ABSTRACT

A novel image registration method is presented for multi-modality, gated cardiac imaging. The motion of the myocardium is registered instead of attributes obtained from image intensities, which may be drastically different. Optical flow methods are used to estimate a set of 3D vector fields for both modalities. The 3D vector fields are assumed to be similar and are rigidly aligned by minimizing a sum-of-squares error objective function. Evaluation of the motion-based (MB) method was performed using simulated cardiac SPECT and CT images of a 4D thorax phantom for registration errors of 1 to 3 cm translation, with and without rotation. The MB method was compared to a mutual information (MI) based method. The MB method was able to register the images with an accuracy of 1-5 mm for an anatomical point in the left ventricle. The MI method required a common background distribution within the two modalities for accurate registration.

Index Terms—Image registration, Image motion analysis, Single Photon Computed Tomography, Computed Tomography, Magnetic Resonance Imaging

1. INTRODUCTION

To accurately determine the degree of coronary artery disease and subsequently a suitable treatment plan, both cardiac anatomy and physiology information is required. Physiological information is typically provided through imaging modalities such as single photon emission computed tomography (SPECT) and positron emission tomography (PET) and presents knowledge of myocardial perfusion and viability. Anatomical information corresponding to wall thickening or thinning may be obtained from cardiac computed tomography angiography (CTA), echocardiography, or magnetic resonance (MR) imaging. Proper interpretation of the physiological and anatomical images requires the images to be spatially registered. In the case of multi-modality imaging systems where anatomical and physiological imaging is performed sequentially and in the same spatial reference frame (PET/CT), image registration is straightforward. However, a potential for patient motion between scans exists and leads to registration error. In the absence of a combined modality imaging system (MR/CT), software registration is required. Due to the problems with patient motion and separate imaging systems, the requirement for software registration of multi-modality images is commonly acknowledged [1].

Current methods of cardiac image registration such as geometric attribute registration, voxel similarity registration, and mutual information (MI) registration are based on the voxel intensities [2]. In inter-modality imaging, appropriate landmarks may not be available in both modalities and voxel intensities are often very dissimilar. This leads to difficulty in implementing landmark and similarity registration methods. MI registration is a method used frequently in inter-modality registration [3,4] and is a measure of the statistical dependence between two random variables. For multi-modality imaging, corresponding pixel intensities are the random variables in MI registration. It has been shown that MI is not capable of accurately registering PET and CT images for attenuation correction [5].

The method presented here is based on the similarity of cardiac motion between modalities, not voxel intensities or landmarks based on voxel intensities. This method may be applied to any gated, cardiac imaging modality. The following describes the motion based method mathematically and an investigation using SPECT and CT images from a simulated thorax phantom.

2. MOTION-BASED REGISTRATION

2.1 Objective function

The registration method requires an estimate of the non-rigid motion between sequential frames of a gated cardiac study. The motion estimation algorithm implemented here is a modification of the Horn and Schunck (HS) optical flow algorithm [6]. The HS algorithm has been incorporated into a sequential quadratic approximation framework [7] and generates a set of 3D vector fields.
\( m^n(r) = (u^n(r), v^n(r), w^n(r)) \)  \( (1) \)

where \( n \) is the frame number and \( u, v, w \) are the orthogonal components of each motion vector at 3D coordinate \( r \).

We assume a rigid-body transform of the motion vectors will register the images. A typical technique to rigidly register similar scalar fields is the minimization of the sum-of-squares error (SSE) of the scalar fields. Here, we rigidly transform a vector field and calculate a similar SSE objective function based on the sum-of-magnitudes-squared of the error vector

\[
E(A, b) = \sum_{r} \left| Am_1^n(r) - m_2^n(Ar + b) \right|^2 \quad (2)
\]

where \( A \) is the rotation matrix, and \( b \) is the 3D translation vector. The rotation matrix \( A \) presents three parameters to estimate in addition to the three for \( b \). The rotation of the 3D coordinate vector by matrix \( A \) requires the same rotation, \( A \), of the motion vectors. Figure 1 illustrates in 2D the affect of rotation on a simple vector field.

![Figure 1. 45 degree rotation of 2D vector field requiring the rotation of \( m_1 \) by \( A \) (dashed line, left image) to match the direction of \( m_2 \) (solid line, right image).](image)

3. EXPERIMENTAL METHODS

3.1. 4D NCAT phantom

Evaluation of the registration method was completed using SPECT and CT images reconstructed from simulated 4D thorax phantom projection data [8]. Noise-free SPECT and CT projection data were created with 8 gated frames over the cardiac cycle using a linear interpolation projector. For the SPECT phantom, relative intensities in the organs were: myocardium and liver-1.0, lung-0.3, soft tissue-0.03, and blood pool-0.03. To simulate a clinical Tc-99m sestamibi SPECT study, noise was added to the projection data with a Poisson random number generator. Ten different noise realizations were generated to acquire a statistical sampling of projection measurements. Ten different noise realizations were generated to acquire a statistical sampling of projection measurements. A simultaneous reconstruction and motion estimation algorithm was performed on the SPECT projection data [9]. The CT images were generated to represent a CTA study with relative organ intensities of: soft tissue, myocardium and liver-1.0, lung-0.3, left ventricle and aorta-1.2, spine bone-1.2, and rib bone-1.6. Reconstructions of the SPECT and CT data are shown in figure 2. Known misregistrations were introduced by rigidly transforming the CT images. Table 1 shows six cases of transformation with and without a rotation of five degrees. In addition to the six cases shown, 12 more cases were evaluated for translations of 2 and 3 cm with and without rotation for a total of 18 test cases.

![Figure 2. Transaxial slice of NCAT reconstructions at end-diastole (left) and end-systole (right). Top row: SPECT, Bottom row: CT.](image)

<table>
<thead>
<tr>
<th>Case</th>
<th>x</th>
<th>y</th>
<th>z</th>
<th>rotation (deg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-5</td>
</tr>
<tr>
<td>6</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-5</td>
</tr>
</tbody>
</table>

3.2. Implementation of motion-based method

The smoothing parameter in the motion estimation algorithm was determined empirically (0.001 for SPECT, 0.01 for CT) using a sample case. The objective function given by equation (2) was minimized using Powell’s multidimensional direction set method and Brent’s one-dimensional optimization algorithm for line minimizations [10]. The direction matrix was initialized as a diagonal matrix. For the line minimizations, the translation parameters were evaluated first followed by the rotation parameters. Tri-linear interpolation was used to determine image intensities at locations between voxel centers. The evaluation of the objective function was limited to a region-of-interest (ROI) shown in figure 3. The ROI included areas where motion vectors were depicted in both image
modalities. The minimization algorithm was terminated when the change in objective function from the previous iteration was less than 0.0001%.

Figure 3. Region of interest considered for registration enclosed by black outlines.

3.3. Measure of registration error

In each mis-registration case of ten noise realizations, the average value of each transformation parameter was determined. For each of the 18 mis-registration cases, the average rigid-body transformation parameters were used to determine the error of the estimated registration. The error metric is defined as the Euclidean distance between the anatomical point in each modality after applying the estimated registration (Figure 4).

Figure 4. Location of anatomical reference point used to calculate registration error.

3.4. Implementation of mutual information method

Evaluation of the MI algorithm was also performed for the 18 test cases. The objective function of MI is

$$E_{MI} (A, b) = \sum_{i,j} h_{A,b}(i,j) \cdot \log \left( \frac{h_{A,b}(i,j)}{h_1(i) \cdot h_2(j)} \right) \quad (3)$$

where $h_{A,b}(i,j)$ is the 2D joint probability density function (pdf) for images 1 and 2, and $h_1(i)$ and $h_2(j)$ are the marginal pdf’s for modalities 1 and 2, respectively. The 2D joint pdf is the joint histogram for images 1 and 2 and is dependent on the rotation and translation parameters. The pdf’s $h_1(i)$ and $h_2(j)$ are represented by the respective image histogram linearly rescaled from 1 to 256 [11]. In the case of large MI, pixels of a given intensity in one image share a narrow range of intensities in the other image. MI is assumed to be large when the images are registered correctly. Therefore, the MI objective function is maximized for the appropriate registration. The maximization of the MI objective function was implemented similarly to the motion-based method using Powell’s multi-dimensional direction set method and Brent’s one-dimensional optimization algorithm for line minimizations [10].

4. RESULTS

For the motion-based method, the average error in the individual components of the translation was 3.0 mm for cases 1-6, 3.1 mm for cases 7-12, and 3.8 mm for cases 13-18. The average error (deg.) in the individual rotation angles was 1.9 for cases 1-6, 2.1 for cases 7-12, and 2.4 for cases 13-18. Figure 5 shows plots of the Euclidean error metric for all 18 cases of mis-registration. The average cpu time per registration on a modern workstation (Penguin Computing Tempest 2100) with a 2.4 GHz processor was 51.8 seconds.

Figure 5. Motion-based registration error as a function of mis-registration distance of the reference point.

The results with the MI method are given in table 2 for Cases 1-6. The top half of the table gives the results using the same SPECT reconstructions that were used for the motion-based methods. For these cases, the accuracy of the registration was excellent and superior to the motion-based method. We suspected the performance of the MI method was supported by the background distribution in the SPECT and CT images. A second set of SPECT reconstructions was obtained from a phantom that had intensity only in the myocardium and was used to test the effect of the background distribution on the MI performance (Figure 6). The bottom half of table 2 gives the MI results using the SPECT reconstructions without background. Compared to the results with SPECT
background intensity, the registration accuracy is worse. Also, there was difficulty in these cases in estimating the rotational component of the mis-registration. Additionally, we found that with larger translational mis-registrations (Cases 7-18), there were many instances in which the MI method failed to produce meaningful results (e.g., translation values greater than the matrix size).

**Figure 6.** (Top left) CT phantom, (Top right) SPECT phantom with display windowing to show background intensities, (Bottom left) SPECT reconstruction with display windowing, (Bottom right) SPECT reconstruction (no background) with display windowing.

**Table 2.**

<table>
<thead>
<tr>
<th>Case</th>
<th>Estimated translation (cm)</th>
<th>Estimated rotation (deg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.00, 1.00, 0.99</td>
<td>0.00, 0.00, 0.00</td>
</tr>
<tr>
<td>2</td>
<td>-0.99, -1.00, -0.97</td>
<td>0.00, 0.43, 0.00</td>
</tr>
<tr>
<td>3</td>
<td>1.16, 0.88, 1.15</td>
<td>4.80, 0.06, 1.91</td>
</tr>
<tr>
<td>4</td>
<td>-1.10, -0.90, -1.00</td>
<td>5.18, 0.01, 0.03</td>
</tr>
<tr>
<td>5</td>
<td>0.97, 1.02, 1.04</td>
<td>-5.06, 0.15, 0.53</td>
</tr>
<tr>
<td>6</td>
<td>-0.88, -1.08, -0.95</td>
<td>-5.01, -0.01, 0.59</td>
</tr>
</tbody>
</table>

### 5. SUMMARY AND DISCUSSION

The method presented has the advantages of not requiring a dependence between the voxel intensities of the two images nor the extraction of surfaces from the two images. The only requirement is the two images depict a common pattern of motion. In evaluation studies using images of simulated, gated cardiac CT and SPECT phantoms with known mis-registration, the motion-based registration method was able to register the images to an accuracy of 1-5 mm for an anatomical point in the vicinity of the myocardium.

With the simulated data in this study, the MI method performed reasonably well except for cases with relatively large displacements. Background structures that are similarly depicted in both images provide critical input to this method. It may be possible to effectively combine the MI and motion-based methods for improved accuracy by using the output from one method as the input to the other. In the future, the effect of variations in motion pattern on the registration accuracy should be studied.

### 6. REFERENCES


