

# Hierarchical Segmentation-Assisted Multimodal Registration for MR Brain Images

Huanxiang Lu<sup>1</sup>, Roland Beisteiner<sup>2</sup>, Lutz-Peter Nolte<sup>1</sup>, Mauricio Reyes<sup>1</sup>

<sup>1</sup>*Institute of Surgical Technologies and Biomechanics, University of Bern, Switzerland*

<sup>2</sup>*Department of Neurology, MR Center of Excellence, Medical University of Vienna, Austria*

---

## Abstract

Information theory-based metric such as mutual information (MI) is widely used as similarity measurement for multimodal registration. Nevertheless, this metric may lead to matching ambiguity for non-rigid registration. Moreover, maximization of MI alone does not necessarily produce an optimal solution. In this paper, we propose a segmentation-assisted similarity metric based on point-wise mutual information (PMI). This similarity metric, termed SPMI, enhances the registration accuracy by considering tissue classification probabilities as prior information, which is generated from an expectation maximization (EM) algorithm. Diffeomorphic demons is then adopted as the registration model and is optimized in a hierarchical framework (H-SPMI) based on different levels of anatomical structure as prior knowledge. The proposed method is evaluated using Brainweb synthetic data and clinical fMRI images. Both qualitative and quantitative assessment were performed as well as a sensitivity analysis to the segmentation error. Compared to the pure intensity-based approaches which only maximize mutual information, we show that the proposed algorithm provides significantly better accuracy on both synthetic and clinical data.

*Keywords:*

Multimodal non-rigid registration, tissue classification, EPI distortion correction

---

## 1. Introduction

Magnetic resonance imaging (MRI) is one of the most popular techniques for human brain imaging due to its versatility and flexibility. In many MR

brain analyses, multimodal non-rigid registration is widely used in applications such as image fusion, EPI distortion correction, etc. Although multimodal non-rigid registration has been an active research area since past few decades, more work is still needed in order to improve the accuracy, the robustness and the computational time. The major difficulty in multimodal registration compared to monomodal registration is the various intensity representations of tissues between different modalities. This fact leads to the non-linearity in the intensity mapping in the joint histogram, hence the cost functions for monomodal registration such as sum of the squared differences (SSD) and cross-correlation (CC) are no longer suitable in this case. Therefore, finding a proper similarity metric is a key task in multimodal registration.

Since it is difficult to extract the intensity relationship directly from multimodal images, information theory-based methods using statistical relations of the intensities are widely employed to tackle multimodal registration problem. Mutual information (MI) is one of the most popular similarity metrics. With its successful performance demonstrated in linear registration [1, 2], MI is also applied in solving non-rigid registration problems. [3] proposed a free-form deformation (FFD) registration using B-spline, in which MI is maximized globally over the entire image space by optimizing iteratively the parameters of the grid nodes. [4] presented a flow-based approach by computing the gradient of MI using a continuous and differentiable joint histogram based on Parzen window. Recently, diffeomorphic demons algorithm was also extended for multimodal registration purpose [5, 6]. Nevertheless, for MRI images where intensity inhomogeneity exists, those algorithms searching to maximize mutual information do not necessarily guarantee the optimal solution. In fact, MI is approximately built on the joint distribution of the intensities between two images. Even though the images are perfectly registered, it is still difficult to know a real intensity relationship, as matching ambiguity often exists as exemplified in Fig. 1. Therefore, a typical maximization of mutual information method only minimizes the global joint entropy without considering the real tissue correspondence.

To address such limitation, different approaches have been proposed that incorporate additional information, such as spatial information [7, 8, 9] and anatomical landmark information [10, 11]. These methods can generally improve the registration accuracy. However, spatial information does not directly provide real correspondence between tissues, whereas anatomical landmark information usually implies either manual annotation or difficulties

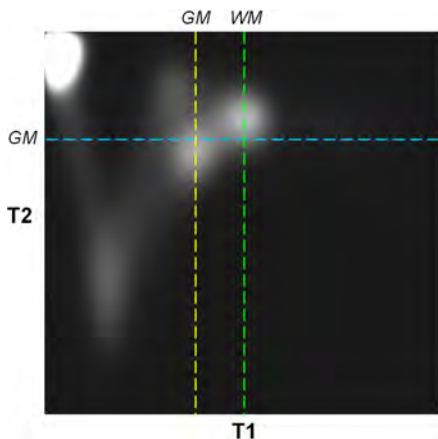


Figure 1: Example of the matching ambiguity problem in a joint histogram between perfectly aligned T1 and T2 brain images. The blue horizontal dashed line represents an intensity value of gray matter (GM) in T2 image. This intensity has the probability of matching both gray matter (GM) (yellow dashed line) and white matter (WM) (green dashed line) in T1 image.

in automatic detection given poor image quality, especially for patient data where the image quality is typically inferior to healthy subject populations due to limited patient compliance and pathological brains.

A similar approach to the work presented hereby was proposed by [12], in which two similarity measures using voxel classification are introduced. One measure takes into account the voxel class probabilities from two images and registers images so that the fuzzy overlap of corresponding voxel object labels becomes similar to the ideal case where the tissue probability maps of both images are identical. While better results compared to pure-intensity method can be obtained, this measure assumes excellent tissue classification on both images. The other measure uses intensities in one image to match the fuzzy class labels in the other image by minimizing the conditional entropy given the class label. However, this measure requires a segmented labeled image, thus is only suitable for applications like atlas-to-image registration. In addition, the sensitivity of these measures to the classification or segmentation accuracy is not provided.

In this work, we investigate a segmentation-assisted similarity metric, termed SPMI, which encodes the tissue classification probability into the point-wise mutual information (PMI) metric introduced by [13]. The measurement is calculated by extending the joint histogram into 3D with the

third dimension representing the probability of two intensity values being the same tissue class. This provides tissue correspondence information that modulates the joint distribution probability and is integrated into the diffeomorphic demons registration model. Note that instead of performing simultaneous segmentation and registration [14, 15], we extract the probabilities of voxel classification at the first step. This is because an initial severe misalignment of the two images may potentially lead to local minima for segmentation due to a larger optimization space, hence indirectly degrading the registration accuracy.

We adopt a hierarchical model, which is originally designed for segmentation [16], to represent our multi-level registration framework. Unlike the conventional multi-level registration approaches in [17] where multi-level refers to different image resolutions, this hierarchical model uses a tree structure with each node representing an anatomical structure computed from the voxel classification results (e.g. see Fig. 2, the region of brain is a node that has child nodes such as gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF)). Therefore, the registration is carried out by incorporating different classification information for different anatomical structure at each level. The reason of using this model is two-fold: 1) Hierarchical structure ensures the alignment at different levels of anatomical region. 2) It provides the flexibility to cope with modalities representing different classifications (e.g. EPI image has similar intensity representation of WM and GM. Therefore the ensemble of these two tissues can be classified to an upper level node as non-CSF, which has its corresponding node in T1).

To evaluate the proposed method, we compare it to two pure intensity-based approaches - PMI Demons [6] and Elastix [18]. We first examine the accuracy and performance on a set of synthetic data created from Brainweb [19], which allows us to perform quantitative analysis on registration accuracy. Furthermore, a sensitivity test is performed to investigate how the voxel classification accuracy affects the final registration results. Then the proposed method is tested with clinical EPI images to correct geometric distortion and evaluated using brain volume recovery ratio and manually-defined landmark distance error.

The details of the method are elaborated in the next section. Then, experiments are presented in Section 3, followed by the discussion in Section 4.

## 2. Method

### 2.1. Registration Model

We adopt the diffeomorphic demons algorithm as our registration model [20]. The main advantage of this approach is the computation efficiency since it uses optical flow as image forces instead of solving partial differential equations or optimizing in a large parametric space. Besides, it also ensures the diffeomorphism of the dense field transformation by mapping the update field on the Lie Group, which has shown to allow accurate statistical morphometry analysis [21]. Given two images  $F$  and  $M$ , the registration model can be summarized by an energy function consisting of a similarity term and a regularization term.

$$E(s) = Sim(F, M \circ s) + \alpha Reg(s), \quad (1)$$

where  $M \circ s$  denotes the moving image  $M$  composed by transformation  $s$ ,  $Sim(F, M \circ s)$  is the similarity criterion measuring the resemblance of two images, the regularization term  $Reg(s)$  regularize the dense field transformation to a specific physical mode (e.g. fluid model, diffusion model, etc) and  $\alpha$  is a weighting factor controlling the amount of regularization.

### 2.2. Similarity Metric

Mutual information (MI) is a similarity metric based on information theory, which is widely used to perform multimodal image registration. It measures how much information is gained about one random variable by the knowledge of another random variable. MI can be computed from the marginal and joint entropies:

$$MI = H(F) + H(M) - H(F, M) = \sum_{i_F, i_M} p(i_F, i_M) \log \left[ \frac{p(i_F, i_M)}{p(i_F)p(i_M)} \right], \quad (2)$$

where  $i_F$  and  $i_M$  are image intensities of image  $F$  and  $M$ ,  $p(i_F, i_M)$  denotes the joint probability, and  $p(i_F)$  and  $p(i_M)$  are the marginal probabilities.

In order to incorporate MI as an energy function into a dense field registration model, in practice one has to transfer this global measurement into a local metric. By rewriting Eq.(2), [13] introduced the point-wise form of MI:

$$MI = \frac{1}{N} \sum_{x \in \Omega} PMI(x) \quad (3)$$
$$PMI(x) = \log \left[ \frac{p(i_F(x), i_M(x))}{p(i_F(x))p(i_M(x))} \right],$$

where  $x$  is the spatial coordinate of the voxel pair in the image space  $\Omega$ . From the above equation, we see that the final summation is taken over the spatial image coordinates instead of the intensities in Eq.(2). Thus, point-wise mutual information (PMI) can be regarded as the contribution from each voxel pair to the global MI.

Note that the probability of intensities in the fixed image  $p(i_F)$  always remains constant. By assuming that the intensity distribution in  $M$  merely changes during the registration,  $p(i_M)$  has also very little influence on the energy function in Eq.(3). Therefore, the optimization of such similarity metric can be approximately regarded as a maximization of the joint probability  $p(i_F, i_M)$ . As shown in Fig. 1, a gray matter (GM) voxel in T2 image might tend to match a white matter (WM) point in T1 image to achieve better joint probability value even though they belong to different tissues. Therefore, this maximization without considering the tissue correspondence may lead the registration to a suboptimal solution.

To address this limitation, we extend the PMI metric by adding a constraint to ensure that two homologous voxels belong to the same biological tissue. This is achieved by incorporating a term  $C_s$  into the joint probability that represents the probability of being the same class, resulting in a 3D joint histogram. The new similarity metric, termed SPMI, can be written as:

$$SPMI(x) = \log \left[ \frac{p(i_F(x), i_M(x), C_s)}{p(i_F(x))p(i_M(x))} \right]. \quad (4)$$

By applying the Bayes rule, the new joint probability becomes

$$p(i_F, i_M, C_s) = p(C_s|i_F, i_M)p(i_F, i_M),$$

where  $p(C_s|i_F, i_M)$  denotes the probability of  $i_F$  and  $i_M$  belonging to the same class. From the above equation,  $p(C_s|i_F, i_M)$  can be regarded as a modulation function that constrains the optimization of the joint entropy. This probability can be computed by summing up the probability of the intensity pair belonging to each class  $k$ :

$$p(C_s|i_F, i_M) = \sum_k^K p(C_k|i_F, i_M).$$

where  $K$  is the number of classes.

Since  $i_F$  and  $i_M$  are from two different images, therefore independent to each other, Eq.(4) can be re-written as:

$$SPMI = \log \left[ \frac{p(i_F, i_M) \sum_k p(C_k|i_F)p(C_k|i_M)}{p(i_F)p(i_M)} \right], \quad (5)$$

where  $p(C_k|i_F)$  and  $p(C_k|i_M)$  are posterior probabilities, which indicate the probabilities that  $i_F$  and  $i_M$  belong to a particular class  $C_k$ , respectively. According to Bayes rule, these posterior probabilities can be calculated as:

$$p(C_k|i)_{\{i=i_F, i_M\}} = \frac{p(C_k, i)}{\sum_{l=0}^K p(C_l, i)} = \frac{p(i|C_k)p(C_k)}{\sum_{l=0}^K p(i|C_l)p(C_l)}, \quad (6)$$

where  $p(i|C_k)$  denotes the probability of intensity  $i$  given class  $C_k$  and the prior  $p(C_k)$  being the proportion of class  $C_k$  in the entire image. As we see from the expanded form of  $p(C_k, i)$ , this term can be expressed as a Gaussian mixture model (GMM)  $G(i, \mu_k, \sigma_k)c_k$  with parameters  $\mu_k$ ,  $\sigma_k$  and  $c_k$  being the mean, the variance and the proportion of the Gaussian model of class  $k$ , respectively. Eq.(6) can then be rewritten as follow:

$$p(C_k|i)_{\{i=i_F, i_M\}} = \frac{G(i, \mu_k, \sigma_k)c_k}{\sum_{l=0}^K G(i, \mu_l, \sigma_l)c_l}, \quad (7)$$

The parameters of the GMM can be computed using a typical expectation-maximization (EM) segmentation algorithm [22], as the one used in our experiments.

### 2.3. Optimization

In the original diffeomorphic demons algorithm, the sum of the squared differences (SSD) is utilized as similarity metric, which allows the energy function to be optimized using a simple Gauss-Newton method. However, as a metric composed of joint intensity probability and joint classification probability, it is not straightforward to apply the same strategy for the optimization of SPMI. Thus, we employ the finite-difference approach presented in [6] for the computation of the gradient to optimize the energy function of SPMI as described below.

The update field  $\mathbf{u}$  of the deformation is computed by averaging a forward force  $F_f$  and a backward force  $F_b$  in order to ensure the registration consistency [23]:

$$\mathbf{u} = k_E(F_f - F_b), \quad (8)$$

where the coefficient  $k_E$  indicates the update step length;  $F_f$  is a gradient of local similarity with respect to the reference image, which tends to move the moving image voxels towards their better match in the fixed image; and  $F_b$  is the reverse of  $F_f$ , which improves matching of the points in the fixed image according to the moving image.  $F_f$  and  $F_b$  are computed using the finite-difference method:

$$\begin{aligned} F_f(x) &= \left. \frac{\partial}{\partial \epsilon} \right|_{\epsilon=0} SPMI(i_F(x + \epsilon), i_{M \circ s}(x)) \\ F_b(x) &= \left. \frac{\partial}{\partial \epsilon} \right|_{\epsilon=0} SPMI(i_F(x), i_{M \circ s}(x + \epsilon)). \end{aligned} \quad (9)$$

where  $\epsilon$  denotes a step of displacement of the voxel for computing the finite-difference.

The resulting transformation  $s$  is then updated by composing  $\mathbf{u}$  on the Lie Group through the exponential operator to ensure the diffeomorphism, and is regularized by a Gaussian kernel [20].

$$s \leftarrow G * (s \circ \exp(\mathbf{u})). \quad (10)$$

#### 2.4. Hierarchical Model Based on Anatomical Region

As described in Sec. 2.2, the anatomical structures are pre-segmented to assist PMI. According to Eq.(5), the SPMI metric requires the same class correspondences in both images. Nevertheless, it might not always be the case for all MRI modalities. As we know, MRI is a flexible imaging technique with different protocols to enhance the visualization of different tissues. One tissue type visible in one modality might not have strong contrast in another modality. This results in different voxel classes for different modalities. Therefore class correspondences can not be ensured for all MR images.

To overcome this limitation, we adopt a tree model, which is originally designed for single image segmentation [16], to represent all anatomical structures in both images for registration at different hierarchical levels. The root of the tree describes the whole image region, the leaves represent all finest structures from the pre-segmentation results and the inner nodes describe intermediate coarser anatomical structures. Using this tree representation, class correspondences can be guaranteed at least at one level of the hierarchical model. Fig. 2 shows an example of the hierarchical model. Even though T1 image and EPI image has different leaves indicating different classification of the finest structure, the class correspondences are identical at level 2.



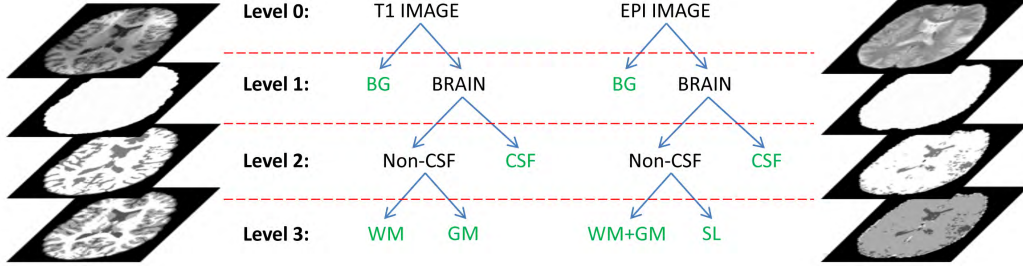


Figure 2: Scheme of the anatomical structure-based hierarchical model. The left tree describes the hierarchical structure of a T1 image. The root represents the whole image. At level 1, the image is divided into background (BG) and brain structure (BRAIN). At level 2, the BRAIN is separated into non-CSF and CSF. Non-CSF is then partitioned into WM and GM at level 3; In the right tree for echo-planar image (EPI), the structure is equivalent to the one of T1 until level 2, whereas at level 3 non-CSF is partitioned into the ensemble of WM and GM as well as the signal loss (SL).

The hierarchical registration, termed H-SPMI, can then be summarized by Algorithm 1. The process starts by using the classification probabilities of the coarsest structure as prior information at the top level, followed by using lower level knowledge of the finer structures until the classes are no longer corresponding in the two images. The classification probabilities of the intermediate nodes can be calculated by summing up all their child nodes' probabilities.

```

Input: Image  $F$  and  $M$ ; Nodes of the tree  $N_F$  and  $N_M$ 

Find nodes  $N_F(l)$  and  $N_M(l)$  at level  $l = 1$ ;
while  $N_F(l) \neq N_M(l)$  do
    foreach voxel  $x$  do
        Compute update field  $\mathbf{u}$  using SPMI according to the
        classification defined by  $N_F(l)$  and  $N_M(l)$ ;
    end
    Compute transformation  $s \leftarrow G * (s \circ \exp(\mathbf{u}))$ ;
    Let  $M \leftarrow M \circ s$ ;
    Increment level  $l$  and refine nodes  $N_F(l)$  and  $N_M(l)$ ;
end

```

**Algorithm 1:** Hierarchical Registration Model

### 3. Results

Validation of non-rigid registration has always been a difficult task since ground truth deformation between the fixed image and moving image is seldom available for real clinical cases. Therefore, to validate this methodology, we conducted several experiments using different types of data varying from phantom to clinical fMRI images. A comparison of different similarity metrics mentioned in Section 2 is presented, as well as Elastix [18] which is an improved version of Free-form Deformation (FFD) [3] method for multimodal registration.

The Elastix software was downloaded from the website (<http://elastix.isi.uu.nl/>). The other compared methods were implemented in C++ using Insight Toolkit library (<http://www.itk.org>). For Elastix, the optimization uses the optimal parameter setting reported in [18], in which the control point spacing is 16 mm in each dimension and the number of resolution levels is four. For PMI, we adopt  $\sigma = 2$  for Gaussian kernel regularization and three levels of resolution as image pyramid. We also compared the differences between SPMI only using the finest classification (i.e. no hierarchy) and using the hierarchical structure (H-SPMI). For these two methods, the number of classes and the hierarchical structure varies depending on the application, which will be presented in each experiment. The regularization schemes are the same as the one used for PMI.

#### 3.1. Evaluation on Synthetic Data

In order to have a quantitative analysis, the algorithm was first validated on simulated MR Brain images from BrainWeb [19], which provides the segmentation ground truth. In this experiment, one T2 image with dimension of  $181 \times 217 \times 181$  and isotropic voxel size of  $1 \times 1 \times 1mm^3$  was used as the fixed image. 20 T1 images with dimension of  $256 \times 256 \times 181$  and isotropic voxel size of  $1 \times 1 \times 1mm^3$  were registered to the reference T2 image. Both T1 and T2 images are classified into four classes using EM algorithm (Background (BG), WM, GM and CSF) and two hierarchical levels (level 1: BG and BRAIN, level 2: BG, WM, GM and CSF) for SPMI and H-SPMI, respectively.

#### *Visual Assessment*

The registration results were first investigated by visual assessment. We extracted the contour of the each brain volume and the ventricle on the fixed

Table 1:  $p$ -value obtained from t-test between the proposed methods and pure intensity-based methods on different tissue types and overall of the whole brain.

	WM	GM	CSF
SPMI vs Elastix	1.0999e-011	0.3394	0.0252
SPMI vs PMI	1.95e-004	0.0013	1.3027e-007
H-SPMI vs Elastix	2.1586e-022	4.3283e-005	3.3476e-008
H-SPMI vs PMI	7.82e-021	1.06e-021	1.88e-014
H-SPMI vs SPMI	2.03e-021	1.87e-019	6.44e-004

T2 images (Fig. 3(a)), and overlaid on all registration results. Fig. 3 shows the results of different methods on one case of the study images. One can see in Fig. 3(b)-(e), Elastix gives fairly poor result compared to the other three methods especially in the contour of the brain and the upper ventricle areas; PMI was able to provide promising results in the ventricle area, while there is a slight misalignment between the brain volume and the background. SPMI and H-SPMI show more superior performance in discriminating the brain volume and background giving almost perfect contour overlay and provide better results around the ventricles. In order to visualize the differences more clearly, we also generated the subtraction images. Due to the different modality between the reference and the results, this is done by subtracting the obtained results to the original Brainweb T1 image, which is perfectly aligned with the T2 reference image. As shown in Fig. 4, similar observation can be found from the subtraction images where stronger differences occur in Elastix, PMI and SPMI compared to H-SPMI.

#### *Quantitative measurement*

We evaluated quantitatively the accuracy of our method by measuring the DICE coefficient on WM, GM and CSF. As shown in the boxplots in Fig. 5, remarkably, the H-SPMI yields best performance in all WM, GM and CSF, indicating the improved accuracy and robustness of this method. More importantly, statistical analysis were conducted showing the significantly better accuracy of SPMI and H-SPMI than pure-intensity methods Elastix and PMI ( $p < 0.01$  in t-test) except for SPMI in GM, as shown in Table 1. In addition, significant difference is also observed between SPMI and H-SPMI.

Furthermore, we also compared the MI value of all results, see Table 2. Interestingly, it is found that H-SPMI and SPMI have systematically higher MI value than PMI, even though PMI is an algorithm greedily searching

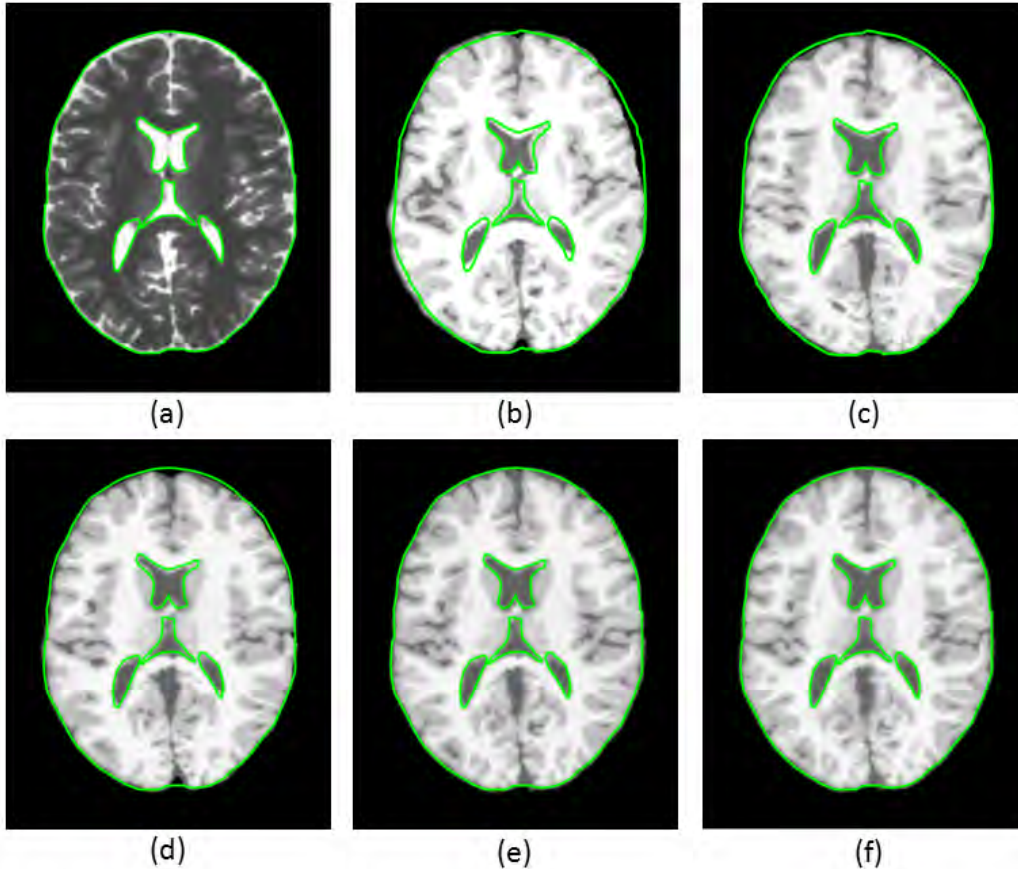


Figure 3: Visual assessment of the result on one case synthetic data. Contours of the brain volume and the ventricle are extracted on the central axial slice of (a) fixed T2 image and then overlaid on the image before registration (b), and different registration results: (c) Elastix, (d) PMI, (e) SPMI and (f) H-SPMI.

higher MI while SPMI and H-SPMI have penalty constraint on the image force. It is because this additional penalty can guide the image force to bypass the potential local minima during the optimization.

The smoothness of the deformation is also an important criteria for assessing the quality of the registration results. To this end, we computed the harmonic energy to measure how smooth the obtained deformations are from different registration methods. As shown in Fig. 6, Elastix produces the least smooth deformation compared to the other three methods since no diffeomorphism is incorporated in the algorithm. On the other hand, SPMI

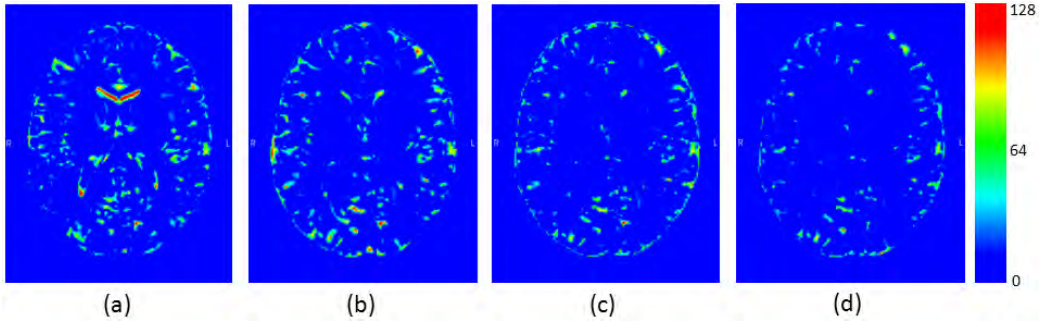


Figure 4: Subtraction image on the central axial slice between the Brainweb T1 image and different registration results: (a) Elastix, (b) PMI, (c) SPMI and (d) H-SPMI.

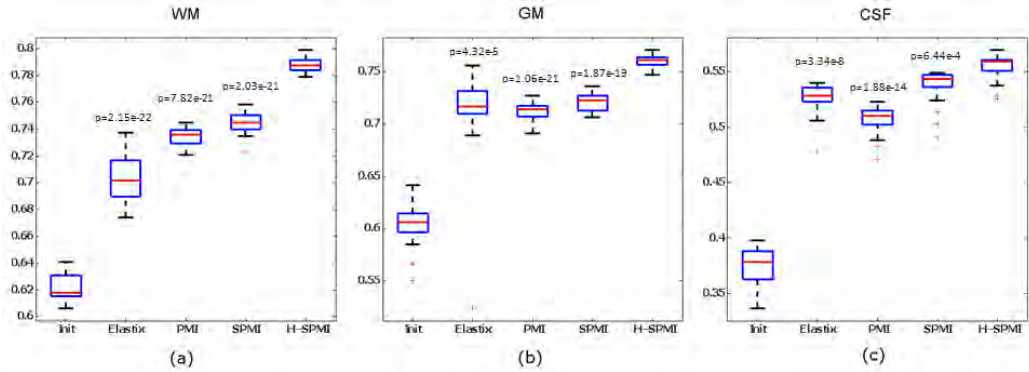


Figure 5: Quantitative evaluation of the experiment on synthetic data. DICE coefficients of Elastix, PMI, SPMI and H-SPMI are grouped according to different regions of interest. (a) WM, (b) GM and (c) CSF.

and H-SPMI maintain similar smoothness of the deformation to the one of PMI while achieving better accuracy.

### Sensitivity Analysis

Various MR hardware and acquisition protocols generate different image qualities. Thus, accurate segmentation of the brain is not always guaranteed. As our method incorporates classification knowledge as prior, the purpose here is to investigate the sensitivity of the proposed method to the segmentation results. To this end, we artificially generated three different segmentation results by adding random noise on the GMM parameters from the EM segmentation. The DICE coefficients of all the tissue types were calculated to assess the segmentation accuracy, as shown in Fig. 7(d). Test

Table 2: MI value obtained with different registration results. Numbers in bold represent the highest values.

Case	Init	Elastix	PMI	SPMI	H-SPMI
1	0.535	0.656	0.675	0.686	<b>0.746</b>
2	0.523	0.687	0.669	0.681	<b>0.744</b>
3	0.496	0.691	0.659	0.684	<b>0.728</b>
4	0.545	0.672	0.674	0.683	<b>0.741</b>
5	0.520	0.611	0.673	0.684	<b>0.741</b>
6	0.526	0.689	0.673	0.683	<b>0.743</b>
7	0.530	0.645	0.672	0.683	<b>0.742</b>
8	0.509	0.687	0.671	0.681	<b>0.735</b>
9	0.473	0.663	0.667	0.657	<b>0.739</b>
10	0.526	0.614	0.671	0.686	<b>0.744</b>
11	0.514	0.657	0.675	0.687	<b>0.747</b>
12	0.512	0.675	0.665	0.679	<b>0.738</b>
13	0.517	0.641	0.668	0.680	<b>0.743</b>
14	0.513	0.663	0.668	0.686	<b>0.744</b>
15	0.488	0.689	0.660	0.679	<b>0.736</b>
16	0.524	0.674	0.674	0.680	<b>0.744</b>
17	0.536	0.676	0.672	0.686	<b>0.741</b>
18	0.526	0.638	0.677	0.674	<b>0.745</b>
19	0.515	0.694	0.668	0.660	<b>0.727</b>
20	0.531	0.639	0.675	0.685	<b>0.743</b>

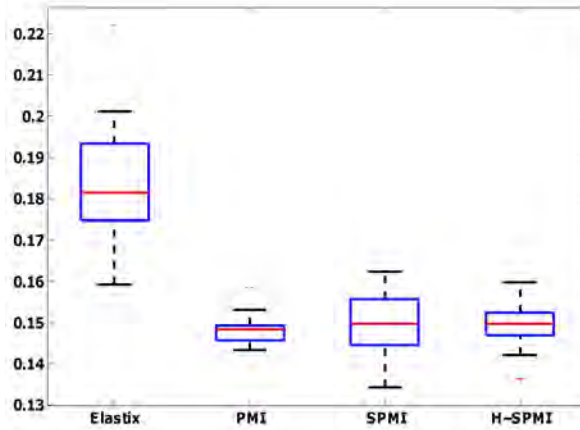


Figure 6: Harmonic energy of the deformation field obtained by Elastix, PMI, SPMI and H-SPMI. Lower harmonic energy indicates smoother deformation.

1 describes the worst scenario with average DICE value of 0.6497. Test 2 and Test 3 achieve intermediate accuracy with DICE of 0.6948 and 0.7461, respectively. Test 4 is the segmentation result without any added noise, whose DICE value is 0.8312. In total 20 registrations, performed by H-SPMI, were carried out in each test. Fig. 7 shows the DICE coefficients of WM, GM and

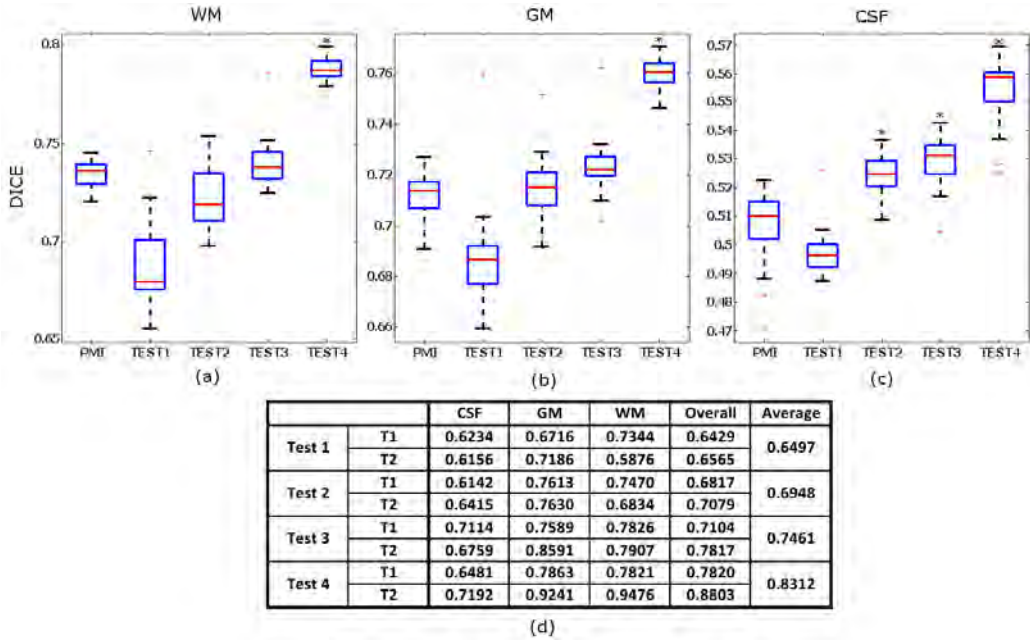


Figure 7: Comparison of DICE coefficients obtained with H-SPMI with different segmentation qualities on (a) WM, (b) GM and (c) CSF. Most left column shows the DICE coefficient obtained from PMI multimodal demons. \* means statistical significance with the PMI results. (d) The DICE coefficient is computed per image modality and for each tissue type in each test. The average of the overall DICE coefficient over both modalities is further considered to facilitate the discussions.

CSF obtained from each test. At the DICE value of 0.7461, the results from H-SPMI yields significant better performance in CSF and are slightly better in WM and GM than the ones from PMI. When the DICE value drops to 0.6948, the performance of H-SPMI is still better in GM and CSF and is comparable with the PMI results in WM. When the segmentation quality decreases to the worst scenario in Test 1, the performance of SPMI drops dramatically with lower DICE values than PMI.

### 3.2. EPI Distortion Correction

To have a more thorough validation, the proposed method was also applied on real clinical images. In this experiment, we employed the proposed approach for correcting echo-planar imaging (EPI) distortion. Since EPI images have geometric distortion along B0 axis due to the fast acquisition protocol, multimodal non-rigid registration is often adopted to correct the

distorted EPI image by registering it to a structural image (e.g. usually T1 or T2 image). The challenges in this application are often the local distortions occurred in sparse slices, rendering the registration task difficult.

In this experiment, two types of dataset were adopted from our clinical partner. Due to largely varying patient compliance and brain pathology with different artifact levels, one type provides generally good image quality with high signal noise ratio ( $\text{SNR} > 50$ ) while the other shows strong signal loss and noise level ( $\text{SNR} < 17$ ). We refer to these two types as “high quality dataset” and “low quality dataset” in the following of the paper, respectively. Each of the dataset contains 10 patient cases including one structural MRI T1 image and one fMRI image acquired using EPI protocol. The T1 image has FOV of  $230 \times 230 \text{ mm}$  with matrix size of  $256 \times 256 \times 256$  and voxel size of  $0.9 \times 0.9 \times 0.9 \text{ mm}^3$ , while the EPI image has the same FOV of  $230 \times 230 \text{ mm}$  with different matrix size of  $128 \times 128$  and 34 axial slices of  $3\text{mm}$  thickness. The “high quality dataset” is acquired with  $\text{TR} = 2500 \text{ ms}$ ,  $\text{TE} = 28 \text{ ms}$ , bandwidth = 2220 Hz, flip angle =  $90^\circ$ , full Fourier encoding; whereas the “low quality dataset” is acquired with  $\text{TR} = 2500 \text{ ms}$ ,  $\text{TE} = 22 \text{ ms}$ , bandwidth = 1445 Hz ( $\text{P1} = 1220 \text{ Hz}$ ,  $\text{P4} = 1395 \text{ Hz}$ ), flip angle  $80^\circ$ , 6/8 partial Fourier factor.

Bias field correction ([24]) was first applied on both T1 and EPI images to remove the intensity inhomogeneity. Skull-stripping was done to facilitate the registration process using BET2 software ([25]) and manual correction. Then, the T1 image was aligned to EPI image using rigid registration to remove motion artifacts and the unnecessary neck region. In this experiment, the proposed H-SPMI method is compared with PMI and Elastix. Two hierarchical levels (level 1: BG and BRAIN, level 2: BG, non-CSF and CSF) is employed. Note that non-hierarchical SPMI is not included hereby due to the inequivalent tree structures between T1 image and EPI image (WM and GM are classified into one class in EPI due to similar intensity values).

For high quality dataset, the registration results show the clear advantage of H-SPMI for all 10 cases. An example comparing different results is shown in Fig. 8, including axial and sagittal views where severe distortion along the anterior-posterior B0 axis can be observed in the frontal lobe. As seen from Fig. 8(c)-(e), H-SPMI was able to recover this local distortion whereas Elastix and PMI both failed.

For low quality dataset, H-SPMI still globally outperforms PMI and Elastix, except for two cases that have rather low image quality resulting in extremely poor classification from EM segmentation. Fig. 9 shows a



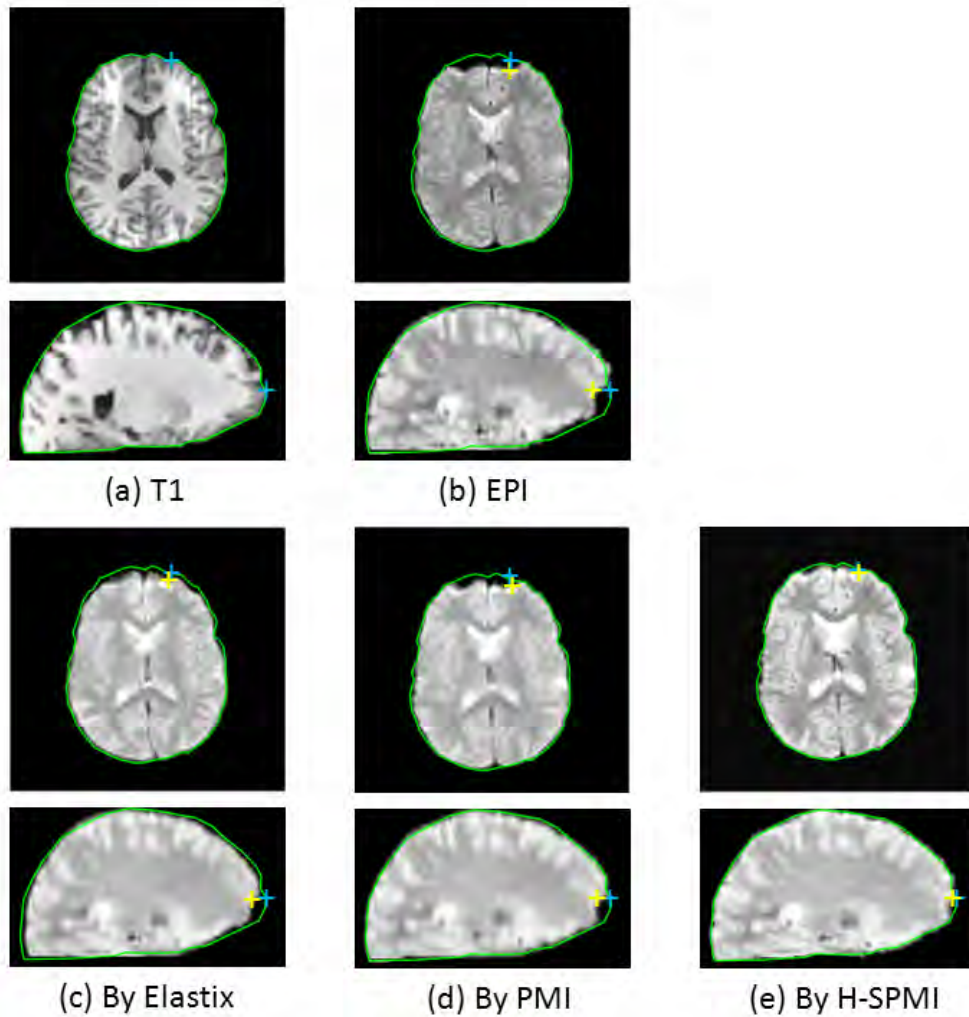


Figure 8: EPI distortion correction results from different methods on one case in the high quality dataset, with axial and sagittal view of (a) structural T1 image, (b) distorted EPI image, as well as recovered image by (c) Elastix, (d) PMI, and (e) H-SPMI. Green line is the contour of reference brain overlaid on each result. Blue cross represents one manually-defined landmark in the T1 reference and yellow cross is the corresponding landmark in the results.

failed case of H-SPMI. While PMI provides a fairly good alignment between the T1 image and the EPI image, one can see that H-SPMI wrongly registered the frontal lobe part of the brain. This is due to the mis-classification between the background and the signal dropout region.

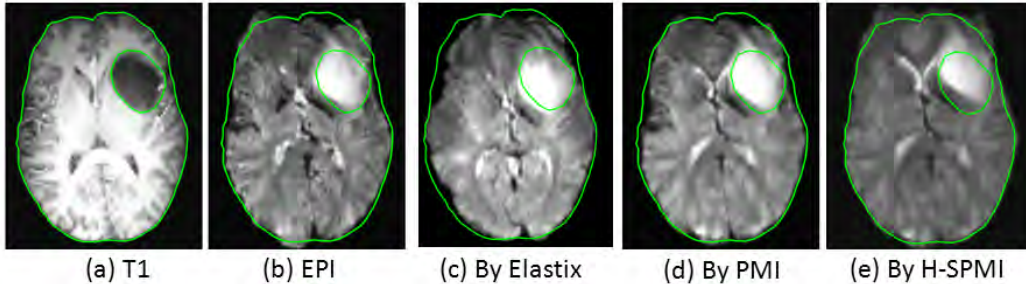


Figure 9: One case in low quality data with strong signal dropout and noise. Green line is the contour of reference brain together with the lesion. It is overlaid on (a) structural T1 image, (b) distorted EPI image, as well as the recovered images by (c) , (d) PMI, and (e) H-SPMI.

Besides visual assessment, the registration accuracy was also quantitatively validated. We first measured brain shape recovery ratio. Ideally, an EPI image without distortion should have 100% overlap over the structural image. Therefore, we computed the DICE coefficient of the brain volume between the T1 image and the EPI image. Note that due to the poor image quality of EPI, segmenting fine structures remains a difficult target. Hence, overlaps between WM, GM and CSF is not computed hereby. As shown in Fig. 10, for high quality dataset, H-SPMI yields higher DICE coefficient in most cases except for case 1 where PMI reaches higher overlap ratio. On the other hand, Elastix has relatively poor results showing limited improvement on brain shape overlapping. For low quality dataset, H-SPMI still outperforms Elastix and PMI in terms of brain shape recovery, except for case 5, 6 and 10 where PMI shows marginal improved overlap ratio.

The second quantitative metric is based on the geometric distance error computed from 10 manually defined landmarks, which are uniformly located in the regions where large distortion occurs as exemplified in Fig. 8. The results are summarized in Fig. 11 and Table 3. For high quality dataset, it can be seen that H-SPMI yields consistently lower error compared to Elastix and PMI. However, in the low quality cases, despite overall better performance, H-SPMI was not able to provide promising registration results for case 6 and case 10 due to the strong signal dropout and noise, which reflects the same observation as for the visual assessment. Statistical analysis reveals that H-SPMI outperforms Elastix and PMI ( $p = 0.0048$  over Elastix and  $p = 0.0236$  over PMI) in high quality dataset. Nevertheless, with low

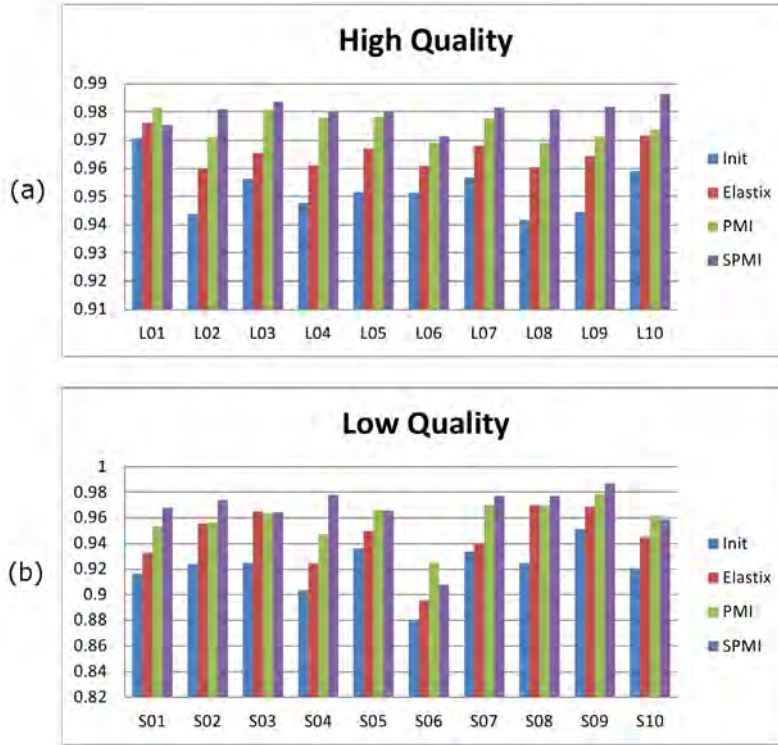


Figure 10: DICE coefficient between the T1 structural images and the recovered images using Elastix, PMI and H-SPMI for both (a) high quality data and (b) low quality data.

quality data, H-SPMI only show statistically significant better performance over Elastix ( $p = 0.0321$ ).

#### 4. Discussion

Dense field multimodal registration using point-wise mutual information metric is often hindered by the correspondence ambiguity where voxel matching is achieved by only maximizing joint probability, even though two given voxels do not necessarily belong to the same biological tissue class. This fact usually causes optimization local minima during the registration process leading to undesired results. In this work we present a registration framework using hierarchical segmentation-assisted point-wise mutual information (H-SPMI) for multimodal brain image registration.

Considering the true tissue type correspondence, prior knowledge of the

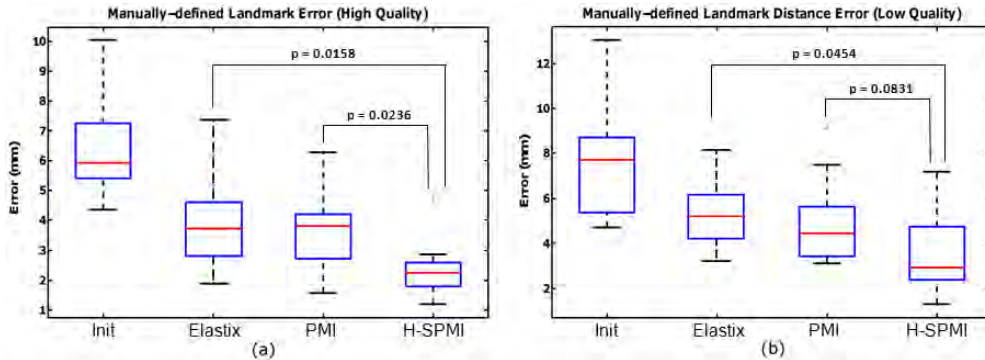


Figure 11: Manually-defined mean geometrical distance errors of registration results from Elastix, PMI and H-SPMI for both high quality data (a) and low quality data (b).

Table 3: Manually-defined geometrical distance errors of registration results from Elastix, PMI and H-SPMI for both high quality data and low quality data. Numbers in bold represent the lowest errors.

	Case	Init	Elastix	PMI	H-SPMI
High Quality	1	4.03 ± 1.07	3.13 ± 1.01	2.34 ± 0.99	<b>1.23 ± 0.82</b>
	2	3.01 ± 1.36	2.82 ± 1.31	2.33 ± 1.58	<b>1.59 ± 1.36</b>
	3	4.81 ± 1.81	3.70 ± 2.05	3.49 ± 1.78	<b>2.62 ± 1.82</b>
	4	3.24 ± 1.12	2.20 ± 1.06	1.42 ± 0.89	<b>0.99 ± 0.94</b>
	5	3.22 ± 1.55	2.05 ± 1.30	2.26 ± 1.20	<b>1.44 ± 1.26</b>
	6	3.35 ± 1.44	2.02 ± 1.61	1.51 ± 1.25	<b>1.23 ± 0.55</b>
	7	3.90 ± 1.52	2.51 ± 1.43	1.97 ± 1.60	<b>1.34 ± 1.37</b>
	8	2.28 ± 0.99	1.77 ± 0.88	0.88 ± 0.68	<b>0.80 ± 0.63</b>
	9	2.71 ± 1.63	1.42 ± 1.13	1.64 ± 1.34	<b>0.60 ± 0.97</b>
	10	5.52 ± 2.91	1.05 ± 1.03	2.43 ± 1.31	<b>1.26 ± 1.18</b>
Low Quality	1	4.56 ± 3.19	4.28 ± 5.05	3.12 ± 3.20	<b>1.24 ± 1.27</b>
	2	4.12 ± 1.74	1.83 ± 1.27	1.73 ± 1.25	<b>1.55 ± 1.06</b>
	3	5.49 ± 4.17	2.18 ± 0.96	4.16 ± 2.96	<b>1.91 ± 1.13</b>
	4	7.24 ± 4.15	4.26 ± 2.77	5.08 ± 3.46	<b>4.00 ± 1.41</b>
	5	2.76 ± 1.49	2.69 ± 1.47	2.14 ± 1.42	<b>1.33 ± 1.08</b>
	6	2.98 ± 1.31	2.84 ± 1.45	<b>1.90 ± 0.73</b>	2.75 ± 1.45
	7	4.84 ± 2.89	3.02 ± 1.84	2.72 ± 2.23	<b>0.72 ± 0.92</b>
	8	4.45 ± 1.88	2.43 ± 1.56	1.88 ± 1.27	<b>1.68 ± 0.98</b>
	9	3.66 ± 1.62	2.59 ± 0.92	2.56 ± 1.09	<b>1.46 ± 1.04</b>
	10	2.62 ± 0.76	2.37 ± 1.41	<b>2.36 ± 1.42</b>	2.62 ± 1.60

voxel class probabilities, which are not directly available from the intensity values, are computed from an EM segmentation algorithm. These probabilities are combined with point-wise mutual information in a 3D joint distribution and serve as penalty terms to the image forces. Incorporation of this classification information can help to reduce the ambiguity of point matching

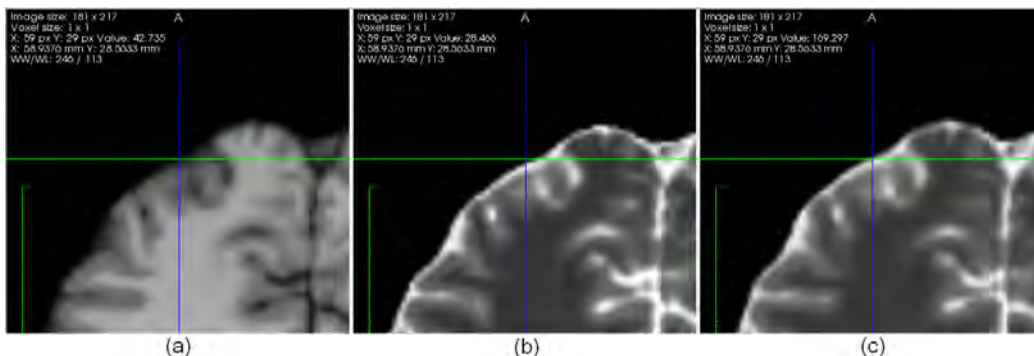


Figure 12: The crosses describe the same voxel on (a) T1 fixed image, (b) PMI result and (c) H-SPMI result. The intensities of this voxel represent CSF, PV and CSF in the three images, respectively.

since the correspondence information introduced by segmentation is not available in the pure intensity-based multimodal registration. As already shown in Fig. 3, H-SPMI provides consistently better alignment on the brain surface. This is because MI-based methods are often confused by the partial volume (PV) interpolation during the registration, whereas H-SPMI is able to distinguish between the PV and the real brain volume. An example is illustrated in Fig. 12, where the crosses describe the same voxel on all three images. In T1 reference image (Fig. 12(a)) and H-SPMI result (Fig. 12(c)), it represents a CSF voxel, whereas in PMI result (Fig. 12(b)) it has the value of a PV voxel. Apparently, in this case PMI confuses the PV voxel as being CSF, while SPMI is able to distinguish correctly.

In addition, using classification information allows us to perform hierarchical optimization based on different levels of anatomical structure, ensuring the registration accuracy and robustness at all levels of anatomical region of interest. In general, large shape variations are recovered at first level, while finer structures are taken into account at next levels. Fig. 13 shows one example of a result of registering a T2 image to a T1 image. At level 1, the registration uses classification between background and brain volume. Therefore, we see perfect contour alignment in result of level 1, while flaws exist in inner structures (see green circles in Fig. 13(b)). This is because at this level the SPMI forces in the inner structures can be regarded as quasi-PMI forces since only all inner structures are classified into only one class. At level 2, we see better alignment in the inner structures using classification of WM, GM and CSF. In addition, the hierarchical structure also allows the

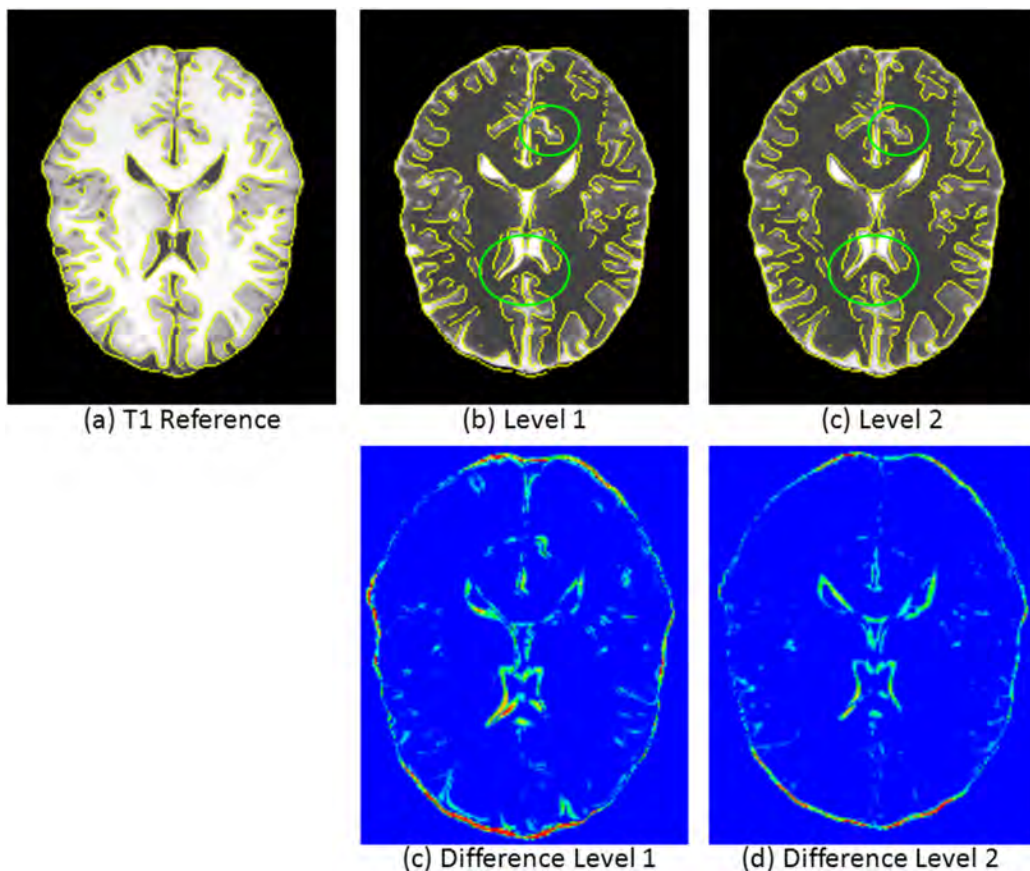


Figure 13: An example showing the registration output at (b) level 1 and (c) level 2 from the Brainweb experiment using H-SPMI. In the first row, the edge from (a) T1 reference image is extracted and overlaid on both outputs. The green circles show the major differences in internal structures between the outputs from two levels. The second row shows subtraction images of the outputs from (d) level 1 and (e) level 2.

use of the proposed method on different clinical applications, as shown in the second experiment where modalities representing various classifications can be registered using the proposed method. In addition, the hierarchical optimization provides the possibilities of increasing the registration accuracy in the lesions or abnormal regions in the brain (e.g. tumors or signal dropout) by adding an additional class.

As incorporation of additional classification information might render the registration sensitive to segmentation errors, we performed sensitivity test to see how the segmentation results affect the final registration accuracy. We

believe that it is an important aspect in order to bring the proposed technique into practice. The results indicate that when using the hierarchical model, the average RMS error increases slightly with worse segmentation, while robustness is still guaranteed showing low standard deviation. Nonetheless, the accuracy remains consistently high even with segmentation DICE coefficients at 0.7223, which can be achieved by the state-of-the-art EM segmentation method given an acceptable image quality. This fact is also confirmed in the EPI distortion correction experiment, in which two scenarios with different image qualities were created. Even though the performance of the proposed method is degraded with poor image quality, it still shows better results compared to PMI. This demonstrates that the proposed algorithm can generally provide significantly better accuracy than a pure MI-based method.

Regarding the optimization of the proposed method, different from the optimization of MI-based methods where intensity probability is estimated using kernel density functions such as Parzen window, we compute the joint probability and marginal probability of the voxel intensity based on the joint histogram. Although kernel density function is able to provide a differentiable solution, the precision of the estimation relies greatly on the number of samples and the parameters of the used function. In this work as the proposed SPMI metric is a combination of the intensity probability and joint classification probability computed from EM algorithm, it is not straightforward to apply the same strategy for the optimization of SPMI. Therefore, symmetric finite-difference based approach is adopted in this work to achieve not only the simplicity but also the consistency for the optimization.

On the other hand, one disadvantage of the proposed method compared to PMI is the computation cost. The PMI took on average 293 s on an Inter Xeon 3.2GHz computer for the synthetic data, whereas H-SPMI took 978 s due to the additional information added in the metric and anatomical region-based hierarchical optimization. However, MI-based Elastix also reaches 745 s in the same experiment.

Our future work will include further clinical validation. Moreover, it would also be interesting to investigate our approach for the problem of missing correspondences, especially in brain oncology where tumors with various shape exist.

## 5. Acknowledgments

This work is funded by the research grant from Swiss National Science Foundation 205321\_135361/1.

## References

- [1] P. Viola, W. M. Wells, Alignment by maximization of mutual information., in: Proc. of ICCV'95, 1995.
- [2] F. Maes, D. Vandermeulen, P. Suetens, Multimodality image registration by maximization of mutual information, *IEEE TMI* 16 (2) (1997) 187–198.
- [3] D. Rueckert, L. I. Sonoda, D. L. G. H. C. Hayes, M. O. Leach, D. J. Hawkes, Non-rigid registration using free-form deformations: Application to breast mr images., *IEEE TMI* 18 (8) (1999) 712C721.
- [4] G. Hermosillo, C. Chefd'hotel, O. Faugeras, Variational methods for multimodal image matching., *IJCV* 50 (3) (2002) 329–334.
- [5] M. Modat, T. Vercauteren, G. R. Ridgway, D. J. Hawkes, N. C. Fox, S. Ourselin, A diffeomorphic demons using normalized mutual information, evaluation on multimodal brain mr images, in: Proc. of SPIE'10, 2010.
- [6] H. Lu, M. Reyes, A. Serifovic, S. Weber, Y. Sakurai, H. Yamagata, P. C. Cattin, Multi-modal diffeomorphic demons registration based on point-wise mutual information, in: In Proc. of ISBI'10, 2010, pp. 372–375.
- [7] D. Loeckx, P. Slagmolen, F. Maes, D. Vandermeulen, P. Suetens, Non-rigid image registration using conditional mutual information, *IEEE TMI* 29 (1) (2010) 19–29.
- [8] C. Studholme, C. Drapaca, B. Iordanova, V. Cardenas, Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change, *IEEE TMI* 25 (5) (2006) 626.
- [9] X. Zhuang, D. J. Hawkes, S. Ourselin, Unifying encoding of spatial information in mutual information for nonrigid registration., in: Proc. of IPMI'09, 2009.



- [10] A. Biesdorf, S. Wörz, H.-J. Kaiser, C. Stippich, K. Rohr, Hybrid spline-based multimodal registration using local measures for joint entropy and mutual information, in: Proc. of MICCAI'09, 2009, pp. 607–615.
- [11] H. Lu, P. C. Cattin, M. Reyes, A hybrid multimodal non-rigid registration of mr images based on diffeomorphic demons, in: Proc. of EMBC'10, 2010.
- [12] E. D'Agostino, F. Maes, D. Vandermeulen, P. Suetens, An information theoretic approach for non-rigid image registration using voxel class probabilities., *MedIA* 10 (3) (2006) 413–431.
- [13] P. Rogelj, S. Kovacic, J. C. Gee, Point similarity measures for non-rigid registration of multi-modal data, *Computer Vision and Image Understanding* 92 (1) (2003) 112 – 140.
- [14] X. Chen, M. Brady, D. Rueckert, Simultaneous segmentation and registration for medical image., in: Proc. of MICCAI'04, 2004.
- [15] K. M. Pohl, J. Fisher, W. E. Grimson, A bayesian model for joint segmentation and registration., *Neuroimage* 31 (1) (2006) 228–39.
- [16] K. M. Pohl, S. Bouix, M. Nakamura, T. Rohlfing, R. W. McCarley, R. Kikinis, W. E. L. Grimson, M. E. Shenton, W. M. Wells, A hierarchical algorithm for MR brain image parcellation., *IEEE transactions on medical imaging* 26 (9) (2007) 1201–12.
- [17] P. Thevenaz, M. Unser, Optimization of mutual information for multiresolution image registration, *Image Processing, IEEE Transactions on* 9 (12) (2000) 2083 – 2099.
- [18] K. M. M. V. J. P. S. Klein, M. Staring, elastix: a toolbox for intensity based medical image registration, *Medical Imaging, IEEE Transactions on* 29 (1) (2010) 196 – 205.
- [19] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, A. C. Evans, BrainWeb: Online Interface to a 3D MRI Simulated Brain Database, *NeuroImage* 5 (1997) 425.
- [20] T. Vercauteren, X. Pennec, A. Perchant, N. Ayache, Diffeomorphic demons: efficient non-parametric image registration., *Neuroimage* (1 Suppl).

- [21] C. Seiler, X. Pennec, M. Reyes, Parametric regression of 3d medical images through the exploration of non-parametric regression models, in: Proc. of ISBI'10, 2010, pp. 452–455.
- [22] K. Van Leemput, F. Maes, D. Vandermeulen, P. Suetens, Automated model-based tissue classification of MR images of the brain, Medical Imaging, IEEE Transactions on 18 (10) (1999) 897–908.
- [23] G. E. Christensen, H. J. Johnson, Consistent image registration, IEEE Transaction on Medical Imaging 7 (2001) 568–582.
- [24] N. Tustison, B. Avants, P. Cook, Y. Zheng, A. Egan, P. Yushkevich, J. Gee, N4ITK: Improved N3 bias correction, Medical Imaging, IEEE Transactions on 29 (6) (2010) 1310–1320.
- [25] S. M. Smith, Fast robust automated brain extraction, Human Brain Mapping 17 (3) (2002) 143–155.