Spectral Embedding-Based Registration (SERg) for Multimodal Fusion of Prostate Histology and MRI

Eileen Hwuang\textsuperscript{a,†}, Mirabela Rusu\textsuperscript{b}, Sudha Karthikeyan\textsuperscript{c}, Shannon C. Agner\textsuperscript{d}, Rachel Sparks\textsuperscript{e}, Natalie Shih\textsuperscript{f}, John E. Tomaszewski\textsuperscript{g}, Mark Rosen\textsuperscript{h}, Michael Feldman\textsuperscript{f}, and Anant Madabhushi\textsuperscript{b,‡}

\textsuperscript{a}Department of Biomedical Engineering, Rutgers University, \textsuperscript{b}Department of Biomedical Engineering, Case Western Reserve University, \textsuperscript{c}Department of Biomedical Engineering, Duke University, \textsuperscript{d}Department of Neurology, Washington University School of Medicine in St. Louis, \textsuperscript{e}Department of Medical Physics and Bioengineering, University College of London, \textsuperscript{f}Department of Pathology, University of Pennsylvania, \textsuperscript{g}Department of Pathology & Anatomical Sciences, University of Buffalo, \textsuperscript{h}Department of Radiology, University of Pennsylvania

ABSTRACT

Multi-modal image registration is needed to align medical images collected from different protocols or imaging sources, thereby allowing the mapping of complementary information between images. One challenge of multimodal image registration is that typical similarity measures rely on statistical correlations between image intensities to determine anatomical alignment. The use of alternate image representations could allow for mapping of intensities into a space or representation such that the multimodal images appear more similar, thus facilitating their co-registration. In this work, we present a spectral embedding based registration (SERg) method that uses non-linearly embedded representations obtained from independent components of statistical texture maps of the original images to facilitate multimodal image registration. Our methodology comprises the following main steps: 1) image-derived textural representation of the original images, 2) dimensionality reduction using independent component analysis (ICA), 3) spectral embedding to generate the alternate representations, and 4) image registration. The rationale behind our approach is that SERg yields embedded representations that can allow for very different looking images to appear more similar, thereby facilitating improved co-registration. Statistical texture features are derived from the image intensities and then reduced to a smaller set by using independent component analysis to remove redundant information. Spectral embedding generates a new representation by eigendecomposition from which only the most important eigenvectors are selected. This helps to accentuate areas of salience based on modality-invariant structural information and therefore better identifies corresponding regions in both the template and target images. The spirit behind SERg is that image registration driven by these areas of salience and correspondence should improve alignment accuracy. In this work, SERg is implemented using Demons to allow the algorithm to more effectively register multimodal images. SERg is also tested within the free-form deformation framework driven by mutual information. Nine pairs of synthetic T1-weighted to T2-weighted brain MRI were registered under the following conditions: five levels of noise (0\%, 1\%, 3\%, 5\%, and 7\%) and two levels of bias field (20\% and 40\%) each with and without noise. We demonstrate that across all of these conditions, SERg yields a mean squared error that is 81.51\% lower than that of Demons driven by MRI intensity alone. We also spatially align twenty-six \textit{ex vivo} histology sections and \textit{in vivo} prostate MRI in order to map the spatial extent of prostate cancer onto corresponding radiologic imaging. SERg performs better than intensity registration by decreasing the root mean squared distance of annotated landmarks in the prostate gland via both Demons algorithm and mutual information-driven free-form deformation. In both synthetic and clinical experiments, the observed improvement in alignment of the template and target images suggest the utility of parametric eigenvector representations and hence SERg for multimodal image registration.

1. INTRODUCTION

Multi-modal medical imaging data helps provide complementary information for improved disease diagnosis, prognosis, treatment guidance, and evaluation. Multimodal imaging modalities can identify different types of
disease cues and also could allow for mapping and enriching disease specific information from one modality to another. An example of this is co-registration of pre-operative in vivo imaging with ex vivo surgically excised histopathology in order to map the true, ground truth extent of cancer from the pathology onto the imaging. This could enable the precise identification and hence enrichment of in vivo imaging modality to detect disease. Challenges arise in identifying corresponding regions for quantitative comparison between different modalities because of the different types of image content and accompanying acquisition artifacts. Multimodal image registration could have significant impact on the accuracy of subsequent comparative assessments involved in diagnosis and treatment evaluation.

Several co-registration algorithms rely on optimization of a similarity measure to increase statistical correlations in image appearance and therefore require the corresponding image intensities to be similar. Measures that utilize intensity differences such as that found in Thirion’s Demons algorithm are unsuitable for multimodal registration owing to the substantial differences in imaging appearances of the modalities. Previous studies have attempted to address this flaw in Demons algorithm by using a different similarity measure, mutual information (MI). Compared to computing the squared difference in pixel intensity between two images, MI can be more favorable in multimodal image registration because it measures entropy, the distribution of intensities in the histogram of the image. Maximizing MI can align two images by maximizing the individual entropies of each of the template and target images while minimizing their joint entropy in search of one-to-one correspondences between pixels. Kroon, et al. attempted to rectify the modality inhomogeneity problem in Demons algorithm by using local MI, an intensity histogram-based similarity measure. In their study, the T1-weighted (T1-w) was converted to a T2-weighted (T2-w) MRI before they aligned the two T2-w modalities. Instead of altering the input images, other approaches replaced the intensity difference-based similarity measure in Demons algorithm with a point-wise MI metric based on the histogram peaks. These approaches address the sensitivity of intensity-based similarity measures on different image modalities that can lead to inadequate registration. To mitigate this problem, these interventions involve either ensuring that the images are in same modality or using an non-intensity difference-based similarity measure instead.

While MRI has been shown to have diagnostic utility in identifying prostate cancer in vivo, it is still not specific enough to be able to distinguish aggressive from indolent prostate cancer in vivo. Definitive validation for presence and extent of prostate cancer is reliant on histopathology. In patients undergoing a radical prostatectomy, co-registering ex vivo pathology with pre-operative staging MRI can allow for mapping of disease extent and disease severity onto the in vivo imaging. This “ground truth” mapping of disease extent onto the corresponding in vivo imaging can lead to learning of the imaging signatures associated with more and less aggressive disease, allowing for development of computer-assisted detection algorithms for prostate cancer. The multimodal prostate histology-MRI registration task is challenging because of the large differences in appearance between histology and MRI. During MRI acquisition, artifacts such as endorectal coil-induced deformations are known to distort the natural shape of the prostate. Following radical prostatectomy, tissue loss and gland shrinkage during histology specimen preparation are also induced.

Consequently one of the strategies to improving multimodal image alignment is to find alternative image representations. In this work we present a new multimodal co-registration method called spectral embedding-based registration (SERg) which generates and employs alternate representations to make the images more similar to each other compared to the original representations. The spirit behind our approach is that the alternate representations obtained via application of spectral embedding (SE) on independent components of textural representations can not only distinguish salient from non-salient areas but also identify corresponding regions that are more similar between images. The registration is then driven by these areas of correspondence and salience. The SERg framework is demonstrated for multimodal registration, particularly focusing on prostate histology and MRI alignment by Demons algorithm and free-form deformation (FFD) driven by MI.

The remainder of the paper is organized in the following manner. Section 2 describes previous work in registration based on alternate representations of the images to be aligned, along with a summary of the novel contributions of this work. Section 3 details the methodology comprising SERg. Section 4 showcases the results of registration experiments performed on multimodal synthetic brain images and corresponding prostate histology and MRI slices. Section 5 describes the experimental results. Concluding remarks are presented in Section 6.
2. PREVIOUS WORK AND NOVEL CONTRIBUTIONS

Alternate data representations have been previously explored to improve registration accuracy and efficiency. This approach is based on the idea that images from different modalities can be more similar in the transformed feature space than in the original space. One way of performing this transformation is by local scale, which is a method of automatically detecting meaningful local structures in an image from a multi-scale representation without a priori information about what image structures are expected. Local scale methods seek to detect image structures of interest by defining a quantitative parameterization, such as multiplying the spatial derivative of any scale with a scale parameter and selecting the scales that have local maxima of the derivative. Another example is extraction of local scale determined by the inverse of the local spatial frequency after wavelet decomposition of a textured plane. Later variations of local scale known as locally adaptive scale allow the size of the scale to alter spatially such that regions containing fine details have smaller locally adaptive scales than a large homogeneous region. Local scale at each location is dependent on a spatial distance function and a homogeneity distance function to find locally connected homogeneous regions based on the presence of edges along object boundaries. Largely homogeneous regions can be processed at a low resolution whereas heterogeneous regions can undergo computations at relatively higher resolutions. Locally adaptive scale has been found to be amenable to various image processing tasks including bias field correction, image segmentation, and image registration. For example, local scale-based data transformation has also been explored in image segmentation because they can be used to generate an alternative yield representation that accentuates object boundaries. Moreover, the local scale features common to both images lead to representations that are more similar to each other than their original forms. Nyul, et al. utilized ball-scale, a locally adaptive scale notion by finding the ball with the largest radius to parameterize each pixel location according to a predefined homogeneity criterion. Transformed image representations were shown to improve multiprotocol image registration with cross-correlation, normalized cross-correlation, and MI based similarity measures.

Other researchers have attempted to use all of the available image data to facilitate co-registration. For instance, Chappelow, et al. presented co-registration of prostate histology and multiprotocol MRI by concurrently leveraging all available image data including T2-w, dynamic contrast enhanced (DCE), and apparent diffusion coefficient (ADC) MRI. Each of these protocols provide orthogonal information such as structural (T2-w), functional (DCE), and physical/physiological (ADC) attributes. These images can be registered in a groupwise manner using multiattribute combined mutual information (MACMI), a higher order version of MI, as a similarity measure. This approach was shown to improve alignment of important structures such as the cancerous lesion and urethra compared with traditional pairwise registration of histology to T2-w MRI and histology to DCE MRI.

Another approach to utilizing multiple image data simultaneously in co-registration has been using a computer-generated set of low level texture features which were coallesced via combined feature ensemble mutual information (COFEMI). This has been shown to improve alignment of very different looking images. Since prostate histology and MRI do not share enough structural and functional information and suffer from tissue deformation during acquisition, a major challenge in co-registration performance is hypersensitivity to multimodal image intensity differences. To mitigate this problem, in Chappelow, et al., texture features, which highlight the spatial arrangement of intensities, were extracted. These texture features included gradient, first order statistical, and second order statistical features which were identified to be robust to artifacts and noise, thereby complementing the original image intensities with structural information. The results from COFEMI demonstrated significant improvement over MI-driven B-spline registration in terms of alignment accuracy measured by L2 distance and MI. Kirsch, Sobel, Gaussian, variance, and median texture features were also found to be useful for identifying prostate boundaries when building active shape models for improving prostate segmentation in T2-w MRI. These studies suggest that using low level texture feature representations of images may play a role in improving the performance of image processing tasks such as segmentation and co-registration.

However, using multiple features may introduce redundancy. Additionally, a large number of features can greatly increase computational complexity. To mitigate these two problems, dimensionality reduction (DR) methods have been previously used to reduce features into a lower dimensional space while also preserving salient information. Linear DR methods including independent component analysis (ICA) and principal component analysis (PCA) have been previously applied in the context of co-registration. In particular, ICA is a...
method of obtaining low dimensional projections of multichannel data by minimizing mutual information and maximizing non-Gaussianity to find orientations of the data that provide the most salient information. Li, et al. demonstrated the usage of ICA on multichannel image registration to fuse multimodal information.\textsuperscript{9} Gabor features in various scales and orientations were extracted from multimodal images, but each channel contained a mixture of various tissue types in various distributions. ICA was then employed to transform the multifeature data into independent component images that each preferentially captures a specific tissue type. It was suggested that ICA repartitions the data from individual feature channels into independent components, each of which is a collection of information relevant to a particular tissue type gathered from all the feature channels.\textsuperscript{9} Using ICA to obtain image representations of multimodal images preserves complementary information from different modalities and removes redundancy such that structural detail independent of intensities of a particular modality may be captured.

Linear DR methods typically encounter difficulties in capturing the nonlinear relationships inherent in biological data such as bias field and noise.\textsuperscript{18} SE, a nonlinear DR method, has the ability to take such nonlinearities into account by emphasizing areas of saliency that are robust to complex correspondences between images.\textsuperscript{17} SE maps high-dimensional data to a two- or three-dimensional space and is based on the normalized cut criterion for graph clustering.\textsuperscript{19} The technique takes the dissimilarity between two images represented in the form of a weighted affinity matrix and performs eigenvalue decomposition, after which the eigenvectors that correspond to the minimum eigenvalues compose the repartitioned feature space. Agner, et al.\textsuperscript{17} applies SE to improve segmentation based on the premise that SE better highlights strong boundary gradients than PCA, a linear DR method. By being able to preserve global and local similarities when decomposing multiple features into a few eigenvectors, SE may present images in a form that better describes the relationship between a pixel and its neighbors. These region statistics allow for simplified image processing tasks including improved edge detection. Previous work has shown that SE-generated alternative data representations are advantageous for segmenting breast lesions in dynamic contrast enhanced MRI.\textsuperscript{17} Now the driving hypothesis for our paper is that the region statistics resulting from SE can also help emphasize structural correspondences between images that can be captured by intensity-based similarity measures for an accurate and robust registration.

In this work, we generate alternative representations of multimodal data that are as similar as possible by combining texture feature extraction, ICA, and SE into a single framework for improving multimodal registration accuracy. Statistical texture features are extracted from the original images since it has been demonstrated that they can be robust to noise and artifacts.\textsuperscript{15} This produces a large collection of features with varying degrees of saliency in terms of image content and may also contain redundancies. Therefore, instead of redefining the MI similarity measure, i.e. COFEMI,\textsuperscript{15} to be used directly on the texture features for co-registration, we employ ICA to automatically select a subset of independent feature combinations that contain less repeated information. While Li, et al.\textsuperscript{9} used ICA on Gabor features to enhance tissue characterization for multimodal image fusion, a customized high-order similarity measure was defined to handle the large set of independent components during image registration. In this paper, we show that performing SE on the ICA components further distills the saliency information into three parametric eigenvector images to serve as alternate representations of the original images. The low-dimensionality and strong similarity of these alternate representations would allow the use of traditional intensity-based similarity measures for improved multimodal registration.

### 3. METHODOLOGY

#### 3.1 Terminology

The goal of SERg framework is to align a target image \( A \) to a template image \( B \). \( A = (C, f) \) in which \( C \) is a 2D grid of spatial locations and \( f(c) \) is the intensity value at location \( c \in C \). The number of pixels is \( |C| \), where \( | \cdot | \) denotes the cardinality of \( C \). The template image \( A \) and target image \( B \) are concatenated and processed as one image for Modules 1 to 3. \( D \) and \( D' \) are sizes of feature vectors denoted by \( | \cdot | \) after texture feature extraction and ICA, respectively. \( d \) is the number of eigenvectors after SE. Feature extraction produces the feature set \( F \), each location of which is referred to as \( F(c) \). Similarly, ICA produces the feature set \( G \) with \( G(c) \) referring to a particular location \( c \). SE takes the weight matrix \( W \) and computes the parametric eigenvector set \( V \). Table 1 summarizes the symbols used in the methodology.
3.2 Brief Overview of SERg

The registration framework (Figure 1) is composed of four modules:

*Module 1) Texture Feature Extraction:* We compute a set of \( D \) low level textural feature representations \( F \) from the concatenation of two multimodal images to capture structural detail. Figure 2 depicts some examples of texture features.

*Module 2) ICA on Texture Features:* We perform ICA to repartition the features such that redundant information is removed while complementary information is preserved. This results in \( D' \) independent components in set \( G \). Examples of the independent component images is presented in Figure 3.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( A )</td>
<td>Target image scene.</td>
<td>(</td>
<td>\cdot</td>
</tr>
<tr>
<td>( B )</td>
<td>Template image scene.</td>
<td>(</td>
<td>\cdot</td>
</tr>
<tr>
<td>( C )</td>
<td>2D grid of spatial locations.</td>
<td>( c )</td>
<td>Single pixel in ( C ).</td>
</tr>
<tr>
<td>( f(c) )</td>
<td>Intensity value at location ( c \in C ).</td>
<td>( v )</td>
<td>Eigenvector at a location ( c ).</td>
</tr>
<tr>
<td>( D )</td>
<td>Number of features.</td>
<td>( V )</td>
<td>Eigenvector set of all ( v ).</td>
</tr>
<tr>
<td>( D' )</td>
<td>Reduced number of features after ICA.</td>
<td>( W )</td>
<td>Weight matrix.</td>
</tr>
<tr>
<td>( d )</td>
<td>Reduced number of features after SE.</td>
<td>( w )</td>
<td>Element of weight matrix ( W ).</td>
</tr>
<tr>
<td>( F(c) )</td>
<td>Set of ( D ) features specified at location ( c ).</td>
<td>( r, s )</td>
<td>Coordinate locations in ( W ).</td>
</tr>
<tr>
<td>( G(c) )</td>
<td>Set of ( D' ) features specified at location ( c ).</td>
<td>( d(r) )</td>
<td>Sum of weight elements at location ( r ).</td>
</tr>
<tr>
<td>( G )</td>
<td>Set of ( D' ) feature images.</td>
<td>( \mathbb{R}^{R \times C} )</td>
<td>Dimension size.</td>
</tr>
<tr>
<td>( \Omega )</td>
<td>Specific region of interest ( A ) and ( B )</td>
<td>( i )</td>
<td>Intensity value or bin range.</td>
</tr>
<tr>
<td>( \omega )</td>
<td>Specific grid location in ( \Omega )</td>
<td>( p )</td>
<td>Number of pixels per bin range.</td>
</tr>
<tr>
<td>( j )</td>
<td>Landmark index</td>
<td>( k, l )</td>
<td>Pixel index of landmark ( j ) in ( A, B )</td>
</tr>
</tbody>
</table>

Table 1. Description of notation used in Section 3 and 4.3.
Figure 2. Examples of texture features extracted from concatenated template and target images are identified as follows: (a) mean, (b) median, and (c) standard deviation.

Module 3) SE on Independent Components: We perform SE on the reduced features resulting in a parametric eigenvector (PE) image representation with $d$ eigenvectors.

Module 4) Image Registration of Alternate Representations: We register the PE representations and apply the deformation field to the original images.

The next section describes each module in detail.

3.3 Detailed Explanation of Individual Modules

3.3.1 Module 1: Texture Features

Let $F(c) = [f_1(c), f_2(c), ..., f_D(c)]$ where $f_1, f_2, ..., f_D$ compose a set of $D$ feature images, the original image, formed by concatenating $A$ and $B$, and $D-1$ computer-extracted grey level texture features. Our features include Kirsch, Sobel, diagonal, and gradient filters in various orientations. In addition, first order statistical features, including mean, median, standard deviation, and range, were calculated with each of three window sizes: 5, 15, and 35. Figure 2 shows a few examples of the extracted texture features for the concatenated T1-w and T2-w synthetic brain images.

3.3.2 Module 2: Independent Component Analysis

This paper uses the FastICA implementation first presented in Hyvarinen, et al. to perform ICA. This method iteratively maximizes non-Gaussianity when identifying the top statistically independent components. $A$ and $B$ have already been concatenated, so texture feature extraction and ICA are performed on both images at once, ensuring that the independent components are linked to both images simultaneously. As a result, the independent component images in $G$ are able to highlight salient structures by displaying distinct complementary information with minimal redundancy (Figure 3). Dimensionality reduction from $D$ to $D'$ with $D' < D$ via ICA also improves the efficiency of the spectral embedding in Module 3.

3.3.3 Module 3: Spectral Embedding

The aim of SE is to project $G^A(c)$ from $D'$ to $d$, where $d < D'$. This creates a non-linear embedding from which only the first $d$ most important features are employed further. SE allows the content of the final $d$ features to be optimized by finding a mix of all the $D'$ features. Let $G^A$ be the entire set of $D'$ features such that $G^A = G^A(c), \forall c \in C$, where $G \in \mathbb{R}^{|C| \times D'}$. Projecting $G \in \mathbb{R}^{|C| \times D'}$ to a low $d$-dimensional space yields $V^A = [v_1, v_2, ..., v_{|C|}] \in \mathbb{R}^{|C| \times d}$. The associated eigenvectors are $v_i = [v_1, ..., v_d]$ where $i \in 1, ..., |C|$. The optimal $V^A$ can be obtained by solving,

$$V^A = \text{argmin} \left( \sum_r \sum_s \| v(r) - v(s) \|^2 w_{rs} \right) \sum_r v^2(r) d(r),$$

where $w_{rs}$ is the $(r, s)$ element of the weight matrix $W = [w_{rs}] \in \mathbb{R}^{|C| \times |C|}$ and $r, s \in C$, which assigns edge weights to characterize similarities between the pairwise observations $r$ and $s$ and $d(r) = \sum_s w_{rs}$. The graph
Figure 3. Examples of the independent component images after performing ICA on the texture features including those presented in Figure 2. (a) and (b) correspond to the top two components and (c) correspond to one of the last components ranked by the FastICA algorithm. The distinct structural detail captured by each independent component image reflects the statistical indepedencies of the identified projections. Figure (a) appears to highlight edges between gray and white matter. Figure (b) identifies interior homogeneous regions. However, Figure (c), which was ranked lower, appears to be more noisy and therefore less able to clearly capture structural detail when compared with the highest-ranked independent components ((a) and (b)).

In this work, Thirion’s Demons registration scheme and MI-driven FFD are used to evaluate the utility of the PEs derived from $V^A$ and $V^B$. Demons registration is driven by the similarity measure

$$D(c) = \frac{\langle V^A(c) - V^B(c) \rangle \langle \nabla V^A(c) \rangle}{\langle \nabla V^A(c) \rangle^2 + \langle (V^A(c) - V^B(c))^2 \rangle},$$

which is based on intensity difference $V^A(c) - V^B(c)$ and the gradient $\nabla V^A(c)$. The resulting $D(c)$ is the calculated deformation field vector at location $c$. We use $\langle \cdot \rangle$ to denote the arithmetic mean computed across the three PE image channels for each term in Equation 3. For example, $\langle V^A(c) - V^B(c) \rangle = \frac{1}{d} \sum_{i} v_i^A(c) - v_i^B(c)$.

MI is the similarity measure used in the FFD registration. Currently the PE images are converted to grayscale images from the CIE 1976 color space. This reduces the eigenvectors into a single channel to employ MI for deformable FFD registration. After the PEs have been registered, the deformation field is applied to the original target image to generate the transformed target image that aligns to the original template image.

4. DATASET DESCRIPTION AND EXPERIMENTAL DESIGN

4.1 Dataset Description

4.1.1 Synthetic BrainWeb Images

We present registration of nine pairs of $181 \times 217$ T1-w and T2-w MRI slices (Figures 1 extracted from the MNI BrainWeb simulated brain database under various conditions: noise (0%, 1%, 3%, 5%, and 7%), bias field (20% and 40%), and a combination of both noise and bias field (7% noise with 20% and 40% bias field). A synthetic local shrinkage deformation using B-splines is induced on the T1-w image for all cases. The undeformed T2-w image is the reference template image and the T1-w image is the target image. Table 3 summarizes the various conditions used in each pair of BrainWeb images for the synthetic registration experiment.
### Synthetic Experiment

**Data**
- 9 pairs of T1 and T2 image with varied noise and bias field (Table 3).

**Experiment**
- Demons Intensity Registration, Demons SERg

**Quantitative Evaluation**
- CR, MI, MSE

### Clinical Experiment

**Data**
- 26 pairs of 2D prostate histology and MRI slices from 4 patients

**Experiment**
- Demons Intensity Registration, Demons SERg, FFD Intensity Registration, FFD SERg

**Quantitative Evaluation**
- CR, MI, DSC, RMSD (mm)

#### Table 2. Summary of experimental protocol for synthetic T1-w and T2-w BrainWeb and clinical prostate histology and MRI data.

#### Table 3. List of various noise and bias field conditions for synthetic T1-w and T2-w BrainWeb co-registration experiments.

### 4.1 Clinical Prostate Histology and MRI

Our validation cohort included *ex vivo* prostate histology and 3 Tesla T2-w axial MRI for twenty-six corresponding slices collected from four patients. Histology specimens were acquired following radical prostatectomy and sliced at 3 mm intervals. Each slice was stained with Hematoxylin and Eosin to highlight internal structures such as nuclei, cytoskeleton, and extracellular matrix. 2D slice correspondences and landmarks between histology and MRI were identified by an expert pathologist. The histology images were originally collected in quadrants and subsequently stitched into pseudo whole mount slices via Histostitcher® software. The MRI is the reference template image and the histology is the target image.

### 4.2 Experimental Description

#### 4.2.1 Applying SERg to Co-Registration of Synthetic T1-w and T2-w BrainWeb Images

The goal of the synthetic registration task is to recover the synthetic shrinkage deformation. The ideal registration should introduce a slight deformable expansion in that region to revert to its original configuration in the undeformed T1-w image. The experimental protocol for co-registration of T1-w and T2-w brain images is outlined as follows:

1. Each pair of T1-w and T2-w brain images are concatenated and features are extracted and transformed according to Modules 1-3 (Sections 3.3.1-3.3.3). The resulting PE images are then separated to prepare for co-registration.

2. In Module 4 (Section 3.3.4), the induced deformation in the T1-w brain image is recovered using Demons registration in the following configurations:
   - (a) **Demons Intensity Registration**: Demons co-registration of the original undeformed T2-w template and deformed T1-w target images.
   - (b) **Demons SERg**: Demons co-registration of the PE images corresponding with the undeformed T2-w template and deformed T1-w target images. The computed deformation field is applied onto the original T1-w image to obtain the final registration result.

3. For registration evaluation of the two methods, Demons Intensity Registration and Demons SERg, the registered T1-w image is compared against the ground truth undeformed T1-w image to assess accuracy of deformation recovery and structural alignment.
4.2.2 Applying SERg to Co-registration of Clinical Prostate Histology and MRI

The goal of the clinical registration task is to obtain an accurate alignment of prostate histology and MRI in terms of boundaries and internal structures in the prostate. The experimental protocol for co-registration of prostate histology and MRI is outlined as follows:

1. During preprocessing of the MRI slices, the prostate regions are manually segmented using 3D Slicer\textsuperscript{22} on T2-w MRI.

2. During preprocessing of the histology slices, the prostate images are converted to grayscale, padded, and downsampled to about 10 \(\mu\)m/pixel resolution to match the aspect ratio and resolution of the reference MRI.

3. The histology image is reoriented through rotation and anisotropic scaling along both axes via MI-driven affine registration to match the MRI.

4. Each pair of preprocessed MRI and histology slices are concatenated and features are extracted and transformed according to Modules 1-3 (Sections 3.3.1-3.3.3). The resulting PE images are then separated to prepare for co-registration.

5. In Module 4 (Section 3.3.4), the prostate histology is deformably registered to the MRI using Demons and MI-driven FFD registration\textsuperscript{1} in the following configurations:
   (a) **Demons Intensity Registration**: Demons co-registration of the original T2-w MRI template and histology target images resulting from Step 3.
   (b) **Demons SERg**: Demons co-registration of the PE images corresponding with the T2-w MRI template and histology target images. The computed deformation field is applied onto the original histology to obtain the final registration result.
   (c) **FFD Intensity Registration**: FFD co-registration of the original T2-w MRI template and histology target images resulting from Step 3.
   (d) **FFD SERg**: FFD co-registration of the PE images corresponding with the T2-w MRI template and histology target images. The computed deformation field is applied onto the original histology to obtain the final registration result.

6. For registration evaluation of the four methods, Demons Intensity Registration, Demons SERg, FFD Intensity Registration, and FFD SERg, the registered histology image is compared against the original T2-w MRI to assess accuracy of prostate border and internal landmark alignment.

4.3 Registration Evaluation

4.3.1 Validating SERg in Synthetic Data

Accuracy of Demons registration is assessed quantitatively by computing correlation ratio (CR), mutual information (MI), and mean squared error (MSE) between the registered T1-w image and the undeformed T1-w image. CR and MI measure the similarity between two images based on intensity distributions. The definition of CR came from Roche, et al.\textsuperscript{23}

\[
CR(A, B) = 1 - \frac{1}{N\sigma^2} \sum_i N_i \sigma^2_i, \tag{4}
\]

in which \(N = |\Omega|\) given that \(\Omega\) is the overlapping region between images \(A\) and \(B\). Also for each intensity value \(i\) in \(\Omega\), the overlapping region is defined such that \(\Omega_i = \{\omega \in \Omega, A(\omega) = i\}\) and \(N_i = |\Omega_i|\). This means that

\[
\sigma^2 = \frac{1}{N} \sum_{\omega \in \Omega} \left( B(\omega)^2 - \left( \frac{1}{N} \sum_{\omega \in \Omega} B(\omega) \right)^2 \right) \tag{5}
\]
and similarly
\[ \sigma_i^2 = \frac{1}{N_i} \sum_{\omega \in \Omega_i} \left( B(\omega)^2 - \left( \frac{1}{N_i} \sum_{\omega \in \Omega_i} B(\omega) \right)^2 \right). \] (6)

The formula for MI used in this work is based on the classic definition of Shannon entropy as described in Pluim, et al.:\(^{24}\)
\[ MI(A, B) = S(A) + S(B) - S(A, B), \] (7)
in which \( S(A) = \sum_i p_i \log \frac{1}{p_i} \) and \( p \) is the number of pixels per bin range \( i \) and is similarly defined for target image \( B \).

The registered T1-w images resulting from Demons intensity registration and SERg are also compared visually to evaluate success in recovering the synthetic deformation by aligning with the template T2-w image. The undeformed target T1-w image serves as ground truth of perfect registration and is compared with the deformed and registered T1-w image. MSE validates alignment success based on raw intensity values at each pixel:
\[ MSE(A, B) = \frac{1}{N} \sum_{c \in C} ((B(c) - A(c))^2). \] (8)

### 4.3.2 Validating SERg in Clinical Prostate Histology and MRI

All four registration methods on the clinical data (i.e. Demons Intensity Registration, Demons SERg, FFD Intensity Registration, FFD SERg) are assessed for accuracy via computation of CR, MI, and Dice similarity coefficient (DSC)\(^{25}\) between the registered histology image and template MRI. Additionally, root mean squared distance (RMSD) is computed between corresponding landmarks in both images. DSC is defined as
\[ DSC(A, B) = \frac{2|\Omega^A \cap \Omega^B|}{|\Omega^A| + |\Omega^B|}. \] (9)
In this work, DSC measures the amount of overlap between the masks of the registered histology image and template MRI to evaluate alignment of their contours.

RMSD measures the distance in millimeters between the corresponding anatomic landmarks in histology and MRI annotated by a pathologist. Consider \( j = 1, ..., J \) with \( J \) represent the number of annotated landmarks in a single image. Also, let \( k_j \) and \( l_j \) represent the pixel index of landmark \( j \) in images \( A \) and \( B \), respectively. RMSD is defined as
\[ RMSD(A, B) = \sqrt{\frac{1}{N} \sum_k \| A(k_j) - B(l_j) \|_2^2}, \] (10)
where \( \| \cdot \|_2 \) represents the L2 norm distance between landmarks. RMSD measures internal alignment accuracy in terms of anatomical features such as the urethra, nodules, and ducts.

### 4.4 Implementation Details for Synthetic and Clinical Experiments

Segmentations and annotations were performed interactively using 3D Slicer\(^{22}\) software. The SE was implemented in MATLAB and took about 5-10 minutes to generate the PEs for each concatenated template-target image with each about 200 × 400 pixels. The Demons algorithm was implemented in ITK and ran for 2000 iterations, taking about 2 minutes to complete each registration. The FFD algorithm was implemented in MATLAB also for 2000 iterations, taking about 4 minutes to perform the registration.
5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Synthetic BrainWeb Registration Results

The synthetic T1-T2 registration experiments demonstrate the consistency of Demons registration improvement provided by SERg when the images are subject to various bias field and noise conditions (Figure 4). It appears that the largest differences between SERg and intensity registration tend to occur when bias field and noise are introduced simultaneously. Table 4 reports the three measures computed to evaluate accuracy of alignment. On average, SERg changes CR, MI, and MSE by 32.01 ± 11.03%, 55.60 ± 53.87%, and −81.51 ± 6.40%, respectively. The marked increase in CR and MI and decrease in MSE suggest that the registered images are more closely aligned with the template images after attempting to recover the synthetic deformation. In attempt to recover the local artificial shrinkage deformation, the intensity registration introduces artifacts in the rest of the image. While SERg does not fully recover the synthetically-induced shrinkage deformation, it successfully prevents the formation of artifacts. This may be attributed to the similar intensities of the PE representations whereas the corresponding anatomy in the original intensity images are indicated by more dissimilar intensities. Since Demons registration computes the squared difference of these corresponding intensities, the more similar images, i.e. PE representations, effectively drive the registration because of their compatibility.

![Figure 4](http://example.com/figure4.png)

Figure 4. Graphs of (a) CR, (b) MI, and (c) MSE plotted for each of the nine Demons registration experiments performed on BrainWeb images with varied noise and bias field. The studies are abbreviated as follows: no noise and no bias field (Control), 20% bias field (BF20), 40% bias field (BF40), 20% bias field with 7% noise, 40% bias field with 7% noise, 1% noise, 3% noise, 5% noise, and 7% noise.

<table>
<thead>
<tr>
<th></th>
<th>CR</th>
<th>MI</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demons Intensity</td>
<td>0.696 ± 0.053</td>
<td>0.407 ± 0.208</td>
<td>1773.040 ± 126.739</td>
</tr>
<tr>
<td>SERg</td>
<td>0.919 ± 0.025</td>
<td>0.634 ± 0.168</td>
<td>327.795 ± 107.971</td>
</tr>
<tr>
<td>% Change by SERg</td>
<td>32.01 ± 11.03%</td>
<td>55.60 ± 53.87%</td>
<td>−81.51 ± 6.40%</td>
</tr>
</tbody>
</table>

Table 4. Average post-Demons registration measures computed between the registered image and the template image for multimodal brain registration: correlation ratio (CR), mutual information (MI), and mean squared error (MSE). The bolded values highlight optimal performance. Alignment improvement is signified by increase in CR and MI and decrease in MSE.

5.2 Clinical ex vivo post-operative prostate histology and in vivo pre-operative MRI Registration Results

After affine registration of the prostate histology to MRI, the images were registered using intensity registration and SERg within both Demons registration scheme and FFD. Figure 5 and Table 5 show the average values across all twenty-six slices for the metrics computed on the registered histology image relative to the template MRI. For both Demons and FFD registration, SERg appears to perform slightly better than raw intensity registration with a 1% increase in DSC and more than 8% decrease RMSD. This suggests some improvement in alignment of
both the contours and internal features. There are minimal differences in DSC across the different registration experiments because the Demons and FFD registration almost completely superimposes the histology and MRI, leaving only slight room for improvement by SERg.

In Demons registration, SERg yields inferior similarity measures (lower CR and MI) but superior prostate overlap (higher DSC) and superior alignment of internal landmarks (lower RMSD) when compared to Demons registration on the original images. Since Demons registration computes statistical correlations between intensities on a per-pixel basis tends to improve CR and MI more effectively when operating directly on the raw intensity images than on the PE representations. However, this occurs at the expense of misaligning the images, as suggested by a lower DSC and higher RMSD compared with SERg.

In FFD registration, SERg performs better than intensity registration across all four measures: higher CR, MI, and DSC as well as lower RMSD. This is possibly because of the altered representations of the prostate shown in Figure 6(b) where the central gland is better separated from the peripheral zone in the PEs than in their original forms. Figure 6(c),(d),(e),(f) show the qualitative registration results from prostate histology and T2-w MRI for one patient axial slice. An improvement in alignment after FFD with SERg is apparent along the bottom edge of the prostate in Figure 6(j).

![Graphs of mean (a) CR, (b) MI, (c) DSC, and (d) RMSD and their standard deviations across all twenty-six registered slices.](image)

Figure 5. Graphs of mean (a) CR, (b) MI, (c) DSC, and (d) RMSD and their standard deviations across all twenty-six registered slices.

<table>
<thead>
<tr>
<th></th>
<th>CR</th>
<th>MI</th>
<th>DSC</th>
<th>RMSD (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demons</td>
<td>0.776 ± 0.039</td>
<td>0.668 ± 0.078</td>
<td>0.940 ± 0.018</td>
<td>4.505 ± 1.179</td>
</tr>
<tr>
<td>SERg</td>
<td>0.692 ± 0.059</td>
<td>0.665 ± 0.075</td>
<td>0.953 ± 0.017</td>
<td>4.278 ± 1.227</td>
</tr>
<tr>
<td>% Change by SERg</td>
<td>-8.364 ± 3.649%</td>
<td>-0.311 ± 0.993%</td>
<td>1.219 ± 0.733%</td>
<td>-22.687 ± 33.474%</td>
</tr>
<tr>
<td>FFD</td>
<td>0.667 ± 0.074</td>
<td>0.674 ± 0.087</td>
<td>0.968 ± 0.022</td>
<td>4.347 ± 1.264</td>
</tr>
<tr>
<td>SERg</td>
<td>0.688 ± 0.069</td>
<td>0.687 ± 0.082</td>
<td>0.975 ± 0.018</td>
<td>4.259 ± 1.270</td>
</tr>
<tr>
<td>% Change by SERg</td>
<td>2.132 ± 4.384%</td>
<td>1.305 ± 2.664%</td>
<td>0.703 ± 2.027%</td>
<td>-8.790 ± 93.628%</td>
</tr>
</tbody>
</table>

Table 5. Registration results for multimodal histology-MRI fusion: correlation ratio (CR), mutual information (MI), Dice similarity coefficient (DSC), and root mean squared distance (RMSD) measured in millimeters. The bolded values highlight optimal performance. Alignment improvement is signified by increase in CR, MI, and DSC and decrease in RMSD.
6. CONCLUDING REMARKS

In this work we presented spectral embedding based registration (SERg), a novel framework for multimodal image alignment. The approach was successfully applied to synthetic brain images and prostate histology-MRI fusion. In the synthetic experiments using Demons registration, artifacts appeared as a result of intensity registration while SERg prevented the emergence of the artifacts. In the clinical experiments, SERg performed on prostate MRI and histology led to better alignment of the borders (higher Dice similarity coefficient) and internal structures (lower root mean squared distance between expert-annotated landmarks). It is possible that spectral embedding captured areas of saliency and correspondence based on extracted image features, generating alternate image representations with increased statistical correlation between images for more accurate multimodal registration.

In our registration scheme, we performed spectral embedding (SE) on the independent components resulting from independent component analysis (ICA) of low level statistical texture features extracted from multimodal images. Although co-registration using multivariate data including texture features and ICA have been demonstrated in the past, previous studies have customized the similarity measure to accommodate high-dimensional data. We show that the combination of feature extraction, ICA, and SE allows the direct use of typical intensity-based similarity measures to provide an effective alignment of multimodal images. This study suggests that extraction of low level statistical texture features from multimodal images followed by ICA to decompose the features into independent components and final distillation to parametric eigenvectors produce alternative representations that are more similar to each other than their original representations. Therefore, this enhanced
similarity increases the driving ability of the similarity measures to align multimodal images. Future work will involve a comparison different combinations of extracted features, ICA, ICA+SE, and SE alone to assess how each of these modules influence the appearance of the PE representations and ultimately the registration accuracy. Further validation of this method will include more multimodal registration problems such as 3D registration and time-series data such as aligning brain MRI and apparent diffusion coefficient images to evaluate glioma treatment changes.

7. ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers R01CA136535-01, R01CA140772-01, and R21CA167811-01; the Department of Defense under award numbers W81XWH-12-PCRP-PTA and W81XWH-11-1-0179; the National Institute of Diabetes and Digestive and Kidney Diseases under award number R01DK098503-02, the DOD Prostate Cancer Synergistic Idea Development Award (PC120857); the QED award from the University City Science Center and Rutgers University, the Ohio Third Frontier Technology Development Grant. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES


