Estimating Consistent Myocardial Motion by 4D Image Warping

Hari Sundar[1], *Harold Litt*[2], *Dinggang Shen*[1]
[1] *Section for Biomedical Image Analysis,*
[2] *Cardiovascular Imaging Section,*
Department of Radiology, University of Pennsylvania School of Medicine, Philadelphia, PA
{hari.sundar, harold.litt, dinggang.shen}@uphs.upenn.edu

Abstract

A method for spatio-temporally smooth and consistent estimation of cardiac motion from MR cine sequences is proposed. Myocardial motion is estimated within a 4-dimensional (4D) registration framework, in which all of the 3D cardiac images obtained at different phases of the cardiac cycle are registered simultaneously. This facilitates spatio-temporally consistent estimation of motion as opposed to other registration-based algorithms which estimate the motion by sequentially registering one frame to another. To facilitate image matching, an attribute vector (AV) is constructed for each point in the image, and is intended to serve as a "morphological signature" of that point. The AV includes intensity, boundary, and geometric moment invariants (GMIs). Hierarchical registration of two image sequences is achieved by using the most distinctive points for initial registration of two sequences and gradually adding less-distinctive points to refine the registration. Experimental results on real data demonstrate good performance of the proposed method for cardiac image registration and motion estimation. The motion estimation is validated via comparisons with motion estimates obtained from MR images with myocardial tagging.

***Index Terms*—Image Registration, Cardiac Motion Estimation, Spatio-Temporal Normalization**

1. Introduction

According to WHO estimates, 16.7 million people around the world die of cardiovascular diseases (CVD) each year [1]. CVDs which primarily affect the myocardium are called cardiomyopathies, and these can be generally classified into two major groups, extrinsic and intrinsic cardiomyopathies. Extrinsic cardiomyopathies are those where the primary pathology is outside the myocardium itself, e.g. ischemic cardiomyopathy or cardiomyopathy secondary to sarcoidosis. Intrinsic cardiomyopathies are not due to an identifiable external cause, and are

usually diffuse and may be difficult to diagnose. Detecting abnormalities in myocardial function can greatly help in the early diagnosis of intrinsic cardiomyopathies.

Intrinsic cardiomyopathies present themselves in different forms, including structural changes such as wall thinning, fat deposition or fibrosis, and functional changes, e.g. variations in ventricular wall motion, ejection fraction, and perfusion. Both of these types of information need to be extracted from the image before an accurate diagnosis can be made. Much work has been done in the area of feature extractors for structural characterization of disease. This has focused primarily on image features such as intensities and gradients [2], moments [3, 4], Gabor features [5, 6], and local frequency representations [7]. One difficulty in analysis of diffuse cardiomyopathies is that not all of the pathology can be characterized in terms of structural changes. Function at rest may be abnormal as a result of ischemic heart disease (including ischemia, infarction, and stunned or hibernating myocardium) or related to myocardial structural changes from infiltrative or genetic cardiomyopathies. During stress testing, new or worsening wall motion abnormalities are indicative of functionally significant coronary artery stenosis [8]. In addition, wall motion imaging to detect regional contractile reserve is an accurate measure of myocardial viability in ischemic heart disease, and the results can help guide coronary revascularization therapy. Thus, characterizing cardiomyopathies based on both structural and functional changes will make the diagnostic algorithm more accurate and robust.

Most clinical modalities used to image myocardial function evaluate passive wall motion (ventriculography) or wall thickening (echocardiography, gated single-photon emission computed tomography, or cine MR imaging). Some types of imaging, such as MRI with myocardial tagging and echocardiography with tissue Doppler imaging, also allow quantitative measurement of regional intramyocardial motion and, subsequently, strain, which can be more sensitive to wall motion abnormalities than just wall thickening alone. MR imaging methods for quantification of intramyocardial wall motion can be loosely classified into two approaches, i.e., those relying on specially developed MR imaging pulse sequences to help in the estimation of myocardial motion,

and those relying on image analysis techniques to extract motion estimates from conventional cine MR sequences.

A. Specialized MR Protocols

Myocardial Tagging

Myocardial tagging was developed to provide non-invasive virtual markers inside the myocardium, which deform with myocardial motion [9, 10]. MR imaging and especially tagged MR are currently the reference modalities to estimate dense cardiac displacement fields with high spatial resolution. The deformation fields, as well as derived motion parameters such as myocardial strain, can be determined within an accuracy of 2mm x 2mm [11, 12]. The primary disadvantage of tagging is the reduced spatial resolution of strain relative to the image spatial resolution. In tagging, after the displaced tag lines are detected [13], the displacement field can be estimated and intramyocardial strain can be computed in a variety of ways [9]. With this approach, although strain may be interpolated to any desired spatial resolution, the fundamental spatial resolution of strain is nominally determined by the distance between the tag lines, which is typically several pixels. Tag detection has an additional disadvantage in that it typically requires substantial manual intervention and is therefore a time-consuming task. Harmonic phase analysis will likely obviate tag detection [14], but the spatial resolution of the resultant strain maps will not necessarily improve. The spatial resolution of strain maps obtained from tagged images after harmonic phase analysis is determined by the *k*-space filter of the analysis; in practice with single breath-hold acquisitions, the resolution has been relatively poor [15]. Additionally, since the right ventricle (RV) is much thinner than the left ventricle (LV), it is difficult to place more than a single tag within the RV, making the estimation of RV motion extremely difficult and inaccurate.

Another problem with tagging is that the tag lines are not very well resolved, especially in the later frames of the cardiac cycle, as tag-fading occurs because of the relaxation of induced magnetization. As a result, the contrast between the tags and the tissue decreases during the later

phases, which in turn increases the error in the estimation of myocardial motion, especially during diastole.

Phase Contrast Imaging

A *second* approach uses MR phase contrast imaging [16], which is based on the concept that spins which are moving in the same direction as a magnetic field gradient develop a phase shift that is proportional to the velocity of the spins. This information can be used directly to determine the velocity of the spins, or in the cardiac case the velocity of any point within the myocardium. The velocity field can then be integrated to yield displacements, which in turn can be used to calculate the strain tensor. However, the principal difficulty with this approach is that four acquisitions have to be performed for each cardiac location, i.e., the regular cine MR sequence and one phase contrast acquisition each for the velocity components in the x, y, and z directions. Consequently, MR phase contrast imaging is not used much in a clinical setting.

DENSE & HARP

Displacement-encoded imaging with stimulated echoes (DENSE) [17] and harmonic phase imaging (HARP) [14] employ 1-1 spatial modulation of magnetization to cosine modulate the longitudinal magnetization as a function of position at end diastole. Later in the cardiac cycle the cosine-modulated signal is sampled and used to compute myocardial strain from the signal phase. The sampled signal generally includes three distinct echoes: a displacement-encoded stimulated echo, the complex conjugate of the displacement-encoded echo, and an echo arising from T1 relaxation. If the T1-relaxation and complex conjugate echoes are suppressed, then a phase image representing just the displacement-encoded echo can be reconstructed. However, data-acquisition in single-breath-hold DENSE MR imaging has been limited to only one cardiac phase. Multiple breath-hold DENSE produces images at multiple phases of the cardiac cycle, but the spatial resolution has been suboptimal; additionally, current implementations use a 2D approach with poor estimation of through-plane displacement.

B. Extracting motion from Cine MR images

An alternate approach is to estimate myocardial motion from cine MR sequences. Cine MR images are acquired in a clinical setting at sub millimeter resolution (in-plane), with slice thickness in the range of 6-10mm. The temporal resolution generally varies between 25 and 70ms. Many methods have been developed for extracting cardiac motion fields from cine MR sequences [11, 18-23]; these can be classified into two main categories. The first approach uses segmentation of the myocardial wall, followed by geometric and mechanical modeling using active contours or surfaces to extract the displacement field and to perform the motion analysis [11, 20, 23]. For matching two contours or surfaces, curvatures are frequently used to establish initial sparse correspondences, followed by dense correspondence interpolation in other myocardial positions by regularization or mechanical modeling [11, 19]. The lack of distinct landmarks on the myocardial wall makes it difficult to estimate the wall motion based on surface tracking. In addition, this approach is very sensitive to the accuracy with which the myocardium can be segmented, and it performs poorly in regions within the myocardium, generally aligning only the myocardial boundaries.

The other approach uses energy-based warping or optical flow techniques to compute the displacement of the myocardium [18, 21, 22]. Perperidis *et al* [21] use a regular grid with a B-spline basis to model the deformation and use normalized mutual information as the similarity metric which is calculated over the whole image. Although this produces reasonable motion estimates that have been validated against tagged MR images [21], there are two major issues that need more attention. *Firstly*, since the similarity function is evaluated over the entire image, it is not very effective at capturing localized small deformations in the myocardial wall. These localized deformations can theoretically be captured by refining the B-Spline grid, but that increases the computation cost of the method. In addition, increasing the grid resolution in regions outside the myocardium, like the blood pool, results in over registration, and creates

unwanted artifacts. Since our ultimate goal is to be able to characterize cardiomyopathies, these subtle variations are very important. Again, these approaches are limited by the dimensionality of the transformation that they can estimate. The B-spline based approach would require a very fine grid in order to capture these subtle variations. However, registering using such a fine grid makes the approach very slow and also causes artifacts to appear, especially in noisy areas. *Secondly*, since most motion estimation approaches are built up from standard 3D deformable registration techniques, they do not take temporal smoothness and consistency into account, which is an important part of the estimation process. For example, while registering two cine MR sequences, Perperidis [21] enforces temporal smoothness by considering a 4D B-spline grid to model the transformation and therefore produces better results than those obtained earlier. We hypothesize that incorporating temporal smoothing into the motion estimation process will produce better estimates of myocardial motion. A more detailed and thorough review of cardiac image registration methods can be found in [24], and a more general review of image registration methods can be found in [25].

2. Method

In this paper, we propose a new method for estimation of spatio-temporally consistent myocardial motion fields from MR images, based on a 4D elastic registration algorithm. Unlike other approaches that estimate myocardial motion by performing a series of deformable 3D-registrations, we instead present it as a 4D registration problem, which makes the motion estimates temporally smooth and consistent. We describe our algorithm in detail in the following three subsections. First, we describe the formulation of the myocardial motion estimation problem as one of 4D registration. Then, we describe the use of attribute vectors for correspondence detection, which is fundamental to the registration process. Finally, we present the actual 4D registration algorithm, which is formulated as an energy minimization problem.

A. Motion estimation using a 4D registration framework

Cardiac motion estimation is the problem of determining a transformation that captures the motion of every point, **x,** in the myocardium over the cardiac cycle ($1 \leq t \leq N$). If we can register 3D images acquired at different phases, t, of the cardiac cycle to each other, we can then have an estimate of cardiac motion. This is the approach taken in most cardiac motion estimation algorithms, i.e., they register all the frames in the sequence to a reference frame, typically the end-diastolic frame. Thus, for estimating cardiac motion from end-diastole to other times in a periodic sequence of N 3D images $I(\mathbf{x},t)=\{I_1(\mathbf{x}), \ldots I_t(\mathbf{x}), \ldots I_N(\mathbf{x})\}$, with I_1 as the end-diastolic image, we need to register the end-diastolic image I_1 to images at all other time-points. As mentioned earlier, this approach does not guarantee temporal smoothness or consistency. In particular, a single frame with bad correspondences can add substantial errors to the motion estimation; solving it within a 4D registration framework may obviate these problems.

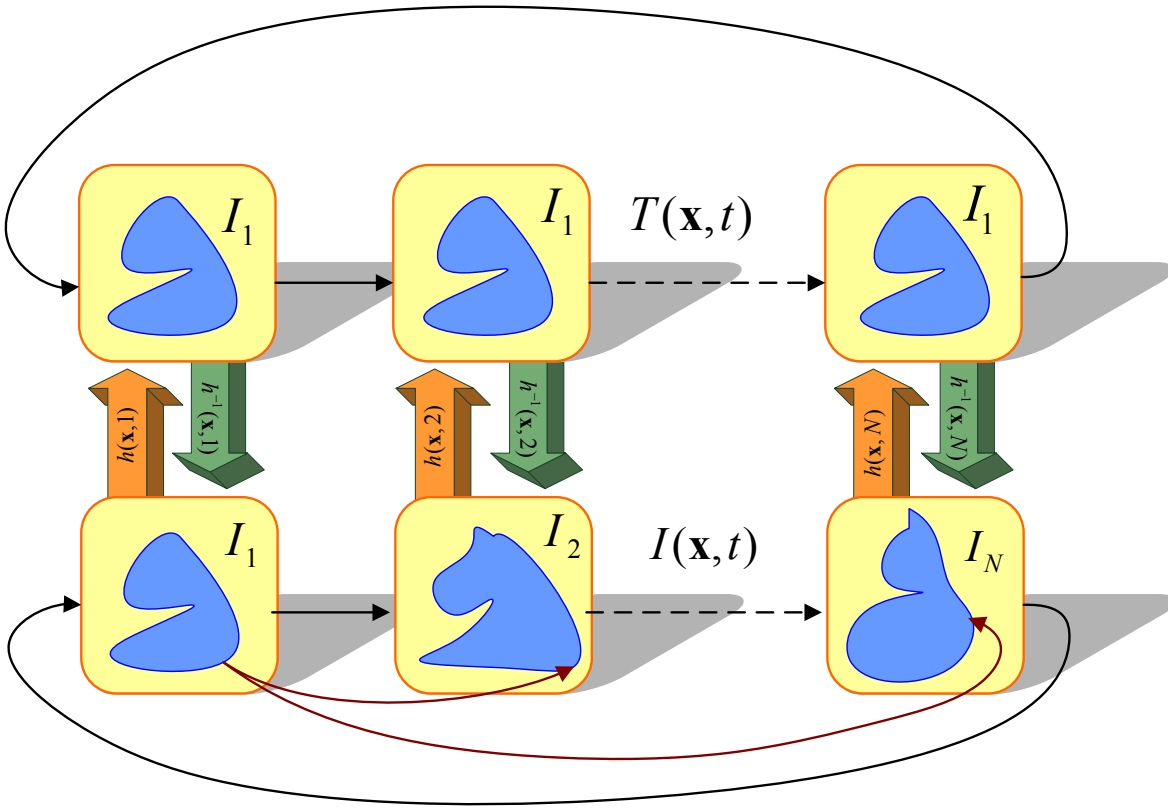


Figure 1. The formulation of myocardial motion estimation as a 4D image registration problem.

In order to use a 4D image registration for simultaneously estimating the motion from the end-diastolic image I_1 to all other time-points $\{I_t\}$, we generate a new 4D image, i.e., a new image sequence $T(\mathbf{x},t)=\{I_1,\cdots,I_1\}$, which repeats the end-diastolic image I_1 as images at *N* different time-points (Figure 1). This image sequence represents the stationary heart of the same patient, and is used as the template sequence. Thus, by registering the 4D image $I(\mathbf{x},t)$ to $T(\mathbf{x},t)$ via a spatial transformation $h(\mathbf{x},t)$, the motion for each point $\mathbf{x}$ in the end-diastolic I_1 to all other time-points in a cardiac sequence can be estimated. Notably, the transformation $h(\mathbf{x},t)$ is restricted to 3D spatial deformations at the same time-point, since no temporal differences exist in the generated image sequence $T(\mathbf{x},t)$ and thus there is no need to consider temporal variations.

B. Attribute Vector

The accuracy and robustness of correspondence detection depends on how effectively the image descriptor can capture the local and global properties of the given point. In a registration context, it is desirable that these descriptors are scale and rotation invariant. The heart does not have many geometrically discernible points and most local descriptors, such as intensity and gradient information, fail to capture the uniqueness of a given voxel. In order to register two image sequences accurately, we design for each point a morphological signature, i.e., an attribute vector **a**(**x**,*t*), for the purpose of minimizing the ambiguity in image matching and correspondence detection during the deformable registration procedure.

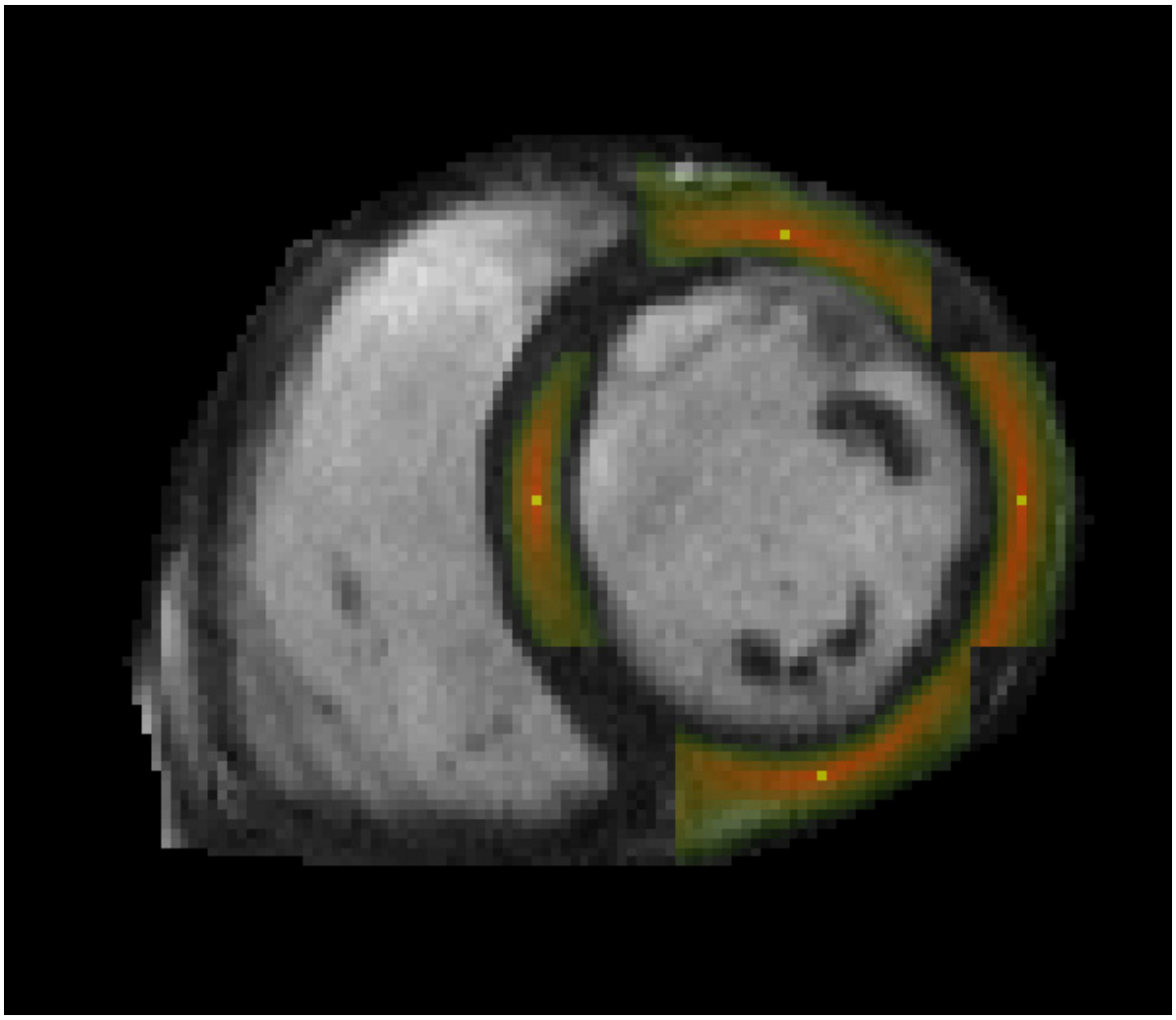


Figure 2. The similarity evaluated in the neighborhood of a voxel based on the attribute vectors. The yellow point marks the voxel in whose neighborhood the similarity is computed. Red represents higher similarity and green lower similarity.

The attribute vector defined at each voxel, $\mathbf{x}$, in the image reflects the underlying structure at different scales. Each attribute vector includes not only image intensity, but also boundary, and Geometric Moment Invariants (GMIs) [3], all of which are computed from the 3D spatial images. The attribute vector, $\mathbf{a(x)}$, is made up of three components $\mathbf{a}_1(\mathbf{x})$, $\mathbf{a}_2(\mathbf{x})$, and $\mathbf{a}_3(\mathbf{x})$. $\mathbf{a}_1(\mathbf{x})$ is a scalar value that represents the edge type of the voxel $\mathbf{x}$ in the image. For the myocardial motion estimation problem, it takes one of 7 discrete values, corresponding to 7 edge

types such as non-edge and six combinations of edges between the myocardium, blood and the surrounding tissue. $\mathbf{a}_2(\mathbf{x})$ is a scalar that represents the gray scale intensity of the voxel $\mathbf{x}$. The vector $\mathbf{a}_3(\mathbf{x})$ comprises of the GMIs of each tissue at different scales. For each scale and each tissue type, there are thirteen rotation invariants that are calculated from the zero-order, second-order and third-order 3D regular moments. GMIs are computed from different neighborhood sizes, and are concatenated into a long attribute vector. GMIs at a particular scale are calculated by placing a spherical neighborhood around each voxel and calculating a number of parameters that are invariant to rotation. GMIs can characterize the geometric properties of objects and are especially useful in distinguishing a voxel from close neighbors which can have similar intensity and edge types. The detailed definitions for attribute vectors are given in [3]. It is worth noting that for the generated image sequence $T(\mathbf{x},t)$, we need only to compute attribute vectors for one 3D image and other identical images will have the same set of attribute vectors.

Cine MR images carry very little information within the myocardium, since generally the blood is bright and the myocardium is dark. Therefore, the blood-myocardium interface provides valuable information for cardiac motion estimation. For this reason, we require that the boundary voxels in the template deform to the same boundary type in the subject. We therefore incorporate the benefits of both the surface-based and image-based motion estimation approaches by requiring the boundaries to match up and allowing the attribute similarities to determine correspondences in other areas. This leads to our definition of the similarity between two voxels $\mathbf{x}$ and $\mathbf{y}$ as,

$$m(\mathbf{a}(\mathbf{x}),\mathbf{a}(\mathbf{y})) = \begin{cases} 0, & \text{if } \mathbf{a}_1(\mathbf{x}) \neq \mathbf{a}_1(\mathbf{y}) \\ c([\mathbf{a}_2(\mathbf{x})\ \mathbf{a}_3(\mathbf{x})],[\mathbf{a}_2(\mathbf{y})\ \mathbf{a}_3(\mathbf{y})]), & \text{otherwise} \end{cases}$$

where $c\left([\mathbf{a}_2(\mathbf{x})\ \mathbf{a}_3(\mathbf{x})],[\mathbf{a}_2(\mathbf{y})\ \mathbf{a}_3(\mathbf{y})]\right)$ is the similarity of the second and third parts of the attribute vectors, which is defined as,

$$c\left([\mathbf{a}_2(\mathbf{x})\ \mathbf{a}_3(\mathbf{x})],[\mathbf{a}_2(\mathbf{y})\ \mathbf{a}_3(\mathbf{y})]\right)=\left(1-\left|\mathbf{a}_2(\mathbf{x})-\mathbf{a}_2(\mathbf{y})\right|\right)\prod_{i=1}^{K}\left(1-\left|\mathbf{a}_3^i(\mathbf{x})-\mathbf{a}_3^i(\mathbf{y})\right|\right),$$

where $\mathbf{a}_3^i(\mathbf{x})$ is the *i*-th element of $\mathbf{a}_3(\mathbf{x})$ that has a total of *K* elements. The effectiveness of the attribute vector for uniquely identifying a given voxel is demonstrated in **Figure 2**.

C. Energy Function

We solve the 4D image registration problem by hierarchically matching attribute vectors in the two image sequences and estimating the transformation that minimizes the difference between these attribute vectors. We model the registration function as an energy minimization problem, where the energy term includes temporal consistency and spatio-temporal smoothness terms in addition to the attribute similarity terms. The energy function can be written as,

$$E = E_F + E_B + E_C + E_S$$

Here, the first two terms, E_F and E_B are the image attribute similarity terms, E_C enforces temporal consistency, and E_S enforces spatio-temporal smoothness. Each of these terms is now explained in detail.

Attribute Similarity: To make the registration independent of which of the two sequences is treated as the template [3, 26], the energy function that evaluates the match of two image sequences should be symmetric with respect to the two images being registered. Therefore, we evaluate both the forward transformation $h(\mathbf{x},t)$ and the backward transformation $h^{-1}(\mathbf{x},t)$ and force them to be consistent[1] with each other. This improves the robustness of the motion estimation algorithm since it enforces inverse-consistency during the correspondence detection. With this consideration, we have two attribute similarity terms, i.e., the forward similarity, E_F, and the backward similarity, E_B, which are respectively defined as,

---

[1] From the perspective of correspondence detection, this merely enforces inverse consistency, i.e., a point **y** is a consistent match for **x** if **x** is also the best match for point **y**.

$$E_F = \sum_{t=1}^{N} \sum_{\mathbf{x}} \omega_T(\mathbf{x},t) \left(\sum_{(\mathbf{z},\tau)\in n(\mathbf{x},t)} d\left(\mathbf{a}_T(\mathbf{z},\tau), \mathbf{a}_I(h(\mathbf{z},\tau))\right)\right)$$

$$E_B = \sum_{t=1}^{N} \sum_{\mathbf{x}} \omega_I(\mathbf{x},t) \left(\sum_{(\mathbf{z},\tau)\in n(\mathbf{x},t)} d\left(\mathbf{a}_T(h^{-1}(\mathbf{z},\tau)), \mathbf{a}_I(\mathbf{z},\tau)\right)\right)$$

The first term E_F is defined on the forward transformation $h(\mathbf{x},t)$, and measures the similarity of attribute vectors between each point in the sequence $T(\mathbf{x},t)$ and its corresponding one in the sequence $I(\mathbf{x},t)$. The second energy term E_B is similar to the first term, but is instead defined on the inverse transformation $h^{-1}(\mathbf{x},t)$ to ensure that each point in the sequence $I(\mathbf{x},t)$ also finds its best matching point in the sequence $T(\mathbf{x},t)$. Specifically, in the forward similarity term, the importance of each point $(\mathbf{x},t)$ in the image registration is determined by its corresponding parameter $\omega_T(\mathbf{x},t)$, which is designed to be proportional to the distinctiveness of this point's attribute vector $\mathbf{a}_T(\mathbf{x},t)$. The match for each point $(\mathbf{x},t)$ is evaluated in its 4D (3D spatial and 1D temporal) neighborhood $n(\mathbf{x},t)$, by integrating all differences between the attribute vector $\mathbf{a}_T(\mathbf{z},\tau)$ of every neighboring point $(\mathbf{z},\tau)$ and the attribute vector $\mathbf{a}_I(h(\mathbf{z},\tau))$ of the corresponding point $h(\mathbf{z},\tau)$ in the sequence $I(\mathbf{x},t)$. The difference of two attribute vectors $d(\cdot,\cdot)$ ranges from 0 to 1 [3]. The size of the neighborhood $n(\mathbf{x},t)$ is large initially and decreases gradually with the progress of the registration, thereby increasing robustness and accuracy of deformable registration.

Consistency: The consistency energy term E_C measures the attribute-vector matching of corresponding points in different time frames of the sequence $I(\mathbf{x},t)$. This can be written as,

$$E_C = \sum_{t=1}^{N} \sum_{\mathbf{x}} \varepsilon_T(\mathbf{x},t) \left(\sum_{(\mathbf{z},\tau)\in n(\mathbf{x},t),\ t\neq\tau} d\left(\mathbf{a}_I(h(\mathbf{z},t)), \mathbf{a}_I(h(\mathbf{z},\tau))\right)\right)$$

For each point $(\mathbf{x},t)$ in the sequence $T(\mathbf{x},t)$, its corresponding point in the sequence $I(\mathbf{x},t)$ is $h(\mathbf{x},t)$. Since the sequence $T(\mathbf{x},t)$ has identical images at different time-points, i.e., same end-diastolic image, points $\{(\mathbf{x},t), 1 \le t \le N\}$ are the N corresponding points in the sequence $T(\mathbf{x},t)$; accordingly, N transformed points $\{h(\mathbf{x},t), 1 \le t \le N\}$ are the established correspondences in the sequence $I(\mathbf{x},t)$.

In this way, we can require the attribute vector $\mathbf{a}_I(h(\mathbf{x},t))$ of a point $h(\mathbf{x},t)$ in the image $I(\mathbf{x},t)$ to be similar to the attribute vector $\mathbf{a}_I(h(\mathbf{x},\tau))$ of its corresponding point $h(\mathbf{x},\tau)$ in the neighboring time-point image $I(\mathbf{x},\tau)$. This requirement is repeated for each position $(\mathbf{z},\tau)$ in a 4D neighborhood $n(\mathbf{x},t)$, and the total attribute-vector difference is weighted by $\varepsilon_I(\mathbf{x},t)$ to reflect the importance of a point $(\mathbf{x},t)$ in the sequence $T(\mathbf{x},t)$. The use of this energy term makes it potentially easier to solve the 4D registration problem, since the registration of cardiac images at neighboring time-points is relatively easier and thus it can provide a good initialization for 4D image registration by initially focusing only on energy terms of E_c and E_S.

Smoothness: The fourth energy term E_S is a smoothness constraint for the transformation $h(\mathbf{x},t)$.

$$E_S = \alpha \cdot E_S^{\text{Spatial}} + \beta \cdot E_S^{\text{Temporal}}$$

For convenience, we separate this smoothness constraint into two components, i.e., a spatial smoothness constraint $E_S^{\textbf{Spatial}}$ and a temporal smoothness constraint $E_S^{\textbf{Temporal}}$, and these two constraints are linearly combined by their own weighting parameters α and β.

For the *spatial* smoothness constraint, we use a Laplacian operator [3] to impose spatial smoothness, which is defined as,

$$E_S^{\text{Spatial}} = \sum_{t=1}^{N} \int_{\mathbf{x}\in\Omega} \left\| \nabla^2 u(\mathbf{x},t) \right\|$$

where $u(\mathbf{x}, t)$ is the displacement which is related to the forward transformation by the relation,

$$h(\mathbf{x},t) = \mathbf{x} + u(\mathbf{x},t)$$

It is important to remember that all images in the sequence $T(\mathbf{x},t)$ are the identical end-diastolic image. Since maximum strain is experienced at end-systole, we have to register the end-diastolic image with the end-systolic image using a large nonlinear transformation, which might be over-smoothed by the Laplacian operator. To avoid this over-smoothing problem, we use a multi-resolution framework, i.e., multi-level transformations, to implement our registration algorithm. Each resolution will estimate its own level of transformation based on the total transformations

estimated from the previous resolutions, and the final transformation used to register two image sequences is the summation of all levels of transformations respectively estimated from all resolutions. Notably, the Laplacian operator is only allowed to smooth the current level of transformation being estimated at the *current resolution*, which effectively avoids smoothing the transformations estimated from the previous resolutions.

For the *temporal* smoothness constraint, we use a Gaussian filter to obtain an average transformation in a 1D temporal neighborhood around each point ($\mathbf{x}$,t), and force the transformation on this point ($\mathbf{x}$,t) to follow its average transformation in the temporal neighborhood.

Constraints: Since the image sequence is periodic, and represents one cardiac cycle that repeats itself, we need to add a constraint for periodicity. This is because the first image I_1 and the last image I_N are temporal neighbors as indicated by solid arrows in Fig 1. In this way, we ensure that each point $\mathbf{x}$ moves smoothly along the temporal direction from the end-diastolic image I_1 to other time-points, and returns to its original position after completing the 4D image registration, since the first images respectively in the two sequences are identical and the transformation between them is thus forced to be exactly zero during the entire registration procedure.

D. Muti-Resolution Approach

We perform the registration in a multi-resolution fashion by adaptively selecting different sets of image points that are used to drive the registration. To do this, we express the total image similarity energy as a weighted sum of the similarity energies of the individual points. The weights are assigned to the points according to the distinctiveness of the attribute vectors, i.e., we assign large weights for the energy terms of points that are distinctive, such as points with high curvature along the left ventricular wall. We assign a weight of zero to all points that are not to be considered for the global energy term at a given registration level. This allows us to focus on the most suitable points for actively driving the image registration.

The adaptive approach using focus points is superior to standard sub-sampling based multi-resolution registration techniques for several reasons. First, it potentially avoids errors in the image similarity terms by using only the most distinctive points at each level, rather than using a regular (or random) sampling of the image space. This increases the robustness of the motion estimation, as it is not affected by the similarity of unreliable points, such as points in the ventricular blood pool. Second, as we pick points adaptively, the estimated transformation captures the actual transformation better than a sub-sampled version, even at the low resolution levels, i.e., we pick more points in regions with more deformation and fewer in regions which are relatively static. Finally, another benefit is that of improved speed of the motion estimation. If we were to solve for the displacement at every grid point on the image, the dimensionality of the cost function that we minimize would be extremely high and the cost function evaluation would be much more expensive. Solving for the displacements at only the focus points reduces the dimensionality of the optimization problem. Thus, the procedure approximates a very high-dimensional cost function (equal to the number of points in the image sequence) by a significantly lower-dimensional cost function of only the active points. This latter function has few local minima, because it is a function of the coordinates of active points, for which relatively unambiguous matches can be found. Therefore, using this strategy, we can speed up the performance of image registration and also reduce the chances of local minima, which in part result from ambiguities in determining the matching pairs of points.

3. Results

Several experiments were performed to validate the accuracy of our motion estimation algorithm, including demonstrating the performance by visual inspection of the resulting warps, motion fields, interface tracking, and ventricular volumes. We also validate the motion fields by comparison to motion estimates obtained using tagged MR data and extracted using the method proposed by Chandrashekara et al [27].

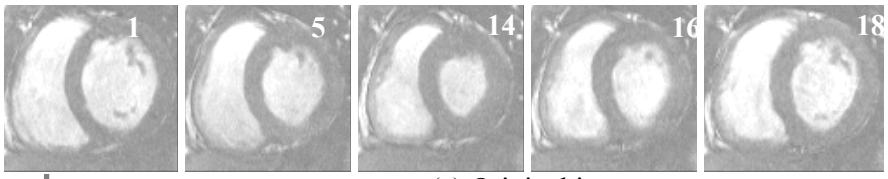
We acquired cine MR sequences with and without myocardial tagging for 3 healthy volunteers in order to validate our motion estimation algorithm. The cine MR and tagged images were acquired on the same scanner during the same scanning session at end-expiration, thus the datasets are assumed self-registered. Short and long axis segmented *k*-space breath-hold cine TrueFISP (SSFP) images were acquired for three healthy volunteers. For validation purposes, we also acquired grid tagged, segmented *k*-space breath-hold cine TurboFLASH images.

A. Visual Inspection of Motion Estimation Results

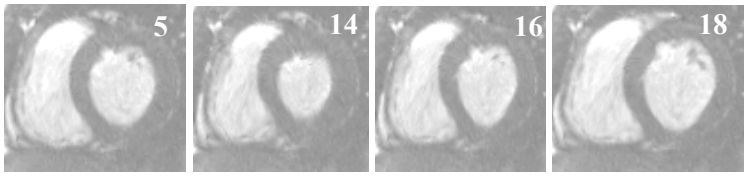
The *first* experiment evaluates the performance of the proposed method by visual inspection of the deformation field and the warped images. We evaluate the smoothness of the deformation field by evaluating the determinant of the Jacobian matrix over the entire myocardium, and verify that this is positive at all voxels. The deformation field, and correspondingly the motion estimates, can additionally be evaluated by warping the end-diastole image and comparing the warped images with the original images. All of the cine sequences in our test datasets had 33 time frames; selected frames along with their warped counterparts are shown in Figure 3. The similarity between the original and the warped images can be seen in these images, even for the end-systolic frames where the deformation is large. The deformation vectors over the left ventricular region are shown in Figure 3(c); note that these are in agreement with expectations of myocardial motion based on other methods of estimation.

An important benefit of motion estimation is that we can use the estimates to perform boundary tracking and segmentation. In other words, if one of the time frames is labeled or segmented, this information can be propagated to the other frames. This can also be used to evaluate the accuracy of the motion estimation algorithm, as we can compare the segmentation results with those performed by hand. The tracking results are shown in Figure 4, where the myocardial walls are marked in red. For the first frame, these contours are obtained by a semi-automatic segmentation procedure using the method in [28]. These semi-automatically segmented

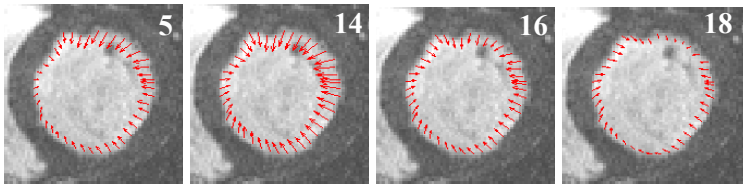
contours can be propagated to the other frames based on the deformation fields obtained from the 4D registration procedure. The deformed contours are shown in both the long and short-axis views of the heart; the blood-myocardium interface is tracked accurately in both views and over different time-frames.



(a) Original images

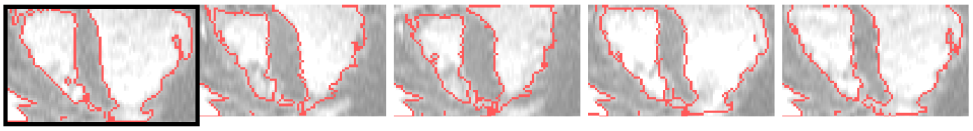


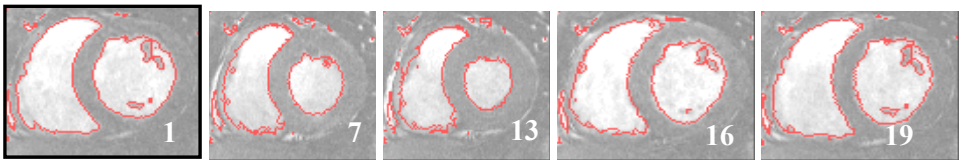
(b) The end-diastolic image warped to other time-points



(c) Estimated deformations around the endocardial border of left ventricle

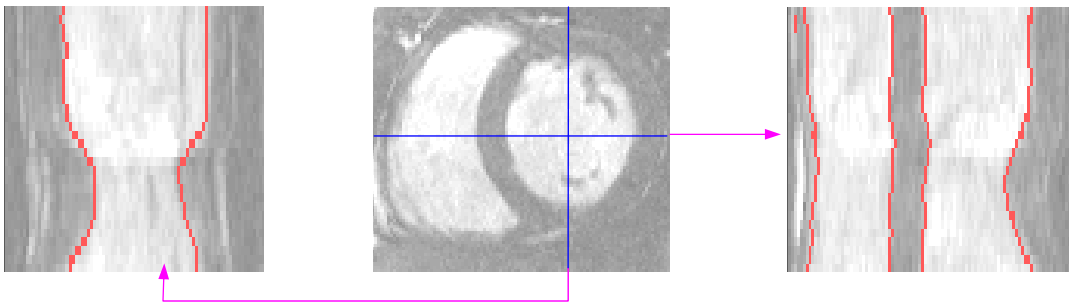
Figure 3. Estimating deformations from end-diastole to other frames. (a) Five selected cardiac images, (b) the results of warping end-diastole to other time-points, (c) deformations around left ventricle, estimated from the end-diastolic image and cropped here for clear display.


(a) Long-axis view



(b) Short-axis view

Figure 4. Tracking/labeling the boundaries of interest in a cardiac sequence.



(a) (b)

Figure 5. Tracking/labeling results in temporal views of two short-axis lines. The red colors indicate the labeling results on the left and right ventricular boundaries.

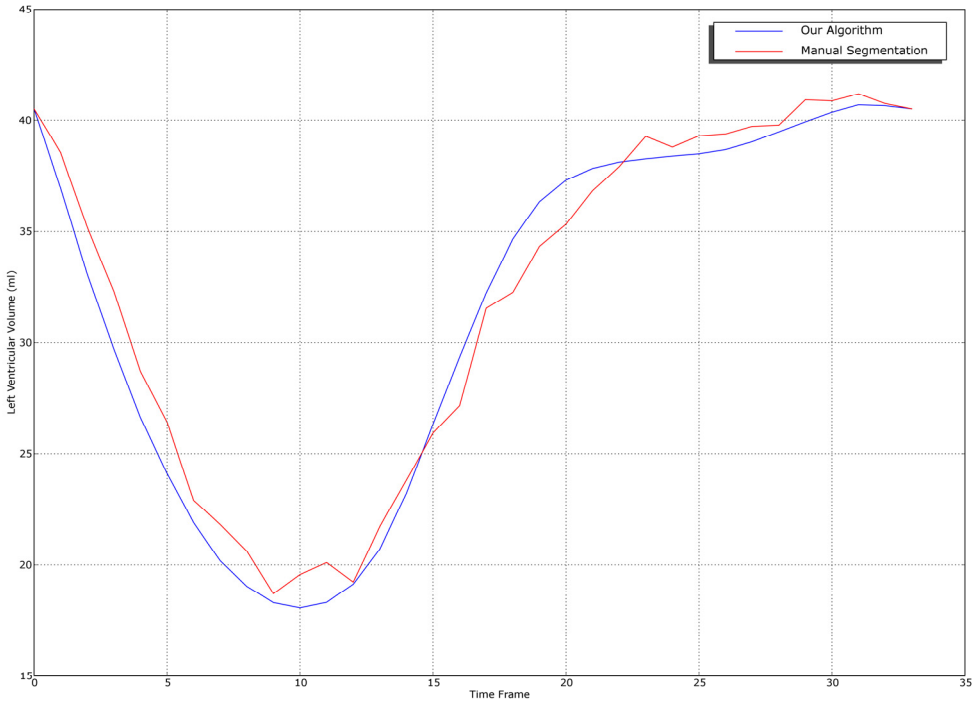


Figure 6. Left ventricular volume of a selected subject, segmented by our algorithm (solid curve) and by hand (dotted curve) over all frames in a cardiac cycle.

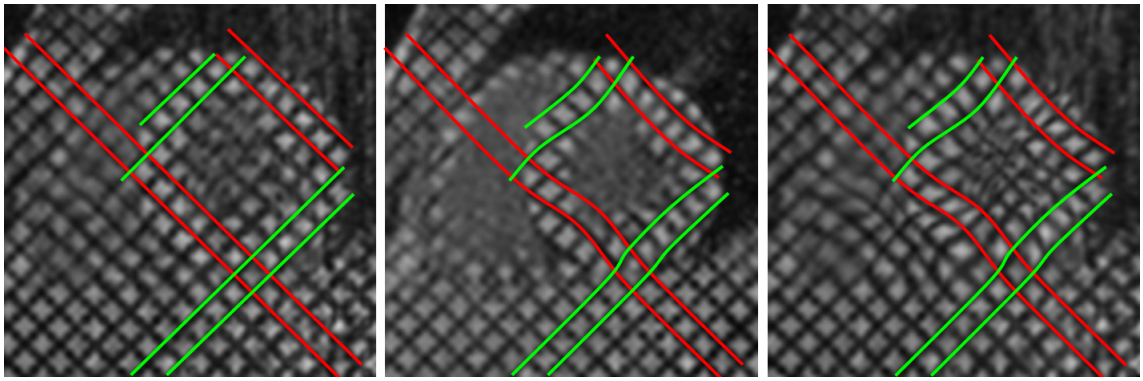
Additionally, a point in the myocardium can also be tracked temporally, as shown in Figure 5. This result demonstrates the temporal smoothness of the motion estimation, which cannot be interpreted easily by observing the motion fields. An alternate way of interpreting this is to analyze the left-ventricular volume, obtained using manual segmentation and by the using

the motion estimates, as seen for one particular subject in Figure 6. The motion estimation based volume curve is smoother than the one obtained by manual segmentation, but is in general agreement with the latter. We manually segmented the left ventricle for 8 subjects, 3 volunteers and 5 with suspected cardiomyopathy and evaluated the difference in ventricular volume over the entire cardiac cycle. The average volume error was 3.37%, with standard deviation 2.56%; and average volume overlap error was 7.04%, with standard deviation 3.28%. The correlation coefficient between the manual and automatic segmentations was r = 0.99.

B. Validation Against Tagged MR Images

Although the above mentioned validation schemes ascertain to a certain extent the accuracy of the obtained motion field, they are limited to validation based on visual inspection of prominent features like the myocardial boundary. Evaluating the accuracy of the registration and consequently the motion estimation in other regions, for example inside the myocardium, is not as straightforward. In order to overcome this, we use the motion estimates obtained to place *virtual taglines* on the tagged images. The virtual taglines are drawn manually on the end-diastolic image and deformed according to the motion estimate to obtain the virtual taglines for the other frames. Visual comparison of these with the actual tags is an effective way of estimating the accuracy of the motion estimation procedure, even in areas without visually discernible features. These results are shown in Figure 7.

Slice 26



Slice 15

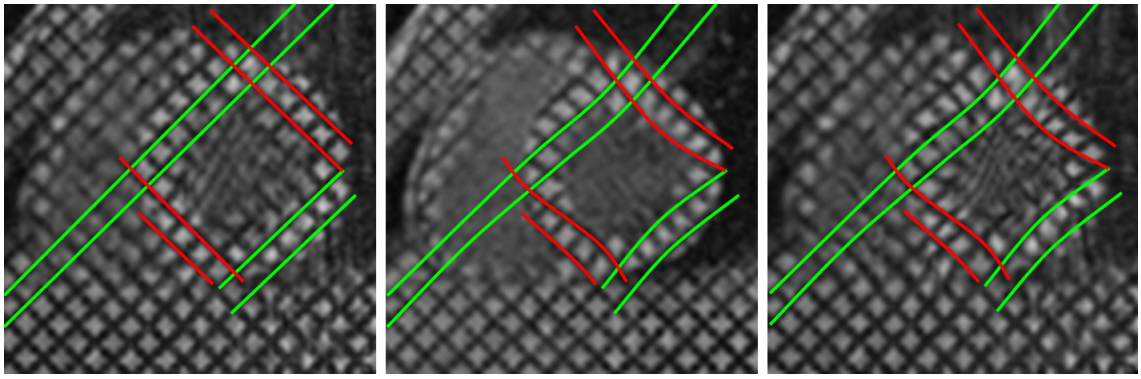


Figure 7. Virtual taglines overlaid on tagged MR images. The virtual taglines were obtained by deforming manually drawn lines (in the end-diastolic image) using the deformation field obtained from the motion estimation procedure. Note good agreement between the virtual and actual taglines in the deformed frames.

To quantize the accuracy of the agreement with the virtual tags, tag-intersection points were manually detected by an observer in three different SA slices and one LA slice for all time frames between end-diastole and end-systole, including 20 points on the SA slices and 14 points on the LA slices. The root mean square error between the in-plane displacements estimated from the registration algorithm and the actual in-plane displacements as measured by the observer are presented in Figure 8. It can be seen that the average error is within acceptable limits (2mm) and support our claim to the estimation of accurate motion estimation.

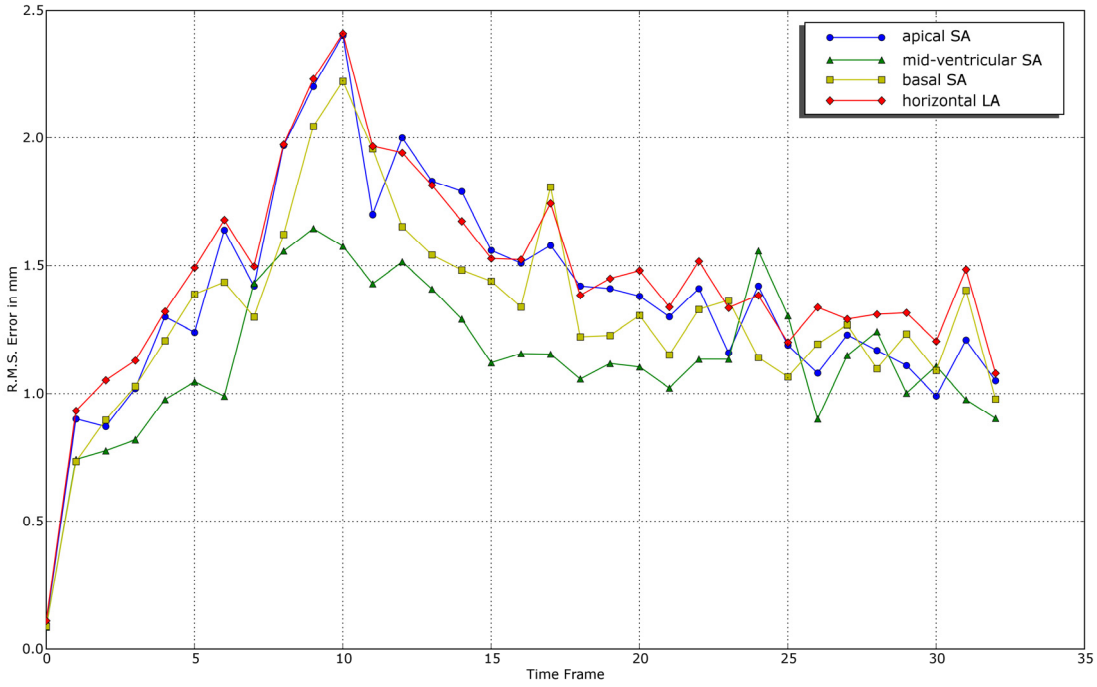


Figure 8. Plot of the rms error for the displacements estimated using the 4D registration algorithm against those obtained via manual tracking of tag intersection points.

In order to further evaluate the accuracy of the motion estimates obtained using our method, we extracted the motion fields from the tagged MR images using the method proposed by Chandrashekara *et al* [27]. We compared the determinant of the Jacobian, calculated in blocks corresponding to the tag spacing (8 mm). The average error in the determinant of the Jacobian between the two methods was 5.63% with a standard deviation of 3.44%.

4. Conclusion

A 4D deformable registration method for estimation of cardiac motion from MR image sequences was presented, and experimentally tested. The cardiac motion estimation was formulated as a 4D image registration problem, which simultaneously considers all images of different time-points and further constrains the spatiotemporal smoothness of estimated motion fields concurrently with the image registration procedure. Also, compared to other motion estimation methods that use relatively simple features such as curvature of the left ventricular border, our method uses a

rich set of attributes, including GMIs, to distinguish the corresponding points across different time-points, thereby maximally reducing ambiguity in image matching. Finally, by selecting the active points hierarchically based on the degree of distinctiveness of their attribute vectors, the proposed registration algorithm has the potential to produce a global solution for motion estimation.

The performance of this 4D registration method for cardiac applications has been evaluated by visual inspection as well as quantitative validation using tagged MR images. Experimental results demonstrate good performance of our method in estimating cardiac motion from cine MR sequences. The motion estimates are statistically similar to the motion of tag lines obtained from tagged MR images. Cine MR images are commonly acquired in a clinical setting, and the use of our algorithm may obviate the need to acquire additional tagged images to estimate myocardial motion. In addition, our method allows for dense estimates of myocardial motion, which is not restricted to the intersection of tag lines, thus helping detect abnormal myocardial motion and also potentially improving early diagnosis and treatment planning of cardiomyopathies.

References:

1. *Heart Disease and Stroke Statistics*. 2004, American Heart Association: Dallas, Texas.
2. Roche, A., et al., *Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient information.* IEEE Transactions on Medical Imaging, 2001. **20**(10): p. 1038-1049.
3. Shen, D. and C. Davatzikos, *HAMMER: Hierarchical attribute matching mechanism for elastic registration.* IEEE Transactions on Medical Imaging, 2002. **21**(11): p. 1421-1439.
4. Teh, C.H. and R.T. Chin, *On image analysis by the method of moments.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1988. **10**(4): p. 496-513.
5. Manjunath, B.S. and W.Y. Ma, *Texture Features for Browsing and Retrieval of Image Data.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996. **18**(8): p. 837-842.
6. Verma, R. and C. Davatzikos. *Matching of Diffusion Tensor Images Using Gabor Features*. in *Proceedings of the IEEE International Symposium on Biomedical Imaging*. 2004. Arlington, Va.

7. Liu, J., B.C. Vemuri, and J.L. Marroquin, *Local frequency representations for robust multimodal image registration.* IEEE Transactions in Medical Imaging, 2002. **21**(5): p. 462–469.
8. Smart, S.C., et al., *Dobutamine-Atropine Stress Echocardiography for the Detection of Coronary Artery Disease in Patients With Left Ventricular Hypertrophy : Importance of Chamber Size and Systolic Wall Stress.* Circulation, 2000. **101**(3): p. 258-263.
9. Zerhouni, E., et al., *Human heart: tagging with MR imaging--a method for noninvasive assessment of myocardial motion.* Radiology, 1988. **169**(1): p. 59-63.
10. Park, J., D. Metaxas, and L. Axel, *Analysis of left ventricular wall motion based on volumetric deformable models and MRI-SPAMM.* Medical Image Analysis, 1996. **1**: p. 53-71.
11. Shi, P., et al., *Volumetric Deformation Analysis Using Mechanics-Based Data Fusion: Applications in Cardiac Motion Recovery.* International Journal of Computer Vision, 1999. **35**(1): p. 87-107.
12. Amini, A.A. and J.L. Prince, *Measurement of Cardiac Deformations from MRI: Physical and Mathematical Models*. Book Chapter, 2001: Kluwer.
13. Guttman, M.A., J.L. Prince, and E.R. McVeigh, *Tag and contour detection in tagged MR images of the left ventricle.* IEEE Transactions on Medical Imaging, 1994. **13**(1): p. 74-88.
14. Osman, N.F., W.S.K., E. R. McVeigh, J.L. Prince,, *Cardiac motion tracking using CINE harmonic phase (HARP) magnetic resonance imaging.* Magnetic Resonance in Medicine, 1999. **42**(6): p. 1048-1060.
15. Garot, J., et al., *Fast Determination of Regional Myocardial Strain Fields From Tagged Cardiac Images Using Harmonic Phase MRI.* Circulation, 2000. **101**(9): p. 981-988.
16. Meyer, F.G., et al., *Tracking myocardial deformation using phase contrast MR velocity fields: a stochastic approach.* IEEE Transactions on Medical Imaging, 1996. **15**(4): p. 453-465.
17. Frederick H. Epstein, W.D.G., *Displacement-encoded cardiac MRI using cosine and sine modulation to eliminate (CANSEL) artifact-generating echoes.* Magnetic Resonance in Medicine, 2004. **52**(4): p. 774-781.
18. Ledesma-Carbayo, M.a.J., et al., *Cardiac Motion Analysis from Ultrasound Sequences Using Non-rigid Registration*. 2208 ed. Lecture Notes in Computer Science. 2001. 889.
19. McEachen, J.C., A. Nehorai, and J.S. Duncan, *Multiframe temporal estimation of cardiac nonrigid motion.*IEEE Transactions on Image Processing, 2000. **9**(4): p. 651-665.
20. Papademetris, X., et al., *Estimation of 3D left ventricular deformation from echocardiography.* Medical Image Analysis, 2001. **5**(1): p. 17-28.
21. Perperidis, D., R.H. Mohiaddin, and D. Rueckert, *Spatio-temporal free-form registration of cardiac MR image sequences* .Medical Image Analysis, 2005. **9**(5): p. 441-456.
22. Song, S.M. and R.M. Leahy, *Computation of 3-D velocity fields from 3-D cine CT images of a human heart.* IEEE Transactions on Medical Imaging, 1991. **10**(3): p. 295-306.
23. Wang, Y.-P., Y. Chen, and A.A. Amini, *Fast LV motion estimation using subspace approximation techniques.* IEEE Transactions on Medical Imaging, 2001. **20**(6): p. 499-513.
24. Makela, T., et al., *A review of cardiac image registration methods.* IEEE Transactions on Medical Imaging, 2002. **21**(9): p. 1011-1021.
25. Zitova, B. and J. Flusser, *Image registration methods: a survey.* Image and Vision Computing, 2003. **21**(11): p. 977-1000.
26. Christensen, G.E. and H.J. Johnson, *Consistent Image Registration.* IEEE Transactions on Medical Imaging, 2001. **20**(7): p. 568-582.

27. Chandrashekara, R., R.H. Mohiaddin, and D. Rueckert, *Analysis of 3-D myocardial motion in tagged MR images using nonrigid image registration.* IEEE Transactions on Medical Imaging, 2004. **23**(10): p. 1245-1250.
28. Pham, D.L. and J.L. Prince, *Adaptive fuzzy segmentation of magnetic resonance images.* IEEE Transactions on Medical Imaging, 1999. **18**(9): p. 737-752.