IEEE Journal of Translational Engineering in Health and Medicine

Received 21 November 2019; revised 26 March 2020; accepted 15 April 2020.
Date of publication 24 April 2020; date of current version 26 May 2020.

Digital Object Identifier 10.1109/JTEHM.2020.2989390

3D Motion Estimation of Left Ventricular Dynamics Using MRI and Track-to-Track Fusion

KUMARADEVAN PUNITHAKUMAR[1,2,3], (Senior Member, IEEE), ISMAIL BEN AYED[4], (Member, IEEE), ABRAAM S. SOLIMAN[5], AASHISH GOELA[6], ALI ISLAM[7], SHUO LI[6], (Senior Member, IEEE), AND MICHELLE NOGA[1,2]

[1]Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, AB T6G 2R3, Canada
[2]Servier Virtual Cardiac Centre, Mazankowski Alberta Heart Institute, Edmonton, AB T6G 2B7, Canada
[3]Department of Computing Science, University of Alberta, Edmonton, AB T6G 2R3, Canada
[4]École de Technologie Supérieure (ÉTS), Montreal, QC H3C 1K3, Canada
[5]MR R&D, Philips Healthcare, 5684 PC Best, The Netherlands
[6]Department of Medical Imaging, Western University, London, ON N6A 3K7, Canada
[7]St. Joseph's Health Care London, London, ON N6A 4V2, Canada

CORRESPONDING AUTHOR: K. PUNITHAKUMAR (punithak@ualberta.ca)

This work was supported in part by an NSERC Discovery Grant, in part by a grant from Servier Canada Inc., in part by the Heart and Stroke Foundation of Alberta, NWT, and Nunavut, through the Research Leadership Fund, and in part by the Mazankowski Alberta Heart Institute.

This article has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the author.

ABSTRACT Objective: This study investigates the estimation of three dimensional (3D) left ventricular (LV) motion using the fusion of different two dimensional (2D) cine magnetic resonance (CMR) sequences acquired during routine imaging sessions. Although standard clinical cine CMR data is inherently 2D, the actual underlying LV dynamics lies in 3D space and cannot be captured entirely using single 2D CMR image sequences. By utilizing the image information from various short-axis and long-axis image sequences, the proposed method intends to estimate the dynamic state vectors consisting of the position and velocity information of the myocardial borders in 3D space. Method: The proposed method comprises two main components: tracking myocardial points in 2D CMR sequences and fusion of multiple trajectories correspond to the tracked points. The tracking which yields the set of corresponding temporal points representing the myocardial points is performed using a diffeomorphic nonrigid image registration approach. The trajectories obtained from each cine CMR sequence is then fused with the corresponding trajectories from other CMR views using an unscented Kalman smoother (UKS) and a track-to-track fusion algorithm. Results: We evaluated the proposed method by comparing the results against CMR imaging with myocardial tagging. We report a quantitative performance analysis by projecting the state vector estimates we obtained onto 2D tagged CMR images acquired from the same subjects and comparing them against harmonic phase estimates. The proposed algorithm yielded a competitive performance with a mean root mean square error of 1.3±0.5 pixels (1.8±0.6 mm) evaluated over 118 image sequences acquired from 30 subjects. Conclusion: This study demonstrates that fusing the information from short and long-axis views of CMR improves the accuracy of cardiac tissue motion estimation. Clinical Impact: The proposed method demonstrates that the fusion of tissue tracking information from long and short-axis views improves the binary classification of the automated regional function assessment.

INDEX TERMS Multiview fusion, cardiac motion estimation, magnetic resonance imaging, nonlinear dynamic state estimation, regional wall motion abnormality.

I. INTRODUCTION

The assessment of the regional left ventricular (LV) function plays an important role in the diagnosis of coronary heart disease [1]–[4]. Cardiac magnetic resonance (CMR) imaging offers a non-invasive solution to capture the LV function. Manual assessment of the LV function over numerous images

produced by the CMR is tedious, time-consuming and expensive. Many automated methods to quantitatively assess the LV regional function have been proposed in the literature [1]–[4]. However, existing methods utilized only two-dimensional (2D) short-axis sequences of the CMR in generating regional function measurements. Due to high inter-slice distance[1] between the short-axis sequences acquired along multiple positions of the LV, these 2D sequences do not entirely capture the LV dynamics that lie in 3D space and consist of a complex combination of movements. Due to lack of through-plane information, standard single view 2D CMR sequences are inadequate for estimating the LV dynamics in 3D space. One can overcome this limitation by exploiting the data from other image sequences that are orthogonal to obtain 3D dynamics and an accurate assessment of the LV function.

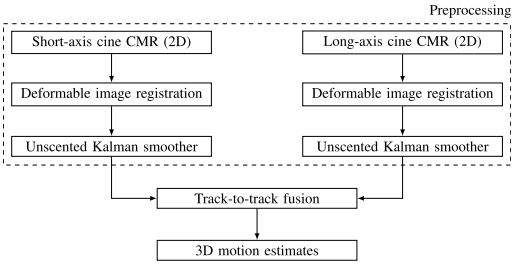


FIGURE 1. The overall system diagram showing the components of the proposed 3D motion estimation algorithm that fuses information from short-axis and long-axis CMR sequences.

This study proposes a novel approach to fuse motion information available in multiple 2D image views, thereby, utilizing the information available in different CMR short-axis sequences as well as long-axis sequences such as 2-chamber, 3-chamber and 4-chamber views. Fig. 1 depicts the overall system diagram of the proposed method which consists of a preprocessing component and a track-to-track fusion algorithm. The multiview fusion approach proposed in this study differs fundamentally from the earlier methods in the literature which, among others, are based on 3D harmonic phase [5], deformable models [6], [7], incompressible models [8] or image registration of short-axis and long-axis sequences [9]. One of the primary benefits of the proposed method is that it uses only anatomical cine sequences that are commonly used in standard clinical examinations due to their shorter processing times. In contrast, previous methods relied on scanning protocols such as myocardial tagging [5], [6] that require longer scan times, or displacement encoding with stimulated echoes (DENSE) CMR images [7] that are not available in routine clinical scans.

The proposed approach first generates a trajectory for each point on the myocardial boundary using a diffeomorphic moving mesh algorithm [10] within each 2D scan, long-axis and short-axis, given delineation of the myocardium on the first frame. The temporal smoothness of the trajectories is obtained by incorporating the prior information corresponds to the LV dynamics via an extension of the nonlinear dynamic state transition model proposed in [11] to 3D. A nonlinear recursive Bayesian framework known as the unscented Kalman smoother (UKS) is proposed to obtain the mean and covariant matrix correspond to each state vector representing the position and velocity of a myocardial point over a cardiac cycle.

The following summarizes the contributions of the current study.

1) A nonlinear state transition model with an augmented dynamic angular frequency component in the state vector is derived by generalizing the model proposed in [12]. We further extend the model to 3D and use the UKS for state estimation.
2) A track-to-track fusion algorithm that is based on maximum-likelihood approach is proposed for combining multiple trajectories. These trajectories are obtained using the UKS from several short-axis and long-axis 2D sequences. To the best of authors' knowledge, this is the first application of the track-to-track fusion in medical imaging.
3) We conducted two independent comprehensive evaluations in cardiac image analysis to demonstrate the advantages of the proposed fusion: (a) Comparing the state estimates obtained with the proposed method using anatomical cine CMR to those obtained with the well-known harmonic phase (HARP) analysis technique [13], which is the most common method for tagged CMR image analysis; (b) Applying the proposed approach to detect regional LV motion abnormality following the 17-segment model recommended by the American Heart Association [14], and evaluating the automated classification results with the assessments by expert radiologists.

In the first experiment, the HARP was implemented as described by Osman *et al.* in [13]. The HARP algorithm utilizes the Fourier domain analysis of the tagged MR images as opposed to their spatial domain. Previous works on estimating cardiac 3D motion from tagged images using the HARP concept were reported in [5], [15], [16]. In this paper, we evaluate the results of the proposed algorithm quantitatively by projecting 3D state vectors onto 2D tagged CMR images acquired from the same subjects.

In the second experiment, the proposed algorithm is evaluated in terms of its classification ability to detect the regional left ventricular functional abnormality. The evaluation was performed over 480 segments obtained from twenty normal subjects and ten patients with abnormal heart function. The classifier used in the experiment was built based on Shannon's differential entropy (SDE) values of two features: normalized radial distance and segment-wise volume/area. Quantitative evaluations were performed for classifier *with and without fusion* in terms of area-under-the-curve (AUC) values of receiver operating characteristic (ROC) curves, and the Bhattacharyya distance metric [17].

[1] Typically the slice distance is 8 – 10 mm in standard clinical acquisitions.

II. PREPROCESSING

A. DEFORMABLE IMAGE REGISTRATION

The objective of the deformable registration is to obtain the point correspondence between two adjacent images T_k and T_{k+1} (for $k = 1, 2 \ldots, K-1$) defined over $\Omega \subset \mathbb{R}^2$ where k and K denote the frame number and the total number of images in a CMR sequence, respectively. Finding the point correspondence between these pair of images can be expressed as the optimization problem given below.

$$\hat{\phi} = \arg\operatorname*{opt}_{\phi} E_{L2}(T_k, T_{k+1}, \phi(x)) \tag{1}$$

where $x \in \Omega$ is the pixel location, $\phi : \Omega \to \Omega$ is a transformation function, and $E_{L2}(\cdot)$ is an L_2-norm based similarity metric. The optimization problem in (1) generally does not lead to a unique solution without additional constraints which can be added by introducing a deformation field with the continuous monitor function $\mu : \Omega \to \mathbb{R}$ and curl of end velocity field $\gamma : \Omega \to \mathbb{R}$ as described below.

1) MOVING MESH GENERATION

The continuous monitor function $\mu(x)$ constrained by

$$\int_{\Omega} \mu = |\Omega| \tag{2}$$

The objective is to obtain a transformation ϕ that meets the following condition.

$$J_{\phi}(x) = \mu(x) \tag{3}$$

A transformation function ϕ that satisfies the condition in (3) can be defined as follows.

Step 1: Generate a vector field $\rho(x)$ which satisfies the following div-curl system

$$\begin{cases} \text{div } \rho(x) = \mu(x) - 1 & \text{(4a)} \\ \text{curl } \rho(x) = \gamma(x) & \text{(4b)} \end{cases}$$

with Dirichlet boundary condition $\rho(x) = 0 \, \forall x \in \partial\Omega$.

Step 2: Construct a velocity vector field using $\rho(x)$:

$$v_t(x) = \frac{\rho(x)}{t + (1-t)\mu(x)}, \qquad t \in [0, 1], \tag{5}$$

where t is an artificial time.

Step 3: By introducing a diffeomorphic function $\psi(x, t)$, solving the following ordinary differential equation

$$\frac{d\psi(x,t)}{dt} = v_t(\psi(x,t)), \quad t \in [0, 1], \psi(x, t = 0) = x, \tag{6}$$

and setting $\phi(x) = \psi(x, t = 1)$, we can obtain the transformation function. We refer the reader to [10] and [18] for the graphic processor and the central processing unit based implementations and other algorithm details of the moving mesh based nonrigid registration algorithm.

B. IMAGE COORDINATE TRANSFORMATION

In order to combine cardiac motion information from different image planes such as short-axis and long-axis sequences, we first need to transform image coordinate based location information into a 3D reference space. We compute the *reference coordinate system* using the following equation:

$$\begin{bmatrix} \mathsf{x} \\ \mathsf{y} \\ \mathsf{z} \end{bmatrix} = \begin{bmatrix} X_{\mathsf{x}}\Delta_i & Y_{\mathsf{x}}\Delta_j & S_{\mathsf{x}} \\ X_{\mathsf{y}}\Delta_i & Y_{\mathsf{y}}\Delta_j & S_{\mathsf{y}} \\ X_{\mathsf{z}}\Delta_i & Y_{\mathsf{z}}\Delta_j & S_{\mathsf{z}} \end{bmatrix} \begin{bmatrix} i \\ j \\ 1 \end{bmatrix} \tag{7}$$

where:

$(\mathsf{x}, \mathsf{y}, \mathsf{z})$	3D coordinates of the pixel location (i, j) in the image plane in units of mm.
S_{xyz}	*Image Position (Patient)* attribute elements.
X_{xyz}	The elements of the row direction cosine of the *Image Orientation (Patient)* attribute.
Y_{xyz}	The elements of the column direction cosine of the *Image Orientation (Patient)* attribute.
(i, j)	Column and row indices to the image plane with zero-indexing.
(Δ_i, Δ_j)	*Pixel Spacing* in column and row directions in mm.

The attributes such as Image Position, Image Orientation, and Pixel Spacing are obtained from digital imaging and communications in medicine (DICOM) headers.

C. TEMPORAL SMOOTHING OF THE DATASET

1) STATE TRANSITION MODEL FOR CYCLIC DYNAMICS

Let $(\mathsf{x}, \mathsf{y}, \mathsf{z})$ be a myocardial point on the 3D reference coordinate system. The point $(\mathsf{x}, \mathsf{y}, \mathsf{z})$ corresponds to (i, j) in the image coordinate system and is computed using (7). Let us define a state vector $s = [\bar{\mathsf{x}} \ \mathsf{x} \ \dot{\mathsf{x}} \ \bar{\mathsf{y}} \ \mathsf{y} \ \dot{\mathsf{y}} \ \bar{\mathsf{z}} \ \mathsf{z} \ \dot{\mathsf{z}} \ \omega]^T$ that describes dynamics of the corresponding myocardial point. Elements $\bar{\mathsf{x}}$ and $\dot{\mathsf{x}}$ denote, respectively, the mean position and the instantaneous velocity over a single heart beat in the x co-ordinate direction. Similar elements in the y and z co-ordinate directions are defined by $\bar{\mathsf{y}}$, $\dot{\mathsf{y}}$, $\bar{\mathsf{z}}$ and $\dot{\mathsf{z}}$. The parameter ω denotes the angular frequency.

We exploit heart's cyclic dynamic property in deriving the state transition model. The discrete-time state transition equation corresponds to the periodic LV tissue dynamics at image frame $k+1$ is given by:

$$s_{k+1} = f_k(s_k) + v_k \tag{8}$$

where $f_k(s_k) = \text{blkdiag}(F_k, F_k, F_k, 1)$, blkdiag$(\cdot)$ denotes the block diagonal matrix, and

$$F_k = \begin{bmatrix} 1 & 0 & 0 \\ 1 - \cos(\omega_k T) & \cos(\omega_k T) & \frac{1}{\omega_k}\sin(\omega_k T) \\ \omega_k \sin(\omega_k T) & -\omega_k \sin(\omega_k T) & \cos(\omega_k T) \end{bmatrix} \tag{9}$$

The parameter T is the time interval between two adjacent frames, and the proposed approach relied on the heart rate information from the DICOM header and the number of cardiac frames in computing T. The noise sequence $\{v_k\}$ is assumed to be normal distributed with zero-mean and covariance Q_k. Dynamic model (8) is an extended version of the 2D

model proposed in [12] to 3D space. Further, the state vector is augmented by introducing a dynamic angular frequency component ω to accommodate the time-varying cyclic motion property of the underlying tissue motion. The measurement equation is given by:

$$z_k = H_k s_k + \eta_k \tag{10}$$

where $H_k = \text{blkdiag}(h, h, h, 1)$, $h = [0\ 1\ 0]$, $\{\eta_k\}$ is normally distributed with zero-mean noise sequence with covariance $R_k = \text{diag}([r\ r\ r])$, $\text{diag}(\cdot)$ denoting the diagonal matrix. The measurement equation describes the relationship between the hidden state-space vector s and the observable parameters, which in this study, are the 3D locations of the endocardial or tag intersection points estimated on the CMR images. The sequence of observations $z_k = [z_{k,\mathrm{x}}\ \ z_{k,\mathrm{y}}\ \ z_{k,\mathrm{z}}]^T$ for $k = 1 : K$ corresponding to the point (i, j) on the image coordinate system is obtained by applying the transformation function $\hat{\phi}$ to (i, j) and mapping them into patient-based coordinate system using (7).

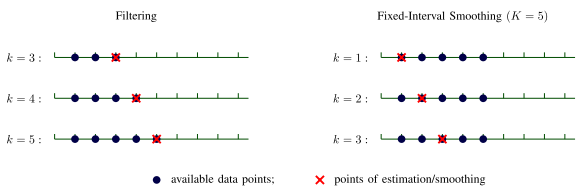


FIGURE 2. An example to illustrate the conceptual difference between filtering and fixed-interval smoothing. The primary difference is that the filtering estimates rely only on past observations at any given time whereas fixed-interval smoothing estimates are based on observations from both past and future which lead to more accurate results.

2) UNSCENTED KALMAN SMOOTHER (UKS)

The state transition model in (8) is nonlinear and, therefore, we employ the UKS [19] to estimate the state vector s_k (for each $k = 1, 2, \ldots, K$). The LV trajectories are estimated as postprocessing which allows for utilizing measurements retrospectively in the estimation algorithm. The UKS used in this study is a retrodiction or smoothing counterpart of the prediction algorithm, unscented Kalman filter (UKF) [20]. In the UKF approach, second-order properties of the underlying probability distributions of the state vector are propagated using a small number of deterministically-selected sample points. This allows for a low computational cost comparable to the extended Kalman filter [21] while providing sufficient estimation accuracy for many nonlinear filtering applications. The primary procedural difference between UKS and UKF is that the UKF uses only the set of observations $z_{1:k}$ to estimate the state s_k at frame k while UKS utilizes both $z_{1:k}$ and $z_{k+1:K}$ to get the state estimate, where K is the total number of frames in the CMR sequence. The performance of the filtering algorithm is improved by the smoothing step that utilizes additional measurement in the UKS. There are several variants of smoothing approach, and in this study, we used *fixed-interval smoothing*, the most common type. The conceptual difference between filtering and fixed-interval smoothing is presented in Fig. 2.

III. TRACK-TO-TRACK FUSION

The objective of track-to-track fusion is to fuse the state estimates s^s and s^l corresponding to the same tissue obtained from short- and long-axis image sequences using UKS, respectively. The state estimates obtained using the UKS consist of the mean and the associated covariance values. In this study, we propose to use a maximum likelihood criterion for the track-to-track fusion [22]. Given s^s and s^l, the likelihood function can be defined as follows.

$$\begin{aligned} L(s) &= -\ln p(s^s, s^l|s) \\ &\propto \left(\begin{bmatrix} s^s \\ s^l \end{bmatrix} - \begin{bmatrix} I \\ I \end{bmatrix} s\right)^T \Sigma^{-1} \left(\begin{bmatrix} s^s \\ s^l \end{bmatrix} - \begin{bmatrix} I \\ I \end{bmatrix} s\right) \end{aligned} \tag{11}$$

where the covariance Σ is given by

$$\Sigma = \begin{bmatrix} \Sigma^s & \Sigma^{sl} \\ \Sigma^{ls} & \Sigma^l \end{bmatrix} \tag{12}$$

The maximum likelihood solution

$$s^{ML} = \arg\max_s L(s), \tag{13}$$

is then computed by solving $\nabla_s L(s) = 0$. This yields:

$$s^{ML} = \left([I\ \ I]\,\Sigma^{-1}\begin{bmatrix} I \\ I \end{bmatrix}\right)^{-1} [I\ \ I]\,\Sigma^{-1}\begin{bmatrix} s^s \\ s^l \end{bmatrix} \tag{14}$$

Define $\mathcal{U} = \Sigma^s$, $\mathcal{V} = \Sigma^{sl}$ and $\mathcal{W} = \Sigma^l$. Apply inversion of a partitioned matrix:

$$\Sigma^{-1} = \begin{bmatrix} \mathcal{U} & \mathcal{V} \\ \mathcal{V}^T & \mathcal{W} \end{bmatrix}^{-1} = \begin{bmatrix} \mathcal{Q} & \mathcal{R} \\ \mathcal{R}^T & \mathcal{S} \end{bmatrix} \tag{15}$$

where

$$\mathcal{Q} = (\mathcal{U} - \mathcal{V}\mathcal{W}^{-1}\mathcal{V}^T)^{-1} \tag{16}$$

$$\mathcal{R} = -\mathcal{Q}\mathcal{V}\mathcal{W}^{-1} \tag{17}$$

$$\mathcal{S} = \mathcal{W}^{-1} + \mathcal{W}^{-1}\mathcal{V}^T\mathcal{Q}\mathcal{V}\mathcal{W}^{-1} \tag{18}$$

Substitute for Σ^{-1} in (14):

$$\begin{aligned} s^{ML} &= (\mathcal{Q} + \mathcal{R}^T + \mathcal{R} + \mathcal{S})^{-1}(\mathcal{Q} + \mathcal{R}^T)s^s \\ &\quad + (\mathcal{Q} + \mathcal{R}^T + \mathcal{R} + \mathcal{S})^{-1}(\mathcal{R} + \mathcal{S})s^l \end{aligned} \tag{19}$$

Substitute $\mathcal{Q}$, $\mathcal{R}$ and $\mathcal{S}$ from (16)-(18) and apply matrix inversion lemma:

$$\begin{aligned} s^{ML} &= (C - B^T)(\mathcal{U} + \mathcal{P} - \mathcal{V} - \mathcal{V}^T)^{-1}s^s \\ &\quad + (\mathcal{U} - \mathcal{V})(\mathcal{U} + \mathcal{W} - \mathcal{V} - \mathcal{V}^T)^{-1}s^l \end{aligned} \tag{20}$$

Substitute Σ^s, Σ^{sl} and Σ^l in (20):

$$\begin{aligned} s^{ML} &= (\Sigma^l - \Sigma^{ls})(\Sigma^s + \Sigma^l - \Sigma^{sl} - \Sigma^{ls})^{-1}s^s \\ &\quad + (\Sigma^s - \Sigma^{sl})(\Sigma^s + \Sigma^l - \Sigma^{sl} - \Sigma^{ls})^{-1}s^l \end{aligned} \tag{21}$$

We can simplify (21) using the assumption that cross-covariance values Σ^{sl} and Σ^{ls} between short-axis and long-axis observations are equal to zero:

$$s^{ML} = \Sigma^l(\Sigma^s + \Sigma^l)^{-1}s^s + \Sigma^s(\Sigma^s + \Sigma^l)^{-1}s^l \tag{22}$$

TABLE 1. Comparison of the overall RMSE (mean and standard deviation) for the methods with and without (*i.e.*, using only the short-axis (SA) or long-axis (LA) view) track-to-track fusion. The RMSE was computed by projecting the state estimates onto SA and LA tagged CMR images and comparing them against harmonic phase estimates. The *p*-values for unpaired t-tests are also reported. The proposed multiview fusion method achieved an RMSE of 1.3 pixels in comparison with the HARP estimates.

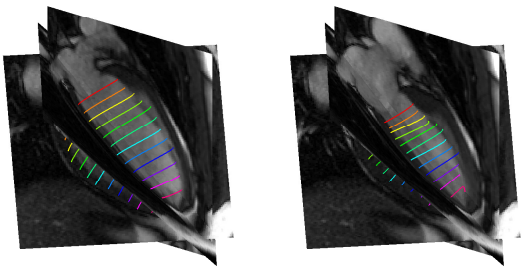
	Proposed method	SA only	LA only	Proposed vs SA	Proposed vs LA
RMSE (pixels)	1.3 ± 0.5	1.6 ± 0.7	1.7 ± 0.5	$p = 0.0296$	$p < 0.001$
RMSE (mm)	1.8 ± 0.6	2.1 ± 0.9	2.3 ± 0.7	$p = 0.1093$	$p < 0.001$

IV. RESULTS

We performed two independent comprehensive experimental studies to prove the benefits of the proposed fusion approach. In the first experiment, we report a quantitative performance analysis by projecting the state vector estimates we obtained onto 2D tagged CMR images acquired from the same subjects and comparing them against harmonic phase estimates. In the second experiment, we applied to the proposed method to detect regional LV functional abnormality following the 17-segment model recommended by the American Heart Association [14] and compared the results with ground truth classifications by expert radiologists.

A. COMPARISONS WITH TAGGED CMR

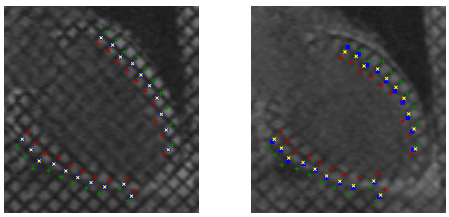
The data consists of 350 cine and 118 tagged CMR image sequences acquired from 30 subjects (22 males, 8 females; mean age: 55.4 ± 14.4 years; age range: 15 to 78 years). The subjects had various underlying cardiac conditions including dilated cardiomyopathy, aortic valve regurgitation, apical aneurysm, myocardial infarction, atrial fibrillation, cardiac sarcoidosis, myocarditis and pericarditis. The images are two dimensional, and they were acquired on various 1.5T CMR scanners using steady-state free precession, such as True-FISP, protocols. The following parameters are used for the acquisition: echo time $= 1.15 - 1.58$ ms; repetition time $= 28.0 - 53.2$ ms; image size $= (150 \times 192) - (256 \times 248)$; in-plane pixel spacing $= (1.09 \times 1.09) - (1.98 \times 1.98)$ mm; flip angle $= 50° - 69°$. Fig. 3 shows representative samples of the fusion results at the end-diastolic and end-systolic phases of the cardiac cycle plotted against 2-chamber, 3-chamber and 4-chamber cine CMR sequences.



(a) End-diastole (b) End-systole

FIGURE 3. Representative examples showing fusion estimates by the proposed approach plotted against long-axis CMR images.

We used the root mean squared error (RMSE) to compare the estimates by the proposed method using cine CMR sequences and the estimates by the HARP using the tagged CMR sequences. The estimates using the proposed method were three dimensional and, therefore, we had to re-project the estimates onto the corresponding tagged CMR images when calculating the RMSE. We used only the first half of the cardiac cycle since the quality of the tag lines becomes poor in the second half of the cardiac cycle, and produces less reliable HARP estimates in our comparison. Further, we used the middle point of the endocardial and epicardial boundary points for comparison as HARP provides more accurate estimates inside the myocardium rather than the boundary.



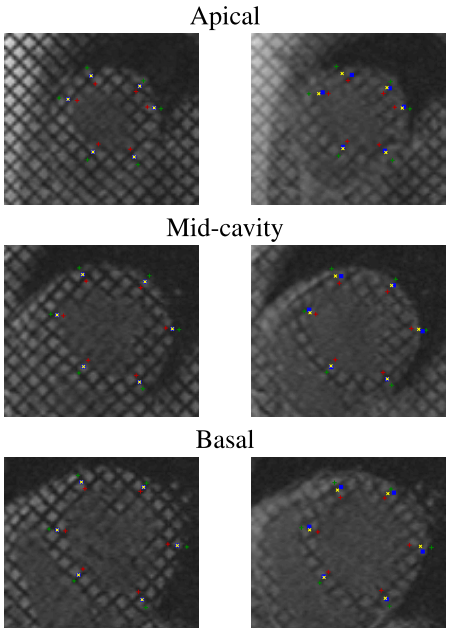
(a) End-diastole (b) End-systole

FIGURE 4. Comparison of the proposed and HARP estimates at end-systolic and end-diastolic phases of the cardiac cycle in 2-chamber view. Green, red and yellow markers, respectively, denote the epicardium, endocardium and corresponding middle points of the projected estimates obtained using the proposed method and cine CMR images; Blue markers indicate the HARP estimates based on tagged CMR images.

Fig. 4 shows representative sample comparisons between the proposed and HARP estimates on a 2-chamber long-axis sequence. Green, red and yellow markers, respectively, denote the epicardium, endocardium, and corresponding middle points of the projected estimates obtained using the proposed method and cine CMR images; Blue markers denote the HARP estimates based on tagged CMR images. A similar comparison on the short-axis images at apical, mid-cavity and basal slices are given in Fig. 5. The RMSE over N number of points is given by, $\text{RMSE} = \sqrt{\frac{1}{\mathsf{N}} \sum_{i=1}^{\mathsf{N}} (\hat{\xi}_i - \tilde{\xi}_i)^2 + (\hat{\phi}_i - \tilde{\phi}_i)^2}$, where $(\hat{\xi}_i, \hat{\phi}_i)$ is the projection of a point on the 2D image coordinate estimated using the proposed method and $(\tilde{\xi}_i, \tilde{\phi}_i)$ is the corresponding point by the HARP using tagged CMR images. Table 1 reports the mean and standard deviation of the overall RMSE of the proposed method and the methods without fusion using only short or long axis images, over 114 image sequences. These RMSE values are calculated based on projections onto short as well as long axis tagged CMR images. The *p*-values for unpaired t-tests are also reported in the table.

TABLE 2. Comparisons of the RMSE (mean and standard deviation) for the methods with and without (*i.e.*, using only the short-axis (SA) or long-axis (LA) view) track-to-track fusion. The RMSE is computed separately by projecting the state estimates onto short (first row) and long (second row) axis tagged CMR images and comparing them against harmonic phase estimates. The *p*-values for unpaired t-tests are also reported. The proposed method yielded better estimates than the methods without fusion in comparisons to both short and long axes tagged CMR sequences.

	Proposed Method	SA only	LA only	Proposed vs SA	Proposed vs LA
RMSE (evaluated on SA tags)	1.6 ± 0.5 mm	1.8 ± 0.7 mm	2.2 ± 0.6 mm	$p = 0.3875$	$p < 0.001$
RMSE (evaluated on LA tags)	2.4 ± 0.6 mm	3.2 ± 0.9 mm	2.5 ± 0.6 mm	$p = 0.0169$	$p = 0.9997$



(a) End-diastole (b) End-systole

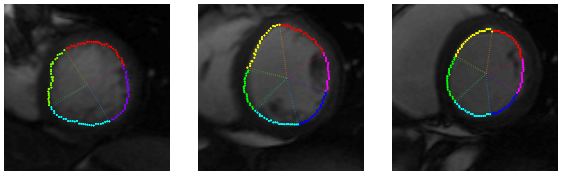
FIGURE 5. Comparison of the proposed and HARP estimates at end-systolic and end-diastolic phases of the cardiac cycle in apical, mid-cavity and basal short-axis views. Green, red and yellow markers, respectively, denote the epicardium, endocardium and corresponding middle points of the projected estimates obtained using the proposed method and cine CMR images; Blue markers indicate the HARP estimates based on tagged CMR images.

The mean RMSE for the proposed method is equal to 1.3 pixels (1.8 mm).

Table 2 reports the RMSE errors evaluated by projecting onto short-axis and long-axis tagged CMR images separately in first and second rows, respectively, as well as the corresponding p-values for unpaired t-tests. Table 2 demonstrate that the proposed method yielded better estimates than the methods without fusion in comparisons to both short and long axes tagged CMR sequences.

B. REGIONAL ABNORMALITY ANALYSIS

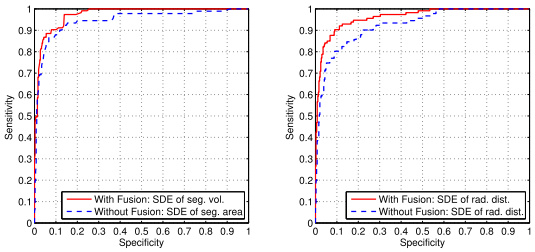
The regional abnormality analysis was performed over datasets acquired from 30 subjects (20 males, 10 females; mean age: 48.7 ± 15.9 years; age range: $16 - 77$ years). The study population composed of 20 normal subjects and 10 patients with abnormal heart function, who had various cardiac diseases including myocardial infarction, ventricular tachycardia, LV pseudoaneurysm, apical aneurysm, cardiac perforation, ascending aortic aneurysm and pericardial effusion. The images were acquired using 1.5T CMR scanners using fast imaging employing steady state acquisition (FIESTA) protocol. The following parameters are used for the acquisitions: echo time $= 0.98 - 1.90$ ms; repetition time $= 2.69 - 4.49$ ms; image size $= (256 \times 256)$; in-plane spacing $= (0.66 \times 0.66) - (1.56 \times 1.56)$ mm; flip angle $= 45^\circ - 55^\circ$. Three representative slices were selected from apical, mid-cavity and basal segments of the LV and the evaluation was performed over 30×3 short-axis image sequences. The regional endocardial borders were automatically segmented following the 17-segmented model recommended by the AHA [14], given the anatomical landmarks, *i.e.,* insertion points denoting the septal wall, on the first frame. A representative example showing the regional segmentation results from apical, mid-cavity and basal short-axis images is given in Fig. 6(a), (b) and (c). In total 480 myocardial segments were analyzed and the classification results correspond to the automated methods were compared with ground truth classifications by radiologists. The manual ground truth classification was built by two experienced radiologists who annotated different portions of the dataset. Binary classification of the motion was used in the evaluation, and a segment was classified as abnormal if the segment is hypokinetic, akinetic, or dyskinetic.



(a) Apical (b) Mid-cavity (c) Basal

FIGURE 6. Representative example showing the results of the regional myocardial segmentation using the proposed approach. Following the recommendation by AHA [14], the myocardium was delineated into four, six and six regional segments for apical, mid-cavity and basal slices, respectively.

Two independent criteria were used in the evaluation of the discriminative power of each classifier features: 1) the ROC curves [23] with corresponding AUC [24] values; and 2) Bhattacharyya measure [17]. We also measured the performance of the classifier features using accuracy, sensitivity and specificity. We utilized a *leave-one-subject-out* method in the evaluation.



(a) (b)

FIGURE 7. Receiver operating characteristic curves of classifier features for the methods with and without (short-axis sequences only) fusion. The proposed method based on fusion demonstrated better classification ability in both classifier features.

TABLE 3. Comparison of the classification accuracy for the methods with and without track-to-track multiview fusion using a leaving-one-subject-out approach. The method without fusion relied only on short-axis image sequences [11]. With fusion, the proposed method outperformed the earlier approach and yielded an overall classification accuracy of 91.9%.

	Proposed method			SA only [11]		
	Acc (%)	Sen (%)	Spe (%)	Acc (%)	Sen (%)	Spe (%)
Apical	90.8	96.9	88.6	92.5	90.0	93.3
Mid-cavity	95.0	95.3	94.9	93.3	93.1	94.1
Base	89.4	97.4	87.3	87.2	100.0	84.9
Overall	**91.9**	**96.5**	**90.5**	**90.8**	**94.5**	**90.0**

Acc=Accuracy, Sen=Sensitivity, Spe=Specificity.

TABLE 4. Comparison of the classification ability of the methods with and without fusion assessed in terms of the area under the ROC curve and Bhattacharyya distance metric. The method without fusion relied only on short-axis image sequences [11]. With fusion, the proposed approach demonstrated improved classification ability in detecting LV functional abnormality.

	Proposed method		SA only [11]	
	AUC	$\mathbb{B}$	AUC	$\mathbb{B}$
SDE of segment area	97.1	0.74	94.3	0.66
SDE of radial distance	95.9	0.70	92.1	0.61

1) ROC AND AUC

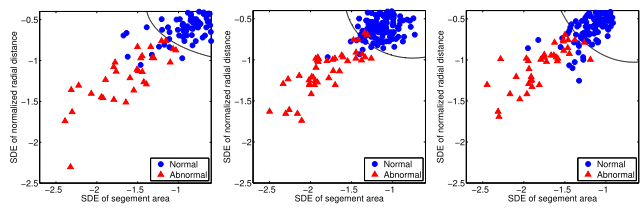
Fig. 7(a) and (b) shows the ROC curves for classifier feature SDE values of segment area/volume and radial distance, respectively. The corresponding AUC values are reported in Table 4. The evaluations based on the ROC curves and the corresponding AUC values for regional cardiac function show that the proposed fusion approach significantly enhances the discriminative ability of each classifier feature over the method that relies only on short-axis sequences in distinguishing the abnormal function from normal function.

2) Bhattacharyya MEASURE

We used Bhattacharyya distance metric to evaluate the overlap between the normal and abnormal functions for each classifier feature. The Bhattacharyya distance metric between the distributions f_N and f_A correspond to the normal function and the abnormal function, respectively, is computed as follows.

$$\mathbb{B} = \sqrt{1 - \sum_{r \in \mathbb{R}} \sqrt{f_N(r) f_A(r)}}, \tag{23}$$

A higher $\mathbb{B}$ indicates the better discriminative ability of the classifier as it is associated with lesser overlap between the distributions. Table 4 also reports the Bhattacharyya distance metric values for each classifier feature, which demonstrate the proposed fusion approach leads to improved classification ability for each classifier feature in comparison to the method that relies only on short-axis image sequences.



(a) Apical (b) Mid-cavity (c) Basal

FIGURE 8. The decision boundaries computed using a Bayesian classifier to detect the normal and abnormal regional myocardial functions based on classifier features. The boundaries were computed separately for apical, mid-cavity and basal slices.

3) CLASSIFICATION PERFORMANCE

We constructed a nonlinear classification boundary using a naive Bayes classifier approach [25] to utilize the SDE values of both classifier features, the segment area/volume and normalized radial distance in distinguishing normal and abnormal regional heart functions. Fig. 8 depicts the separate decision boundaries corresponds to regional segments from apical, mid-cavity and basal slices. The blue circles and red triangles denote the normal and abnormal regional functions, respectively. The classification performance was measured in term of accuracy, sensitivity and specificity using a *leaving-one-subject-out* strategy. Table 3 reports the classification performance results correspond to the proposed method and the method that uses only short-axis images [11].

The proposed method *with multiview fusion* yielded an overall accuracy of 91.9%, a sensitivity of 96.5%, and a specificity of 90.5%, whereas the method that uses only short-axis image sequences [11] yielded an accuracy of 90.8%, a sensitivity of 94.5%, and a specificity of 90.0%. The evaluation demonstrates the proposed fusion method improves the regional function classification accuracy significantly.

The proposed method took an average of 145 seconds to produce the results for a typical dataset with 10 short-axis and 3 long-axis sequences on an Intel Core i7 CPU with 16GB random access memory. The proposed method requires manual delineation of the LV on the first frame in each sequence, and no other manual intervention is needed.

V. DISCUSSION

With images obtained at different views and times, one of the leading challenges in CMR image analysis is the fusion of information. Most of the current methods [3], [4], [26], [27] overlook information fusion, an approach that can be much more informative than what is obtained from individual input images. The proposed Bayesian track-to-track fusion provides a solution to fuse information from multiple sources by analyzing the reliability of each source. To the best of authors knowledge, this is the first attempt to fuse cine CMR images acquired from short and long axes sequences to advance the regional assessment of the left ventricular functional assessment. There are few studies on 3D cardiac motion estimation [28], [29], they relied on tagged MR sequences and not on the standard anatomical cine CMR, the most widely used clinical CMR sequence for LV functional assessment.

The RMSE values reported in tables 1 and 2 demonstrate that the proposed fusion approach has a better agreement with tag intersection estimates by the HARP method than the methods that rely only on short or long-axis. The p-values reported in Table 2 demonstrate that the proposed method performs significantly better than methods that rely only on short or long-axis when evaluated on a different view. This is due to the fact that the methods rely on a single view are incapable of estimating the through-plane motion, which leads to poor performance, whereas the proposed method combines information from different views and incorporates both in-plane and through-plane motions in the motion estimation.

Automated regional assessment of the LV is a challenging problem. Suinesiaputra *et al.* report an accuracy of 63.7% (base), 67.4% (mid-cavity), and 66.7% (apical) for their method when a binary classification of the visual wall motion scoring was used as a reference [3]. Afshin *et al.* report an accuracy of 84.0% (base), 85.7% (mid-cavity) and 89.8% (apical) for their method for similar classification problem [4]. The proposed method with fusion achieved an accuracy of 89.4% (base), 95.0% (mid-cavity) and 90.8% (apical) when binary visual wall motion scoring is used as a reference.

Tracking the normal functioning LV is more challenging due to large deformations and high variations in velocity values than the ventricle with reduced function. However, abnormal motions such as ventricular dyssynchrony might affect the performance of the proposed tracking algorithm due to variations in motion parameters within the cardiac cycle. The datasets used in our study consisted of only three short-axis tagged MRI sequences acquired from a patient with atrial fibrillation, and the proposed method yielded an RMSE of 1.3 ± 0.4 mm. Further studies are warranted to assess the performance of the proposed algorithm on specific cases with severe motion abnormalities.

The proposed method was evaluated against CMR tagging only in terms of RMSE. The proposed approach fuses the motion information from short and long-axis sequences, and therefore, could be used for tracking both endocardial and epicardial boundaries. Future studies will rely on these boundary tracking to compute the regional strain values to compare them against the CMR tagging-based values.

The proposed method applied for regional function assessment classified the cardiac motion as only normal and abnormal (hypokinetic, akinetic and dyskinetic). Out of 480 regional segments used in this study, the normal, hypokinetic, akinetic and dyskinetic motions were observed in 367, 84, 18, and 11 segments, respectively. Due to the low number of available cases for akinetic and dyskinetic motions, the proposed study did not attempt to classify the motion into specific abnormal motions. Future studies will focus on overcoming the class imbalance problem by adding more cases for under-represented classes and attempt to classify the abnormal motions further.

One of the limitations of the proposed approach is that it assumes the patients have not moved between the short and long axes CMR acquisitions. Visual assessments confirmed that this assumption was held for the majority of the patients. In future, we plan to investigate registration approaches to resolve the problem of aligning short-axis and long-axis CMR acquisitions when patients move during the scans.

VI. CONCLUSION

In this study, we proposed a track-to-track fusion approach obtain 3D LV motion estimates from different orthogonal 2D image sequences. The proposed approach uses a deformable image registration method as a preprocessing step to obtain initial trajectories for each myocardial point, given the delineation of the LV on the first frame. Then, a UKS based recursive Bayesian framework was utilized to obtain consistency across the temporal frames in each sequence. The proposed approach was evaluated by projecting the 3D state vector on to 2D tagged CMR images. Our method yielded a competitive performance with a mean RMSE of 1.3 pixels in comparison to well-known HARP estimates. In another independent experimental evaluation, we demonstrated the proposed fusion approach could be used to improve the binary classification of the regional LV function in comparison to earlier methods that do not use fusion and relies only on short-axis CMR sequences.

REFERENCES

[1] J. Garcia-Barnes *et al.*, "A normalized framework for the design of feature spaces assessing the left ventricular function," *IEEE Trans. Med. Imag.*, vol. 29, no. 3, pp. 733–745, Mar. 2010.

[2] K. Punithakumar, I. B. Ayed, I. G. Ross, A. Islam, J. Chong, and S. Li, "Detection of left ventricular motion abnormality via information measures and Bayesian filtering," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 4, pp. 1106–1113, Jul. 2010.

[3] A. Suinesiaputra *et al.*, "Automated detection of regional wall motion abnormalities based on a statistical model applied to multislice short-axis cardiac MR images," *IEEE Trans. Med. Imag.*, vol. 28, no. 4, pp. 595–607, Apr. 2009.

[4] M. Afshin *et al.*, "Regional assessment of cardiac left ventricular myocardial function via MRI statistical features," *IEEE Trans. Med. Imag.*, vol. 33, no. 2, pp. 481–494, Feb. 2014.

[5] L. Pan, J. Prince, J. Lima, and N. Osman, "Fast tracking of cardiac motion using 3D-HARP," *IEEE Trans. Inf. Technol. Biomed.*, vol. 52, no. 8, pp. 1425–1435, Jul. 2005.

[6] T. Chen, X. Wang, S. Chung, D. Metaxas, and L. Axel, "Automated 3D motion tracking using Gabor filter bank, robust point matching, and deformable models," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 1–11, Jan. 2010.

[7] B. S. Spottiswoode, X. Zhong, C. H. Lorenz, B. M. Mayosi, E. M. Meintjes, and F. H. Epstein, "3D myocardial tissue tracking with slice followed cine DENSE MRI," *J. Magn. Reson. Imag.*, vol. 27, no. 5, pp. 1019–1027, May 2008.

[8] A. Bistoquet, J. Oshinski, and O. Skrinjar, "Left ventricular deformation recovery from cine MRI using an incompressible model," *IEEE Trans. Med. Imag.*, vol. 26, no. 9, pp. 1136–1153, Sep. 2007.

[9] J. Lötjönen, D. Smutek, S. Kivistö, and K. Lauerma, "Tracking atria and ventricles simultaneously from cardiac short- and long-axis MR images," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, vol. 2878, R. Ellis and T. Peters, Eds. Berlin, Germany: Springer-Verlag, 2003, pp. 467–474.

[10] K. Punithakumar, P. Boulanger, and M. Noga, "A GPU-accelerated deformable image registration algorithm with applications to right ventricular segmentation," *IEEE Access*, vol. 5, pp. 20374–20382, 2017.

[11] K. Punithakumar, I. Ben Ayed, A. Islam, I. G. Ross, and S. Li, "Regional heart motion abnormality detection via information measures and unscented Kalman filtering," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), vol. 6361, T. Jiang, Ed. Berlin, Germany: Springer-Verlag, 2010, pp. 409–417.

[12] K. Punithakumar, I. Ben Ayed, A. Islam, I. G. Ross, and S. Li, "Tracking endocardial motion via multiple model filtering," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 8, pp. 2001–2010, Aug. 2010.

[13] N. F. Osman, W. S. Kerwin, E. R. McVeigh, and J. L. Prince, "Cardiac motion tracking using CINE harmonic phase (HARP) magnetic resonance imaging," *Magn. Reson. Med.*, vol. 42, no. 6, pp. 1048–1060, Dec. 1999.

[14] M. D. Cerqueira *et al.*, "Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart: A statement for healthcare professionals from the cardiac imaging committee of the council on clinical cardiology of the American heart association," *Circulation*, vol. 105, no. 4, pp. 539–542, 2002.

[15] A. S. Soliman, W. S. Kerwin, and N. F. Osman, "MR harmonic phase surface for tracking cardiac motion," *J. Cardiovascular Magn. Reson.*, vol. 12, no. S1, Jan. 2010, Art. no. P245.

[16] A. S. Soliman and N. F. Osman, "3D motion tracking of the heart using harmonic phase (HARP) isosurfaces," in *Proc. Med. Imag., Image Process.*, vol. 7623, Mar. 2010, Art. no. 76232C.

[17] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 5, pp. 564–577, May 2003.

[18] H.-M. Chen, A. Goela, G. J. Garvin, and S. Li, "A parameterization of deformation fields for diffeomorphic image registration and its application to myocardial delineation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI*, vol. 6361, T. Jiang, Ed. Berlin, Germany: Springer-Verlag, 2010, pp. 340–348.

[19] S. SÄrkkÄ, "Unscented Rauch-Tung-Striebel smoother," *IEEE Trans. Autom. Control*, vol. 53, no. 3, pp. 845–849, Apr. 2008.

[20] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. IEEE*, vol. 92, no. 3, pp. 401–422, Mar. 2004.

[21] Y. Bar-Shalom, T. Kirubarajan, and X.-R. Li, *Estimation With Applications to Tracking and Navigation*. New York, NY, USA: Wiley, 2002.

[22] K. C. Chang, R. K. Saha, and Y. Bar-Shalom, "On optimal track-to-track fusion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 33, no. 4, pp. 1271–1276, Oct. 1997.

[23] K. H. Zou, A. J. O'Malley, and L. Mauri, "Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models," *Circulation*, vol. 115, no. 5, pp. 654–657, Feb. 2007.

[24] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve.," *Radiology*, vol. 143, no. 1, pp. 29–36, Apr. 1982.

[25] G. A. F. Seber, *Multivariate Observations*. Hoboken, NJ, USA: Wiley, 1984.

[26] C. Petitjean and J.-N. Dacher, "A review of segmentation methods in short axis cardiac MR images," *Med. Image Anal.*, vol. 15, no. 2, pp. 169–184, Apr. 2011.

[27] K. Punithakumar *et al.*, "Regional heart motion abnormality detection: An information theoretic approach," *Med. Image Anal.*, vol. 17, no. 3, pp. 311–324, Apr. 2013.

[28] R. Chandrashekara, R. H. Mohiaddin, and D. Rueckert, "Analysis of 3-D myocardial motion in tagged MR images using nonrigid image registration," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1245–1250, Oct. 2004.

[29] E. Oubel *et al.*, "Cardiac motion estimation by joint alignment of tagged MRI sequences," *Med. Image Anal.*, vol. 16, no. 1, pp. 339–350, Jan. 2012.