardiac gated at late diastole by synchronizing image acquisition with the electrocardiogram signal. Over all four subjects, on average, 15 3DtE images were acquired for each sequence, with a minimum of eight and a maximum of 19 cardiac phases.

The use of optical tracking allowed computation of the rigid-body transformation mapping the position of the probe at a given acoustic window with respect to any other acoustic window within a 2-mm error range [19], [20]. This rigid-body transformation was used for validation purposes only and constituted the ground-truth for the experiments described in Section IV-A and IV-C.

IV. EXPERIMENTS

The aim of the experiments was to validate the proposed technique in terms of registration accuracy, capture range, and execution time. In order to achieve this, three separate
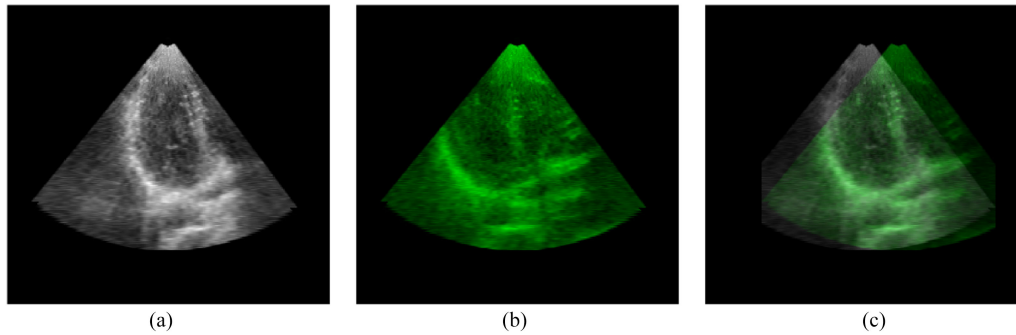
Fig. 3. First cardiac phase of (a) the 2-D target sequence and (b) the 2-D source sequence. (c) Overlay of the registered sequences using the ground-truth transformation provided by the optical tracking. In this example, a translation of $20\,\mathrm{mm}$ in the x-direction aligns the source to the target sequence.

experiments were carried out. In the first experiment, the proposed error metric was tested on a 2-D+t echocardiography (2DtE) sequence registration (see Section IV-A). In the second experiment, for each subject, each 3DtE sequence was transformed with a randomly generated rigid-body transformation and the transformed sequence was registered to the original sequence (see Section IV-B). In the third set of experiments, for each subject, each 3DtE sequence was registered to all other sequences from the same subject (see Section IV-C).

In order to evaluate the proposed method, the following error metrics were compared:

NR: No registration was performed between target and source sequence, meaning that the estimated translations and rotations were set to 0. Given the availability of the ground-truth transformation, this method quantifies the difference in probe location and orientation between the sequences being registered.

SSD: Target and source 3DtE sequences were registered using the intensity-based SSD error metric, computed over all images in each sequence. As mentioned in Section II, this method corresponds to the proposed technique when only the first term of (5) is used and no dimensionality reduction is applied to the input sequences (i.e., $\mathbf{U}_{I_t} = \mathbb{I}$).

PB: Target and source 3DtE sequences were registered using the phase-based error metric proposed by Grau *et al.* [10]. For the computation of this error metric, phase and orientation images need to be computed for each cardiac phase considered. In order to reduce the computational burden, as suggested in [10], only the end-diastolic and end-systolic cardiac phases were considered. For the computation of the monogenic signal, the following wavelength of the log-Gabor filter was employed: $\omega = 4 \cdot 2^2$ [10].

PB$_a$: Same as the above technique, but all cardiac phases in the sequences were considered;

PCA: The proposed error metric. For all experiments, γ was fixed to 0.1. This value was determined empirically on a small randomly chosen subset of 3DtE sequences. The influence of γ on the registration accuracy was also investigated, as detailed in Section IV-C. The PCs retaining 99% of the input variance were employed in all the performed experiments, resulting in $d \approx N$.

The dimension of the lower dimensional subspace d is, therefore, determined by the retained variance and varies for each echo sequence. For the experiments presented in this paper, d ranged from 8 to 19.

In order to allow comparison of the execution times, the same hill climbing optimization algorithm was employed for all the techniques compared. The measured execution time described the time required to optimize the error metric only, with no consideration of the time required for the preprocessing of the sequences. Linear interpolation was employed to interpolate the intensities of the source sequence at the intermediate steps of the optimization. The algorithms were implemented in Python using the SimpleITK open source software library [22]. Details of each experiment are reported in the following sections.

A. 2-D+t Echocardiography Registration

This experiment aimed to assess the smoothness of the variation of the proposed error metric on a simple 2-D sequence registration. Given the ground-truth transformations provided by the optical tracking, two 3DtE sequences differing only by a translational transformation were selected. Two-dimensional slices were extracted from the chosen 3DtE sequences to generate two 2DtE sequences. Fig. 3 shows the first cardiac phase of the 2-D target and 2-D source sequences, along with the ground-truth registration. In this case, the ground-truth transformation was given by a 20-mm translation in the x-direction.

The parameter space $\theta_{2D} = [t_x, t_y, r]$ was evenly sampled and values for the SSD, PB, and PCA error metrics were computed for each combination of $[t_x, t_y, r]$. In this way, the variation of the compared error metrics could be visually assessed. Results of this experiment are reported in Section V-A.

B. Same-View 3-D+t Echocardiography Registration

This experiment aimed to validate the capture range and execution time of the compared error metrics. For each 3DtE sequence of each subject, ten randomly generated transformations were used to transform the sequence to a new spatial position and orientation. Translations in the x, y, z coordinates were randomly selected from a uniform distribution with $[-20\,\mathrm{mm}, 20\,\mathrm{mm}]$ range, while rotations were randomly sampled from a uniform distribution within a $[-30°, 30°]$ range. These ranges represent the range of probe location and orienta-

tion typically found in a clinical setting (e.g., between parasternal and apical views). Over all subjects, 170 registrations were performed. As a preprocessing step, the original and transformed sequences were smoothed using a Gaussian kernel with $\sigma^2 = 1$ mm and subsequently resampled to an isotropic voxel size of 2 mm. An identity transformation was used as starting estimate for all of the compared registrations. This means that the initial overlap between the two image sequences was 100% of the image volume, although the "true" overlap (i.e., after the ground-truth transformation) was ≈ 60–90% for the datasets used in this paper. The results of this experiment are reported in Section V-B.

C. Multiview 3-D+t Echocardiography Registration

This experiment aimed to validate the accuracy and robustness of the proposed technique compared to state-of-the-art registration techniques in a realistic clinical scenario. For each subject, each 3DtE sequence was registered to every other 3DtE sequence, so that any bias toward the chosen target sequence was removed from the analysis. Given the ground-truth transformation provided by the optical tracking of the echo probe (see Section III), we discarded from the analysis the registrations where the "true" overlap (i.e., after ground-truth transformation) between the registered sequences was below 50% of the field of view. This ensured that similar LV features and motion were sufficiently imaged by both the sequences being registered. Following this criterion, over all subjects, four out of 28 image sequence pairs were discarded, resulting in 24 pairs. Registrations for each pair were performed in both directions, making 48 evaluated registrations in total. As for the previous experiment, the target and source sequences were smoothed using a Gaussian kernel with $\sigma^2 = 1$ mm and subsequently resampled to a 2-mm isotropic spacing. In the cases where the number of cardiac phases differed (i.e., $N_t \neq N_s$), the sequences were temporally interpolated to $N = \max\{N_t, N_s\}$ cardiac phases using a nearest neighbor temporal interpolation. Again, an identity transformation was used as starting estimate for all of the compared registrations, meaning that the initial overlap between target and source sequences was 100% of the image volume.

In an additional experiment, the influence of γ on the PCA registration accuracy was investigated using the same multiview 3DtE data by varying the value of γ between 0 and 1 in steps of 0.2.

Results of these experiments are reported in Section V-C.

V. RESULTS

This section reports results of the experiments described in Section IV.

For the second and third experiments, given the availability of the ground-truth rigid-body transformation, the error in estimation $\epsilon_E = \theta_{GT} - \theta_E$ was computed for each of the registrations compared (see Section IV) by subtracting the estimated rigid-body transformation θ_E from the ground-truth rigid-body DOFs θ_{GT}. The registration accuracy was quantified by the norm of the estimation error ϵ_E for translations and rotations separately. Furthermore, for the multiview registration, target reg-

TABLE I
RESULTS OF THE SAME-VIEW 3DtE REGISTRATION VALIDATION

Method	Norm of Translation Error median/75th percentile (mm)	Norm of Rotation Error median/75th percentile (°)	Execution Time mean/std (s)
NR	19.5/22.5	30.8/34.7	NA
SSD	0.6/102.5	0.06/33.7	656/770
PB	1.4/37.5	3.7/30.7	1228/571
PB$_a$	1.6/38.9	4.0/30.3	7802/2063
PCA	0.4/0.6	0.04/0.05	776/265

Due to the skewness of the error distribution (see Fig. 5), the median and 75th percentile of the norms of the translation and rotation errors are reported. Results are reported over all subjects. The norm of the translation error is reported in millimeters, while the norm of the rotation error is in degrees. The mean and standard deviation of the execution time is reported in seconds. NR quantifies the amount of starting misalignment between sequences. The execution time of NR is reported as not applicable.

istration errors [23] were computed using as targets the world coordinates of each voxel in the overlapping domain of the target sequence and the source sequence transformed with the ground-truth transformation.

Results of the visual assessment of the smoothness of the proposed error metric are reported in Section V-A. Results of the same-view registration are reported in Section V-B, while Section V-C reports results of the multiview registration.

A. 2-D+t Echocardiography Registration

Values of the compared error metrics, i.e., SSD, PB, and PCA, were computed over a sample of evenly distributed points in the parameter space $\theta_{2D} = [t_x, t_y, r]$. Translations t_x, t_y were sampled in the range $[-30 \text{ mm}, 30 \text{ mm}]$ with a 5-mm step size, while the rotation was sampled in the range $[-30°, 30°]$ with a $5°$ step size. Images illustrating the error metric values within the parameter space for the SSD error metric, PB error metric, and the proposed PCA error metric are shown in Fig. 4. To allow direct visual comparison, values for all three measures are normalized between 0 and 1.

All three compared error metrics showed a minimum value at the correct transformation $\theta_{2D} = [20 \text{ mm}, 0 \text{ mm}, 0°]$. However, differences can be noted between the distributions of the compared error metrics. SSD shows a smoother distribution compared to both PB and PCA, although areas with low values of the error metric are more localized in the case of PB and PCA. In particular, it can be noted that PCA shows a very localized region of low metric values compared to the other metrics (middle frames of Fig. 4). This is likely to lead to better precision of the technique. Note that, for the dataset shown in Fig. 4, the optimizations for all error metrics converged to the global minimum.

B. Same-View 3-D+t Echocardiography Registration

Results of the same-view registration are reported in Table I and Fig. 5.

In this experiment, the ground-truth rigid-body transformation θ_{GT} was given by the randomly generated transformation. All the compared techniques were employed on the same randomly transformed sequence. The norm of the translation error is reported in millimeters, while the norm of the rotation error is reported in degrees.
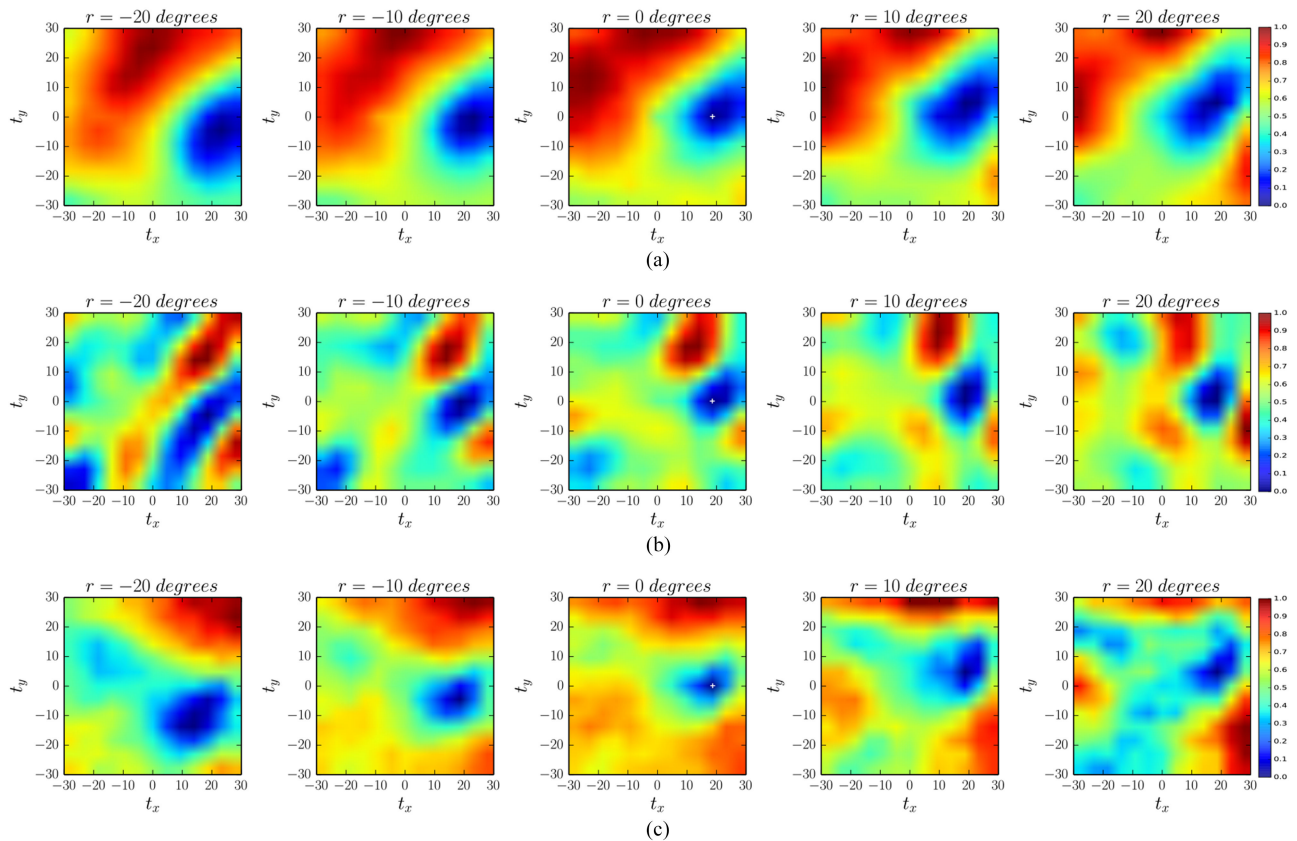
Fig. 4. Images showing the error metrics within the parameter space $\theta_{2D} = [t_x, t_y, r]$ for the 2DtE sequence registration. Values of error metric are normalized within the $[0, 1]$ range. (a) SSD error metric, (b) PB error metric, and (c) PCA error metric. All three metrics correctly show a minimum at the transformation $[20 \text{ mm}, 0 \text{ mm}, 0°]$, as indicated by a white cross.

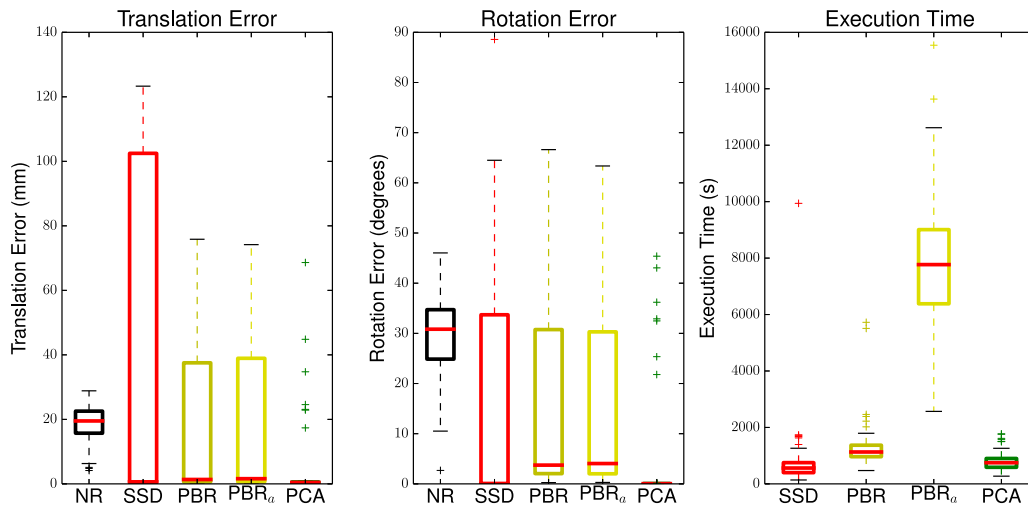Fig. 5. Boxplots of the norm of translation error (mm), rotation error (°), and execution time (s) for the same-view registration validation. NR quantifies the amount of starting misalignment between sequences.

From Table I and Fig. 5, it can be noted that the proposed error metric provides by far the best registration results. Considering as a failed registration one with a norm of the translation error > 10 mm or a norm of the rotation error $> 10°$, out of 170 performed registrations, SSD failed 78 times, PB failed 67 times, PB$_a$ failed 71 times, while PCA failed seven times only. The registration based on SSD proved to be the least robust, in

accordance with the findings in [9]. This is probably due to the influence of noise and image artefacts, which make the optimization of this error metric highly prone to local minima. Similarly, phase-based registration failed in several cases, confirming the fact that this method requires a suitable initialization for a correct outcome, as highlighted in [10]. Even with the use of all cardiac phases, PB$_a$ did not produce better results compared

TABLE II
RESULTS OF THE MULTIVIEW 3DtE REGISTRATION VALIDATION

Method	Norm of Translation Error median/75th perc (mm)	Norm of Rotation Error median/75th perc (°)	Registration Error median/75th perc (mm)	Execution Time mean/std (s)
NR	18.4/23.6	12.4/16.1	20.5/29.9	NA
SSD	7.2/18.9	9.0/16.1	11.2/26.2	435/237
PB	3.4/5.3	4.6/8.2	4.8/8.6	1068/409
PCA	2.9/4.0	3.5/5.6	3.8/5.9	583/281

The median and 75th percentile of the norm of the translation (mm) and rotation (°) errors are reported, as well as of the target registration errors (mm). Results are reported over all subjects. The mean and standard deviation of the execution time is reported in seconds. NR quantifies the amount of starting misalignment between sequences. The execution time of NR is reported as not applicable.
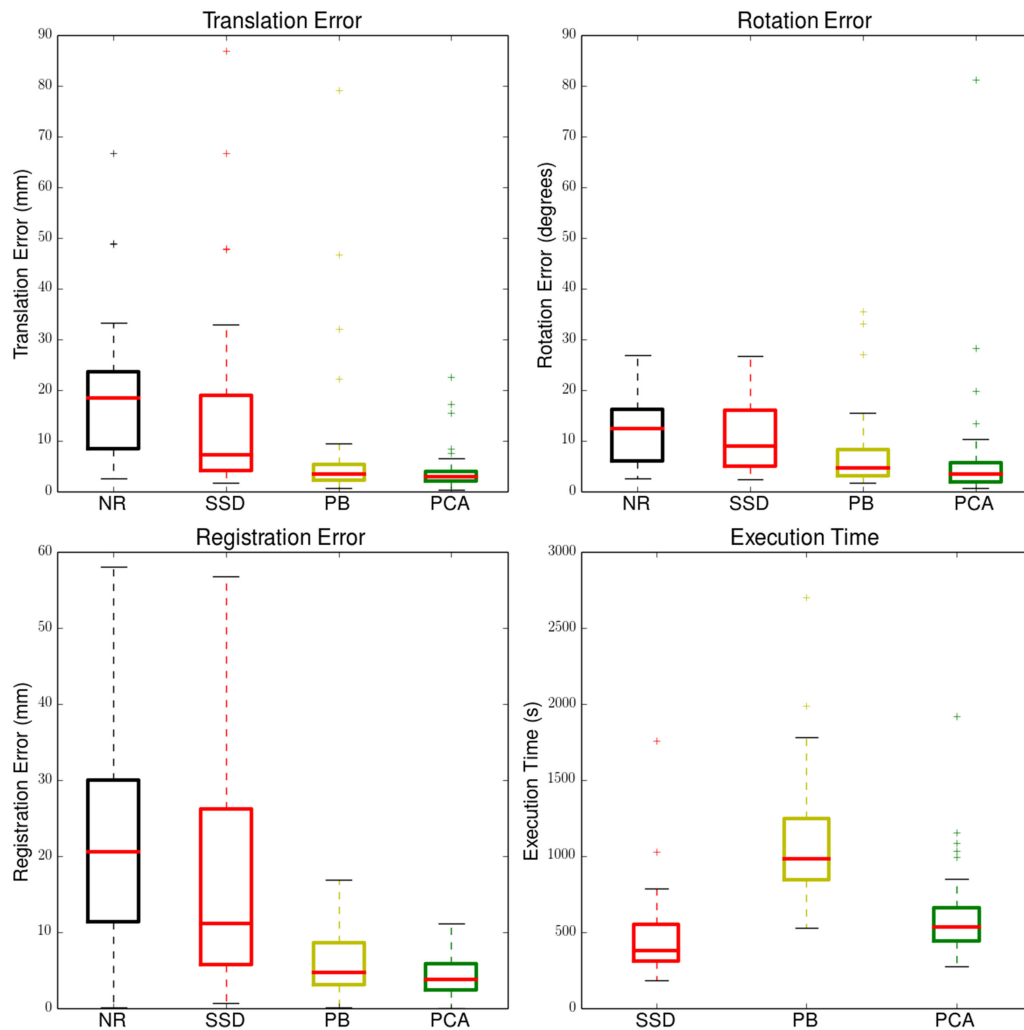

Fig. 6.   Boxplots of the norm of translation error (mm), rotation error (°), target registration error (mm), and execution time (s) for the multiview registration validation. NR quantifies the amount of starting misalignment between sequences.

to the use of the end-diastolic and end-systolic phases only, but did increase the execution time. For this reason, the PB$_a$ method was not considered in the assessment of the multiview registration.

In terms of execution time, the SSD registration was the fastest, although many registrations failed and ended prematurely. The execution time of the proposed PCA error metric was about half that of the PB metric.

C. Multiview 3-D+t Echocardiography Registration

In this experiment, the ground-truth rigid-body transformation θ_{GT} was provided by the optical tracking of the echo probe (see Section III). The norm of the translation error is reported in millimeters, the norm of the rotation error is reported in degrees, while target registration errors are reported in millimeters. Results are reported in Table II and Fig. 6.
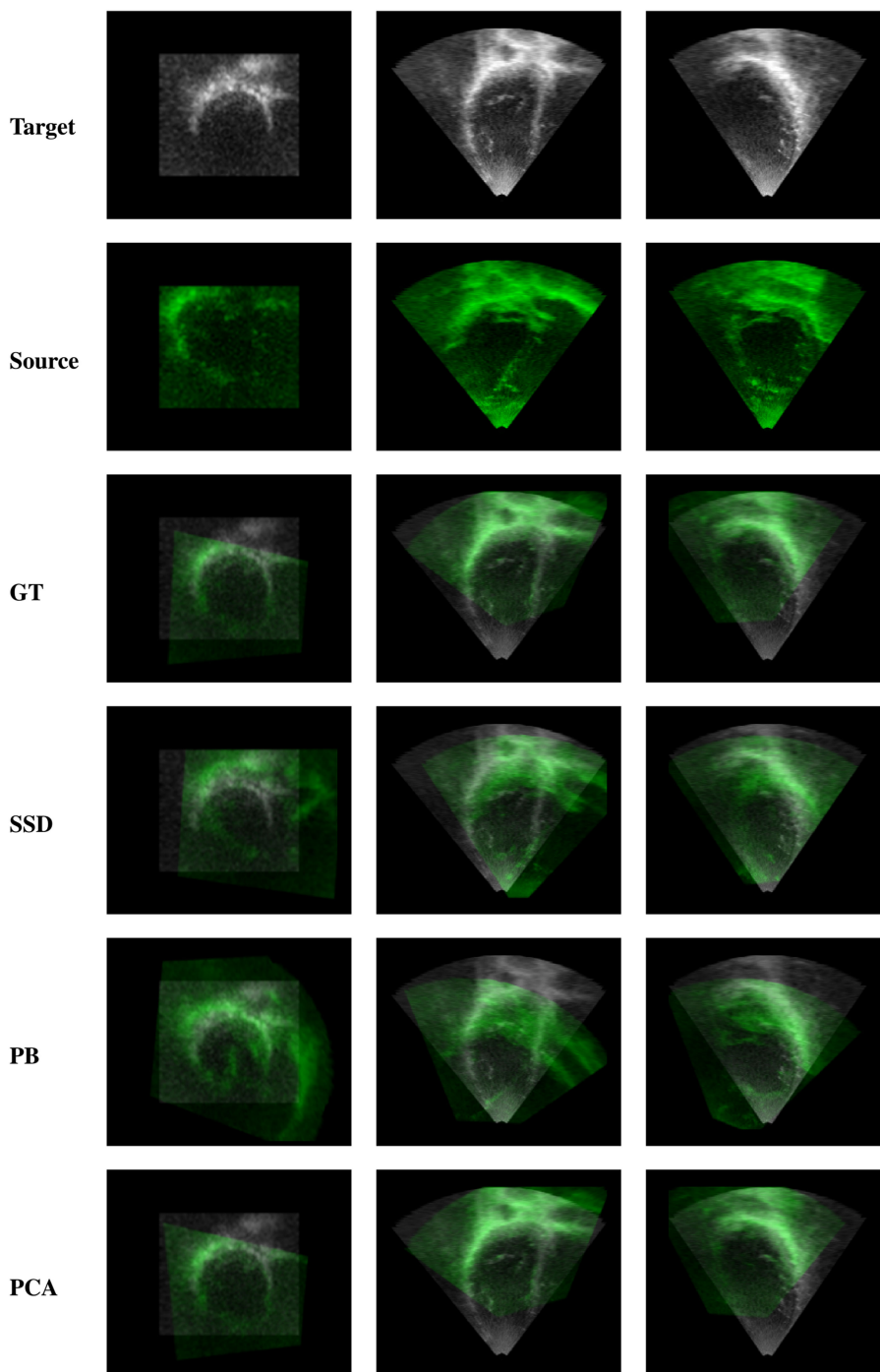
Fig. 7. Example of successful PCA-based registration. The top two rows show three orthogonal views of the end-diastolic phase of the target and source 3DtE sequences, respectively. The other rows show the same views for, from top to bottom, the ground-truth transformation provided by the optical tracking, SSD registration, PB registration, and the proposed PCA registration. Only the PCA-based registration correctly estimated the rigid-body transformation.

Similarly to the results reported in Section V-B, the proposed PCA error metric provided the most accurate registrations with a reduced computational time. Considering as a failed registration one with a norm of the translation error > 10 mm or a norm of the rotation error $> 10°$, out of 48 evaluated registrations, SSD failed 26 times, PB failed 11 times, while PCA failed six times. Compared to PB, the proposed technique showed an improvement in registration accuracy in median/75th percentile of 14.7%/24.5% for the translation, of 23.9%/31.7%

for the rotation, and of 20.8%/31.4% for the target registration error. The improvement in mean execution time was 45.4%. The overall execution time of the SSD-based registration was lower than the proposed PCA registration due to the incorrect registrations that ended prematurely at a local minimum. The average execution time of a single cost function evaluation was 0.66 ms for SSD, 0.69 ms for PCA, and 1.37 ms for PB. For a typical successful registration with a ground-truth translation norm of ≈ 17 mm and norm of rotation of $\approx 5°$, the SSD, PCA,
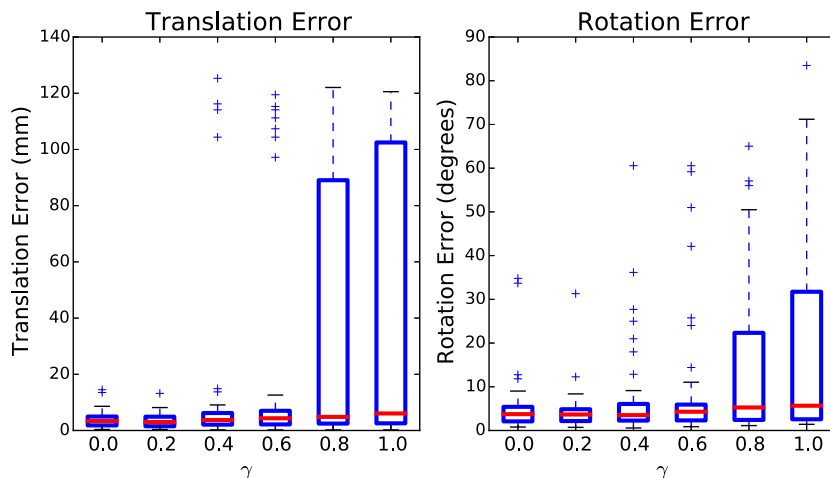
Fig. 8.   Influence of γ on the PCA registration accuracy. Boxplots report the norm of translation error (mm) and rotation error (°) for varying values of γ between 0 and 1 in steps of 0.2. The results reported in Table II and Fig. 6 were achieved using $\gamma = 0.1$.
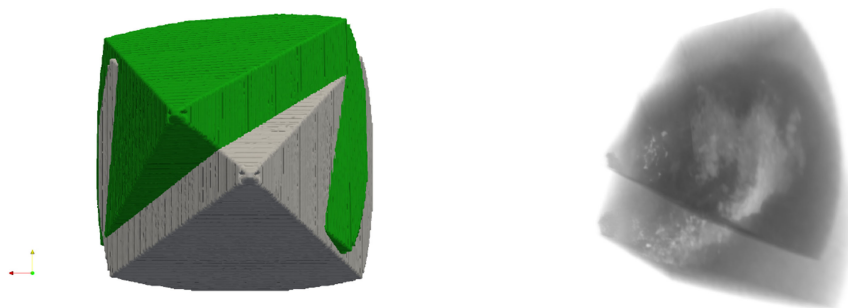

Fig. 9.   Example of capture range of the proposed registration. Registered source 3DtE sequence (green frustum) is shown with respect to the target 3DtE sequence (gray frustum) for the successful PCA-based registration reported in Fig. 7.

and PB-based methods converged to a correct estimation in 23, 25, and 36 iterations, respectively.

The effect of γ on the PCA registration accuracy is illustrated in Fig. 8. As mentioned in Section II, the second term of the proposed error metric ensures that the variation captured by the registered sequences corresponds to the variation due to the cardiac cycle motion. We can see that using high values of γ (i.e., using mostly the first term in (5), which measures the distance between the imaging sequences in the subspace) results in poor registration performance, as it produces transformations that match regions with little or no cardiac cycle motion. Lower values of γ avoid such transformations, and the optimal value can be seen to be close to the 0.1 value used in the main experiments.

An example of a successful PCA registration is shown in Fig. 7, while Fig. 9 shows an example of the capture range of the proposed 3DtE registration.

VI. DISCUSSION

Results of the evaluation showed the proposed PCA-based registration to outperform state-of-the-art registration in terms of accuracy, capture range, and execution time. In the multiview 3DtE registration evaluation, over the 48 evaluated pairwise registrations, the proposed registration showed an improvement in registration accuracy in median/75th percentile of 20.8%/31.4% for the target registration error compared to a phase-based registration. Compared to the same method, the improvement in mean execution time was 45.4%. Results for accuracy and

execution time show that the proposed technique has potential for applications where such requirements are paramount, as in the case of image-guided interventions [24], [25], as well as for the fusion of multiview 3DtE sequences. Furthermore, our results show an improved capture range, meaning that no starting estimate for the transformation is required. We believe that the better capture range is due to the fact that our technique takes advantage of the temporal variation of the image sequences, i.e., it is able to use motion information as well as image information in the registration. When image artefacts are present, similarity metrics based on intensities or on phase may be more likely to get trapped in local minima by mistaking artefacts for boundaries. As these artefacts may not exhibit temporal consistency, the proposed method is less sensitive to the starting estimate.

Since the proposed technique exploits the similarity in the temporal variation of the 3DtE sequences, it can only be applied to image sequences and not to single cardiac phases. However, this does not represent a limit in multiview 3DtE registration, as 3-D$+t$ sequences that image the whole cardiac cycle are routinely acquired in clinical practice. As the temporal covariance between the sequences is maximized, the error metric could also be applied for the registration of multiview free-breathing liver echo images, where the respiratory motion between sequences could be exploited. However, the greater intercycle variation of respiratory motion might cause difficulties. Application of the proposed error metric to other echography imaging, such as obstetrics or fetal echo, might not be straightforward.

Our technique assumes that there has been no extra motion (e.g., due to probe movement or respiratory motion) during the acquisitions of the 3DtE sequences. If this assumption is violated, the configurations of the volumes in the reduced PCA subspace would be altered. This would affect the registration and would likely produce an estimated transformation, which averages over the extra motion.

Future work will investigate the influence of the number of cardiac phases on the registration accuracy. The proposed method was evaluated on healthy volunteer datasets only. In the case of diseased subjects with irregular heartbeats, the PCA error metric could be extended to deal with multiple heartbeats, therefore increasing the registration robustness. Moreover, more sophisticated temporal interpolations will be explored. The proposed method can also be easily extended to provide symmetric transformations, such that the same result is provided regardless of the choice of the target and source sequences. Furthermore, different deformation models (e.g., affine or nonrigid) could be employed instead of a rigid transformation.

A further interesting area of investigation is the role of the PCA subspace. In future work, we plan to test alternative dimensionality reduction techniques and to investigate the effect that different projections have on our error metric. Possible alternatives include independent component analysis or random projections.

Finally, the proposed technique could be extended to deal with intermodality registration of cardiac imaging sequences, such as cine magnetic resonance or CT. For this purpose, intermediate images representing the anatomical structure [15], [26], [27] of the anatomy could be extracted and the proposed error metric applied to these structural image sequences rather than the input sequences. This will be focus of future work.

VII. Conclusion

We have presented a novel PCA-based error metric for the registration of multiview 3DtE sequences. The proposed method exploits the underlying temporal variation of the 3DtE sequences due to LV motion to compute a novel error metric, which is robust to image noise and artefacts. Results show the proposed PCA-based registration to outperform state-of-the-art registration of 3DtE sequences in terms of accuracy, capture range, and execution time, thus showing the potential of replacing standard image error metrics for an automatic and robust registration of 3DtE sequences.

Download

Data and ground-truth rigid transformations used for the validation of the proposed technique in Section IV-C are available to download on the Zenodo open access repository (http://dx.doi.org/10.5281/zenodo.30999). Python and MATLAB implementations of the proposed PCA error metric are available to download on the GitHub open access repository (https://github.com/devisperessutti/Python.git and https://github.com/gomezalberto/Matlab.git).

References

[1] C. Jenkins *et al.*, "Reproducibility and accuracy of echocardiographic measurements of left ventricular parameters using real-time three-dimensional echocardiography," *J. Am. College Cardiol.*, vol. 44, no. 4, pp. 878–886, 2004.

[2] F. W. Kremkau and K. J. Taylor, "Artifacts in ultrasound imaging," *J. Ultrasound Med.*, vol. 5, pp. 227–237, 1986.

[3] A. J. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. Med. Imag.*, vol. 25, no. 8, pp. 987–1010, Aug. 2006.

[4] C. Szmigielski *et al.*, "Real-time 3D fusion echocardiography," *JACC: Cardiovascular Imag.*, vol. 3, no. 7, pp. 682–690, 2010.

[5] K. Rajpoot *et al.*, "The evaluation of single-view and multi-view fusion 3D echocardiography using image-driven segmentation and tracking," *Med. Image Anal.*, vol. 15, no. 4, pp. 514–528, 2011.

[6] K. Rajpoot *et al.*, "Multiview fusion 3-D echocardiography: Improving the information and quality of real-time 3-D echocardiography," *Ultrasound Med. Biol.*, vol. 37, no. 7, pp. 1056–1072, 2011.

[7] C. Yao *et al.*, "Multi-view 3D echocardiography compounding based on feature consistency," *Phys. Med. Biol.*, vol. 56, no. 18, p. 6109, 2011.

[8] G. Piella *et al.*, "Multiview diffeomorphic registration: Application to motion and strain estimation from 3D echocardiography," *Med. Image Anal.*, vol. 17, no. 3, pp. 348–364, 2013.

[9] C. Wachinger *et al.*, "Registration strategies and similarity measures for three-dimensional ultrasound mosaicing," *Acad. Radiol.*, vol. 15, no. 11, pp. 1404–1415, 2008.

[10] V. Grau *et al.*, "Registration of multiview real-time 3-D echocardiographic sequences," *IEEE Trans. Med. Imag.*, vol. 26, pp. 1154–1165, Sep. 2007.

[11] E. Schreibmann *et al.*, "Four-dimensional image registration for image-guided radiotherapy," *Int. J. Radiation Oncol. Biol. Phys.*, vol. 71, no. 2, pp. 578–586, 2008.

[12] J. Peyrat *et al.*, "Registration of 4D cardiac CT sequences under trajectory constraints with multichannel diffeomorphic demons," *IEEE Trans. Med. Imag.*, vol. 29, no. 7, pp. 1351–1368, Jul. 2010.

[13] K. Abd-Elmoniem *et al.*, "Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 9, pp. 997–1014, Sep. 2002.

[14] C. Wachinger and N. Navab, "Entropy and Laplacian images: Structural representations for multi-modal registration," *Med. Image Anal.*, vol. 16, no. 1, pp. 1–17, 2012.

[15] C. Wachinger *et al.*, "Manifold learning for image-based breathing gating in ultrasound and MRI," *Med. Image Anal.*, vol. 16, no. 4, pp. 806–818, 2012.

[16] A. Myronenko *et al.*, "Maximum likelihood motion estimation in 3D echocardiography through non-rigid registration in spherical coordinates," in *Functional Imaging and Modeling of the Heart* (ser. Lecture Notes in Computer Science), vol. 5528, N. Ayache, H. Delingette, and M. Sermesant, Eds. Berlin, Germany: Springer, 2009, pp. 427–436.

[17] D. L. G. Hill *et al.*, "Medical image registration," *Phys. Med. Biol.*, vol. 46, no. 3, p. R1, 2001.

[18] I. Jolliffe, *Principal Component Analysis*. New York, NY, USA: Springer-Verlag, 2002.

[19] K. Rhode *et al.*, "Registration and tracking to integrate X-ray and MR images in an XMR facility," *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1369–1378, Nov. 2003.

[20] K. Rhode *et al.*, "A system for real-time XMR guided cardiovascular intervention," *IEEE Trans. Med. Imag.*, vol. 24, no. 11, pp. 1428–1440, Nov. 2005.

[21] Y. L. Ma *et al.*, "Ultrasound calibration using intensity-based image registration: For application in cardiac catheterization procedures," *Proc. SPIE*, vol. 6918, pp. 69180O-1–69180O-9, 2008.

[22] B. Lowekamp *et al.*, "The design of simpleitk," *Frontiers Neuroinformat.*, vol. 7, p. 45, 2013.

[23] J. Fitzpatrick *et al.*, "Predicting error in rigid-body point-based registration," *IEEE Trans. Med. Imag.*, vol. 17, no. 5, pp. 694–702, Oct. 1998.

[24] A. J. Noble *et al.*, "Ultrasonic image analysis and image-guided interventions," *Interface Focus*, vol. 1, pp. 673–685, 2011.

[25] D. Peressutti *et al.*, "A novel Bayesian respiratory motion model to estimate and resolve uncertainty in image-guided cardiac interventions," *Med. Image Anal.*, vol. 17, no. 4, pp. 488–502, 2013.

[26] G. P. Penney *et al.*, "A novel framework for multi-modal intensity-based similarity measures based on internal similarity," *Proc. SPIE*, vol. 6914, p. 69140X, 2008.

[27] M. P. Heinrich *et al.*, "MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration," *Med. Image Anal.*, vol. 16, no. 7, pp. 1423–1435, 2012.

Authors' photographs and biographies not available at the time of publication.