Deformable Motion Tracking of Cardiac Structures (DEMOTRACS) for Improved MR Imaging

Maneesh Dewan¹ Christine H. Lorenz^{2,3} Gregory D. Hager¹

¹Department of Computer Science, Johns Hopkins University, Baltimore, USA

² Siemens Corporate Research, Baltimore, USA

³Department of Radiology and Radiological Sciences, Johns Hopkins University, Baltimore, USA

 $\label{eq:alpha} \ensuremath{^1\{\text{maneesh,hager}\}@cs.jhu.edu} \ensuremath{^{2,3}\text{christine.lorenz}@siemens.com}$

Abstract

The speed and quality of imaging cardiac structures (coronary arteries, cardiac valves etc) in MR can be improved by tracking and predicting their motion in MR images. The problem is challenging not only due to the complex motion of these structures that significantly changes the appearance of the region of interest, but also the ability to track at different spatial and temporal resolutions depending on the application. We have developed a multipletemplate based tracking approach to track the cardiac structures in MR images. The algorithm has two novel features. First a bidirectional coordinate-descent algorithm is derived to improve accuracy and performance of tracking. Second we propose a method for choosing an optimal set of templates for tracking. The efficacy of the algorithm has been validated by tracking the coronary artery and cardiac valves reliably and accurately in thousands of high resolution cine and low-resolution real-time MR images.

1. Introduction

Heart-related diseases remain one of the major causes of morbidity and mortality in the United States [1], making noninvasive screening an important tool for early detection of the disease. MRI is a good candidate for noninvasive screening as it does not require the use of radiation. However, the speed of MR image acquisition is slow, because MRI data is acquired in the spatial-fourier domain (known as the k-space) in an incremental fashion. The image is then reconstructed by taking the inverse fourier transform of the k-space data. The extent or range of the acquired k-space data governs the resolution of the reconstructed image, i.e. the larger the range, the higher the spatial resolution and the more time necessary for data acquisition. Thus, in MR data acquisition there is a fundamental trade-off between the speed of data acquisition and image resolution.

As a result, data acquisition in cardiovascular MR re-

lies on taking data over multiple heartbeats (known as cine acquisition) either in a breath-hold (volunteer holding his breath) or during free breathing. Cardiac structures like the coronary arteries and valves undergo complex motion induced by both respiratory and cardiac motion [9, 13]. The motion of the target being imaged during MR data acquisition affects the resulting image quality by introducing ghost-like artifacts, blur, and by reducing the image contrast. Furthermore, variability in respiratory and cardiac motion cycles within and across patients makes it difficult to gauge and predict the motion of the cardiac structures, and to compensate for that motion during MR imaging. In MR, when the motion of the object being imaged is not known, motion compensation is done by acquiring data at time points where the object is at (or almost at) the same location. The current state-of-the-art methods thus compensate for cardiac motion by filling the k-space data at the same time point in the cardiac cycle over multiple heart beats.

One of the advantages of MR imaging is that any plane through the object can be chosen for image acquisition. This can be useful for cardiac imaging applications where the imaging plane can be adaptively changed to follow the cardiac structure being imaged, provided the motion of the cardiac structure can be estimated.

The goal of this paper is thus to develop motion tracking algorithms to track cardiac structures in MR images to improve MR imaging speed, quality, and reliability. In MR coronary angiography, our objective is to track coronary artery in high speed, low-resolution MR images, and to use the extracted motion information to "servo" the MR imaging slice to compensate for motion, thereby acquiring the data as if the structure was stationary [6]. In other applications, motion information can be estimated by tracking in MR images acquired during an initial scan (pre-scan), which can then be used to adaptively re-position the acquisition slice during data acquisition. For example, Fig-



Figure 1. Motion of the mitral (top row) and aortic (bottom row) valves through the cardiac cycle in 4-chamber and coronal views respectively. The overlaid cyan line indicates the valve plane.

ure 1 shows the extent of motion of the aortic and mitral valve planes through the cardiac cycle. In order to image the valve, the imaging plane should coincide with the valve plane throughout the cardiac cycle. Thus, if the motion of the valve plane through the cardiac cycle is extracted, it can be used to move the imaging slice during data acquisition such that the valve plane always lies in it. In each case, the key to success lies in the ability to track the cardiac structures in motion-affected MR images across a spectrum of temporal and spatial resolutions, from low-resolution real-time to high-resolution cine images.

1.1. Background - Tracking Methods



Figure 2. Comparison of high resolution cine (left) and realtime low-resolution images (right) in 4-chamber (top row) and short axis views (bottom row). The arrow indicates the location of the left coronary artery. Note that the coronary artery is clearly visible in the high resolution images but blurred out in corresponding realtime images.

As discussed above, our tracking algorithms must be able to track the cardiac structures in both high resolution and low resolution images. One of the challenges of tracking in low-resolution, real-time images is that the coronary arteries and the cardiac valves themselves are not visible as shown in Figures 2 and 3, respectively. Thus, we instead track a region containing the cardiac structure of interest as a substitute for the motion of the structure itself. Also, the magnitude of motion expected between frames is small compared to the size of the template and the reso-



Figure 3. Mitral and aortic valves in high resolution cine (left) and realtime low-resolution images (right) in 4-chamber (top row) and coronal views (bottom row). The overlaid line indicates the valve plane.

lution of the images. In this scenario, region-based tracking methods [2, 8] are usually preferred, as they provide highly accurate (typically to small fractions of a pixel) motion estimates without depending on specific feature extraction or enhancement methods. In region-based tracking approaches, one usually minimizes an objective function that matches the appearance of an area of an image with a stored reference window referred to as a "template". More recently in region-based tracking, the kernel-based methods [5, 7] have become popular primarily due to their broad range of convergence and robustness to small unmodelled spatial deformations.

Another challenge is that, during the cardiac cycle, the appearance of the target region undergoes significant, unpredictable (but repeatable from beat to beat) deformations. As a result, we have seen that a single template deformed using commonly employed parametric models (e.g. affine deformation [2, 3, 8]) is insufficient to track through an entire cardiac cycle. An alternative is to use multiple templates. One simple way to extend the tracking using multiple templates is to run the conventional template tracking [2, 8] with each template and then select the solution that most effectively minimizes the objective function. This approach usually fails in mid-systole and early diastole when both the appearance change and motion are large between frames. Another approach is to compute the appearance of the target using a linear combination of basis images which are either a set of chosen templates, or an orthogonal basis thereof [3]).

When using a multiple template approach, there are two important considerations. First, a large number of templates in a lot of different frames is required to capture the appearance variation. In MR, this is usually done manually which is a time-consuming and error-prone process. A semi-automatic/automatic learning-based algorithm could be used for template selection, but we have found patientto-patient variability is large and as a result automated selection is not very accurate. Second, it is important to note that if some of the selected templates happen to be geometric transformations of other selected templates, the algorithm can (and will) compensate for motion using appearance variation. As a result, without careful choice of templates, multiple template tracking can significantly underestimate motion. We return to this point in section 2.3.

Therefore, we present a new tracking framework that makes use of multiple templates but avoids the shortcomings mentioned above. This method incorporates two novel features. First, we derive a bi-directional, coordinatedescent optimization that simultaneously computes the location and affine mixture in each image. Second, we present a semi-automated selection method for template selection, which ensures that independent templates are chosen from points where the target undergoes large changes. The details of the tracking algorithm are presented in the next section.

2. Multiple-Template Tracking

In this section, we first formulate the multiple template problem as a constrained optimization, and derive an initial solution. We then show that this solution can be extended to a bi-directional coordinate descent algorithm. Finally, we present a method for template selection to ensure reliable and accurate tracking.

2.1. Problem Formulation

Following [3, 8], the basic idea behind the tracking approach is to describe the current region of interest or the target as a geometric transformation of an affine combination of a set of templates. Before expressing this in a mathematical setting, we first define the notation and terminology used. Let \mathbf{G}_t be the target region defined at pixel locations $\mathbf{X} = {\mathbf{x}_i}_{i=1...n}$ in the current image \mathbf{I}_t at time t. The image \mathbf{I}_t is defined on a larger set of pixel locations $\mathbf{Y} = {\mathbf{y}_i}_{i=1...N}$ such that $\mathbf{X} \subseteq \mathbf{Y}$. Here *N* and *n* are the number of pixels in the image and target region respectively. Let $\mathbf{Z}_i; i = 1...m$ be the reference templates or target regions selected from a given sequence of images $\mathbf{I}_i; i = 1...m$. We express the current target region in terms of the reference templates as

$$\mathbf{G}_t = T(\sum_{i=1}^m w_i \mathbf{Z}_i, \tau) \tag{1}$$

where w_i is the weight corresponding to template \mathbf{Z}_i and τ are the transformation parameters. The transformation T is an operator defined on the locations of the image region s.t. $T(\mathbf{Z}_i, \tau) = \mathbf{Z}_i(T(\mathbf{x}_i, \tau)) \quad \forall \mathbf{x}_i \in \mathbf{X}$. For example, for a simple translational motion model, one can write $T(\mathbf{Z}_i, \mathbf{u}) = \mathbf{Z}_i(\mathbf{x}_i + \mathbf{u}) \quad \forall \mathbf{x}_i \in \mathbf{X}$, where the transformation parameter \mathbf{u} is the translation vector. In order to make the equation linear in terms of the unknowns w_i and τ , we can rewrite (1) as follows

$$\hat{T}(\mathbf{G}_t, \tau) = \sum_{i=1}^m w_i \mathbf{Z}_i$$
(2)

where $\hat{T}(.,.) = T^{-1}(.,.)$. It is well known that one can linearize the expression $\hat{T}(\mathbf{G}_t, \tau)$ using Taylor series [2, 8] as $\hat{T}(\mathbf{G}_t, \tau) = \mathbf{G}_t(\hat{T}(\mathbf{x}, \tau)) \approx \mathbf{G}_t(\mathbf{x}) + \frac{\partial \mathbf{G}_t(\hat{T}(\mathbf{x}, \tau))}{\partial \tau} \Big|_{\tau=0} \tau$. This approximation is valid under the assumption that τ is small. In practice, the current template is warped with the parameter value at the previous time step before the optimization step, so effectively one solves for the change in parameter value $\Delta \tau$ [2, 8], which is usually small. Also, the affine linear combination of templates constrains the weights s.t. $\sum_{i=1}^{m} w_i = 1$. Thus the Lagrangian optimization function that needs to be minimized can be written as

$$L(\mathbf{w},\tau,\lambda) = \left\| \mathbf{G}_t(\mathbf{x}) + \mathbf{D}\tau - \sum_{i=1}^m w_i \mathbf{Z}_i \right\|^2 + \lambda \left(\sum_{i=1}^m w_i - 1\right)$$
(3)

where $\mathbf{D} = \frac{\partial \mathbf{G}_t(\hat{T}(\mathbf{x}, \tau))}{\partial \tau} \Big|_{\tau=0}$ and λ is the lagrange multiplier. One usually solves for the unknown parameters (\mathbf{w}, τ) by taking the partial derivatives of $L(\mathbf{w}, \tau, \lambda)$ w.r.t \mathbf{w}, τ and λ , and setting them to zero [4]. Note, $\mathbf{w} = (w_1, w_2, \dots, w_m)^T$ is a vector containing all the weights. After this step, followed by a few rearrangements, one obtains

$$(\mathbf{G}_t(\mathbf{x}) - \mathbf{Z}_m + \mathbf{R}\alpha)^T \mathbf{Q} = 0$$
(4)

where
$$\mathbf{R} = [\mathbf{D}, \mathbf{Z}_m - \mathbf{Z}_1, \mathbf{Z}_m - \mathbf{Z}_2, \dots, \mathbf{Z}_m - \mathbf{Z}_{m-1}]$$

 $\alpha = (\tau, w_1, w_2, \dots, w_{m-1})^T$
 $\mathbf{Q} = [\mathbf{D}, \mathbf{A} - \mathbf{Z}_1, \mathbf{A} - \mathbf{Z}_2, \dots, \mathbf{A} - \mathbf{Z}_{m-1}]$
 $\mathbf{A} = \frac{\sum_{i=1}^m \mathbf{Z}_i}{m}$

Now, (4) can be easily solved to compute α , and hence (τ, \mathbf{w})

$$\boldsymbol{\alpha} = -(\mathbf{Q}^T \mathbf{R})^{-1} \mathbf{Q}^T (\mathbf{G}_t(\mathbf{x}) - \mathbf{Z}_m)$$
(5)

For the solution in (5) to exist, the rank of the matrix $\mathbf{Q}^T \mathbf{R}$ should be full. The matrix $\mathbf{Q}^T \mathbf{R}$ can drop rank under the following two conditions

- 1. One or more templates are linearly dependent on the remaining templates.
- 2. If two templates are related by the relation; $\mathbf{Z}_i = \mathbf{Z}_j + \mathbf{sD}$, for some vector **s** and $i, j \in [1,m]$.

These conditions act as a guideline for selecting the templates (more discussion in section 2.3).

2.2. Extension to Bidirectional Methods

It has been shown that bidirectional gradient methods [10, 11] significantly improve the rate of convergence and convergence range. Therefore, in this section, we extend the multiple-template tracking described in section 2.1 to a bidirectional formulation. In order to do so, equation 1, can be written symmetrically about τ as

$$T(\mathbf{G}_t, -\tau/2) = T(\sum_{i=1}^m w_i \mathbf{Z}_i, \tau/2)$$
(6)

Following the steps in section 2.1, the Lagrangian optimization function becomes

$$L(\mathbf{w}, \tau, \lambda) = \left\| \underbrace{\mathbf{G}_{t}(\mathbf{x}) - \mathbf{D}\tau/2 - \sum_{i=1}^{m} w_{i}(\mathbf{Z}_{i} + \mathbf{D}\mathbf{Z}_{i})}_{\mathbf{P}} \right\|^{2}$$
(7)
$$+ \lambda(\sum_{i=1}^{m} w_{i} - 1)$$

where $\mathbf{D} = \frac{\partial \mathbf{G}_{t}(T(\mathbf{x},\tau))}{\partial \tau}\Big|_{\tau=0}$ and $\mathbf{DZ}_{\mathbf{i}} = \frac{\partial \mathbf{Z}_{i}(T(\mathbf{x},\tau))}{\partial \tau}\Big|_{\tau=0}$. Taking partial derivatives of $L(w, \tau, \lambda)$ w.r.t. \mathbf{w}, τ and λ , and eliminating λ , one gets the partial derivatives w.r.t. w_{i} 's and τ as follows

$$\frac{\partial L}{\partial w_i} = \mathbf{P}^T \left[\mathbf{M} - (\mathbf{Z}_i + \mathbf{D}\mathbf{Z}_i \frac{\tau}{2}) \right] \quad , i = 1, ..., m - 1 \quad (8a)$$

$$\frac{\partial L}{\partial \tau} = \mathbf{P}^T \left[-\frac{\mathbf{D}}{2} - \sum_{i=1}^m w_i \mathbf{D} \mathbf{Z}_i \right]$$
(8b)

where $\mathbf{M} = \frac{\sum_{i=1}^{m} \mathbf{Z}_i + \mathbf{D} \mathbf{Z}_i \tau/2}{m}$ and **P** is the given by underbrace in 7. It can be noted that the partial derivative w.r.t. w_m is not required as it is linearly dependent on the other m - 1partial derivatives w.r.t. $w_i, i = 1, ..., m - 1$. Since the partial derivatives in 8 are non-linear in w_i 's and τ , the optimization cannot be solved in closed form. Another way of solving the optimization is by iteratively solving each subequation in 8 separately by keeping the other variable constant, also known as coordinate descent algorithm [4]. If the other variable is kept constant, the solution now becomes

$$(w_1,\ldots,w_{m-1})^T = (\mathbf{Q}\mathbf{1}^T\mathbf{R}\mathbf{1})^{-1}\mathbf{Q}\mathbf{1}^T\mathbf{E}\mathbf{1}$$
(9a)

$$\boldsymbol{\tau} = (\mathbf{J}^T \mathbf{J})^{-1} \mathbf{J}^T (\mathbf{G}_t - \sum_{i=1}^m w_i \mathbf{Z}_i)$$
(9b)

where
$$\mathbf{E1} = \mathbf{G}_t - \frac{\tau}{2}\mathbf{D} - (\mathbf{Z}_m + \frac{\tau}{2}\mathbf{D}\mathbf{Z}_m)$$

 $\mathbf{J} = \frac{\mathbf{D} + \sum_{i=1}^m \mathbf{D}\mathbf{Z}_i}{2}$

and the i^{th} column of **R1** and **Q1** are given by

$$[\mathbf{R}\mathbf{1}_i] = \left[\mathbf{Z}_i - \mathbf{Z}_m + \frac{\tau}{2}(\mathbf{D}\mathbf{Z}_i - \mathbf{D}\mathbf{Z}_m)\right]$$
$$[\mathbf{Q}\mathbf{1}_i] = \left[\mathbf{M} - (\mathbf{Z}_i + \frac{\tau}{2}\mathbf{D}\mathbf{Z}_i)\right]$$

One of the limitations of the coordinate descent algorithm is that it can get stuck in local minima if not properly initialized. We initialize the algorithm with the τ value obtained from 5 which is usually a good initialization.

2.3. Template Selection

In order to make sure the solution in 5 and 9 is nonsingular, the following matrix should be well conditioned

$$\mathbf{N} = [\mathbf{Z}_1, \mathbf{D}\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{D}\mathbf{Z}_2, \dots, \mathbf{Z}_m, \mathbf{D}\mathbf{Z}_m]$$
(10)

where $\mathbf{DZ}_{i} = \frac{\partial \mathbf{Z}_{i}(T(\mathbf{x}, \tau))}{\partial \tau} \Big|_{\tau=0}$. A large condition number indicates a poor selection of templates indicating that a template is a transformation of some other template or a combination of templates. For the matrix $\mathbf{Q}^{T}\mathbf{R}$, this can also happen when the target region in the current frame is close to one of the two templates. An assumption made here is that the derivative of the current region \mathbf{D} will be close to the linear combination of the derivative of templates. Although, the matrix \mathbf{N} indicates whether a given set of templates would ensure a non-singular solution of the optimization but it does not tell you which templates to select from a given set of templates.

If we take a closer look at the matrix **N**, the two columns $[\mathbf{Z}_i, \mathbf{DZ}_i]$ corresponding to a template \mathbf{Z}_i constitute the transformation tangent plane [14] of the template \mathbf{Z}_i . It is easy to see that if the tangent planes of the selected templates that constitute the matrix N are well separated, the matrix N will be well-conditioned. This notion can be captured by 'tangent distance' [14], that is defined as the minimum distance between the two tangent planes. Thus, to ensure a non-singular solution, one needs to find a set of templates with well separated tangent planes and are a representative set of all the templates. Hence, our algorithm is based on clustering using the tangent distance. We use the spectral clustering method proposed by Ng et. al. [12]. Each entry in the affinity matrix is the tangent distance between the two corresponding templates. The number of clusters is chosen by the user from the eigenvalues of the normalized eigenvector matrix generated from the affinity matrix. The templates that are closest to the cluster mean centers are selected for tracking.

The algorithm was run on manually selected templates chosen at the left end of the mitral valve in the 4-chamber view. In this case, the number of clusters was chosen to be 3, as the first three eigenvalues of the normalized eigenvector matrix were most significant. Figure 4 shows the optimal templates selected at frames 5, 17 and 36. The tracking was performed using these templates. To compare it with a non-optimal choice of templates, the tracking was also run with equally spaced templates at frames (5,20,35) and (1,20,40) respectively. The tracking results are compared in Figure 5. The tracking result for optimal templates was found to be visually the best. It is important to note that the manual selection is not ground truth as it is very difficult to select templates with continuous motion. The tracking results for 'equal-spaced 1' and optimal templates are close as the templates lie in similar clusters. Another point to note is that in the current formulation, we pick the template



Figure 4. Optimal Templates selected (enclosed in rectangle) from the manually selected 40 templates.



Figure 5. Tracking comparison with optimal and non-optimal templates. The templates used for equal-spaced 1 and 2 are (5,20,35) and (1,20,40) respectively.

closest to the mean as the optimal template in that cluster which might not be the best template corresponding to that cluster. Also, increasing the number of optimal templates can further improve the tracking.

3. Demonstrations

3.1. Data and Parameters

The left coronary artery is visible in the 4-chamber, short axis and coronal views whereas the mitral and aortic valves are visible in 4-chamber and coronal views (see Figure 1) respectively. Thus, the acquired data consisted of both low-resolution real-time and high resolution cine images in the short axis, 4-chamber and coronal views taken from 5 volunteers. The high resolution cine data had the following parameters: In-plane reconstructed resolution = 1.6 - 1.92mm, Slice thickness = 6mm, Acquisition matrix size = $120 \times 192 - 156 \times 192$ pixels, Temporal resolution = 21 - 33msec, on a 1.5T scanner (Espree, Siemens). The real-time SSFP images were acquired during both free-breathing and breath-hold with the following parameters: In-plane reconstructed resolution = 2.76 - 2.89mm, Slice thickness = 6



Figure 6. Process flow for the proposed cardiac valve imaging approach

- 8mm, Acquisition matrix size = $50 \times 128 - 68 \times 128$ pixels, Interpolated matrix size = $88 \times 128 - 112 \times 128$ pixels. The breath-hold real-time data was acquired with the subjects holding their breath at end-expiration over a duration of 6-8 cardiac beats (≈ 130 frames). It is important to note that the high resolution cine images are acquired by filling the fourier k-space over multiple heart beats whereas the low resolution real-time images are acquired by filling the k-space in a single heart beat. For the tracking results presented in the next subsections, a 2D translation model was found to be sufficient for modeling the transformation.

3.2. Tracking in High-Resolution

Valve Tracking: As discussed in the introduction section, if the location of the valve plane is estimated at different time points in the cardiac cycle, the imaging plane can be repositioned to lie in the valve plane. Figure 6 gives the different steps involved in the approach. Note that the third step involves tracking in real-time images. We will defer discussion regarding this step until the next section.

The valve planes through the cardiac cycle are estimated by tracking the region around the end-points of the valve. The valve is then localized by the line segment joining the end-points (center of the tracked region) of the valve. The algorithm requires user input for the selection of templates that can be time-consuming and tedious. If the user is guided through a few frames (key frames) to select the template or reference locations of the valve end-points, the template selection process can be significantly simplified and sped up. These key frame locations (in terms of percent cardiac cycle length) should contain optimal/near optimal templates and should be invariant over different volunteers. In order to find such key frames, we use the template selection process described in section 2.3 for every volunteer. The valve end-points were manually located in all frames for all volunteers. The manual locations were smoothed by filtering and chosen as center locations for the templates with size 25x25 pixels. The number of clusters was set to 5. All the templates for each volunteer were segmented into 5 clusters using the algorithm in section 2.3. The key-frames



Motion comparison between tracking

Figure 8. Error between tracked and manually selected locations across all volunteers.



Figure 9. Process flow for the proposed coronary imaging approach.

were picked such that they span all the 5 clusters in all volunteers. For the mitral valve, the key-frames were selected at 0, 17.5, 37.5, 52.5 and 87.5 percent of the cardiac cycle. The key-frames for the aortic valve were at 0,17.5, 37.5, 60, and 82.5 percent of the cardiac cycle. The tracking of both mitral and aortic valves was then performed in all volunteers using templates selected just from these key frames.

Figure 7 shows the valve tracking results and its comparison to manual selection in one volunteer. It is easy to see that the tracking results match very closely with the manual selection. As the orientation of the heart is volunteerspecific, the valve motion data from all volunteers in the 4-chamber view was transformed into a consistent heart coordinate system (see Figure 7), that was computed using the 4-chamber and short axis slice orientations. The error between the tracked and manually selected mid-point of the valves was also computed. Figure 8 shows the mean and standard deviation of the total error between the tracked and manually selected location of the mid-point of the valves in all 5 volunteers. The isotropic pixel dimension in these high resolution images range from 2.5-2.7mm across all volunteers. Thus, the error is well within sub-pixel accuracy.

3.3. Tracking in Low-Resolution

Coronary Artery Tracking: As mentioned in introduction, an approach was recently proposed by [6] to reduce



Figure 10. Tracked positions of left coronary artery in 4-chamber (1st row), short axis (2nd row) and coronal (3rd row) views at different points in the cardiac cycle. The cyan box represents the region being tracked and the green dot denotes the coronary location.

the effects of motion variability by tracking the coronary artery in low-resolution realtime images in each heart cycle, and then using the motion information to guide the high resolution image acquisition (see Figure 9). The feasibility of the approach relies on the ability to track the coronary motion reliably in low-resolution MR images taken during free-breathing. The tracking of the coronary artery was performed in 4-chamber, short axis and coronal views. As the number of templates per cardiac cycle is low (10-15), the templates were chosen at end-systole, mid-diastole and enddiastolic time points in the cardiac cycle. It was made sure that these templates satisfied the full rank condition. Figure 10 shows the tracked position of the left coronary artery in free breathing low resolution MR images in all the three views.

Validation with high resolution: The variability in the cardiac component of the coronary motion can be attributed to three factors: a) cardiac cycle length (heart rate) variability, b) cardiac motion variability and c) motion hysteresis between beats. This makes it difficult to obtain ground truth motion of the coronary arteries on a beat-to-beat basis and hence validate the tracked coronary motion. We therefore validated the tracking by comparing the coronary motion in 'high resolution cine' with tracked coronary motion in 'low resolution real-time' images, both taken during breath-hold. The tracked coronary motion in the real-time breath hold images in all three views was segmented into systolic and diastolic periods for each cardiac cycle. The first and second order derivatives were used to detect end-systolic (ES) and end-diastolic (ED) time points in each cardiac cycle. Note that the motion data obtained by tracking was not filtered. The coronary locations in the 'high resolution cine' images were manually selected. To reduce the effect of cardiac cy-



Figure 7. Valve tracking and its comparison to manual selection in a single volunteer. First and second rows correspond to the mitral valve and aortic valves respectively. The blue and magenta lines corresponds to the tracked and manually selected valve planes. The first image in the top row shows the heart coordinate system. The bottom row shows the comparison of motion of the mid-point of both mitral and aortic valves.



Figure 11. Comparison of tracked coronary motion in realtime with the motion in high resolution images in coronal (left), four chamber (center) and short axis views (right) respectively.

cle length variability the systolic and diastolic durations for the motion data in both high resolution and low resolution were normalized and interpolated before comparison. Also, the motion data during each cardiac beat was made zeromean to reduce the effect of motion hysteresis. Figure 11 shows the comparison for one volunteer. The low resolution motion tracking is in good agreement with the high resolution coronary motion except in transformed x direction of both 4-chamber and short axis views, where the range of motion is very small. In general, for all the volunteers, the total error (error in both x and y directions) between high resolution and average low resolution tracked motion during the mid-diastolic to end-diastolic period where the image acquisition is done was quite small. The total error across 5 volunteers is shown in Figure 12. The error is usually less than 2-3 mm which is much less than a pixel of total error in

low resolution images (\approx 4mm). The isotropic pixel dimensions in high resolution images and low-resolution images varied from 1.77 to 1.9 mm and 2.76 to 2.9 mm respectively for all 5 volunteers. Keeping in mind the variability in coronary motion, manual coronary selection and tracking, the error values are small and within acceptable limits.

Valve Tracking: As mentioned earlier, one of the steps in the process flow for valve imaging (see Figure 6) involves real-time tracking in one or two breath-holds before switching to the data acquisition. This is important because a different breath-hold is used for pre-scan and data acquisition, and there could be an unknown offset in valve plane positioning due to either hysteresis effects or shift between different breath-holds, that can be corrected by real-time tracking. Thus, here we present the tracking of mitral and aortic valve planes in low-resolution images in 4-chamber and



■ 4-chamber view ■ short axis view ■ coronal view

Figure 12. Total error between high resolution and average low resolution realtime tracked motion



Figure 13. Tracking of mitral (top row) and aortic (bottom row) valves in low-resolution real-time images. Three frames from the 128 frame long sequence for each view are shown.

coronal views (see Figure 13) respectively. The template selection was similar to that used for coronary artery tracking. The tracking accuracy in real-time images was assessed visually and were found to be in good agreement.

4. Conclusions

We have presented a multiple template-based tracking method to track the cardiac structures like coronary arteries and cardiac valves reliably and accurately in MR images. The tracking algorithm has been validated on MR images of varying temporal and spatial resolution. The ability to track these cardiac structures in MR images will help greatly improve the speed and quality of imaging these structures in MR.

The selection of a large number of multiple templates is time-consuming and often error-prone that can lead to underestimation of motion parameters. Thus, we present an algorithm for selection of a few optimal templates where the target undergoes significant appearance changes. We also present a bidirectional gradient optimization that improves the range of convergence and accuracy of tracking. The approach also provides a framework for updating multiple templates online and use a subset of templates depending on the point in the cardiac cycle to further improve tracking accuracy.

Currently, we are working on integrating the tracking approach into a MR Scanner and evaluate the effect of motion compensation in human volunteers. In future, we plan to further validate the approach using cardiac motion phantoms undergoing known motion transformations. In addition, we would like to integrate estimation techniques like Kalman filter to remove jitter and further improve the tracking.

References

- [1] Heart disease and stroke statistics 2006 Update. A report from the american heart association statistics committee and stroke statistics subcommittee, 2006. 1
- [2] Baker S, Matthews I. Lucas-kanade 20 years on: a unifying framework. *Intl. Journal Comp. Vision*, 56(3):221 – 255, March 2004. 2, 3
- [3] Black MJ, Jepson AD. Eigentracking: robust matching and tracking of articulated objects using a view-based representation. *Intl. Journal Comp. Vision*, 26:63–84, 1998. 2, 3
- [4] Boyd S, Vandenberghe L. *Convex optimization*. Cambridge Univ. Press, 2004. 1st edition. 3, 4
- [5] D. Comaniciu and V. Ramesh and P. Meer . Kernel-Based Object Tracking. *IEEE Trans. Patt. Anal. Mach. Intelligence*, 25(5):564–575, 2003. 2
- [6] Dewan M, Hager GD, Shea SM, Lorenz CH. Compensating for beat-to-beat variation in coronary motion improves image quality in coronary MR. *Proc 13th Intl Soc Magn Reson Med*, page 2159, 2006. 1, 6
- [7] G. D. Hager, M. Dewan, and C. V. Stewart. Multiple Kernel Tracking with SSD. Proc. IEEE Conf. Comp. Vision and Patt. Recog., 1:790–797, 2004. 2
- [8] Hager GD, Belhumeur PN. Efficient region tracking with parametric models of illumination and geometry. *IEEE Trans. Patt. Anal. Mach. Intelligence*, 20(10):1025–1039, 1998. 2, 3
- [9] Karwatowski SP et al. Assessment of regional left ventricular long-axis motion with MR velocity mapping in healthy subjects. J. Magn. Reson. Imaging, 4:151–55, 1994. 1
- [10] Keller Y, Averbuch A. Fast motion estimation using bidirectional gradient methods. *IEEE Transactions on Image Processing*, 13(8):1042–1054, 2004. 3
- [11] Malis E. Improving vision-based control using efficient second-order minimization techniques . *IEEE International Conference on Robotics and Automation*, 2004. 3
- [12] Ng A, Jordan M, Weiss Y. On spectral clustering: Analysis and an algorithm. In Advances in Neural Information Processing Systems (NIPS) 14, 2002. 4
- [13] Shechter G et. al. Three-dimensional motion tracking of coronary arteries in biplane cineangiograms. *IEEE Trans. Med. Imaging*, 22:493–503, 2003. 1
- [14] Simard PY, LeCun Y, Denker JS, Victorri B. Transformation invariance in pattern recognition – tangent distance and tangent propagation. *International Journal of Imaging Systems* and Technology, 11(3), 2001. 4