Research Article

Joint Segmentation and Groupwise Registration of Cardiac Perfusion Images Using Temporal Information

Dwarikanath Mahapatra

Department of Computer Science, Swiss Federal Institute of Technology, 8092 Zurich, Switzerland

Correspondence should be addressed to Dwarikanath Mahapatra, dwarikanath.mahapatra@inf.ethz.ch

We propose a joint segmentation and groupwise registration method for cardiac perfusion images by using temporal information. The nature of perfusion images makes groupwise registration especially attractive as the temporal information from the entire image sequence can be used. Registration aims to maximize the smoothness of the intensity signal, while segmentation minimizes a pixel's dissimilarity with other pixels having the same segmentation label. The cost function is optimized in an iterative fashion using B-splines. Tests on real patient datasets show that compared to two other methods, our method shows lower registration error and higher segmentation accuracy. This is attributed to the use of temporal information for groupwise registration and mutually complementary registration and segmentation information in one framework, while other methods solve the two problems separately.

1. Introduction

Dynamic contrast-enhanced (DCE) magnetic resonance (MR) images (or perfusion MRI) have developed as a popular noninvasive tool for the functional analysis of internal organs. Contrast agent is injected intravenously into the patient, and a series of MR images are acquired over a period of time. As the contrast agent flows through the blood stream, the intensity of corresponding regions increases. Since the image acquisition process can take up to 20 minutes, patient movement is inevitable. Additionally, elastic deformation of cardiac tissues due to patient breathing needs to be corrected. Perfusion images are characterized by rapid intensity change over time, low spatial resolution, and noise and make registration challenging. Previous techniques mostly employed a pairwise registration approach, that is, all images of a sequence are individually registered to a fixed reference image [1–3]. The success of such approaches depends upon the robustness of the cost function to intensity change. Although intensity change due to contrast agent flow poses challenges, it also provides important temporal information

for registration and segmentation of the perfusion image sequence. In this work, we propose a method which makes use of the temporal dynamics of contrast agent flow to achieve joint segmentation and groupwise registration of an MR cardiac image sequence.

The changing intensity due to contrast agent flow provides valuable segmentation by highlighting the organ of interest. It is interesting to note that different regions of a cardiac image sequence have different intensity-time characteristics (Figure 4). Temporal flow information was used for registration [4] and segmentation [5] of perfusion images. In [6], the registration framework constrained the deformation field such that different regions follow a particular intensity-time profile. But it is an accepted fact that improved registration leads to accurate segmentation and vice versa. In previous works [7, 8], we have approached the problem of joint registration and segmentation of cardiac MRI. However, we did not make use of the temporal information from the image sequence, rather focusing on pairwise registration of images. In this paper, we make use of temporal information to achieve registration and segmentation of the entire image

sequence. Instead of pairwise registration, we solve the problem using groupwise registration as it allows us to impose constraints based on temporal information.

Of late, groupwise registration methods have gained popularity because of a need to register large number of datasets for atlas construction [9, 10]. Groupwise registration is generally approached by two techniques. The first approach uses pairwise registration between a template and all other images in the population as in [11–13]. Yang et al. in [13] utilize the voxelwise geometry and orientation information for registration. Different measurements and rotation invariant features are used in a deformable matching mechanism for image alignment. However, pairwise registration has two limitations. First, selecting a fixed template may not accurately represent the population. Secondly, pairwise registration is not very effective when registering two images with significant anatomical differences. Only those subjects that are close to the template are registered properly.

To overcome the above limitations, there are many methods that achieve registration using all images from the population. This approach is more faithful to the term "groupwise registration." The goal is to warp all subjects in a population towards a hidden common space simultaneously within a single framework [14, 15]. The groupwise registration problem is formulated as one of optimization, with a global cost function defined on all aligned images [16, 17]. The cost function in [16, 17] is defined as the stack entropy or the entropy of corresponding voxels in different volumes. The justification is that if images are properly aligned, intensity values at corresponding voxels from all volumes will form a low entropy distribution. With this constraint, groupwise registration is achieved within a B-spline-based free-form deformation framework.

When across-subject variation is very large, it is generally difficult to achieve good registration by simply registering each image to a template image. Therefore, many approaches make us of intermediate templates for registration [18–21]. In [18, 21], an intermediate template which does not belong to the original dataset is created to aid registration between two images. Tang et al. in [21] warp the template image with a set of simulated deformation fields learnt using principal component analysis (PCA) on a set of training deformation fields. Such approaches do not guarantee that the intermediate template is realistic which may affect registration results. Jia et al. in [19] construct a tree structure where each node of the tree is represented by an image, and similar images are represented by connected nodes. Each image is registered with the help of intermediate templates determined along its own path with respect to the final template.

Groupwise registration methods are particularly suitable for perfusion DCE-MRI. Each region of the scanned organ is characterized by a different intensity profile over time. For example, in cardiac perfusion MRI, the contrast agent first flows into the right ventricle (RV) and then into the left ventricle (LV) before being flushed out of the cardiovascular system. Thus, a pixel within the RV shows a peak intensity magnitude early in the scanning sequence, while for the LV blood pool, the intensity peak occurs later (Figure 4). With this available information, we can formulate the cost function such that, after registration, pixels from certain regions follow a particular intensity time profile. This is achieved by joint segmentation and groupwise registration of DCE-MRI. There are not many works dealing with joint segmentation and groupwise registration of cardiac perfusion MRI, although Zhang et al. in [22] describe a method for the rigid registration of brain perfusion images. The cost function is derived from the total quadratic variation of image intensity. Metz et al. in [23] propose a method for groupwise registration of dynamic lung data using both spatial and temporal constraints where groupwise optimization of B-splines is used.

The temporal intensity patterns of pixels also determine their segmentation labels (i.e., RV, LV blood pool, myocardium background, etc.). Thus, the image sequence can be segmented along with groupwise registration. It is a well-known fact that registration and segmentation are mutually complementary approaches. Many works have combined them in a joint registration and segmentation framework [24–26]. Including segmentation information into the cost function reduces registration error and improves segmentation accuracy. In this paper, we propose a joint segmentation and groupwise registration (JSGR) approach for cardiac perfusion MRI. Our method combines intensity information from the entire image sequence (for registration) and maximizes the similarity between a pixel and other pixels belonging to the same class (for segmentation). We describe our method in Section 2, present experimental results in Section 3, and conclude with Section 4.

2. Theory

JSGR aims to find the transformation for every image to minimize a cost function. We do not have an explicitly defined reference image but constrain the transformations such that the registered images approach a common image space (which is the reference image) and are approximated to be the center of the images being registered. Wu et al. in [27] have highlighted the importance of a sharp mean image for accurate groupwise registration. Their observations are derived from constructing atlases for a large population of brain images. Such datasets show a lot of variability, and a fixed reference image is sure to introduce bias. Our method is used to achieve registration and segmentation using images from the same dataset, where each patient has multiple frames from different time intervals.

The general objective function for JSGR consists of two terms, that is,

$$E = E_{\text{data}} + E_{\text{smooth}}, \qquad (1)$$

where E_{data} is the data cost and E_{smooth} is the smoothness cost. The data cost depends upon the type of images being registered, and the smoothness cost depends upon the optimization framework.

Perfusion images are characterized by rapid intensity change over time. Instead of relying on low-level information, we aim to exploit the temporal information for groupwise registration. First, we give a brief description of
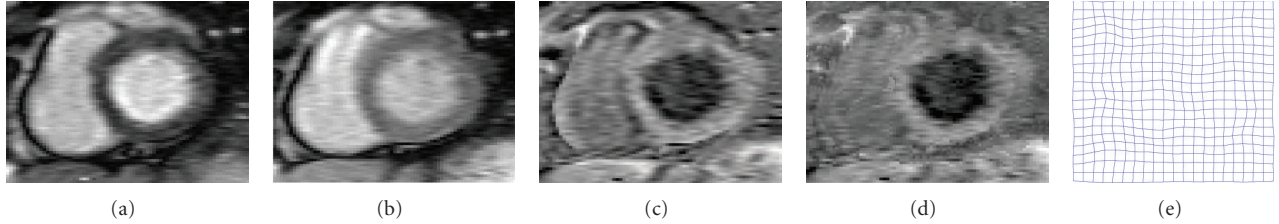
FIGURE 1: B-spline registration results for cardiac perfusion images. (a) Reference image; (b) floating image; difference image; (c) before registration; (d) after registration using *JRGS*; (e) deformed grids obtained from *JRGS*.

B-splines and their optimization. Then we explain the formulation of our data cost (E_{data}).

2.1. B-Spline-Based Registration. A B-spline-based freeform deformation (FFD) transformation model was presented in [28] for the elastic registration of breast images. The basic idea of FFDs is to deform an object by manipulating an underlying mesh of control points. The resulting deformation controls the shape of the 3D (or 2D) object and produces a smooth and continuous transformation. The transformation field consists of a global and local component and is defined as

$$T(\mathbf{x}) = T_{\text{local}}\big(T_{\text{global}}(\mathbf{x})\big), \tag{2}$$

where T_{global} is an affine transform obtained using [1], and T_{local} is the deformation based on B-splines. The DCE-MR images are first affinely registered to a chosen reference image. Note that this reference image is only for the purpose of rigid registration and is usually the scan showing all tissues without ambiguity. For elastic registration, there is no explicitly defined reference image. Further discussion is restricted to T_{local}.

We define an initial $n_x \times n_y \times n_z$ grid of control points denoted as Φ. The grid points are denoted as $\Phi_{i,j,k}$ and have uniform spacing. The free-form deformation can be written as the 3D tensor product of 1D cubic B-splines,

$$T_{\text{local}}(\mathbf{x}) = \mathbf{x} + \sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3} B_l(u)B_m(v)B_n(w)\Phi_{i+l,j+m,k+n}, \tag{3}$$

where $\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3})$ is the displacement vector, $i = \lfloor x/n_x \rfloor - 1$, $j = \lfloor y/n_y \rfloor - 1$, $k = \lfloor z/n_z \rfloor - 1$, $u = x/n_x - \lfloor x/n_x \rfloor$, $v = y/n_y - \lfloor y/n_y \rfloor$, $w = z/n_z - \lfloor z/n_z \rfloor$, and B_l is the lth cubic B-spline basis function given by the following equations:

$$B_0(u) = \frac{(1-u)^3}{6},$$

$$B_1(u) = \frac{(3u^3 - 6u^2 + 4)}{6},$$

$$B_2(u) = \frac{(-3u^3 + 3u^2 + 3u + 1)}{6}, \tag{4}$$

$$B_3(u) = \frac{(u^3)}{6}.$$

Since our datasets are in 2D, the corresponding equations are

$$T_{\text{local}}(\mathbf{x}) = \mathbf{x} + \sum_{l=0}^{2}\sum_{m=0}^{2} B_l(u)B_m(v)\Phi_{i+l,j+m}. \tag{5}$$

B-splines are locally controlled, which makes them computationally efficient even for a large number of control points. In particular, the basis function of cubic B-splines have a limited support, that is, changing control point $\Phi_{i,j}$ affects the transformation only in the local neighborhood of that control point.

The local deformation of the cardiac tissues should be characterized by a smooth transformation. To constrain the spline-based FFD transformation to be smooth, one can introduce a penalty term which regularizes the transformation. The general form of such a transformation in 3D takes the following form:

$$
\begin{aligned}
&E_{\text{smooth}} \\
&= \frac{1}{V}\int_0^X\int_0^Y\int_0^Z\left[\begin{array}{l}\left(\dfrac{\partial^2 T}{\partial x^2}\right)^2 + \left(\dfrac{\partial^2 T}{\partial y^2}\right)^2 + \left(\dfrac{\partial^2 T}{\partial z^2}\right)^2 + \\[2ex] 2\left(\dfrac{\partial^2 T}{\partial xy}\right)^2 + 2\left(\dfrac{\partial^2 T}{\partial xz}\right)^2 + 2\left(\dfrac{\partial^2 T}{\partial yz}\right)^2\end{array}\right]dx\,dy\,dz,
\end{aligned}
\tag{6}
$$

where V denotes the volume of the image domain. Our registration algorithm deals with 2D images in the spatial domain. Therefore, the smoothness cost can be formulated as

$$
\begin{aligned}
&E_{\text{smooth}} \\
&= \frac{1}{A}\int_0^X\int_0^Y\left[\left(\frac{\partial^2 T}{\partial x^2}\right)^2 + \left(\frac{\partial^2 T}{\partial y^2}\right)^2 + 2\left(\frac{\partial^2 T}{\partial xy}\right)^2\right]dx\,dy,
\end{aligned}
\tag{7}
$$

where A is the area of the image domain.

2.2. Similarity Measure. The images are first affinely aligned with respect to a reference image using [1]. Note that this reference image is only for rigid alignment and is not used for groupwise registration. Seed points belonging to RV, LV blood pool, myocardium, and background are identified (as shown by red arrows in the first image of the first row of Figure 3), and the labels of other pixels are determined
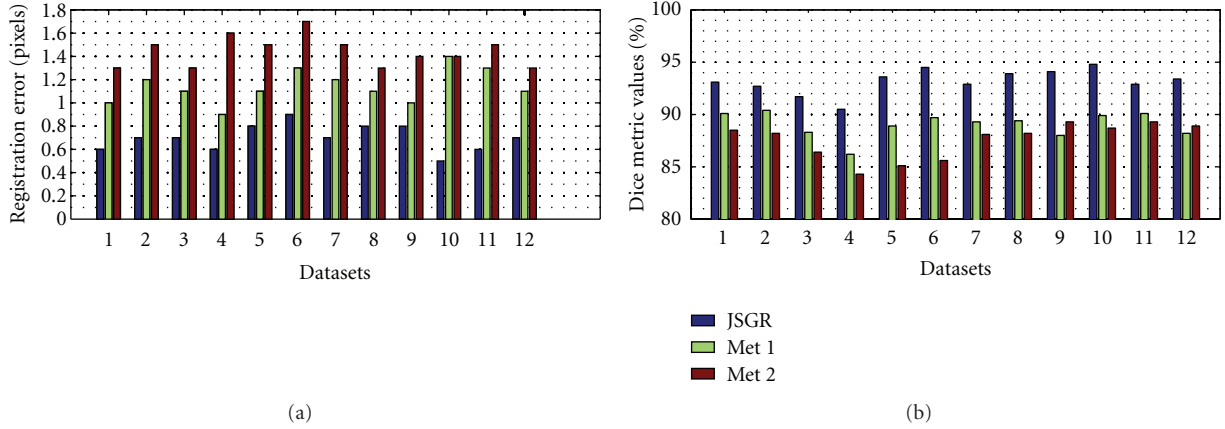
(a)

(b)

FIGURE 2: Registration and segmentation results of the LV for each of the 12 datasets. (a) Registration error (in pixels); (b) segmentation accuracy in %.
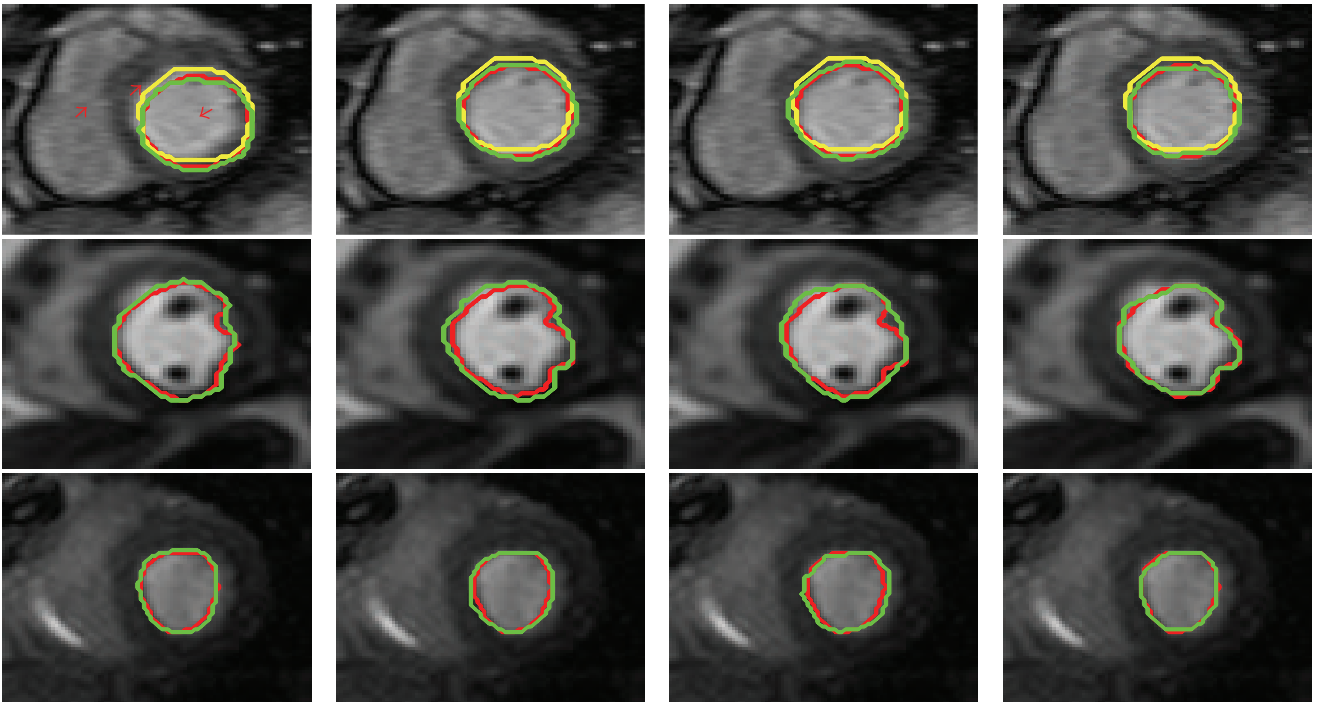

FIGURE 3: Contours of segmented regions overlaid on images from the dataset. First column shows results for *JSGR*, second column for *Met* 1, third column for *Met* 2, and fourth column for *JRS*. Each row corresponds to a different dataset.

using graph cuts. The initial labeling is used to calculate the cost functions in the first round of iteration. After the B-spline grid of each image is updated, the images are transformed and the segmentation labels are also updated based on the transformed images. We define the data cost as a combination of two terms which, individually, exploit the different characteristics of the perfusion datasets. Thus, E_{data} is defined as

$$E_{\text{data}} = w_1 \times E_W + w_2 \times E_Q, \qquad (8)$$

where E_Q calculates the *total quadratic variation* of the dataset, E_W calculates the *within-class distance* of each

pixel, and w_1, w_2 are weights that determine the relative contribution of each term. $w_1 = .4$ and $w_2 = 1$. Two weights are used in order to examine the relative contribution of each term to the results (discussed in Section 3.4). Below we explain each term in greater detail.

2.3. Total Quadratic Variation (E_Q). After registration of DCE-MRI, we expect the intensity time variation of pixels to be smooth after motion correction. The *total quadratic variation*, E_Q, measures the smoothness of the intensity signal by a combination of its first- and second-order derivatives. The sum of first derivatives over the entire
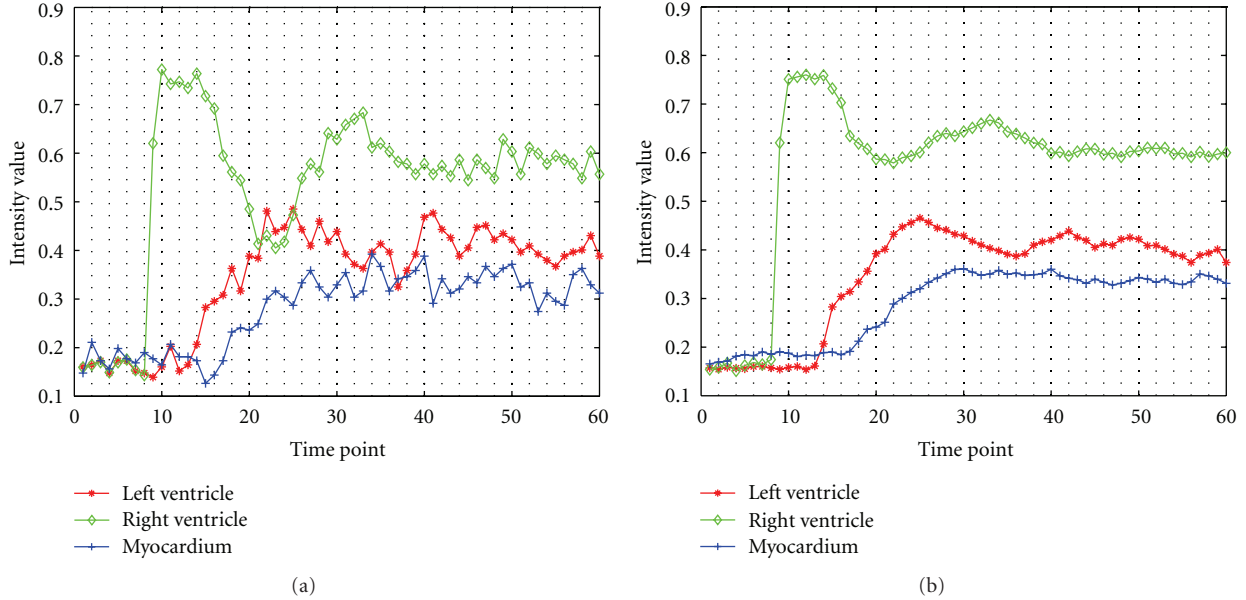
FIGURE 4: Intensity change with time for pixels on epicardium, RV, and endocardium. (a) Before groupwise registration; (b) after groupwise registration.

sequence contributes to a smooth signal, while the sum of second derivatives favours a piecewise linear model of the signal. During the precontrast and postcontrast stages, or for regions without contrast enhancement, the first-order derivative is relevant because we expect the intensity of the same tissue to remain constant (i.e., the first-order derivative of a constant signal is zero). During the wash-in and wash-out stages, we expect the intensity of the same tissue to increase or decrease approximately with a constant rate. In an unregistered image sequence, the derivatives during the wash-in and wash-out stage will alternate between positive and negative values. However, as the image sequence is registered, the intensity changes are gradual and the sum of derivative values is minimal. We also use the second-order derivative of the intensity vector to encourage a piecewise linear intensity signal, as the second derivative of a linear signal is zero. Except for position of peak enhancement, the second derivative is zero at other time points.

Let I_n denote the image at the nth time point (or the nth frame in the dynamic sequence. The intensity of its ith pixel at time t is given by $I_{n,i}(t)$, where $t = 1, \ldots, T$. Thus E_Q is given by

$$E_Q = \sum_{n=1}^{N} \sum_{i} \sum_{t=0}^{T} \left(I'_{n,i}(t) + I''_{n,i}(t) \right) dt. \tag{9}$$

Here, I' and I'' are, respectively, the first and second derivatives of the intensity signal, and N is the total number of pixels in each image. Note that both the first and second derivatives are used for the entire image sequence.

2.3.1. Within-Class Distance (E_W).

The within class distance, E_W, integrates segmentation information into the JSGR framework. E_W ensures that pixels of the same class have similar intensity time profiles. In other words, pixels with the same label are made similar to a representative signal from that class. We use the mean intensity vector for representing each class of pixels. For a pixel i with known label l, its intensity vector provides greater information about the class labels. The within-class distance is calculated as the difference with respect to the mean intensity vector of class $l (\bar{\mathbf{I}}_l)$ and is given by

$$E_W = \sum_{i} ||\mathbf{I}_i - \bar{\mathbf{I}}_l||. \tag{10}$$

Note that if pixel i belongs to class l (as determined from the current labels), then its difference only with respect to the mean intensity vector of class l is calculated; $|| \cdot ||$ denotes the Euclidean distance of the vectors. If a pixel has been correctly labeled as LV blood pool (or RV), then the residual error from the mean of LV blood pool (or RV) will be low. On the other hand, if the labeling is wrong, then the corresponding error is high. Initially, due to many unregistered pixels, the mean intensity vector may not be a very accurate representation of the class. But after every iteration, the mean intensity vector becomes smooth with the update of segmentation labels and starts to truly represent the particular class. In an iterative method, these constraints (E_Q and E_W) ensure that the labels converge correctly. Although the blood pool shows a lot of intensity change, it does not affect our method. By combining the contributions of the registration and segmentation terms, we overcome the effects of intensity change.

Here, we need a representative intensity vector for a particular class. Although the mean vector is not necessarily the best representative vector at the beginning of registration, it is the best choice for a balance between registration accuracy and computational complexity. PCA could be a more

accurate choice for the representative vector but significantly increases computation complexity. Moreover, we observe that with increasing number of iterations, the mean vector does converge to the representative vector of the class. This is supported by the fact that the registration and segmentation accuracies are the same when using the mean vector (after convergence) and principal components. But by using PCA, the computation time increases by more than 1.5 times.

2.4. Optimization. B-splines are used to optimize the cost function in (1). A uniform grid is initialized for all images. The first image's grid coefficients are updated based on the present value of the energy function. The image is transformed, and the segmentation labels are immediately updated. The next image's grid coefficients are then updated followed by image transformation and update of segmentation labels. This is repeated for all images of the sequence corresponding to different time points. This constitutes one iteration for *JSGR*. The segmentation labels are updated after the iteration to be used as the starting point for the next iteration. We repeat the process until the cost function does not decrease further.

The advantage of dynamic update of segmentation labels is that it reflects the value of the energy function of the updated image sequence based on the latest transformations. Thus, if the energy function value does not change while updating the grid coefficients of three consecutive images, then JSGR can be immediately terminated. Once the process has been terminated, the final segmentation labels can be immediately obtained using graph cuts.

In order that the images are registered to a common image space, the average deformation of a pixel is constrained to be zero. This is achieved by making the sum of coefficients for the corresponding grid point over all images to be zero. After the grid of every image is updated, the images are transformed, and the segmentation labels are dynamically updated, thus ensuring that the change in grid coefficients reflects immediately in the registration process. It also leads to more accurate registration than updating the segmentation labels at the end of each iteration. The algorithm is summarized in Algorithm 1.

3. Experiments and Results

Cardiac images were acquired on a 1.5T Siemens Sonata MR scanner following bolus injection of Gd-DTPA contrast agent. There are 12 sequences acquired from 12 patients in whom it was important to look at myocardial perfusion. Each sequence comprised of 60 frames with a total of 720 frames. The datasets were acquired with electrocardiographic (ECG) gating such that the images were acquired during the same phase of the cardiac cycle. This minimized cardiac motion, but some deformations were still observed due to patient breathing. The pixel spacing ranges from (1.5×1.5) to (2.8×2.8) mm^2. The acquired images were from the same midcavity slice. The images were corrected for rotation and translation motion before segmentation. The initial labeling was obtained using graph cuts [5]. The B-spline grid was of

size 10×10 with the spacing between grid control points varying from 7 to 9 pixels.

3.1. B-Spline Grids. An example of B-spline-based illustration is shown in Figure 1. A reference image, floating image, difference images before and after registration, and the deformed grid obtained after registration are shown. The initial grid is uniform and without any deformations. This example has been used for illustration purpose in which a reference image was fixed and all other images were registered using *JRGS*. From the difference images and deformed grid, we can conclude that the deformations of the floating image are captured by B-spline registration.

3.2. Registration Results. We compare the registration error from *JSGR* with [16] (*Met* 1), [28] (*Met* 2), and our pairwise joint registration and segmentation method in [7] (*JRS*). *Met* 1 does not define an explicit reference image and uses the entropy of the pixel stack for the cost function. *Met* 2 uses pairwise registration with normalized mutual information (NMI) as the similarity measure. It has been used before for registering contrast-enhanced breast MRI. *JRS* is a joint registration and segmentation method, but it only uses information from a pair of images and not from the entire image sequence. Our algorithm converged after 6 iterations. The threshold cost difference above which registration continues is set at 0.1. We present qualitative and quantitative results of our method (*JSGR*) in terms of registration error and segmentation accuracy. For *Met* 1, the number of iterations required for convergence was 5, while on an average, *Met* 2 needed 4 iterations for convergence while registering a pair of images. *JRS* is a one-time optimization method and is not iterative. The reported results are after convergence of all algorithms.

The outline of the LV (epicardium and endocardium) and RV is manually identified by expert observers in the original image sequence. These contours are denoted as C_{org}. The transformation of each image is used to map the contours to the registered image space. Let these contours be denoted as C_{trans}. Since it is practically impossible to have ground truth value for elastic registration, we take a different approach to calculate registration error. In a perfusion image sequence, the shape of the heart should be same for ideal registration of all images. To calculate the registration error, we fix one image as reference (usually the first image of the sequence) and calculate the contour distance between the LV (or any object of interest) in all other images and the reference image. At the end of our *JSGR* method, not only has the sequence been registered but the segmentation labels of each frame have also been determined. Thus, we just need to find the contour distance between regions with the same labels. By contour, we refer to the outline of the segmented region. Note that the same procedure as above can be used to calculate the registration error before registration. The contour distance is calculated using the following steps:

(1) let the reference contour for the registered image sequence be denoted as $C_{\text{trans}}^{\text{ref}}$,

```
(1) Determine the initial labeling $L_f$ of the image sequence.
(2) Set $t = 1$ for first registration round.
(3) Initialize uniform B-spline grid for all images.
(4) Calculate the value of energy function using the data cost formulation ((9) and (10)) and the smoothness cost (7).
(5) Update the B-spline grid for the image, transform the image, and update the segmentation labels.
(6) Repeat the above steps for all images of the sequence.
(7) Repeat for all images until convergence.
```

ALGORITHM 1: Joint segmentation and groupwise registration framework.

TABLE 1: Summary of registration and segmentation performance on cardiac perfusion datasets. The values indicate average and standard deviations for all datasets.

| | Registration error (mm) | | | | | Segmentation result | | | |
| | Before | After registration | | | | DM (%) | | | |
	Registration	JSGR	Met 1	Met 2	JRS	JSGR	Met 1	Met 2	JRS
LV epicardium	2.2 ± 1.2	0.7 ± 0.1	1.0 ± 0.2	1.1 ± 0.1	0.8 ± 0.4	93.6 ± 0.7	90.1 ± 0.3	90.6 ± 0.6	92.1 ± 1.1
LV endocardium	2.8 ± 1.0	0.5 ± 0.2	0.9 ± 0.1	0.9 ± 0.2	0.7 ± 0.3	92.8 ± 0.6	88.9 ± 0.1	88.4 ± 0.3	92.1 ± 0.8
LV overall	2.6 ± 1.1	0.5 ± 0.2	0.9 ± 0.3	1.0 ± 0.1	0.7 ± 0.4	93.1 ± 0.5	89.4 ± 0.2	88.8 ± 0.8	92.1 ± 0.9

(2) let the ith point on C_{trans} be denoted by $C_{\text{trans}}(i)$ and let the jth point on $C_{\text{trans}}^{\text{ref}}$ be denoted by $C_{\text{trans}}^{\text{ref}}(j)$,

(3) for every $C_{\text{trans}}(i)$ find the point on $C_{\text{trans}}^{\text{ref}}(j)$ such that the distance between $C_{\text{trans}}(i)$ and $C_{\text{trans}}^{\text{ref}}(j)$ is minimum

$$d(i, j) = \min_j \left\| C_{\text{trans}}(i) - C_{\text{trans}}^{\text{ref}}(j) \right\|, \qquad (11)$$

(4) for each point $C_{\text{trans}}(i)$, the corresponding $d(i, j)$ is calculated. The contour distance (CD) is the average distance and is defined as

$$\text{CD} = \frac{\sum_i d(i, j)}{\sum i}. \qquad (12)$$

The average error measures for 12 datasets are given in Table 1. Lower the registration error better is the method's performance. We observe that JSGR has the lowest registration, while Met 2 has the highest error. Met 2 registers all images to a fixed reference image. Intensity change is common between two images of the dataset. Consequently, registration is prone to error while using a simple NMI-based similarity measure. Groupwise registration has the advantage that information from the entire sequence can be exploited for registration which is particularly important for DCE-MRI. The importance of dynamic information is highlighted by the fact that JRGS shows slightly higher accuracy than JRS which is a pairwise joint registration and segmentation method. JRS's performance is better than Met 1 although Met 1 is a groupwise registration method. This indicates that inclusion of segmentation information improves registration more than the inclusion of dynamic information. Between Met 1 and JSGR, the latter performs better because it combines segmentation and registration information in the cost function. Subsequently, the final registration errors for JSGR are lower than Met 1. In Figure 2(a), we show the error measures for the LV for each of the 12 datasets.

The time taken for registering one full dataset (including 60 images) is 1 hr 33 minutes using JSGR, 58 mins using Met 1, 45 mins using Met 2, and 55 mins using JRS. JSGR and JRS were implemented using MATLAB 7.5 on a PC having a Pentium 4, 3 GhZ processor. Met 1 was implemented by the authors using ITK and thus has low execution time. Met 2 was also implemented in MATLAB for pairwise registration. Since JSGR is a joint segmentation and groupwise registration method, it takes more than twice the time compared to Met 2.

3.3. Segmentation Results. Segmentation accuracy is calculated based on Dice metric (DM) values between manual segmentation and automatic segmentation for different methods. Manual segmentations for each slice are obtained after registration by each method. After registration is complete for each method, the segmentation labels are obtained by applying graph cuts on the intensity vectors. Now, the segmentation labels for corresponding pixels on different slices will be the same. These labels are compared with the manual segmentations to get the segmentation accuracy using DM. For JSGR, the labels are already obtained after the registration process.

The average DM values for the LV over all 12 datasets are shown in Table 1. Figure 2(b) shows the average DM values of the LV for each dataset over all frames of the sequence. The DM values are highest for JSGR thus indicating maximum accuracy amongst the three methods.

Figure 3 shows the segmented contours for LV endocardium overlaid on a representative image of the database. The representative image is chosen such that the blood pool is visible without any ambiguities. The manual segmentations are shown in red, while the automatic segmentations

TABLE 2: Average registration and segmentation accuracy of LV with change in w_1.

w_1	0	0.1	0.2	0.4	1	1.5	2	2.5	3	4	5
Registration error (pix)	1	0.9	0.7	0.6	0.6	0.7	0.7	0.8	0.9	1.2	1.5
DM values (%)	89.3	90.6	91.9	93.8	93.7	93.1	92.9	91	89.9	89.3	88.8

TABLE 3: Average registration and segmentation accuracy of LV with change in w_2.

w_2	0	0.1	0.2	0.4	1	1.5	2	2.5	3	4	5
Registration error (pix)	1.4	1.2	1	0.8	0.6	0.6	0.7	0.75	0.8	1.1	1.4
DM values (%)	88.2	89.5	90.4	91.7	93.8	93.5	92.7	92.1	91.8	89.2	88.3

are shown in green. The first row also shows the initial segmentation in yellow. The initial segmentation is a result of applying graphcut to the unregistered image sequence and using the intensity vector of each pixel. Note that the automatic segmentations are the average contours over all frames of the sequence and the manual segmentations are also the average of the manually drawn contours. Since the segmentation labels are calculated from the intensity vectors, the labels will be the same for corresponding pixels in all frames. The first column shows results for *JSGR*, second column shows results for *Met* 1, the third column shows results for *Met* 2, and the fourth column shows results for *JRS*. Each row shows results for different datasets. Again we observe that *JSGR* shows the best agreement with manual segmentations due to the combination of registration and segmentation information. The other two methods being solely focused on registration perform inadequately for cardiac perfusion images. The accuracy measures highlight the importance of integrating registration and segmentation information. This combination is particularly important when images have low contrast, and also when segmentation information is available to be exploited for registration. Although the *DM* and registration error for *JRGS* and *JRS* are similar, a *t*-test gives $P < 0.032$, thus indicating statistically different results, and hence improvement in results by using dynamic information.

Figure 4 shows the intensity variations with time for pixels on the LV blood pool, RV, and endocardium. Edge pixels are chosen for epicardium and RV, and therefore, the intensity change before registration (Figure 4(a)) is noisy. After registration using *JSGR*, the intensity variation is smoother (Figure 4(b)) and highlights the success of our method for time-varying data.

3.4. Importance of E_Q and E_W. It is important to look at the contribution of E_Q and E_W to the overall registration procedure. E_Q can be termed as the registration energy, while E_W is the segmentation energy. We examine the improvement brought about by E_W to registration accuracy, and also the improvement in segmentation accuracy due to E_Q. We vary w_1 (8) from 0 to 5 (keeping $w_2 = 1$ fixed) and calculate the registration and segmentation accuracy values for the LV (shown in Table 2). It is observed that as w_1 increases from 0, the registration and segmentation performance both improve. In fact, when $w_1 \leq 0.4$, the registration accuracy is less than or comparable to *Met* 1 but improves with greater contribution of E_W. However, if $w_1 > 2$, the registration accuracy starts to degrade (evident from higher registration error), and the *DM* values also decrease due to unbalancing of each terms contribution.

Similarly, when we increase the value of w_2 (with $w_1 = 0.4$) from 0 to 5, we observe low *DM* values when $w_2 < 1$. The best segmentation accuracy is obtained for $1 \leq w_2 \leq 3$. However, when $w_2 > 3$, the *DM* values start to decrease and the registration error also increases. Table 3 shows the change in registration error and *DM* values with change in w_2. The best results are obtained for $w_2 = 1$ and $w_1 = 0.4$.

4. Conclusion

We have proposed a novel method for joint segmentation and groupwise registration of the LV in cardiac perfusion images. By maximizing the smoothness of the temporal intensity signal, our method uses available temporal information from the entire image sequence. This helps to overcome the effects of intensity change. Segmentation information is incorporated by minimizing the error between a pixel's intensity vector and mean intensity vector of the same class. Compared to manual segmentations, our method gives higher segmentation accuracy than other methods. The registration errors for our method are the lowest from amongst the three methods. Our being a *joint segmentation and groupwise registration* approach, both registration and segmentation performance are better than conventional methods. This is because of two factors: (1) exploiting temporal information from DCE-MRI sequence in a *groupwise registration* framework and (2) use of mutually complementary registration and segmentation information, while most other methods solve registration and segmentation separately. Our method has the potential to be used for other data types having time-varying characteristics, and in future, we aim to use our method on other dynamic datasets.

References

[1] Y. Sun, M. P. Jolly, and J. M. F. Moura, "Contrast-invariant registration of cardiac and renal MR perfusion images," in *Proceedings of the 7th International Conference of Medical Image Computing and Computer-Assisted Intervention (MICCAI '04)*, pp. 903–910, September 2004.

[2] D. Mahapatra and Y. Sun, "MRF-based intensity invariant elastic registration of cardiac perfusion images using saliency information," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 4, pp. 991–1000, 2011.

[3] T. Song, V. S. Lee, H. Rusinek, M. Kaur, and A. F. Laine, "Automatic 4-D registration in dynamic mr renography based on over-complete dyadic wavelet and Fourier transforms,' in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '04)*, vol. 3750, pp. 205–213, 2005.

[4] V. S. Lee, H. Rusinek, L. Bokacheva et al., "Renal function measurements from MR renography and a simplified multi-compartmental model," *American Journal of Physiology*, vol. 292, no. 5, pp. F1548–F1559, 2007.

[5] Y. Boykov, V. S. Lee, H. Rusinek, and R. Bansal, "Segmentation of dynamic N-D data sets via graph cuts using markov models," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '01)*, pp. 1058–1066, 2001.

[6] N. Hackstein, J. Heckrodt, and W. S. Rau, "Measurement of single-kidney glomerular filtration rate using a contrast enhanced dynamic gradient-echo sequence and the rutland-patlak plot technique," *Journal of Magnetic Resonance Imaging*, vol. 18, no. 6, pp. 714–725, 2003.

[7] D. Mahapatra and Y. Sun, "Joint registration and segmentation of dynamic cardiac perfusion images using mrfs," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '10)*, pp. 493–501, 2010.

[8] D. Mahapatra and Y. Sun, "Integrating segmentation information for improved mrfbased elastic image registration," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 170–183, 2012.

[9] S. Joshi, B. Davis, M. Jomier, and G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," *NeuroImage*, vol. 23, no. 1, pp. S151–S160, 2004.

[10] D. W. Shattuck, M. Mirza, V. Adisetiyo et al., "Construction of a 3D probabilistic atlas of human cortical structures," *NeuroImage*, vol. 30, no. 3, pp. 1064–1080, 2008.

[11] N. Lepore, C. Brun, X. Pennec et al., "Mean template for tensor based morphometry using deformation tensors," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '07)*, vol. 2, pp. 826–833, 2007.

[12] L. Zöllei, A. Stevens, K. Huber, S. Kakunoori, and B. Fischl, "Improved tractography alignment using combined volumetric and surface registration," *NeuroImage*, vol. 51, no. 1, pp. 206–213, 2010.

[13] J. Yang, D. Shen, C. Davatzikos, and R. Verma, "Diffusion tensor image registration using tensor geometry and orientation features," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '08)*, pp. 905–913, 2008.

[14] A. Barmpoutis and B. Vemuri, "Groupwise registration and atlas reconstruction of 4th order tensor fields using the R+ Riemannian metric," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '09)*, p. 64, 2009.

[15] H. Zhang, B. B. Avants, P. A. Yushkevich et al., "High-dimensional spatial normalization of diffusion tensor images improves the detection of white matter differences: An example study using amyotrophic lateral sclerosis," *IEEE Transactions on Medical Imaging*, vol. 26, no. 11, pp. 1585–1597, 2007.

[16] S. Balci, P. Golland, M. Shenton, and W. Wells, "Free-form b-spline deformation model for groupwise registration," in *Proceedings of the Statistical Registration Workshop (MICCAI '07)*, pp. 23–30, 2007.

[17] E. G. Learned-Miller, "Data driven image models through continuous joint alignment," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 2, pp. 236–250, 2006.

[18] S. Baloch, R. Verma, and C. Davatzikos, "An anatomical equivalence class based joint transformation-residual descriptor for morphological analysis." *Information Processing in Medical Imaging*, vol. 20, pp. 594–606, 2007.

[19] H. Jia, P. T. Yap, G. Wu, Q. Wang, and D. Shen, "Intermediate templates guided groupwise registration of diffusion tensor images," *NeuroImage*, vol. 54, no. 2, pp. 928–939, 2011.

[20] J. Hamm, C. Davatzikos, and R. Verma, "Efficient large deformation registration via geodesics and learned manifold of images," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '09)*, pp. 680–687, 2009.

[21] S. Tang, Y. Fan, G. Wu, M. Kim, and D. Shen, "RABBIT: Rapid alignment of brains by building intermediate templates," *NeuroImage*, vol. 47, no. 4, pp. 1277–1287, 2009.

[22] L. Zhang, C. Chefd'hotel, and G. Bousquet, "Group-wise motion correction of brain perfusion images," in *Proceedings of the 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI '10)*, pp. 832–835, April 2010.

[23] C. T. Metz, S. Klein, M. Schaap, T. van Walsum, and W. J. Niessen, "Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach," *Medical Image Analysis*, vol. 15, no. 2, pp. 238–249, 2011.

[24] A. Yezzi, L. Zöllei, and T. Kapur, "A variational framework for joint segmentation and registration," in *Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA '01)*, pp. 44–51, December 2001.

[25] J. An, Y. Chen, F. Huang, D. Wilson, and E. Geiser, "A variational PDE based level set method for a simultaneous segmentation and non-rigid registration," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '05)*, pp. 286–293, 2005.

[26] K. M. Pohl, J. Fisher, W. E. L. Grimson, R. Kikinis, and W. M. Wells, "A Bayesian model for joint segmentation and registration," *NeuroImage*, vol. 31, no. 1, pp. 228–239, 2006.

[27] G. Wu, H. Jia, Q. Wang, and D. Shen, "Groupwise registration with sharp mean," in *Proceedings of the Medical Image Computing and Computer-Assisted Intervention (MICCAI '10)*, pp. 570–577, 2010.

[28] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, pp. 712–721, 1999.