



Deformable multimodal registration for navigation in beating-heart cardiac surgery

Jacob J. Peoples¹ · Gianluigi Bisleri¹ · Randy E. Ellis¹

Received: 27 January 2019 / Accepted: 1 March 2019 / Published online: 19 March 2019
© CARS 2019

Abstract

Purpose: Minimally invasive beating-heart surgery is currently performed using endoscopes and without navigation. Registration of intraoperative ultrasound to a preoperative cardiac CT scan is a valuable step toward image-guided navigation.

Methods: The registration was achieved by first extracting a representative point set from each ultrasound image in the sequence using a deformable registration. A template shape representing the cardiac chambers was deformed through a hierarchy of affine transformations to match each ultrasound image using a generalized expectation maximization algorithm. These extracted point sets were matched to the CT by exhaustively searching over a large number of precomputed slices of 3D geometry. The result is a similarity transformation mapping the intraoperative ultrasound to preoperative CT.

Results: Complete data sets were acquired for four patients. Transesophageal echocardiography ultrasound sequences were deformably registered to a model of oriented points with a mean error of 2.3 mm. Ultrasound and CT scans were registered to a mean of 3 mm, which is comparable to the error of 2.8 mm expected by merging ultrasound registration with uncertainty of cardiac CT.

Conclusion: The proposed algorithm registered 3D CT with dynamic 2D intraoperative imaging. The algorithm aligned the images in both space and time, needing neither dynamic CT imaging nor intraoperative electrocardiograms. The accuracy was sufficient for navigation in thoroscopically guided beating-heart surgery.

Keywords Cardiac surgery · Surgical navigation · Image registration · Transesophageal echocardiography · Cardiac CT

Introduction

Minimally invasive beating-heart surgery, such as ablation for persistent atrial fibrillation (AF) [1], is a technically challenging procedure that is currently performed without computer navigation. Hybrid treatment of surgical AF ablation has been shown to be effective and durable [2]. The procedure is performed by thoroscopic access from the right chest wall, so the surgeon has no direct vision when accessing the left atrium. An opportunity for navigation arises because of the extensive medical imaging involved: A preoperative cardiac computed tomography (CT) scan is routine, as are the intraoperative use of transesophageal echocardiog-

raphy (TEE) and fluoroscopy. The purpose of this work is to register two of these common data sources—TEE and CT—in support of beating-heart navigation. The complexity of the surgical procedure and examples of routinely available images are illustrated in Fig. 1.

We propose a method to register 2D TEE to 3D CT. Much existing work has relied on dynamic CT [3–5] that used electrocardiography (ECG) data captured with the TEE to temporally align the two image sources. Desires to reduce radiation exposure prompted interest in methods using static CT images. Li et al. [6] proposed synthetic dynamic CT, which fused a static CT with a preoperative dynamic 3D TEE, to construct a temporally varying CT model; this produced synthetic images with registration results of similar quality to those using dynamic CT, but required both preoperative and intraoperative 3D TEE. Other methods using static CT typically have relied on CT gating and then used ECG data collected simultaneously with TEE to achieve temporal alignment [7,8].

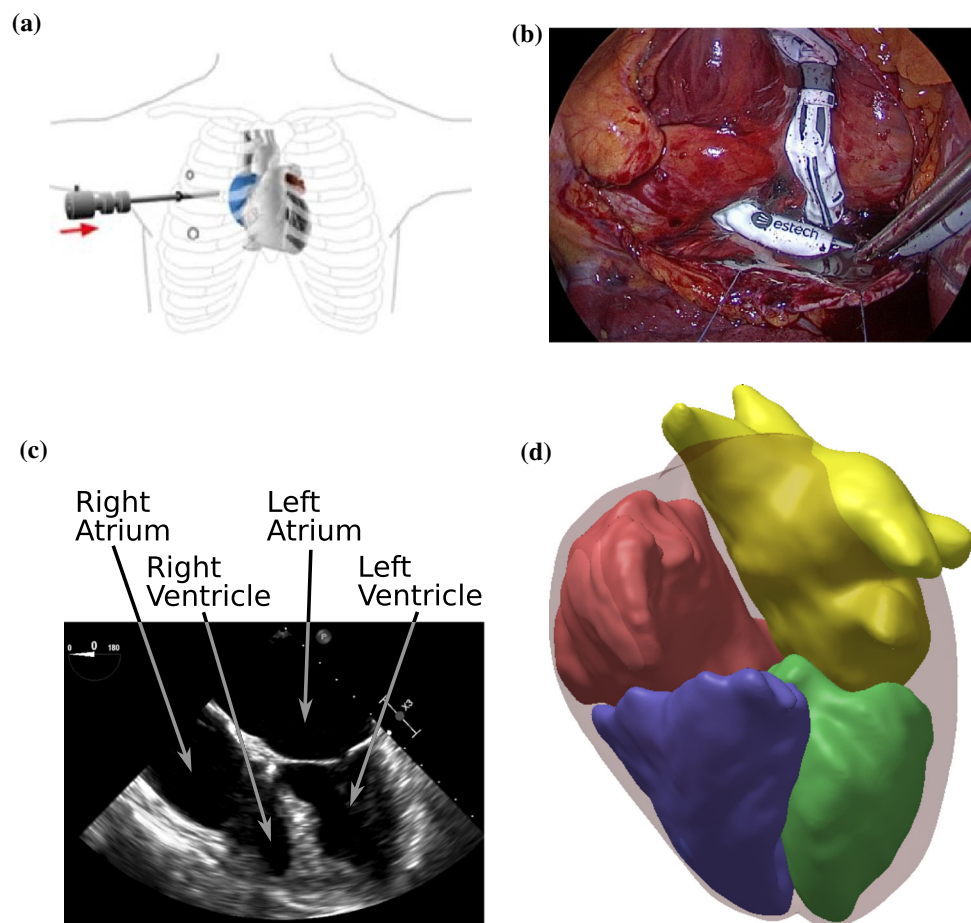
Instead, our approach relied on only two data sources: a single static CT and a 2D TEE sequence. A 3D mesh model

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11548-019-01932-2>) contains supplementary material, which is available to authorized users.

✉ Randy E. Ellis
ellis@queensu.ca

¹ Queen's University, Kingston K7L 3N6, Canada

Fig. 1 Surgical approach and routinely available image data. **a** Drawing of a thoracoscopic approach via three incisions in the right chest wall. **b** Image from a thoracoscope during beating-heart surgery, showing the ablation catheters and minimal visible landmarks. **c** TEE ultrasound image, acquired intraoperatively. **d** Surface models of the four cardiac chambers



of the four cardiac chambers was derived by segmenting the CT scan. The chamber walls were manually delineated in a single image from the TEE sequence. Given these inputs, our algorithm computed a similarity transformation that mapped the 2D ultrasound image space to 3D CT space. This problem represents a 2D/3D slice-to-volume registration with an added complication due to the dynamic nature of the 2D imaging. It is worth noting that our reliance on segmentations of the preoperative image means that other modalities, such as magnetic resonance imaging, could be used in place of CT, as long as they can be segmented with sufficient accuracy.

Our algorithm first extracted a point set that represented the visible cardiac chambers in each image of the TEE sequence by deforming a template to model variations from image to image. Using this deformed template, an ultrasound image was registered to CT space by solving a set of rigid 2D/2D registration problems. This registration was done by precomputing, from the 3D chamber models, a large number of 2D slices along different axes and positions. The optimal transformation from ultrasound to CT space was found by an exhaustive search that registered each ultrasound image to each precomputed 2D slice.

Methods in the literature for segmenting or registering heart chambers in echocardiographic images include methods based on level sets, statistical shape or appearance models, machine learning, optical flow, and nonrigid registration [9]. Our ultrasound data sets had properties that limited the applicability of existing methods. Not all cardiac chambers were visible in all images of a sequence, and some chambers were only partly visible. The TEE ultrasound views differed significantly between patients, making it difficult to generalize the appearance of the images across patients. Also, because we were working with data sets from only a few patients, we could not rely on statistical methods that required large amounts of training data.

We used a template-based nonrigid registration technique that was initialized for a single image. The template was a deformable model that consisted of a set of oriented points, with each point labeled as coming from a particular cardiac chamber. The template could be deformed by a hierarchy of affine transformations to capture the motion of the beating heart: There was a transformation for the overall template, written as A , followed by one independent affine transformation for each cardiac chamber, written as $C_1 \dots C_4$. The hierarchy of transformations is illustrated in Fig. 2. Registration through this hierarchical transformation was formulated

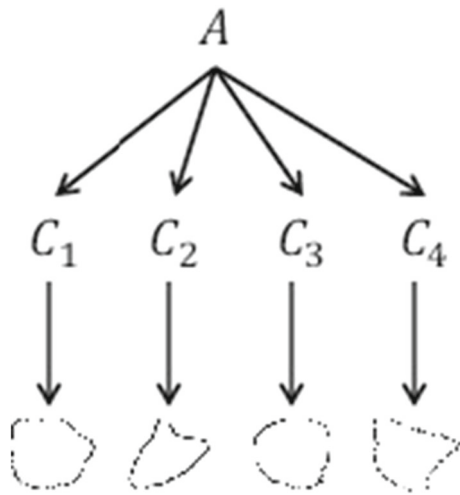


Fig. 2 A hierarchy of deformable models. The transformation A captures global deformation; A is applied to partitioned transformations $C_1 \dots C_4$ that each model one cardiac chamber

with a hybrid mixture model similar to that used in other recent work on oriented point registration [10,11], though these works were limited to rigid/similarity transformations. We optimized the transformation parameters using a novel generalized expectation maximization (GEM) algorithm that imposed temporal coherence on the extracted point sets by regularizing the transformations between adjacent images in the sequence.

These extracted ultrasound point sets were then exhaustively registered to a set of precomputed slices of the preoperative CT-derived geometry via a variant of the iterative closest point (ICP) [12] algorithm operating in 2D. This approach is in some sense a hybrid method, combining a discrete search over the space of planes with a continuous in-plane registration to obtain the complete 3D similarity transform mapping ultrasound to CT.

Discrete methods of slice-to-volume 2D/3D registration are enticing as they tend to be less prone to local minima than continuous alternatives [13]. Typically, discrete methods are formulated as labeling problems on Markov random fields (MRF). For nonrigid registration, the typical approach is for the nodes of the MRF to represent control points of a grid which are labeled by discrete parameters encoding local details of the transformation [14–17]. These methods interpret the deformed grid to infer a rigid transformation and a 2D in-plane deformation field. Discrete methods for linear transformations (such as rigid, similarity, and affine transformations) are less easily encoded this way, because all transformation parameters are inherently global. To solve this problem, Zikic et al. [18] proposed an approximation to the linear transform registration energy consisting entirely of second-order terms on a fully connected graph with one node per transformation parameter; however, their framework was intended for monodimensional registration, or 2D/3D regis-

tration with projective images, such as fluoroscopy. Porchetto et al. [19] subsequently proposed a framework for rigid slice-to-volume registration using a similar MRF encoding. In spite of their promising results, none of the above-described slice-to-volume registration methods also account for the temporal alignment that we consider in the present work.

The primary contribution of this paper is a novel CT to TEE registration pipeline relying on neither ECG for temporal registration nor on ECG-gated dynamic CT. To achieve this, we present a novel GEM algorithm for nonrigid registration of ultrasound image sequences, and a simple discretized approach to dynamic slice-to-volume registration. We present the results of our registration on data sets from four patients who underwent minimally invasive epicardial ablation.

Methods

This study was conducted in accordance with the principles outlined in the Declaration of Helsinki. With institutional review board (IRB) approval for retrospective image analysis, preoperative CT images and intraoperative TEE sequences were acquired from four patients.

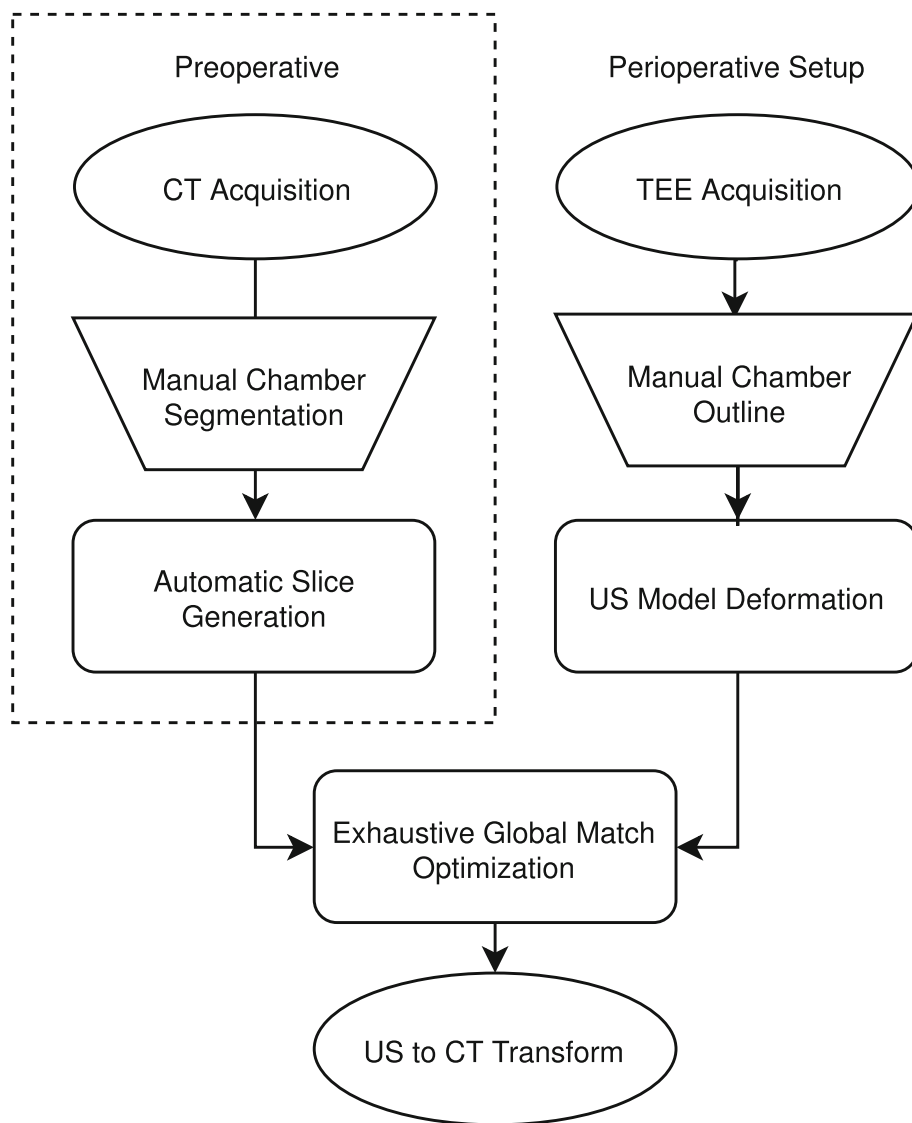
The CT images were acquired, during breath-hold, using electrocardiography (ECG) gating at approximately 75% of the RR cycle with an Aquilion ONE 320-slice scanner (Canon Medical Systems, Tustin, CA, USA). The TEE ultrasound images were acquired using a Vivid 9 scanner (GE Healthcare, Milwaukee, WI, USA). The CT images in DICOM format were imported into Slicer open-source software; segmented; and the resulting surfaces were exported as STL meshes. These meshes, and the TEE images in AVI format, were imported into MATLAB (MathWorks, Natick, MA, USA) and processed using custom software.

An overview of our entire registration pipeline is shown in Fig. 3. Each preoperative CT scan was manually segmented and used to generate a large number of precomputed slices of 3D geometry. We then extracted point sets from a sequence of ultrasound images representing the visible parts of the endocardium in each chamber across the cardiac cycle. We exhaustively registered the ultrasound point sets to the precomputed slices of the CT to find the globally best matching ultrasound–CT slice pair, ultimately retrieving a similarity transform mapping from ultrasound to CT coordinates.

3D CT preprocessing

The four cardiac chambers were manually segmented from the preoperative CT scan, which produced a mesh model of each chamber. The 2D/3D registration found the optimal feasible scan plane S that was in the CT coordinate frame. To find such feasible scan planes, an initial estimate of a scan

Fig. 3 Our registration pipeline consisted of processing the preoperative CT images to precompute a large number of slices of 3D geometry. We then extracted the motion from an ultrasound sequence and registered the resulting ultrasound frames to the preoperative CT space. In an actual clinical workflow, processes inside the dotted box would occur preoperatively, while all other processes would occur during perioperative setup. *Note* trapezoidal processes require manual input



plane was computed. Feasible scan planes were found by varying the plane's parameters from the initial feasible plane.

The mesh points of the cardiac chambers were assembled into a single $3 \times N_C$ matrix. An initial feasible plane was computed from principal components analysis of this matrix. The pose of each other feasible plane S was calculated by varying along the normal direction of the initial feasible plane in 2 mm increments and by varying the initial normal direction in 1° increments. This produced a set of N_P feasible ultrasound planes. We used $N_P = 30 \times 30 \times 30 = 27,000$ feasible planes in this study.

For subsequent registration, each feasible plane S was intersected with each 3D surface model of each of the four cardiac chambers. The precomputation produced a data set that contained, for each slice S :

- $\{x_i^S\}$ —a set of points obtained by intersecting each plane S with the four chamber meshes

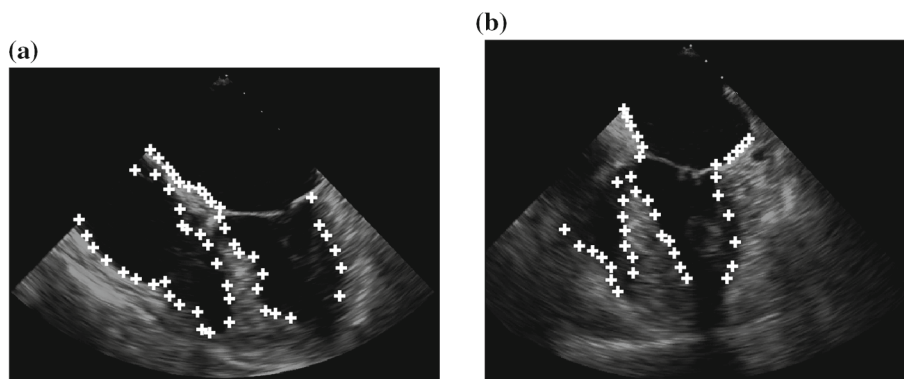
- R_S^{CT} —a rigid transformation from the local coordinates of slice S to the 3D CT space

Ultrasound processing by template deformation

We formulated the ultrasound processing problem as an application-specific nonrigid registration between sets of oriented points. For each t of N_E TEE ultrasound images, we extracted a point set $\{x_i^t\}_{i=1}^{N_t}$ using an edge detector. The orientation \hat{x}_i of each extracted point x_i was computed as the normalized image gradient at that point, giving a set of oriented points $X^t = \{(x_i, \hat{x}_i)^t\}$ for each ultrasound image t .

We used a deformable template shape to represent the chambers visible in the ultrasound image. The template shape consisted of oriented point set $Y = \{(y_j, \hat{y}_j)\}_{j=1}^M$ where M is the total number of points in the model. The points $\{y_j\}$ were initialized by manually placing a set of points delineating each chamber on the first ultrasound image in the sequence.

Fig. 4 Manually placed points that delineated the visible chamber walls were used to initialize the deformable model. In the second example, the right atrium was not fully imaged, so $K = 3$



The orientation \hat{y}_j at each y_j was automatically computed as an outward facing normal vector. Letting K be the number of visible chambers, a label map $\ell_{US} : \{1, 2, \dots, M\} \rightarrow \{1, 2, \dots, K\}$ mapped the index j of an oriented point to a chamber index $\ell_{US}(j)$. An example of two manually instantiated models is shown in Fig. 4.

Due to the nature of ultrasound imaging, the point sets extracted by edge detection are noisy. To prevent overfitting, and improve robustness, we opted to use a model of deformation that limited the total degrees of freedom, while still providing a reasonably high degree of deformability. In particular, to capture the beating-heart rhythm in the ultrasound image sequence, we used hierarchical affine transformations of the deformable template shape. A global affine transformation A modeled the overall movement of the heart and there were separate affine transformations $\{C_k\}$ for the oriented points in each cardiac chamber, so the total transformation $G = C_1 \circ \dots \circ C_K \circ A$ partitioned the oriented points into K sets. Each affine transformation T was represented as $T = (B_T, d_T)$, where B_T is a 2×2 matrix and d_T is a translation vector.

The transformation G , applied to an oriented point, had the effect $G(y_j)$ on position and $G\star\hat{y}_j$ on the normal \hat{y}_j . The transformed deformable template model was thus $G(Y) = \{(G(y_j), G\star\hat{y}_j)\}$. The task was, for each ultrasound image t , to compute the transformation G that best aligned the model's oriented point set $G(Y)$ to the target's oriented point set X^t .

Formulating the registration The registration framework was based on mixture models and the expectation maximization (EM) algorithm [20], which is common in point set registration [21]. As is done in an expectation conditional maximization (ECM) algorithm [22], our M-step separately optimized the transformation parameters and the other statistical parameters. However, to make the transformation update tractable, we only required that the algorithm be a GEM algorithm [20].

We chose the mixture components to take the form of a product of a Gaussian distribution a von Mises distribution. Let $X = \{(x_i, \hat{x}_i)\}_{i=1}^N$ be the N oriented points from the

current ultrasound image, $\mathcal{N}(\cdot|\mu, \sigma^2)$ represent a normal distribution centered at μ with isotropic variance σ^2 , and $f(\cdot|v, \kappa)$ represent a von Mises distribution with mean v and concentration parameter κ . An oriented point (x, \hat{x}) had the distribution

$$p(x, \hat{x}|G, \kappa, \sigma^2) = \frac{\omega}{2\pi A} + \frac{1-\omega}{M} \sum_{j=1}^M \mathcal{N}(q|G(p_j), \sigma^2) f(\hat{x}|G\star\hat{y}_j, \kappa) \tag{1}$$

where A is the area of a region containing all the points in the target set and I_0 is the zeroth-order modified Bessel function of the first kind. The first term of Eq. (1) is a uniform distribution over the registration domain to capture outliers in the target point set, and the terms of the sum correspond to the mixture model. The hyper-parameter ω represents the expected proportion of points in the target point set that are outliers.

To avoid excessive deformations of the shape model between successive ultrasound images, a regularization function was incorporated by adding the prior $P(G) \propto \exp(-\Phi(G))$. The regularization function affected only the chamber transformations and took the form

$$\Phi(G) = \lambda \sum_{k=1}^K \|B_{C_k} - B_{k0}\|_F^2$$

where B_{C_k} is the 2×2 transformation matrix from affine transform C_k and B_{k0} is the initial value of B_{C_k} at the beginning of the registration. Registration of the deformable template to the target oriented point set X could then be achieved by maximizing the posterior

$$p(G|X, \sigma^2, \kappa) = P(G) \prod_{i=1}^N p(x_i, \hat{x}_i|G, \sigma^2, \kappa) \tag{2}$$

with respect to the transformation G and parameters σ^2 and κ .

We used an EM approach to optimize Eq. (2) by introducing a “hidden” variable $Z = \{z_i\}_{i=1}^N$, representing from which mixture component each oriented point in the target set was sampled. For each i , $z_i = j$ meant the i th data point was sampled from the j th cluster (or the outlier term if $z_i = 0$). The expected complete-data negative log-likelihood was

$$Q(G, \sigma^2, \kappa) = \sum_{i,j=1}^{N,M} \alpha_{ji} \left(\frac{\|x_i - G(y_j)\|^2}{2\sigma^2} - \kappa \hat{x}_i^T (G \star \hat{y}_j) \right) + N_\alpha \log(\sigma^2) + N_\alpha \log(I_0(\kappa)) + \Phi(G) \tag{3}$$

where $N_\alpha = \sum_{i,j} \alpha_{ji}$ and constant terms have been removed. The $M \times N$ matrix of coefficients α_{ji} were posteriors, computed during the E-step as

$$\alpha_{ji} = p(z_i = j | x_i, G, \sigma^2, \kappa) = \frac{\mathcal{N}(x_i | G(y_j), \sigma^2) f(\hat{x}_i | G \star \hat{y}_j, \kappa)}{\frac{2M\pi I_0(\kappa)\sigma^2\omega}{A(1-\omega)} + \sum_{j=1}^M \mathcal{N}(x_i | G(y_j), \sigma^2) f(\hat{x}_i | G \star \hat{y}_j, \kappa)}$$

We divided the M-step into two separate conditional M-steps. The first optimized the distribution parameters σ^2 and κ ; the second optimized the transformation G .

The distribution M-step Equation (3) could be optimized independently with respect to σ^2 and κ . The variance σ^2 was optimized analytically giving update

$$\sigma^2 = \frac{\sum_{i,j} \alpha_{ji} \|x_i - G(y_j)\|^2}{N_\alpha}$$

We adapted the truncated Newton approximation maximum likelihood estimation for κ from hypersphere clustering methods [23], giving the iterative solution

$$\bar{R} = \frac{\sum_{i,j} \alpha_{ji} \hat{x}_i^T (G \star \hat{y}_j)}{N_\alpha}, \quad A(\kappa) = I_1(\kappa)/I_0(\kappa),$$

$$\kappa' = \kappa - \frac{A(\kappa) - \bar{R}}{1 - A(\kappa)^2 - A(\kappa)/\kappa}$$

The transformation M-step To simplify the mathematical expressions, the concept of virtual observations can be used to derive an equivalent paired-point expression that can be optimized in place of Eq. (3) when updating the transformation [24]. We extended the concept to the case of oriented points, defining virtual weights $W = \{(w_j, \hat{w}_j)\}$, and virtual oriented points $V = \{(v_j, \hat{v}_j)\}$ as

$$w_j = \frac{1}{2\sigma^2} \sum_i \alpha_{ji}; \quad \hat{w}_j = \kappa \sum_i \alpha_{ji}; \quad v_j = \frac{\sum_i \alpha_{ji} x_i}{N_\alpha};$$

$$\hat{v}_j = \frac{\sum_i \alpha_{ji} \hat{x}_i}{N_\alpha}$$

With these definitions, optimizing Eq. (3) with respect to G is equivalent to optimizing

$$Q'(G) = \sum_j [w_j \|v_j - G(y_j)\|^2 - \hat{w}_j \hat{v}_j^T (G \star \hat{y}_j)] + \Phi(G) \tag{4}$$

Notice that this is effectively a weighted, regularized, paired-point registration between the model point set Y and target virtual point set V through transformation G .

We further decomposed the transformation M-step into two conditional M-steps: The first updated only the global affine transformation A , holding the others constant, and the second independently updated each chamber transformation C_k .

We did this decomposition in such a way that each step was a solution to a weighted, regularized, paired-point registration between oriented point sets through a single affine transformation. Symbolically, let $P = \{(p_j, \hat{p}_j)\}$ and $Q = \{(q_j, \hat{q}_j)\}$ be corresponding oriented point sets; let $U = \{(u_j, \hat{u}_j)\}$ be a set of weights; let β be a regularization weight; and let B_0 be an initial matrix. Define an algorithm to do a regularized paired-point affine registration between these point sets as

$$\mathcal{A}(P, Q, U, \beta, B_0) = \arg \min_T \sum_j [u_j \|p_j - T(q_j)\|^2 - \hat{u}_j \hat{p}_j^T (T \star \hat{q}_j)] + \beta \|B_T - B_0\|_F^2. \tag{5}$$

Assuming a smooth underlying shape from which the oriented points are sampled, an affine transformation T will act on a normal \hat{q} as

$$T \star \hat{p} = \frac{B_T^{-T} \hat{q}}{\|B_T^{-T} \hat{q}\|}$$

Equation (5) can be solved for the optimal transformation matrix B_T using a quasi-Newton solver directly on the matrix components, with an analytically computed gradient. Given the optimal matrix B_T , the optimal translation can be computed as

$$d_T = \frac{\sum_j u_j p_j}{\sum_j u_j} - B_T \frac{\sum_j u_j q_j}{\sum_j u_j}$$

We will now describe our decomposed transformation M-step in terms of applications of this algorithm \mathcal{A} . First, consider the update of the chamber transformations. Sup-

posing the global transformation has already been updated to some A' , $Q'(C_1 \circ \dots \circ C_K \circ A')$ can be decomposed into separate terms for each chamber transformation

$$Q'_k(C_k) = \sum_{j \in \mathcal{L}_{US}^{-1}(k)} [w_j \|v_j - C_k \circ A'(y_j)\|^2 - \hat{w}_j \hat{v}_j^T ((C_k \circ A') \star \hat{y}_j)] + \lambda \|B_{C_k} - B_{0k}\|_F^2$$

$$Q'(C_1 \circ \dots \circ C_K \circ A') = \sum_k Q'_k(C_k)$$

With this decomposition, the optimal update for each chamber transformation would be

$$C'_k = \arg \min_{C_k} Q'_k(C_k) = \mathcal{A}(V_k, A'(Y_k), W_k, \lambda, B_{k0})$$

where V_k , Y_k , and W_k are, respectively, the set of virtual points, model points, and virtual weights belonging to the k th chamber.

In an ECM formulation, we would need to update the global transformation by optimizing Eq. (4) with respect to A , holding $\{C_k\}$ fixed. However, the composition of the chamber transformations with A makes such an optimization challenging. We instead only required our algorithm to be a GEM algorithm. In the present context, this only required an update from G to G' guaranteeing $Q'(G') < Q'(G)$.

This guarantee was achieved through a simple strategy. We updated the global transformation, completely ignoring $\{C_k\}$ via $A' = \mathcal{A}(V, Y, W, 0, 0)$. We then updated the chamber transformations as described above. If this strategy failed to improve Q' , we simply reverted the global transformation to its initial value A and recomputed the chamber transformations. This was guaranteed to improve Q' , since the updates to C_k optimize $Q'_k(C_k)$. The complete transformation M-step is given in pseudocode in Algorithm 1.

Algorithm 1 An algorithm to update G , wherein the loop is guaranteed to run at most twice

Input: data oriented points X , model oriented points Y , current transform $G = C_1 \circ C_2 \circ \dots \circ C_K \circ A$
Output: New transformation G'

```

Compute virtual transformations  $V$ 
 $E \leftarrow Q'(T)$ 
 $E' \leftarrow E$ 
 $A' \leftarrow \mathcal{A}(V, Y, W, 0, 0)$ 
while  $E' \geq E$  do
  for all  $C_k$  do
     $C'_k \leftarrow \mathcal{A}(V_k, A'(Y_k), W_k, \lambda, B_{k0})$ 
  end for
   $G' \leftarrow C'_1 \circ C'_2 \circ \dots \circ C'_K \circ A'$ 
   $Q' \leftarrow Q'(T')$ 
  if  $E' \geq E$  then
     $A' \leftarrow A$ 
  end if
end while

```

Cardiac deformation over time After initializing the model by manually placing points on the first ultrasound image, for each image in the sequence we registered the model using our described GEM algorithm, initializing the global transformation G with the transformation that had been calculated from the preceding image. For the first image, the regularization matrices $\{B_{k0}\}$ were set to the identity matrix. Subsequently, for each TEE ultrasound image, the regularization matrices were initialized to the transformation matrices B_{C_k} from the converged computations of the previous image. This initialization produced a relatively stable set of local affine transformations C_k , which acted primarily as small corrections to the global affine transformation A . The result of this process was a set of N_E transformations $\{G^t\}$ and deformed point sets $\{U^t = G^t(Y)\}$ that represented each TEE ultrasound image.

Managing poorly registered images Some TEE ultrasound images were not well registered by our automated algorithm. To detect these poorly registered images, we computed the root mean square error between the ultrasound virtual points and model points as

$$RMSE_t = \sqrt{\frac{\sum_j \|v_j - G^t(y_j)\|^2}{M}}$$

Given this sequence, we used a peak detection algorithm to compute the local maxima in RMSE. We found all outlier peaks, determined as peaks lying more than three scaled median absolute deviations from the median peak, and used them to compute a threshold RMSE as

$$RMSE_{\text{thresh}} = \min_{t \text{ an outlier}} RMSE_t.$$

At the end of this ultrasound image registration process, we had a set of $N'_E < N_E$ ultrasound images that were registered with RMSE below the threshold. We kept only the positional data from the registered images, which produced a data set of N'_E point sets $U^t = \{u_j^t\}_{j=1}^M$ representing the ultrasound sequence.

Matching 3D slices to ultrasound images

We performed an exhaustive search that matched the N_P precomputed slices from the CT model with the N'_E TEE ultrasound point sets, finding the pair producing the globally minimal RMSE. The match was computed by registering the ultrasound point set $\{u_j^t\}$ to the CT slice $\{x_i^S\}$ with a similarity transformation.

The registration was initialized by creating a correspondence between the center of each cardiac chamber in the

feasible CT slice and the deformed ultrasound model. This provided an approximate translation, orientation, and uniform scale factor between the point sets. Using this initial estimate, the similarity transform was optimized using a modified ICP algorithm [12], in which the nearest-neighbor search was limited to points from the same cardiac chamber. The registration was scored by the RMSE between the registered ultrasound point set and the converged nearest neighbors from the CT slice.

The exhaustive search resulted in a slice \hat{S} and ultrasound image \hat{U} representing the optimal match, as defined by least RMSE. The registration between this optimal ultrasound image and slice pair gave a similarity transform $T_{US}^{\hat{S}}$ mapping the ultrasound to the coordinates of slice \hat{S} . Combining this similarity transformation with the rigid transform $R^{\hat{S}}$, we were able to compute the total transformation from ultrasound to CT space

$$T_{US}^{CT} = R_{\hat{S}}^{CT} \circ T_{US}^{\hat{S}}$$

Data analysis

The RMS error of the deformed cardiac model to the edge-detected points in the ultrasound image was tabulated for each patient's TEE sequence. The RMS error of the transformed deformed model to the CT segmentation was also tabulated. This transformation encoded scale factors for the size of pixels in the ultrasound image, which could be used to express the ultrasound RMS errors in millimeters.

The expected error was computed from the ultrasound RMSE and a value from the literature. In CT coronary angiography with ECG gating, blood vessels of a diameter less than 1.5 mm are routinely excluded [25]. The ultrasound RMS error was the square root of a variance, as was the minimum blood vessel diameter, so we computed the expected error by adding the values in quadrature $\sqrt{\text{RMSE}_t^2 + 1.5^2}$. We pooled two RMS errors for all patients: the overall RMS error of 2D/3D registration and the computed expected error. The pools were compared with a two-sided *t*-test.

Results

We processed a total of four data sets, manually segmenting each patient's preoperative CT scan and using it to generate a total of 27,000 slices. Of these slices, those which did not contain a cross section from all four chambers were discarded. For each patient, we sought to deform our manually initialized 2D model to each ultrasound image. The free parameters of the registration, namely outlier weight ω and registration weight λ , were tuned for each patient, but all varied within a similar range. The initial values for σ^2 and κ were consistently set to 250 and 1.5, respectively. After deforming the model to each ultrasound image in a sequence, and automatically detecting and removing individual poorly registered frames, we computed the RMS error between the deformed model in each image and the points extracted from the image. Table 1 summarizes these results for each patient and documents the various settings used.

The exhaustive search produced a registration of the TEE ultrasound sequence to the CT scan by seeking an optimally matched 3D slice and ultrasound point set. The total number of slices and images compared, along with the RMS error between the best matching 3D slice and ultrasound image, is reported in Table 2. The registration results are illustrated for each patient in Fig. 5, which show the 2D slices and 3D models plotted in ultrasound space.

From a *t*-test of the pooled RMS values in Table 2, there was no statistically significant difference ($p \approx 0.58$) between the overall registration error and the error expected from the combination of the ultrasound model deformation error and the uncertainty in the cardiac CT scan. A paired *t*-test also showed no statistically significant difference ($p \approx 0.60$).

Discussion

Minimally invasive arrhythmia surgery has gained popularity over the past decade, especially in the most complex subset of patients, such as those with persistent atrial fibrillation [1]. In particular, an epicardial approach has provided high rates of success even when compared to a transcatheter one in persistent atrial fibrillation [26] but with an increased incidence

Table 1 Model deformation settings and results for each of the four patients are shown

Patient index	Total images N_E	ω	λ	Kept images N'_E	RMSE (mm)	RMSE (pixels)	All images RMSE (pixels)
1	104	0.7	2	88	2.1 ± 0.5	7.8 ± 1.7	8.3 ± 2.1
2	61	0.85	3	45	2.6 ± 0.8	6.7 ± 2.0	9.5 ± 5.6
3	111	0.9	2	110	2.7 ± 0.3	5.7 ± 0.6	5.7 ± 0.6
4	105	0.9	3	104	1.8 ± 0.3	6.1 ± 1.1	6.1 ± 1.1

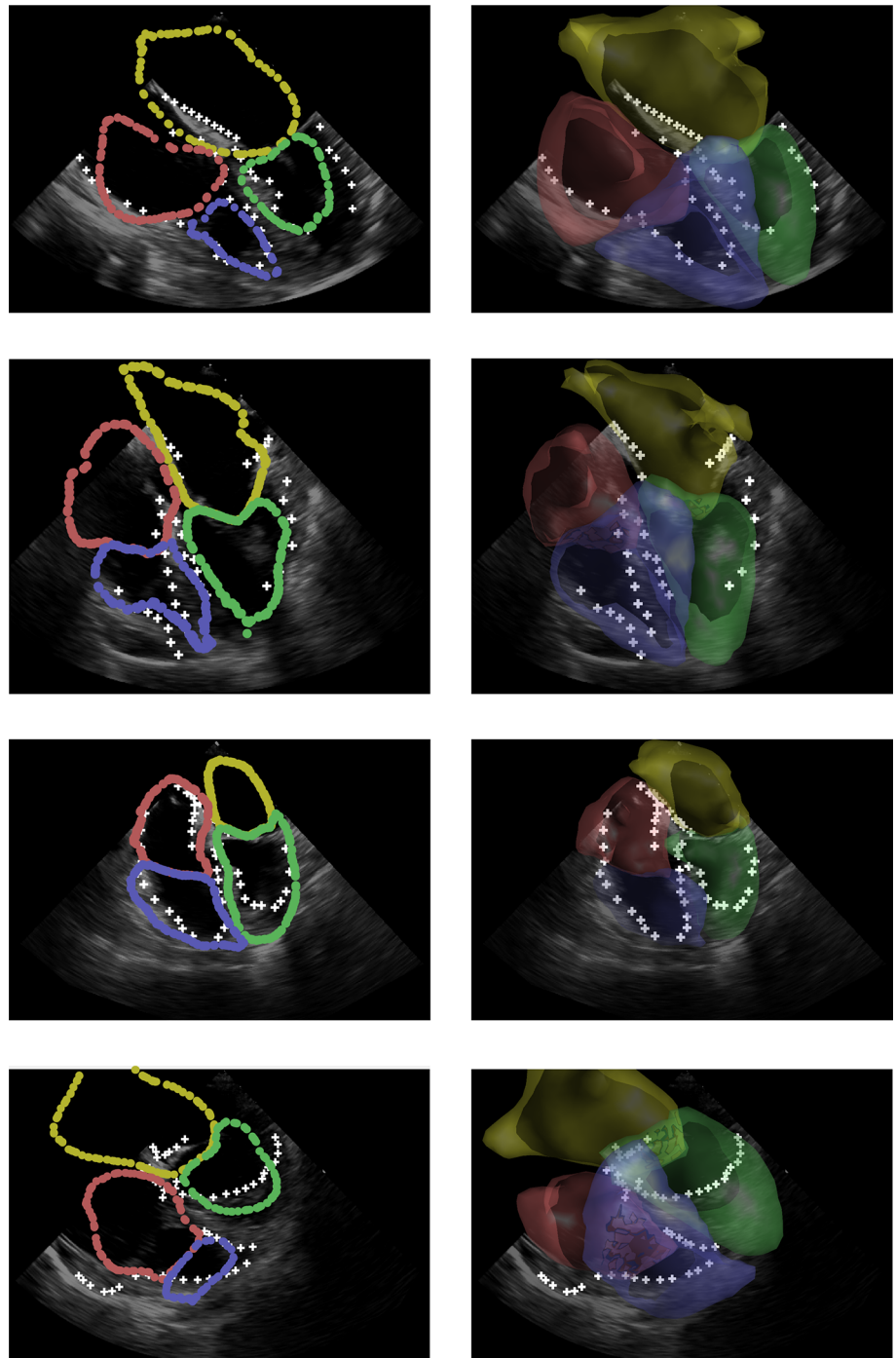
RMSE values are reported in pixels and millimeters as the mean, plus or minus standard deviation, across all well-registered images in the sequence. Millimeter values were estimated using the scale factors computed when transforming ultrasound to CT space. The final column reports the RMSE when automatically rejected images are kept

Table 2 The total number of searched slices and images are reported

Patient index	No. slices N_P	No. of images N'_E	RMS (mm)	Expected (mm)
1	21,972	104	2.5	2.6
2	24,215	61	2.3	3.0
3	24,235	111	4.2	3.1
4	22,515	105	3.0	2.3

The RMS error was for the best match. The expected error was calculated from the ultrasound RMS error and the lower limit of 1.5 mm for vessels in CT coronary angiography, which represented the uncertainty in CT cardiac imaging

Fig. 5 The best matching slices and ultrasound images allowed us to generate a transform for each 3D model from CT space to ultrasound space. The left-hand side shows the best matching 3D cross sections (circular markers), with the ultrasound image and model (plus markers). On the right-hand side are the same matches, with the 3D chamber models shown in translucent 3D. Each row corresponds to one patient



of perioperative complications. In particular, the occurrence of intraoperative bleeding complications has been reported in a range from 0 to 10% across several reports [27]: Hence, the possibility to identify strategies to improve surgical navigation and minimize damages to the heart or the nearby vascular structures appears to be critical for a safer application of such less invasive techniques.

Our ultrasound processing used a novel hierarchical affine registration algorithm with oriented points, which provided advantages for our data sets compared to many existing methods for segmentation or registration of TEE ultrasound image sequences. The method made no assumptions about the TEE viewing angles, or about the number of visible cardiac chambers. Level set methods may be less well suited to partial views because they are constrained to producing closed contours. Meanwhile, methods based on statistical shape models are typically constrained to a single view or topological configuration as represented in the training data, or else require a large amount of user interaction to construct a patient specific model [28]. Unlike a statistical or machine learning method, our method also did not require large amounts of training data.

Our ultrasound registration made use of oriented points. The addition of orientation information to point set registration has been found to benefit accuracy in both ICP-like methods [29,30] and probabilistic methods [10,11,31]. In our case, the primary benefit of orientations came from an enhanced ability to distinguish between spatially close but oppositely oriented points. For example, nearby points such as those on either side of either septum were easily distinguished when using orientation information.

We registered the deformable model to each image using a mixture model formulation. Although ICP-like formulations may be more efficient—and can use probabilistic noise models with oriented points [29]—our optimization step for the transformation fundamentally relied on the formulation as a GEM algorithm. It is possible that a sparse or incremental approach to EM [32], or a threshold on the matching likelihood [33], could be used to improve computational efficiency. Such optimizations could be realized using a PD-tree lookup for oriented points [29]. Furthermore, although GEM algorithms offer weaker theoretical guarantees than the ECM algorithms typically used for registration, in all our experiments the algorithm converged in fewer than 70 iterations. This experience is in line with recent work on statistical shape model registration using a GEM algorithm [34].

To facilitate our oriented point-based registration, we relied on a simple edge detector to extract features from each ultrasound image in a sequence. In a noisy modality such as ultrasound, this approach inevitably produces noisy data and spurious features. We addressed this by using a probabilistic registration method that is intended to be robust to outliers, along with a transformation model that is inherently

limited in its total degrees of freedom to prevent overfitting. Our mixture model might become more robust to outliers by using a Student's t distribution for the positional data, rather than a Gaussian [10,35,36]. Alternatively, an improved feature detection method may reduce outliers and improve robustness. Scale-invariant features, such as SIFT [37] or SURF [38] features or their 3D variants [39], have been shown to provide high-quality correspondences for ultrasound images [40,41] and have been successfully applied in the domain of cardiac imaging [42].

The main computational bottleneck in our method was the exhaustive search for the 2D/3D match which, in our unoptimized MATLAB implementation on an older PC, took a few hours to run for each trial. This process is, however, amenable to a number of possible optimizations to be considered in future work. The nature of this search makes it very simple to be processed in parallel to whatever extent the available hardware allows. The regular structure of our generated slices, essentially forming a 3D grid along axis height and two angle dimensions, leaves the approach open to a coarse-to-fine search to reduce the number of required 2D registrations. Alternatively, it may be possible to adopt more efficient discrete optimization methods by formulating our planar search in an MRF framework using a similar approach to that of Porchetto et al. [19] adapted to feature-based matching criteria. Restrictions on the TEE view could also be used to reduce the degrees of freedom of possible transformations [8], thereby reducing the dimensionality of the search space. Finally, integration with ECG might reduce the number of ultrasound frames needing to be considered by allowing us to select ultrasound images near the phase of the cardiac cycle at which the preoperative CT was gated [8]. In fact, because the search time is proportional to the number of ultrasound images in the sequence—which in our data sets was between 45 and 110 (see Table 1)—ECG integration alone could speed the search by between one and two orders of magnitude, which may be fast enough for execution during perioperative setup.

For a deformable organ such as the heart, the restriction of our method to the computation of a similarity transform may seem limiting; however, allowing in-plane deformation of the slice would make it difficult to differentiate frames of the TEE sequence. Rigid slice-to-volume registration appears to perform as well as nonrigid methods when the deformation of the anatomy is low between the two image acquisitions and less so otherwise [19]. Our algorithm relies on this disparity to be able to compute both the temporal and spatial registration simultaneously; it assumes that temporally aligned images should result in lower spatial registration error than those that are temporally misaligned. However, additional deformations may be present even in temporally well-registered images due to factors such as irregular heart beats, respiration, and positional differences. Though nonrigid reg-

istration may help accommodate these differences, nonrigid slice-to-volume registrations can also produce results which are anatomically unsound, despite having low geometric error [13]. Instead, remaining deformations along with uncertainty in the computed transformation may be partially ameliorated by adaptive visualization techniques [43].

After computing a similarity transform mapping ultrasound to CT coordinates perioperatively, it could be necessary to refine the transform intraoperatively to account for small changes in the relative placement of the probe. The use of a tracked ultrasound probe would enable us to automatically account for these movements without requiring further registration computations. In this case, our perioperative registration would serve to compute the mapping from intraoperative world coordinates to preoperative CT coordinates. Alternatively, the perioperative registration could be used as an initialization for a more efficient intraoperative registration to be executed online. The static nature of our preoperative CT limits the opportunity for dynamic image fusion beyond a simple similarity transform because we lack a model of the 3D heart motion; however, in AF patients, motion captured by preoperative dynamic CT may not be representative of the motion observed intraoperatively due to the irregular heart beats characterizing the condition.

It is important to note that the registration errors reported in this study are based on the nearest neighbor between the model and target point sets, as opposed to being based on specific anatomical landmarks. It is therefore possible to have cases of poor registration with very low error values, especially when dealing with nonrigid transformations. Our goal in this work was to demonstrate the overall feasibility of our approach; in future work, it will be important to validate our method with objective, landmark-based error metrics, possibly via synthetic data and/or a phantom study. In such a study, we could better analyze which sources of error in our pipeline, if any, need to be improved to reach clinical applicability, and verify the strength and correctness of the global minimum in our search.

Though most of our process is automated, we do rely on human input for segmentation purposes both preoperatively and during perioperative setup. In particular, we require manual segmentation of the four cardiac chambers from the preoperative CT, as well as perioperative manual delineation of points outlining the chambers in the intraoperative ultrasound. It may be possible to automate some or all of this work using modern automatic segmentation methods, such as those based on deep learning.

Conclusion

We presented a way to register a preoperative 3D cardiac CT scan to intraoperative 2D TEE ultrasound images. This

was done by first optimizing a hierarchy of affine transformations using a novel generalized expectation maximization formulation and then solving a discretized slice-to-volume registration. Our approach computed the registration in both time and space, not relying on ECG gating for temporal alignment. Results, computed from data of four surgical patients, were quantitatively and qualitatively promising. Future work may include ground-truth validation of our method, optimization of our computations to reach sufficient speed for clinical applications, and integration into a navigation system for minimally invasive beating-heart cardiac surgery.

Acknowledgements This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grant RGPIN-2018-04430.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. This article does not contain any studies with animals performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Bisleri G, Rosati F, Bontempi L, Curnis A, Muneretto C (2013) Hybrid approach for the treatment of long-standing persistent atrial fibrillation: electrophysiological findings and clinical results. *Eur J Cardiothorac Surg* 44(5):919–923
2. Muneretto C, Bisleri G, Bontempi L, Curnis A (2012) Durable staged hybrid ablation with thoracoscopic and percutaneous approach for treatment of long-standing atrial fibrillation: a 30-month assessment with continuous monitoring. *J Thorac Cardiovasc Surg* 144(6):1460–1465
3. Huang X, Moore J, Guiraudon G, Jones DL, Bainbridge D, Ren J, Peters TM (2009) Dynamic 2D ultrasound and 3D CT image registration of the beating heart. *IEEE Trans Med Imaging* 28(8):1179–1189
4. Luo Z, Cai J, Peters TM, Gu L (2013) Intra-operative 2-D ultrasound and dynamic 3-D aortic model registration for magnetic navigation of transcatheter aortic valve implantation. *IEEE Trans Med Imaging* 32(11):2152–2165
5. Tavard F, Simon A, Leclercq C, Donal E, Hernandez AI, Garreau M (2014) Multimodal registration and data fusion for cardiac resynchronization therapy optimization. *IEEE Trans Med Imaging* 33(6):1363–1372
6. Li FP, Rajchl M, White JA, Goela A, Peters TM (2015) Ultrasound guidance for beating heart mitral valve repair augmented by synthetic dynamic CT. *IEEE Trans Med Imaging* 34(10):2025–2035
7. Khalil A, Faisal A, Lai KW, Ng SC, Liew YM (2017) 2D to 3D fusion of echocardiography and cardiac CT for TAVR and TAVI image guidance. *Med Biol Eng Comput* 55(8):1317–1326

8. Sandoval Z, Castro M, Alirezaie J, Bessire F, Lafon C, Dillenseger JL (2018) Transesophageal 2D ultrasound to 3D computed tomography registration for the guidance of a cardiac arrhythmia therapy. *Phys Med Biol* 63(15):155007
9. Tavakoli V, Amini AA (2013) A survey of shaped-based registration and segmentation techniques for cardiac images. *Comput Vis Image Underst* 117(9):966–989
10. Ravikumar N, Gooya A, Frangi AF, Taylor ZA (2017) Generalised coherent point drift for group-wise registration of multi-dimensional point sets. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S (eds) *Medical image computing and computer-assisted intervention*, vol 10433. LNCS. Springer, Cham, pp 309–316
11. Min Z, Wang J, Meng MQH (2018) Robust generalized point cloud registration using hybrid mixture model. In: *IEEE Int Conf Robot Autom*, pp 4812–4818
12. Besl PJ, McKay ND (1992) A method for registration of 3-D shapes. *IEEE Trans Pattern Anal Mach Intell* 14(2):239–256
13. Ferrante E, Paragios N (2017) Slice-to-volume medical image registration: a survey. *Med Image Anal* 39:101–123
14. Ferrante E, Paragios N (2013) Non-rigid 2D–3D medical image registration using Markov random fields. In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N (eds) *Medical image computing and computer-assisted intervention*, vol 8151. LNCS. Springer, Heidelberg, pp 163–170
15. Ferrante E, Fecamp V, Paragios N (2015) Slice-to-volume deformable registration: efficient one-shot consensus between plane selection and in-plane deformation. *Int J Comput Assist Radiol Surg* 10(6):791–800
16. Ferrante E, Fecamp V, Paragios N (2015) Implicit planar and in-plane deformable mapping in medical images through high order graphs. In: *Proc IEEE Int Symp Biomed Imaging*, pp 721–724
17. Ferrante E, Paragios N (2018) Graph-based slice-to-volume deformable registration. *Int J Comput Vis* 126(1):36–58
18. Zikic D, Glocker B, Kutter O, Groher M, Komodakis N, Kamen A, Paragios N, Navab N (2010) Linear intensity-based image registration by Markov random fields and discrete optimization. *Med Image Anal* 14(4):550–562
19. Porchetto R, Stramana F, Paragios N, Ferrante E (2017) Rigid slice-to-volume medical image registration through Markov random fields. *Med Comput Vis Bayesian Graph Models Biomed Imaging* 2016:172–185
20. Dempster AP, Laird NM, Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Series B Methodol* 39(1):1–38
21. Maiseli B, Gu Y, Gao H (2017) Recent developments and trends in point set registration methods. *J Vis Commun Image Represent* 46(C):95–106
22. Meng XL, Rubin DB (1993) Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika* 80(2):267–278
23. Sra S (2012) A short note on parameter approximation for von Mises–Fisher distributions: and a fast implementation of $I_s(x)$. *Comput Stat* 27(1):177–190
24. Horaud R, Forbes F, Yguel M, Dewaele G, Zhang J (2011) Rigid and articulated point registration with expectation conditional maximization. *IEEE Trans Pattern Anal Mach Intell* 33(3):587–602
25. Buechel RR, Husmann L, Herzog BA, Pazhenkottil AP, Nkoulou R, Ghadri JR, Treyer V, von Schulthess P, Kaufmann PA (2011) Low-dose computed tomography coronary angiography with prospective electrocardiogram triggering: feasibility in a large population. *J Am Coll Cardiol* 57(3):332–336
26. Boersma LV, Castella M, van Boven W, Berruzo A, Yilmaz A, Nadal M, Sandoval E, Calvo N, Brugada J, Kelder J, Wijffels M, Mont L (2012) Atrial fibrillation catheter ablation versus surgical ablation treatment (fast): a 2-center randomized clinical trial. *Circulation* 125(1):23–30
27. Khojnehzad A, Ellenbogen KA, Al-Atassi T, Wang PJ, Kasirajan V, Wang X, Edgerton JR (2017) Hybrid atrial fibrillation ablation: current status and a look ahead. *Circ Arrhythm Electrophysiol* 10(10):e005263
28. Jacob G, Noble JA, Behrenbruch C, Kelion AD, Banning AP (2002) A shape-space-based approach to tracking myocardial borders and quantifying regional left-ventricular function applied in echocardiography. *IEEE Trans Med Imaging* 21(3):226–238
29. Billings S, Taylor R (2015) Generalized iterative most likely oriented-point (G-IMLOP) registration. *Int J Comput Assist Radiol Surg* 10:1213–1226
30. Serafin J, Grisetti G (2015) NIPC: dense normal based point cloud registration. *Rep U S*, pp 742–749
31. Baka N, Metz CT, Schultz CJ, van Geuns R, Niessen WJ, van Walsum T (2014) Oriented Gaussian mixture models for non-rigid 2D/3D coronary artery registration. *IEEE Trans Med Imaging* 33(5):1023–1034
32. Neal RM, Hinton GE (1998) A view of the EM algorithm that justifies incremental, sparse, and other variants. In: Jordan MI (ed) *Learning in graphical models*. Springer, Dordrecht, pp 355–368
33. Granger S, Pennec X (2002) Multi-scale EM-ICP: a fast and robust approach for surface registration. *Comput Vis ECCV*, pp 418–432
34. Bernard F, Salamanca L, Thunberg J, Tack A, Jentsch D, Lamecker H, Zachow S, Hertel F, Goncalves J, Gemmar P (2017) Shape-aware surface reconstruction from sparse 3D point-clouds. *Med Image Anal* 38:77–89
35. Zhou Z, Zheng J, Dai Y, Zhou Z, Chen S (2014) Robust non-rigid point set registration using Student's- t mixture model. *PLoS One* 9(3):1–11
36. Ravikumar N, Gooya A, Çimen S, Frangi AF, Taylor ZA (2018) Group-wise similarity registration of point sets using Student's t -mixture model for statistical shape models. *Med Image Anal* 44:156–176
37. Lowe DG (1999) Object recognition from local scale-invariant features. In: *Proc IEEE Int Conf Comput Vis*, vol 2, pp 1150–1157
38. Bay H, Ess A, Tuytelaars T, Gool LV (2008) Speeded-up robust features (surf). *Comput Vis Image Underst* 110(3):346–359
39. Toews M, Wells WM III (2013) Efficient and robust model-to-image alignment using 3D scale-invariant features. *Med Image Anal* 17(3):271–282
40. Toews M, Wells WM III (2018) Phantomless auto-calibration and online calibration assessment for a tracked freehand 2-D ultrasound probe. *IEEE Trans Med Imaging* 37(1):262–272
41. Machado I, Toews M, Luo J, Unadkat P, Essayed W, George E, Teodoro P, Carvalho H, Martins J, Golland P, Pieper S, Frisken S, Golby A, Wells W III (2018) Non-rigid registration of 3D ultrasound for neurosurgery using automatic feature detection and matching. *Int J Comput Assist Radiol Surg* 13(10):1525–1538
42. Bersvendsen J, Toews M, Danudibrotto A, Wells III WM, Urheim S, Estépar RSJ, Samset E (2016) Robust spatio-temporal registration of 4D cardiac ultrasound sequences. *Proc SPIE Int Soc Opt Eng*
43. Lasowski R, Benhimane S, Vogel J, Jakobs TF, Zech CJ, Trumm C, Clason C, Navab N (2008) Adaptive visualization for needle guidance in RF liver ablation: taking organ deformation into account. *Proc SPIE Int Soc Opt Eng* 6918:69180A

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.