Automated Segmentation of the Right Ventricle in 3D Echocardiography: A Kalman Filter State Estimation Approach

Jørn Bersvendsen*, Fredrik Orderud, Richard John Massey, Kristian Fosså, Olivier Gerard,
Stig Urheim, and Eigil Samset

Abstract—As the right ventricle's (RV) role in cardiovascular diseases is being more widely recognized, interest in RV imaging, function and quantification is growing. However, there are currently few RV quantification methods for 3D echocardiography presented in the literature or commercially available. In this paper we propose an automated RV segmentation method for 3D echocardiographic images. We represent the RV geometry by a Doo-Sabin subdivision surface with deformation modes derived from a training set of manual segmentations. The segmentation is then represented as a state estimation problem and solved with an extended Kalman filter by combining the RV geometry with a motion model and edge detection. Validation was performed by comparing surface-surface distances, volumes and ejection fractions in 17 patients with aortic insufficiency between the proposed method, magnetic resonance imaging (MRI), and a manual echocardiographic reference. The algorithm was efficient with a mean computation time of 2.0 s. The mean absolute distances between the proposed and manual segmentations were 3.6 ± 0.7 mm. Good agreements of end diastolic volume, end systolic volume and ejection fraction with respect to MRI (-26 ± 24 mL, -16 ± 26 mL and $0 \pm 10\%$, respectively) and a manual echocardiographic reference (7 ± 30 mL, 13 ± 17 mL and $-5 \pm 7\%$, respectively) were observed.

Index Terms—Image segmentation, statistical shape models, 3D echocardiography.

I. INTRODUCTION

COMPARED to the Left Ventricle (LV), image processing methods of the Right Ventricle (RV) are infrequently reported in the literature. However, as the RV's role in cardiovascular diseases is being more widely recognized, interest in RV

*J. Bersvendsen is with GE Vingmed Ultrasound AS, 3183 Horten, Norway, and with the Department of Informatics, University of Oslo, 0373 Oslo, Norway, and also with the Center for Cardiological Innovation, 0372 Oslo, Norway (e-mail: jorn.bersvendsen@ge.com).

F. Orderud and O. Gerard are with GE Vingmed Ultrasound AS, 3183 Horten, Norway (e-mail: fredrik.orderud@ge.com; gerard.olivier@ge.com).

R. J. Massey is with the Department of Cardiology, Oslo University Hospital, 0372 Oslo, Norway (e-mail: rmassay@ous-hf.no).

K. Fosså is with the Department of Radiology and Nuclear Medicine, Oslo University Hospital, 0372 Oslo, Norway (e-mail: krifos@ous-hf.no).

S. Urheim is with the Department of Cardiology and Institute for Surgical Research, Oslo University Hospital, 0372 Oslo, Norway (e-mail: surheim@ous-hf.no).

E. Samset is with GE Vingmed Ultrasound AS, 3183 Horten, Norway, and with the Department of Informatics, University of Oslo, 0373 Oslo, Norway, and also with the Center for Cardiological Innovation, 0372 Oslo, Norway (e-mail: eigil.samset@ge.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2015.2453551

function and imaging is growing. For instance, accurate assessment of the RV has been shown to be important in patients with pulmonary hypertension or LV dysfunction [1], [2].

In current clinical practice, manual measurements in Magnetic Resonance Imaging (MRI) are considered the gold standard for quantitative assessment of RV volumes and ejection fraction [1]. However, MRI is seldom routinely available, and in some cases contraindicated. Furthermore, since manual assessment is time consuming and subject to inter-observer variability, automated methods are preferred. In the last decade, Three-Dimensional Transthoracic Echocardiography (3DTTE) has emerged as an alternative for assessing the RV.

Automated segmentation of the RV is largely an unsolved problem, and sparsely reported in the literature for both 3DTTE and MRI. For MRI, the problem is typically solved with a bi-ventricular segmentation method [3]. Methods using a strong prior, such as deformable models [4], active shape models [5] or atlases [6]–[9], and image based methods with and without anatomical priors [10]–[13] have been proposed.

For 3DTTE, Angelini *et al.* [14] have presented a level-set framework for segmentation of both LV and RV. The LV and RV surfaces are implicitly represented as the zero-intersection of a higher order function, which is fitted to the image by energy minimization. Stebbing *et al.* [15] has described a segmentation method using an explicit Loop subdivision surface model of the RV and solving the fitting problem with energy minimization. Missing edges and information of RV shape in the target population is implicitly handled by solving the energy minimization simultaneously in either multiple views of the same patient, or across multiple patients. Commercially, the only widely available tool for RV segmentation in 3DTTE is the TomTec 4D RV-Function software (TomTec Imaging Systems, Munich, Germany). Although the tool has been extensively validated in the literature [16]–[21], the details of the underlying method have not been published.

Segmentation of the LV in 3DTTE has received much more attention in the literature, and methods based on deformable models [22], [23], active shape models [24], active appearence models [25], level-set methods [14], and classification [26] have been presented. However, these methods are not immediately applicable to the RV, as there are several challenges that are especially demanding in this case, such as an increased anatomical complexity and larger inter-patient variation. For echocardiography in particular, challenges such as weaker myocardial borders, thin walls, pronounced trabeculations and reduced image

quality make the problem even harder. Furthermore, 3DTTE acquisition of the whole RV is generally challenging. For instance, the anterior part of the RV outflow tract (RVOT) and anterio-lateral RV free wall is typically shadowed by the sternum or lung tissue [21].

In this paper we present a computationally efficient RV segmentation algorithm for 3DTTE images. The algorithm is based on a real-time segmentation framework previously applied to the LV [22]. We represent the RV anatomy using a compact geometric model with modes of variation derived from manual segmentations in MRI sequences. The segmentation problem is then solved by state estimation using a Kalman filter to combine the geometric model with edge detections and a motion model. Evaluation of the method was performed by comparing volumes and ejection fractions in 17 clinical cases to MRI and a state-of-the-art commercial 3DTTE RV assessment tool, as well as surface-surface distance metrics with respect to manual segmentations in 3DTTE.

II. METHODS

The method presented here is an application of a previously published real-time volumetric segmentation framework that has been applied to the left ventricle [22], [27] and aortic root [28]. In this framework, a compact geometric representation of the target anatomy is fitted to the image by fusing information from the shape prior, edge detection, and motion prediction using an extended Kalman filter. The underlying RV geometry is represented by a subdivision surface and extended with a Statistical Shape Model (SSM) derived from manual segmentations in cine MRI images.

A. Model

1) Surface Model: We represent the underlying RV geometry with a Doo-Sabin subdivision surface. This type of surface is a generalization of a uniform quadratic B-spline for arbitrary topology, defined by a control polyhedron which is interpolated to create a smooth surface.

Orderud *et al.* [27] have described a computationally efficient method of calculating the basis functions at arbitrary surface locations. For a location (u, v) on the surface defined by a control polyhedron with vertices $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_N]$, the transformed point $\mathbf{p} \in \mathbb{R}^3$ is given by

$$\mathbf{p}(u, v) = \sum_i^N b_i(u, v)\mathbf{u}_i = \mathbf{U}\mathbf{b}(u, v). \tag{1}$$

where $\mathbf{b}$ is a vector of basis functions. This surface deformation is denoted $\mathbf{T}_s(\mathbf{U})$. Note that the basis functions are uniquely defined by the topology of the control polyhedron, which means they can be precomputed for a given surface distribution. This is important for computational efficiency [27]. The RV subdivision surface is defined by a control polyhedron consisting of 32 vertices; 12 on the lateral free wall, 9 at the septum, 5 at the Tricuspid Valve (TV) annulus, 5 at the Pulmonary Valve (PV) annulus and 1 at the apex. Fig. 1 shows the control polyhedron and the resulting surface.

2) Hinge Transform: One of the anatomical variations of the RV is the orientation of the Right Ventricle Outflow Tract (RVOT) with respect to the RV long axis. In order to capture
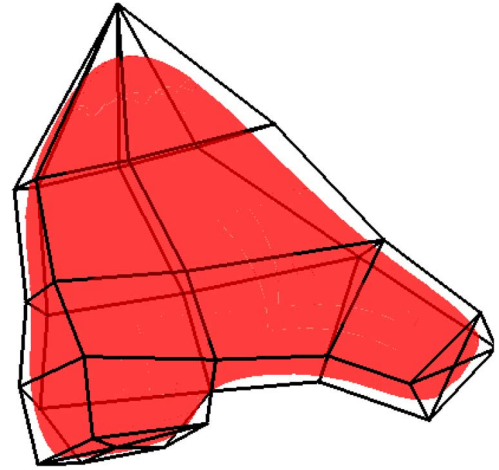

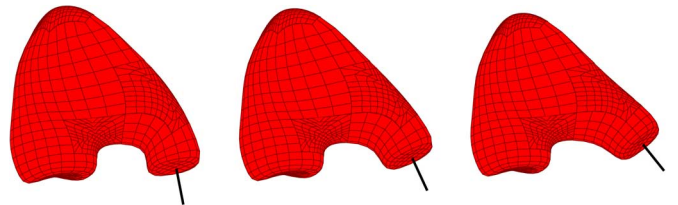Fig. 1. Control polyhedron and resulting Doo-Sabin surface representation of the right ventricle.

Fig. 2. Mean model with changing hinge transform parameter. From left to right: $-15°$, 0 and $15°$ rotations. The black line illustrates the normal direction of the pulmonary valve, for reference.

this variability in a natural way, we introduce a hinge transform for the pulmonary valve control vertices. This transform is a rotation about the apex-PV-TV plane normal direction and a displacement in the RVOT long axis, as demonstrated in Fig. 2.

The hinge transform $\mathbf{V} = \mathbf{T}_h(\mathbf{x}_h, \mathbf{U})$ is applied by transforming the control vertices $\mathbf{U}$,

$$\mathbf{v}_i = \begin{cases} \mathbf{R}(\theta_h)\mathbf{u}_i + d_h\mathbf{d}_{\mathrm{RVOT}} & \text{if } i \in \Omega_H \\ \mathbf{u}_i & \text{otherwise,} \end{cases} \tag{2}$$

where $\mathbf{x}_h = [\theta_h, d_h]$ are the rotation and displacement parameters and Ω_H is the set of control vertices for which the hinge applies.

3) Statistical Shape Model: Statistical Shape Models (SSM) have been established as a robust tool for modeling cardiac structures [29]. By building a statistical model from manual segmentations, one can introduce strong prior shape information while simultaneously handling shape variation in a natural way. SSMs are normally constructed by performing Principal Component Analysis (PCA) on a training matrix whose rows are coordinates for all landmarks in manually segmented point distribution models. This results in a set of modes of variation ϕ_m such that any valid shape $\mathbf{x}$ can be approximated by

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \sum_i \beta_i \boldsymbol{\phi}_i, \tag{3}$$

where $\bar{\mathbf{x}}$ is the mean model and β_i are the shape parameters.

We derived the mean model and statistical modes of variation from manual segmentations of short axis cine-MRI recordings from an open access database [30]. 14 patients with high

quality RV images were selected from the database, constituting 5 patients with myocardial infarction, 4 with non-infarct heart failure, 3 with hypertrophy and 2 healthy subjects. Stack misalignment was corrected by fitting the LV center points to a second order polynomial [31]. All recordings consisted of 20 phases resulting in a total of 280 images.

For each reference MRI image, the RV endocardial borders were manually traced in all slices, and the pulmonary valve center, tricuspid valve center and RV apex landmarks were identified, resulting in a set of manual trace points $\mathcal{P} = \{\mathbf{p}_i\}$. The Doo-Sabin surface was fitted to each trace set by minimizing the sum of Euclidean distances between $\mathcal{P}$ and corresponding surface points,

$$(\hat{\mathbf{U}}, \hat{\mathbf{x}}_h) = \arg\min_{\mathbf{U},\mathbf{x}_h} \frac{1}{2} \sum_{i \in \mathcal{P}} \|\mathbf{p}_i - \mathbf{q}_i\|^2, \tag{4}$$

where $\mathbf{q}_i$ is the projection onto the surface $\mathbf{T}_s(\mathbf{T}_h(\mathbf{U}, \mathbf{x}_h))$ of the trace point $\mathbf{p}_i$. The optimization was solved using gradient descent,

$$\hat{\mathbf{y}}_{k+1} = \hat{\mathbf{y}}_k + \lambda \sum_i (\mathbf{p}_i - \mathbf{q}_i)^\top \mathbf{J}_i|_{\hat{\mathbf{y}}_k}, \tag{5}$$

where

$$\mathbf{y} = \left[\mathbf{x}_h^\top, \mathbf{u}_1^\top, \mathbf{u}_2^\top, \cdots, \mathbf{u}_N^\top\right]^\top \tag{6}$$

is a vector of optimization variables, and $\mathbf{J}_i$ is the associated Jacobian matrix

$$\mathbf{J}_i = \frac{\partial \mathbf{q}_i}{\partial \mathbf{y}} = \left[\frac{\partial \mathbf{T}_h}{\partial \mathbf{x}_h}^\top, \left(\frac{\partial \mathbf{T}_s}{\partial \mathbf{T}_h}\frac{\partial \mathbf{T}_h}{\partial \mathbf{U}}\right)^\top\right]^\top. \tag{7}$$

The hinge was removed by applying $\mathbf{T}_h^{-1}$ on the control vertices of each training surface, before rigidly registering to a randomly selected surface, allowing for anisotropic scaling. In order to perform PCA on surface points as opposed to control vertices, the surface points $\mathbf{Q}$ corresponding to the control vertices $\mathbf{P}$ were calculated by $\mathbf{Q} = \mathbf{B}\mathbf{P}$ where $\mathbf{B}$ is a matrix of Doo-Sabin basis functions at the centroid of each surface patch.

After performing PCA on the surface locations, each mode $\boldsymbol{\phi}_i$ was projected back into the control vertex space by $\boldsymbol{\psi}_i = \mathbf{B}^{-1}\boldsymbol{\phi}_i$. To reduce the parameter space, the first N_m modes constituting at least 95% of the total variation were kept. The local control vertices of every valid shape $\mathbf{U}$ can then be approximated by a linear combination of the modes of variation,

$$\hat{\mathbf{U}} = \bar{\mathbf{U}} + \sum_{i=1}^{N_m} \mathbf{x}_{s,i} \boldsymbol{\psi}_i, \tag{8}$$

where $\mathbf{x}_s$ is the shape parameters and $\bar{\mathbf{U}}$ is the mean model. This transform is denoted $\mathbf{T}_{ssm}(\mathbf{x}_s)$.

4) Global Transform: In order to align the model in the image space, we introduce a global transform allowing translation, anisotropic scaling, and rotation. The global transform is denoted $\mathbf{T}_g(\mathbf{x}_g)$ where $\mathbf{x}_g$ is the 9-degrees-of-freedom transform state vector.

5) Motion Model: In order to capture the movement of the Atrio-Ventricular (AV) plane during the cardiac cycle, we introduce a motion model. The hinge and global transforms can be

defined by three landmarks; valve center, tricuspid valve center, and the RV apex position. We model the trajectory $\mathbf{p}(t)$ of each of these landmarks as a simple linear motion

$$\hat{\mathbf{p}}(t) = \mathbf{p}_{ED} + \bar{\alpha}(t)(\mathbf{p}_{ES} - \mathbf{p}_{ED}), \tag{9}$$

where $\mathbf{p}_{ES}$ and $\mathbf{p}_{ED}$ are the landmark positions in the End Diastolic (ED) and End Systolic (ES) frames.

For each of the 14 MRI images used for statistical analysis, the landmarks $\mathbf{p}(t)$ were manually traced during the cardiac cycle and the coefficient $\alpha(t)$ was calculated by

$$\alpha(t) = \frac{(\mathbf{p}(t) - \mathbf{p}_{ED}) \cdot (\mathbf{p}_{ES} - \mathbf{p}_{ED})}{\|\mathbf{p}_{ES} - \mathbf{p}_{ED}\|^2}. \tag{10}$$

Each $\bar{\alpha}$ was then calculated as the mean of all training trajectories.

6) Complete Model: The combined transform between a local point $\mathbf{q} = (u, v)$ on the surface to the global space is

$$\mathbf{T}(\mathbf{q}, \mathbf{x}) = \mathbf{T}_g\left(\mathbf{T}_s\left(\mathbf{T}_h\left(\mathbf{T}_{ssm}(\mathbf{x}_s), \mathbf{x}_h\right), \mathbf{q}\right), \mathbf{x}_g\right). \tag{11}$$

B. Segmentation

The method presented here is an application of a previously presented real-time volumetric segmentation framework [22], [27]. The segmentation is represented as a state estimation problem and solved with an extended Kalman filter with the following processing chain.

1) The model is initialized in the image by aligning it with manually identified landmarks.
2) Iteratively for each consecutive frame in the cardiac cycle, starting at ED,
   a) The next state is predicted according to the motion model.
   b) Edges are detected normal to the predicted surface location.
   c) The prediction and edge detections are fused to form an updated state estimate.
   d) The shape parameters are normalized to a plausible shape.
3) The previous iteration is repeated, iterating backwards in time.
4) Forwards and backwards iterations are fused, resulting in the final segmentation.

Following an approach previously applied to segmentation of the aortic root [28], we perform the whole processing cycle in two passes. First, the segmentation is performed using a learned prediction step and a restricted model allowing no local deformations, which results in a rough segmentation and tracking of the RV during the cardiac cycle. This result is then used as the motion model for the second pass, in which local model deformations are allowed, resulting in the final segmentation.

1) Initialization: The segmentation is initialized by manually identifying three landmarks in both ED and ES; RV apex, Pulmonary Valve (PV) center, and Tricuspid Valve (TV) center.

Based on the landmarks $\mathcal{L}_I = \{\mathbf{p}_{apex}, \mathbf{p}_{tv}, \mathbf{p}_{pv}\}$ in the image space, an initial alignment state $\mathbf{x}_0$ for the hinge and global transforms is calculated by aligning $\mathcal{L}_I$ with the corresponding landmarks $\mathcal{L}_M = \{\mathbf{q}_{apex}, \mathbf{q}_{tv}, \mathbf{q}_{pv}\}$ in the model space.

2) Prediction: The time domain dynamics of the model is controlled in the Kalman filter prediction step, similar to previous works [28]. We use a combination of the previous estimate $\hat{\mathbf{x}}_{k-1|k-1}, \hat{\mathbf{P}}_{k-1|k-1}$ and a regularization state $\hat{\mathbf{x}}_{0,k}$ to predict the next estimate by

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}\hat{\mathbf{x}}_{k-1|k-1} + (\mathbf{I} - \mathbf{A})\mathbf{x}_{0,k} \tag{12}$$

$$\hat{\mathbf{P}}_{k|k-1} = \mathbf{A}\hat{\mathbf{P}}_{k-1|k-1}\mathbf{A}^{\top} + \mathbf{Q}_k, \tag{13}$$

where $\mathbf{A}$ is a diagonal matrix specifying the regularization strength for each state and $\mathbf{Q}_k$ is the estimated prediction noise. The diagonal elements of $\mathbf{A}$ were chosen separately for translation, scaling, rotation and deformation states.

In the first pass, $\mathbf{x}_{0,k}$ is calculated using the initialization method described in Section II-B-1 using the motion model landmark points described in Section II-A-5. In the second pass, $\mathbf{x}_{0,k}$ is the result of the first segmentation pass.

3) Edge Detection: In order to drive the surface towards the image borders, edge detection is performed normal to the predicted surface at 650 evenly distributed points. Each local edge point $\mathbf{p}_l$ with associated unit normal $\mathbf{n}_l$ is transformed to the global space $\mathbf{p}_g$, $\mathbf{n}_g$, and the corresponding Jacobian matrix $\mathbf{J}_g$ is calculated by

$$\mathbf{p}_g = \mathbf{T}_g(\mathbf{p}_l; \hat{\mathbf{x}}) \tag{14}$$

$$\mathbf{n}_g = |\mathbf{M}|(\mathbf{M}^{-1})^{\top}\mathbf{n}_l \text{ where } \mathbf{M} = \left.\frac{\partial \mathbf{T}_g(\mathbf{p}; \mathbf{x})}{\partial \mathbf{p}}\right|_{\substack{\mathbf{p}_l \\ \hat{\mathbf{x}}}} \tag{15}$$

$$\mathbf{J}_g = \left.\frac{\partial \mathbf{T}}{\partial \mathbf{x}}\right|_{\substack{\mathbf{p}_l \\ \hat{\mathbf{x}}}} = \left[\frac{\partial \mathbf{p}}{\partial \mathbf{x}_g}, \frac{\partial \mathbf{p}}{\partial \mathbf{x}_h}, \frac{\partial \mathbf{p}}{\partial \mathbf{x}_s}\right], \tag{16}$$

where $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{k|k-1}$ and

$$\frac{\partial \mathbf{p}}{\partial \mathbf{x}_g} = \frac{\partial \mathbf{T}_g(\mathbf{x}_g)}{\partial \mathbf{x}_g} \tag{17}$$

$$\frac{\partial \mathbf{p}}{\partial \mathbf{x}_h} = \frac{\partial \mathbf{T}_g(\mathbf{q})}{\partial \mathbf{q}}\frac{\partial \mathbf{T}_s(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{T}_h(\mathbf{x}_h)}{\partial \mathbf{x}_h} \tag{18}$$

$$\frac{\partial \mathbf{p}}{\partial \mathbf{x}_s} = \frac{\partial \mathbf{T}_g(\mathbf{q})}{\partial \mathbf{q}}\frac{\partial \mathbf{T}_s(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{T}_h(\mathbf{U})}{\partial \mathbf{U}}\frac{\partial \mathbf{T}_{ssm}(\mathbf{x}_s)}{\partial \mathbf{x}_s}, \tag{19}$$

due to the chain rule.

Edge displacements v_i are detected by searching along $\mathbf{n}_g$ around $\mathbf{p}_g$ using the least mean squares fit to an intensity step function. We use 40 samples with 1 mm spacing for each edge profile, resulting in a capture range of 4 cm. Outlier edges are rejected based on the intensity step height and differences between neighboring edges.

The edge points are grouped into 5 regions; Septum, posterior LV-RV attachment, posterolateral free wall, anterolateral free wall including the outflow tract and anterior LV-RV attachment, and base, as shown in Fig. 3. Each group has different *a priori* uncertainties of the edge measurements, used to express the common appearance of RV 3DTTE images. The uncertainties are highest for the anterolateral free wall and lowest for the septum and posterolateral free wall.

For each region k, each measured edge displacement $v_{k,i}$ has an associated measurement noise with estimated variance $r_{k,i}$. It is generally hard to properly estimate the variance $r_{k,i}$ of each
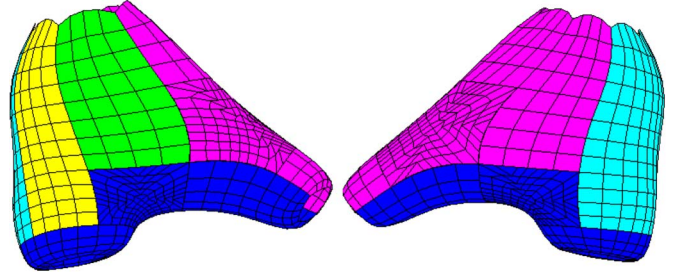

Fig. 3. Model regions as seen from medial (left) and lateral (right) sides; septum (green), posterior LV-RV attachment (yellow), posterolateral free wall (cyan), anterolateral free wall (magenta), and base (blue).
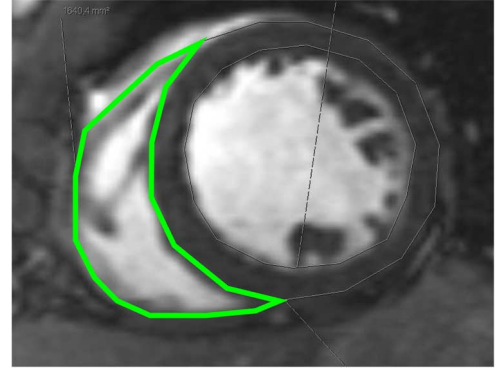

Fig. 4. Example of manual segmentation in MRI short axis view.

edge profile in the ultrasound image. To simplify, we calculate the quality-of-fit $q_{k,i}$ for each edge profile as the sum of squared deviations in the intensity fit. The variances are then normalized by

$$r_{k,i} = R_k q_{k,i} \sum_j \frac{1}{q_{k,j}} \tag{20}$$

such that the amount of measurement influence $\sum_i r_{k,i}^{-1} = R_k^{-1}$ is consistent across datasets, where R_k is the *a priori* variance of the edge detection of the k'th region.

4) State Update: The update step in the Kalman filter fuses the previous estimate, state prediction, and detected edges. In order to speed up the calculations, the Kalman gain is simplified by assuming independent measurement noises [22]. With this simplification, the update step can be written as

$$\mathbf{h}_i^{\top} = \mathbf{n}_i^{\top}\mathbf{J}_i \tag{21}$$

$$\mathbf{P}_{k|k}^{-1} = \mathbf{P}_{k|k-1}^{-1} + \sum_i \mathbf{h}_i r_i^{-1}\mathbf{h}_i^{\top} \tag{22}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{P}_{k|k}\sum_i \mathbf{h}_i r_i^{-1}v_i, \tag{23}$$

where $\hat{\mathbf{x}}_{k|k}$ and $\mathbf{P}_{k|k}^{-1}$ are the updated state and covariance matrix, and v_i, r_i, $\mathbf{n}_i$ and $\mathbf{J}_i$ are the edge displacement, edge noise variance, normal vector, and surface Jacobian matrix respectively.

This computation is efficient as it does not require inversion of matrices with size dependent on the number of measurements.

5) Normalization: The deformation states $\boldsymbol{\beta}$ were modeled as a multi-variate gaussian distribution,

$$M(\boldsymbol{\beta}) = \sum_{i=1}^{N_m} \frac{\beta_i^2}{\lambda_i} \sim \chi^2. \tag{24}$$

To constrain the allowed deformation to plausable shapes, the deformation states $\boldsymbol{\beta}_{k|k}$ for which $M(\boldsymbol{\beta}_{k|k}) \geq M_t$ are projected onto the hyperellipsoid $M(\boldsymbol{\beta}_{k|k}) = M_t$ where M_t is the 95% threshold of the χ^2 distribution according to normal SSM practice [29].

6) Kalman Smoother: In order to prevent the segmentation from lagging behind the image borders and to regularize towards a cyclic volume curve, we apply Kalman smoothing by tracking forwards and backwards in time as previously described [28].

The Kalman filter is iterated forward over frames $k = 1, 2, \ldots, N$ to produce estimate $\mathbf{x}_{f,k}$ with estimated covariance $\mathbf{P}_{f,k}$. Backwards iteration over frames $k = N, N-1, \ldots, 1$ produces $\mathbf{x}_{b,k}$ and $\mathbf{P}_{b,k}$. The forward and backward state estimates are then assimilated by

$$\mathbf{P}_k = \left(\mathbf{P}_{f,k}^{-1} + \mathbf{P}_{b,k}^{-1} \right)^{-1} \tag{25}$$

$$\hat{\mathbf{x}}_k = \mathbf{P}_k \left(\mathbf{P}_{f,k}^{-1} \hat{\mathbf{x}}_{f,k} + \mathbf{P}_{b,k}^{-1} \hat{\mathbf{x}}_{b,k} \right). \tag{26}$$

C. Validation

1) Model: To ensure that the subdivision surface was able to represent the RV geometry with the limited degrees of freedom introduced in Section II-A-1, the mean and maximum (Hausdorff) euclidean distances between all points on the manual endocardial contour and corresponding points on the fitted surface were calculated.

To further ensure that the subspace consisting of the selected modes of variation was suitable to represent the observed range of anatomies, a leave-one-patient-out statistical shape model reconstruction validation was performed. For each patient, the mean model parameters $\bar{\mathbf{x}}$ and modes of variation $\boldsymbol{\Phi}$ were constructed for the remaining 13 patients as described in Section II-A-3, selecting the N_m modes constituting at least 95% of the variation. For each surface of the unseen patient, the reconstructed surface was calculated using parameters

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \boldsymbol{\Phi}\hat{\boldsymbol{\beta}}, \tag{27}$$

where $\hat{\boldsymbol{\beta}} = \arg\min_{\boldsymbol{\beta}} \|\hat{\mathbf{x}}(\boldsymbol{\beta}) - \mathbf{x}\|$ and $\mathbf{x}$ are the parameters of the unseen surface. The mean signed and unsigned euclidean distances between points on the unseen MRI traces and reconstructed surfaces were calculated, as well as the Hausdorff distances.

2) Segmentation: The segmentation was validated retrospectively on 3DTTE recordings of 17 patients with aortic insufficiency. The patient characteristics are given in Table I. Segmentation quality was evaluated by comparing End Diastolic Volume (EDV), End Systolic Volume (ESV), Stroke Volume (SV) and Ejection Fraction (EF) between the proposed method and manual measurements in MRI, as well as

TABLE I
PATIENT CHARACTERISTICS OF THE 17 CLINICAL
CASES USED FOR VALIDATION

Age, yr	47 ± 15
Male, n	15 (88 %)
Weight, kg	84 ± 14
Height, cm	178 ± 9
LV hypertension, n	1 (6 %)
LV hypertrophy, n	10 (63 %)

a state-of-the-art commercial segmentation tool for 3DTTE (TomTec 4D RV-Function, version 2.0, TomTec Imaging Systems, Munich, Germany) with and without manual contour correction.

In addition, the following surface-surface distance metrics between the automated and manual 3DTTE segmentations were calculated: mean and median absolute distance, mean and median signed distance, and Hausdorff distance.

a) 3DTTE acquisition: 3DTTE recordings of the RV were acquired from an apical position. The probe was tilted slightly in order to get as much as possible of the RV into the sector. The images were recorded using ECG-gated multi-beat acquisitions under breath hold. All images were acquired on a Vivid E9 scanner using a 4 V probe (GE Vingmed Ultrasound AS, Horten, Norway). The mean temporal resolution and sector size were 28 ms and $63 \times 63°$ respectively. 15 recordings were acquired from 6 beats, and 2 recordings from 4 beats.

b) 3DTTE automated segmentation: The RV apex, pulmonary valve, and tricuspid valve were manually identified in ES and ED using a dedicated software. The proposed automated segmentation method was then run on the image without any further manual involvement.

c) 3DTTE reference segmentation: The TomTec 4D RV-Function software was used as a reference for RV assessment in 3DTTE. For each image, the analysis starts by manually identifying the following 10 landmarks in a dedicated software; mitral and tricuspid valve center, LV and RV apex, anterior and posterior RV-LV attachment points as well as medial and lateral RV endocardial border points in a single short axis slice, and two points spanning the aortic valve annulus. The software then automatically tracks the endocardial borders during the cardiac cycle and generates corresponding volume loops. After tracking, all contours can be manually adjusted. This software was run twice for each image, by an experienced cardiologist. Firstly, extensive manual contour corrections were performed to provide a ground-truth manual 3DTTE reference. Secondly, the tool was run without any manual interaction after initialization, in order to provide a comparable algorithmic reference for the proposed method.

d) MRI acquisition: Magnetic resonance images were acquired with Siemens 1.5 tesla scanners (Siemens Avanto and Siemens Sonata; Siemens Medical Systems, Erlangen, Germany), using a breath-hold, prospectively ECG-triggered, segmented, balanced steady-state free precession gradient-echo cine sequence with minimum echo and repetition times. Slices were 6 mm thick with a 4 mm short-axis interslice gap, a spatial resolution of 1.9×1.3 mm, and a temporal resolution of 30–35 ms. Endocardial borders were traced manually at a PACS work station (Sectra Medical Systems AB, Linköping, Sweden).

$\mu - 3\sigma$　　　　　　　$\mu + 3\sigma$

Mode 1

Mode 2

Mode 3

Mode 4

Fig. 5.   The most dominant modes of variation of the statistical shape model.

Right and left ventricular volumes and ejection fractions were calculated by short axis slice summation.

3) Statistical Analysis: Statistical analysis was performed using MedCalc, version 15.2 (MedCalc Software, Ostend, Belgium). Correlation between measurements were assessed by the Intra-class Correlation Coefficient (ICC) using a two-way model with absolute agreement. Continuous variables are presented as mean $\pm$ SD.

III. RESULTS

A. Model

For the 280 fitted MRI segmentations, the mean distance between the manual trace and fitted model was 1.8 ± 0.37 mm with a mean signed distance of 0.30 ± 0.12 mm and Hausdorff distance 7.6 ± 2.3 mm.

12 deformation modes were necessary to cover at least 95% of the observed variation. The first modes of variation contained the expansion, elongation, and curvature around the LV, shown in Fig. 5. The leave-one-patient-out reconstruction resulted in a surface-surface mean absolute distance error of 2.1 ± 0.4 mm, signed distance error 0.4 ± 0.4 mm, and Hausdorff distance 8.7 ± 3.0 mm.

B. Segmentation

All recordings were successfully segmented with computation time 2.0 ± 0.33 s (53 ± 4.1 ms per frame) on a standard laptop. The mean signed and unsigned surface-surface distances between the automated and 3DTTE reference segmentations were 0.5 ± 1.4 mm and 3.6 ± 0.7 mm respectively. The corresponding median surface-surface distances were 0.2

± 0.7 mm and 3.0 ± 0.6 mm, and the Hausdorff distances were 11.6 ± 2.0 mm. A comparison of volumes and ejection fractions for all methods is reported in Table II. Figs. 6 and 7 show a Bland-Altman analysis, and clustering and agreement between the methods respectively. Example segmentations are shown in Figs. 4 and 8.

IV. DISCUSSION

Fitting the proposed subdivision surface to manual segmentations in MRI images for statistical shape analysis resulted in an acceptable residual distance between the fitted surface and manual trace, and an associated mean signed distance close to zero. Although this error could be reduced by increasing the number of control vertices of the RV model, this would come at the cost of reduced segmentation robustness. We conclude that the surface was a sufficiently accurate representation of the observed RV geometries, and the compact representation did not significantly restrict the expressivity of the surface. Furthermore, the statistical shape model using only 12 modes of variation was able to accurately represent the observed geometrical distribution in a leave-one-patient-out experiment.

The method achieved a Mean Absolute Distance (MAD) error of 3.6 mm and Hausdorff Distance (HD) of 11.6 mm. It is natural to compare this to to LV segmentation methods, where MAD and HD errors of about 2.3 mm and 8.5 mm have been reported for current state-of-the-art methods [33], including state-estimation approaches similar to the proposed method [34]. The noticeably higher surface errors illustrate how challenging the RV segmentation problem is in comparison to the LV, because of the increased shape complexity and reduced

TABLE II
COMPARISON BETWEEN MANUAL MEASUREMENTS IN MRI, REFERENCE METHOD IN 3DTTE WITH AND WITHOUT MANUAL
CONTOUR CORRECTION, AND THE PROPOSED METHOD FOR EDV, ESV, SV AND EF. (a) End Diastolic Volume (mL);
(b) End Systolic Volume (mL); (c) Stroke Volume (mL); (d) Ejection Fraction (%)

(a)

	MRI	3DTTE reference (manual)	3DTTE reference (automatic)
Reference man.	−32.4 ± 36.3 (0.63)		
Reference auto.	−58.8 ± 40.6* (0.39)	−26.4 ± 29.1 (0.67)	
Proposed	−25.8 ± 23.7 (0.79)	6.6 ± 29.7 (0.78)	33.1 ± 32.9* (0.56)

(b)

	MRI	3DTTE reference (manual)	3DTTE reference (automatic)
Reference man.	−28.4 ± 28.6* (0.54)		
Reference auto.	−23.3 ± 31.7 (0.52)	5.1 ± 24.0 (0.67)	
Proposed	−15.7 ± 25.5 (0.73)	12.7 ± 17.1 (0.80)	7.6 ± 29.1 (0.56)

(c)

	MRI	3DTTE reference (manual)	3DTTE reference (automatic)
Reference man.	−4.0 ± 19.4 (0.54)		
Reference auto.	−35.5 ± 20.6* (0.13)	−31.5 ± 17.3* (0.26)	
Proposed	−10.0 ± 19.1 (0.38)	−6.0 ± 20.4 (0.42)	25.5 ± 17.8* (0.22)

(d)

	MRI	3DTTE reference (manual)	3DTTE reference (automatic)
Reference man.	4.8 ± 7.6 (0.39)		
Reference auto.	−9.5 ± 10.4* (0.03)	−14.3 ± 8.5* (0.03)	
Proposed	0.0 ± 10.1 (0.29)	−4.8 ± 6.9 (0.47)	9.5 ± 11.5* (-0.10)

Values are mean ± SD (ICC). * $p < 0.05$ by two-tailed Student's t-test. Negative values indicate an underestimation of the row relative to the column.
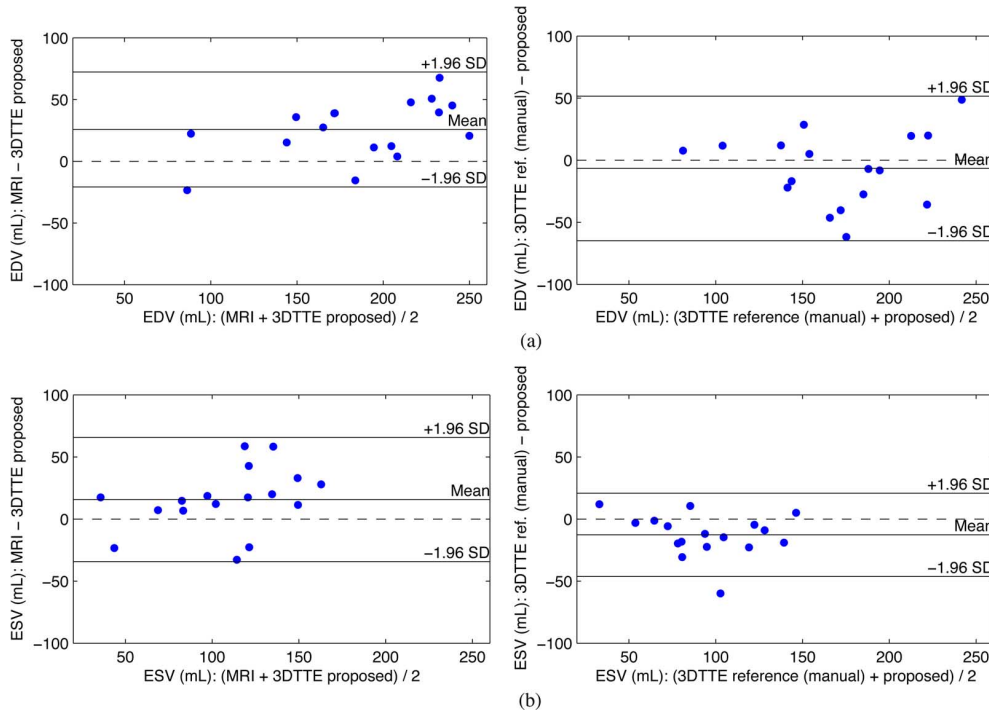

Fig. 6. End diastolic and end systolic volumes quantified by the proposed automated method in 3DTTE compared to MRI and 3DTTE reference (manual). (a) End diastolic volume; (b) end systolic volume.

image quality. The majority of the surface error was observed in the RVOT region. This is where the image quality is poorest, and both the automated and manual segmentations are most uncertain. Additionally, the tool used to generate the reference segmentations has limited flexibility in editing the RVOT.

The observed HD of 11.6 mm is comparable to what was recently reported from the MICCAI'12 MRI RV segmentation challenge, where the lowest HD was 7.3 mm and 9.3 mm for two separate data sets [3]. Recently, Stebbing *et al.* [15] presented a RV segmentation method that, instead of using statistical shape information directly, performs segmentation simultaneously in either multiple images from different views of the same patient, or in images of multiple patients. The method achieved median

signed trace-surface distances of about 1.5 mm (median over 4 cases) for multiple images of a single patient and 1.7 mm (median over 12 cases) for multiple patients. These error metrics are significantly lower than what was shown in our validation (3.0 mm mean). However, the metrics were based on traces in three short axis and two long axis slices which did not cover the RVOT. They are therefore not directly comparable to the surface-surface distances reported in our validation. Finally, our proposed method was significantly faster, requiring on average 2 s per patient compared to about 60 minutes for simultaneous segmentation of 12 patients.

The model contains a hinge transform used to capture differences in the RVOT orientation. This transform was not con-
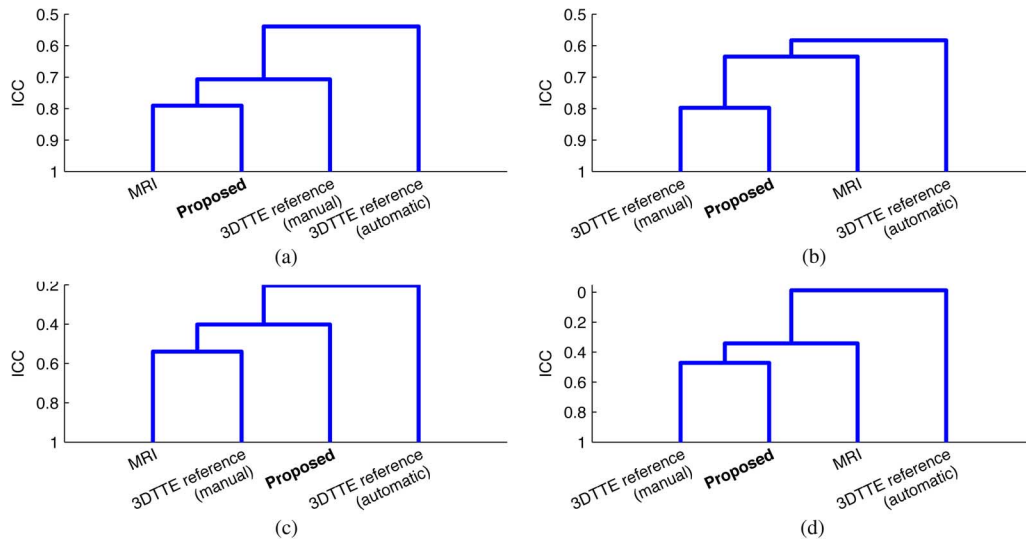
Fig. 7. Dendrograms from agglomerative hierarchical clustering using average linkage [32] of Intra-class Correlation Coefficients (ICC) between manual measurements in MRI, reference method in 3DTTE with and without manual contour correction, and the proposed method. (a) EDV (mL); (b) ESV (mL); (c) SV (mL); (d) EF (%).
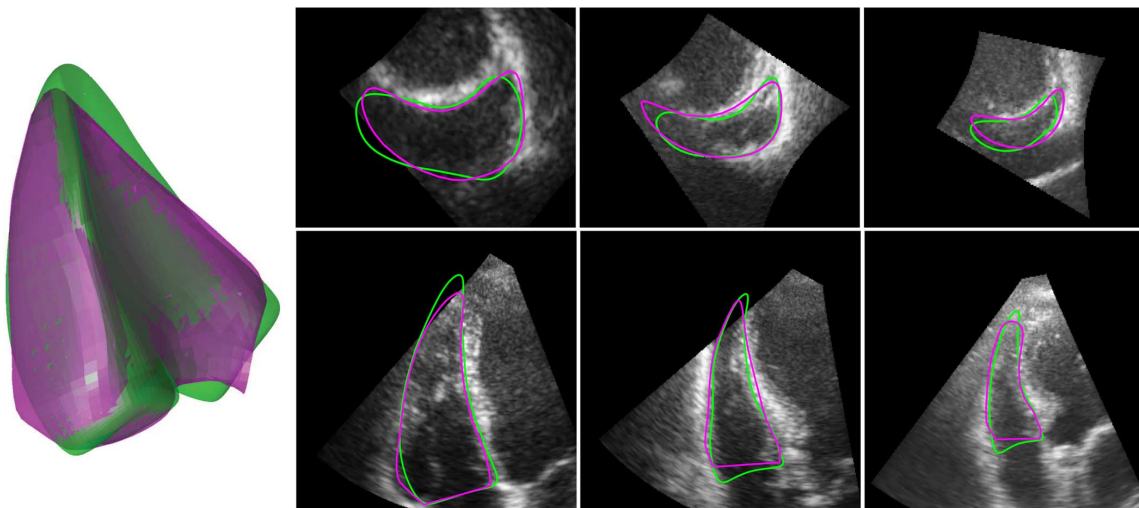

Fig. 8. Example of segmentation by the proposed method (green) compared to the manual reference (magenta). *Left*: Segmented meshes for a single case. *Right*: short axis (top) and long axis (bottom) planes for three different cases.

sidered when building the statistical shape model, as PCA applied in euclidean space is generally not suited to capture rotational movement. This means that the model will lack deformation modes relating the RVOT orientation with RV shape. However, the hinge introduces the necessary degrees of freedom to place the pulmonary valve alignment landmark in a natural way, without any surface deformations. This is both desirable for the user, and results in a good initialization which is important for the Kalman filter driven segmentation.

We observed an underestimation with respect to MRI of the EDV, ESV and SV of both the manual 3DTTE reference and the proposed method, while the EF was unbiased for the proposed method and overestimated for the manual reference. This trend is consistent with results reported in the literature for manual and semi-automatic measurements, except for the manual EF which is typically slightly underestimated [21], [35]. Because of the large differences in methodology between TTE and MRI, perfect correlation should not be expected [17]. For instance, the RV volume in the proposed 3DTTE segmentation method is strictly bounded by the tricuspid and pulmonary valve planes. In contrast, using disk summation in MRI short axis slices, the volume is bounded by a slice approximately in the atrioventricular plane. Compared to TomTec 4D RV-Function without manual contour correction, the proposed method resulted in higher intra-class correlation coefficients and smaller biases compared to both MRI and 3DTTE reference for all quantification variables, EDV, ESV, SV and EF. In addition, the proposed method requires less manual user interaction; identification of 3 landmarks in two phases compared to 10 landmarks in a single phase.

Segmentation in ultrasound images is generally challenging, because of the low spatial resolution and the presence of artifacts such as dropouts, speckles and reverberations. For the RV in particular, the pulmonary valve, RVOT and anterolateral free wall is often acoustically shadowed by the sternum and lungs [21]. Furthermore, irregular trabeculations are typically very pronounced in the apical region of the RV [3]. Robustness of the segmentation method is therefore essential for solving this

problem. The proposed RV surface model has a compact representation and inherently enforces regularization. By combining this with a statistical shape model we are mitigating problems related to edges disappearing due to acoustic shadowing or out of sector motion, by assuming in this region a statistically plausible shape, influenced by the general shape of the ventricle. The simple line search edge detectors are robust to noise in the ultrasound image, and because we combine several of them, we limit the impact of individual errors on the final segmentation. Similarly, using a strong prior geometry and motion model in the Kalman filter ensures a robust segmentation.

An echocardiography procedure is interactive in nature, and a computationally efficient method is therefore vital for widespread usage in clinical practice. The Kalman filter approach is very computationally efficient, with a mean computation time of 2.0 s. Further improvements could be made as the method lends itself to GPU parallelization.

In a clinical setting, it is important that the user is able to adjust the segmentation. The proposed framework naturally extends to user input, as manually annotated points and traces can be introduced as measurements (23). If these measurements are modeled with a low variance, the surface will interpolate the annotations and affect the segmentation during the whole cardiac cycle.

The effectiveness of the statistical shape model approach is dependent on the quality and quantity of the training data. In this study, we used a limited training set of only 14 patients. This can be justified by the fact that the training set contained patients with a variety of diseases. However, the method remains to be proven on different patient populations with diseases relevant to the RV.

A fundamental limitation to the statistical shape method is the inability of the surface model to adequately represent anatomies very different from what was observed in the training set. For some diseases, this problem can be mitigated by increasing the training set without any further changes in methodology. However, in some variations of structural congenital heart disease, such as double outlet right ventricle, the method proposed here is not directly suited, unless the disease is known prior to segmentation and a separate model and training set specific to the disease is used.

V. CONCLUSION

In this study an automated method for segmenting the right ventricle in 3D echocardiography has been described and validated against MRI and manual echocardiographic segmentation. The method is robust and computationally efficient, and resulted in good correlation with both MRI and manual 3DTTE reference in 17 clinical cases.

REFERENCES

[1] L. L. Mertens and M. K. Friedberg, "Imaging the right ventricle-current state of the art," *Nat. Rev. Cardiol.*, vol. 7, no. 10, pp. 551–563, Oct. 2010.

[2] E. R. Valsangiacomo Buechel and L. L. Mertens, "Imaging the right heart: The use of integrated multimodality imaging," *Eur. Heart J.*, vol. 33, no. 8, pp. 949–960, Apr. 2012.

[3] C. Petitjean *et al.*, "Right ventricle segmentation from cardiac MRI: A collation study," *Med. Image Anal.*, vol. 19, no. 1, pp. 187–202, Jan. 2015.

[4] J. Peters *et al.*, "Automatic whole heart segmentation in static magnetic resonance image volumes," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007*. Berlin, Germany: Springer, 2007, vol. 4792, pp. 402–410.

[5] M. S. ElBaz and A. S. Fahmy, "Active shape model with inter-profile modeling paradigm for cardiac right ventricle segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012*. Berlin, Germany: Springer, 2012, vol. 15, pp. 691–698.

[6] Y. Ou, J. Doshi, G. Erus, and C. Davatzikos, "Multi-atlas segmentation of the cardiac MR right ventricle," in *Proc. 3D Cardiovasc. Imag.,*, 2012.

[7] M. Zuluaga, M. Cardoso, M. Modat, and S. Ourselin, "Multi-atlas propagation whole heart segmentation from MRI and CTA using a local normalised correlation coefficient criterion," in *Functional Imaging and Modeling of the Heart*. Berlin, Germany: Springer, vol. 7945, pp. 174–181.

[8] W. Bai, W. Shi, H. Wang, N. Peters, and D. Rueckert, "Multi-atlas based segmentation with local label fusion for right ventricle MR images," in *Proc. 3D Cardiovasc. Imag.*, 2012.

[9] W. Bai *et al.*, "A probabilistic patch-based label fusion model for multi-atlas segmentation with registration refinement: Application to cardiac MR images," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1302–1315, Jul. 2013.

[10] C. M. S. Nambakhsh *et al.*, "Left ventricle segmentation in MRI via convex relaxed distribution matching," *Med. Image Anal.*, vol. 17, no. 8, pp. 1010–1024, Dec. 2013.

[11] D. Grosgeorge, C. Petitjean, J.-N. Dacher, and S. Ruan, "Graph cut segmentation with a statistical shape model in cardiac MRI," *Comput. Vis. Image Understand.*, vol. 117, no. 9, pp. 1027–1035, Sep. 2013.

[12] C. W. Wang, C. W. Peng, and H. C. Chen, "A simple and fully automatic right ventricle segmentation method for 4-dimensional cardiac MR images," in *Proc. 3D Cardiovasc. Imag.*, 2012.

[13] O. Maier, D. Jimenez, A. Santos, and M. Ledesma-Carbayo, "Segmentation of RV in 4d cardiac MR volumes using region-merging graph cuts," in *Comput. Cardiol.*, Sep. 2012, pp. 697–700.

[14] E. D. Angelini *et al.*, "Segmentation of real-time three-dimensional ultrasound for quantification of ventricular function: A clinical study on right and left ventricles," *Ultrasound Med. Biol.*, vol. 31, no. 9, pp. 1143–1158, Sep. 2005.

[15] R. V. Stebbing, A. I. Namburete, R. Upton, P. Leeson, and J. A. Noble, "Data-driven shape parameterization for segmentation of the right ventricle from 3D+t echocardiography," *Med. Image Anal.*, vol. 21, no. 1, pp. 29–39, Apr. 2015.

[16] P. S. Niemann *et al.*, "Anatomically oriented right ventricular volume measurements with dynamic three-dimensional echocardiography validated by 3-Tesla magnetic resonance imaging," *J. Am. Coll. Cardiol.*, vol. 50, no. 17, pp. 1668–1676, Oct. 2007.

[17] G. Leibundgut *et al.*, "Dynamic assessment of right ventricular volumes and function by real-time three-dimensional echocardiography: A comparison study with magnetic resonance imaging in 100 adult patients," *J. Am. Soc. Echocardiogr.*, vol. 23, no. 2, pp. 116–126, Feb. 2010.

[18] H. B. van der Zwaan *et al.*, "Clinical value of real-time three-dimensional echocardiography for right ventricular quantification in congenital heart disease: Validation with cardiac magnetic resonance imaging," *J. Am. Soc. Echocardiogr.*, vol. 23, no. 2, pp. 134–140, Feb. 2010.

[19] A. M. Crean *et al.*, "3D Echo systematically underestimates right ventricular volumes compared to cardiovascular magnetic resonance in adult congenital heart disease patients with moderate or severe RV dilatation," *J. Cardiovasc. Magn. Reson.*, vol. 13, no. 1, p. 78, Jan. 2011.

[20] A. Dragulescu, L. Grosse-wortmann, C. Fackoury, and L. Mertens, "Echocardiographic assessment of right ventricular volumes: A comparison of different techniques in children after surgical repair of tetralogy of Fallot," *J. Am. Soc. Echocardiogr.*, pp. 596–604, 2012.

[21] E. Ostenfeld *et al.*, "Manual correction of semi-automatic three-dimensional echocardiography is needed for right ventricular assessment in adults; validation with cardiac magnetic resonance," *Cardiovasc. Ultrasound*, vol. 10, no. 1, p. 1, Jan. 2012.

[22] F. Orderud, "A framework for real-time left ventricular tracking in 3D+T echocardiography, using nonlinear deformable contours and Kalman filter based tracking," in *Comput. Cardiol.*, Sep. 2006, pp. 125–128.

[23] J. Hansegard, S. Urheim, K. Lunde, S. Malm, and S. Rabben, "Semi-automated quantification of left ventricular volumes and ejection fraction by real-time three-dimensional echocardiography," *Cardiovasc. Ultrasound*, vol. 7, no. 1, p. 18, 2009.

[24] J. Hansegard, F. Orderud, and S. I. Rabben, "Real-time active shape models for segmentation of 3d cardiac ultrasound," in *Computer Analysis of Images and Patterns*. Berlin, Germany: Springer, 2007, vol. 4673, LNCS, pp. 157–164.

[25] K. Y. E. Leung *et al.*, "Improving 3d active appearance model segmentation of the left ventricle with Jacobian tuning," in *Proc. SPIE Med. Imag., Image Process.*, 2008, vol. 6914.

[26] B. Georgescu, X. Zhou, D. Comaniciu, and A. Gupta, "Database-guided segmentation of anatomical structures with complex appearance," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2005, vol. 2, pp. 429–436.

[27] F. Orderud and S. I. Rabben, "Real-time 3D segmentation of the left ventricle using deformable subdivision surfaces," in *IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2008, pp. 1–8.

[28] J. Bersvendsen, J. O. Beitnes, S. Urheim, S. Aakhus, and E. Samset, "Automatic measurement of aortic annulus diameter in 3-dimensional Transoesophageal echocardiography," *BMC Med. Imag.*, vol. 14, no. 1, p. 31, Sep. 2014.

[29] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: A review," *Med. Image Anal.*, vol. 13, no. 4, pp. 543–563, Aug. 2009.

[30] P. Radau *et al.*, "Evaluation framework for algorithms segmenting short axis cardiac MRI," *MIDAS J. Cardiac MR Left Ventricle Segmentation Challenge*, vol. 7, 2009.

[31] C. Swingen *et al.*, "An approach to the three-dimensional display of left ventricular function and viability using MRI," *Int. J. Cardiovasc. Imag.*, vol. 19, no. 4, pp. 325–336, Aug. 2003.

[32] J. Rislien and E. Samset, "A non-parametric permutation method for assessing agreement for distance matrix observations," *Stat. Med.*, vol. 33, no. 2, pp. 319–329, 2014.

[33] D. Barbosa, D. Friboulet, J. D'hooge, and O. Bernard, "Fast tracking of the left ventricle using global anatomical affine optical flow and local recursive block matching," *MIDAS J.*, 10 2014.

[34] E. Smistad and F. Lindseth, "Real-time tracking of the left ventricle in 3d ultrasound using Kalman filter and mean value coordinates," *The MIDAS J.*, 10 2014.

[35] Y. J. Shimada *et al.*, "Accuracy of right ventricular volumes and function determined by three-dimensional echocardiography in comparison with magnetic resonance imaging: A meta-analysis study," *J. Am. Soc. Echocardiogr.*, vol. 23, no. 9, pp. 943–953, Sep. 2010.