DIFFUSION WEIGHTED IMAGING DISTORTION CORRECTION USING HYBRID MULTIMODAL IMAGE REGISTRATION

Huanxiang Lu[1], *Philippe C. Cattin*[2], *Lutz-Peter Nolte*[1], *Mauricio Reyes*[1]

[1]Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland
[2]Medical Image Analysis Center, University of Basel, Switzerland

ABSTRACT

In this paper, we introduce a hybrid image registration approach for diffusion weighted image (DWI) distortion correction. General intensity-based multimodal registration uses mutual information (MI) as the similarity metric, which can cause matching ambiguities due to the intensity correspondence uncertainty in some anatomical regions. We propose to overcome such limitations by enhancing the registration framework with automatically detected landmarks. These landmarks are then integrated naturally into multimodal diffeomorphic demons algorithm using Gaussian radial basis functions. The proposed algorithm was tested with clinical DWI data, with results demonstrating that better distortion correction can be achieved using the hybrid algorithm as compared to using a pure intensity-based approach.

Index Terms— Image Registration, Diffusion Weighted Imaging, Landmark Detection

1. INTRODUCTION

Diffusion-weighted imaging (DWI) is a powerful imaging technique which is capable of providing variational information of the biological tissues through detection of water molecular diffusion. Most magnetic resonance (MR) diffusion studies require multiple images with which diffusion sensitization is in noncollinear directions. Such an analysis requires all the images to be spatially well aligned. However, the accuracy of the analysis is often restricted by geometrical distortions inherent to the pulse sequence called motion probing gradient (MPG). The eddy currents iduced by this sequence cause field inhomogeneities that lead to these spatial distrortions.

In order to facilitate further DWI analysis, image registration based distortion correction [1, 2, 3] has been widely researched and showed to be superior over the acquisition protocol-based methods and the field map-based methods. The principle is to align the images at each gradient direction to the undistorted B0 image. Since B0 image are often acquired with T2w sequences, multimodal registration is required. Most of the proposed methods are focused on affine registration using mutual information (MI). However, the eddy current distortion is not necessarily linear. Therefore, non-rigid registration is indispensable in this case.

In this paper, we extend our previous work on multimodal diffeomorphic demons registration [4] by incorporating salient landmarks. The landmarks are extracted automatically and serve as guidance for the voxel-wise matching. The energy function is governed by point-wise mutual information, landmark information and a regularization term. The optimization of this energy results in a combination of a dense vector field and Gaussian basis functions for landmarks in a space of diffeomorphic transformations. The methodology of the whole framework, including the detection and integration of the landmarks, is elaborated in Section 2, followed by the experiments and results on clinical DWI datasets in Section 3. Advantages and limitations of this work are discussed in Section 4.

2. HYBRID MULTIMODAL IMAGE REGISTRATION

2.1. Why Hybrid Registration?

In general, non-rigid registration algorithms can be classified into two categories in terms of the used image information: landmark-based and intensity-based. The landmark-based methods define a unique smooth transformation based on corresponding landmarks, which are usually located in anatomically salient regions. The correspondence of the points away from the landmarks is defined by a certain interpolation method. Landmark-based approaches ensure accurate registration in salient regions, and they are usually computational efficient and capable of handling large geometrical deformations. Nevertheless, they do not guarantee overall voxel to voxel correspondences and the result can be significantly affected by the choice of the landmarks. Therefore, more attention has been recently paid to intensity-based approaches, where the transformation is directly computed from the intensity information of all the voxels in the image data. Thus, the intensity-based methods can often achieve globally better accuracy. However, one major limitation is the insufficient information used in this approach for voxel-wise matching, which in general causes matching ambiguity. The ambiguity is especially severe with mutual information-based multimodal registration, where voxel matching is established only depending on the statistical relation of the intensity distribution in two images. Therefore, more information needs to be incorporated to make the voxel distinct to match.

Consequently, it is natural to think of combining the advantages of those two approaches. Although substantial efforts has been put into hybrid registration approaches during the past years [5, 6, 7, 8, 9, 10, 11], only few [8, 9] can deal with multimodal images. However, these methods require manually annotated landmarks, which is a tedious, time-consuming task. In this work, automatic landmark detection is achieved for the sake of efficiency and clinical applicability. The detected landmarks are integrated to the multimodal demons registration algorithm in a diffeomorphic space. Details are given in the following subsections.

2.2. Landmark Detection

Given two images F and M, the goal of landmark detection is to find a set of points $\mathbf{q} \in F$ and their correspondences $\mathbf{r} \in M$, which should be anatomically meaningful in order to serve as a guide to improve the registration accuracy. The proposed landmark detection scheme is divided into two steps: (1) detecting of salient landmarks in the reference image F, (2) locating their corresponding points in target image M.

Scale saliency [12] is adopted to detect landmarks in F. It is a geometrical measure to detect salient regions with a complex structure and bases on information theoretical idea that higher entropy indicates more complex structures. Scale saliency is defined by the weighted entropy across scale and position:

$$Y(\mathbf{z}, x) = H(\mathbf{z}, x) \times W(\mathbf{z}, x), \tag{1}$$

where $Y(\mathbf{z}, x)$ is the scale saliency value of a point x at scale $\mathbf{z}$; H and W denote the entropy and the weighting function respectively, which are defined as:

$$H(z, x) = -\sum_{x \in F} p(z, x) \log_2 p(z, x) \tag{2}$$

$$W(z, x) = z \cdot \sum_{x \in F} |p(z, x) - p(z-1, x)| \tag{3}$$

where p is the probability density as a function of point x at scale z. The scale vector $\mathbf{z}$ at which entropy is peaked is defined as:

$$\mathbf{z} = \left\{ z : \frac{\partial H(z, x)}{\partial z} = 0, \frac{\partial^2 H(z, x)}{\partial z^2} < 0 \right\} \tag{4}$$

In brain images, anatomical landmarks are usually located at the regions with complex structure such as sulci, vertices of the ventricle, etc. Thus, the points in F with high scale salient responses are considered as the anatomical landmarks. Moreover, to encourage uniform registration accuracy in the image, it is desirable that the detected landmarks are homogeneously distributed within the image space. Hence, instead of a global thresholding and clustering as in [12], we first partition the image space into N regions, in each region the K most salient candidates were selected as shown in Figure 1. Then, in order to keep each landmark distant enough from the others, a k-means clustering is performed on those candidates. For each cluster, the most scale salient point is taken as a landmark.

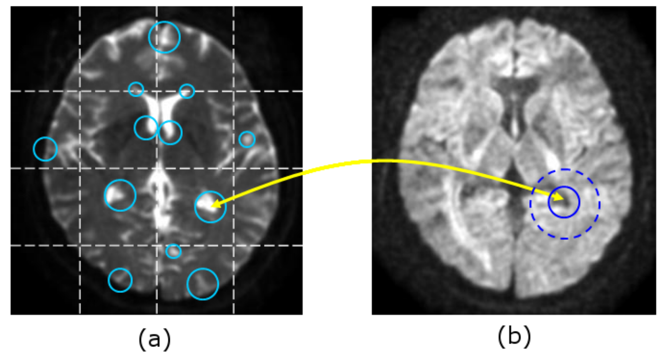

Fig. 1: (a) The reference image space is partitioned into regions where different scales of the salient landmarks are detected. (b) The corresponding landmark in the target image is detected by block matching in a local search space (dashed circle)

To find the correspondences of the landmarks $\mathbf{q}$ in the target image M, we employ block matching algorithm [13] due to its robustness and efficiency. As the salient scale z of each landmark $\mathbf{q}$ is known, the correspondence pairing is done by optimizing a cost function to find a set of n transformations $\mathbf{T}_i$, $i = 1, \ldots, n$ between the window functions of $\mathbf{q}_i$ and $\mathbf{r}_i$:

$$Q(\mathbf{T}_i) = MI(D_F(\mathbf{q}_i, z), \mathbf{T} \circ D_M(\mathbf{q}_i, z)), \tag{5}$$

where D_F and D_M are the window functions of $\mathbf{q}$ and $\mathbf{r}$; MI denotes mutual information measurement for its capability of coping with multimodal matching. The correspondence $\mathbf{r}_i$ can then be easily obtained by transforming $\mathbf{q}_i$ by $\mathbf{T}_i$.

2.3. Integration to Intensity-Based Registration

The detected landmarks are then integrated with multimodal diffeomorphic demons algorithm, which is herein utilized as the intensity-based method. This registration algorithm can be summarized by a model with an energy consisting of a similarity function, a transformation error function and a regularization term. The diffeomorphism is ensured by mapping the update field at each iteration through the exponential operation on the Lie group. In order to cope with multimodal image registration, we previously extended this framework by adopting point-wise mutual information (PMI) as the similarity metric [4], since PMI calculates the voxel-wise contribution of the global MI and thus is easy to integrate in a dense field approach. Given a deformation field s, the energy function can be described as follows:

$$E(c,s) = Sim(F, M \circ c) + \sigma \|s - c\|^2 + \sigma\lambda Reg(s), \quad (6)$$

where Sim is the intensity similarity metric defined as:

$$Sim(F, M \circ c) = -\log\left(\frac{p(i_F, i_{M\circ c})}{p(i_F)p(i_{M\circ c})}\right), \quad (7)$$

Reg is the regularization term which is typically a Gaussian kernel, and c is the estimated transformation according to the metric.

To incorporate the detected landmarks, we reformulate the demons energy function by introducing the energy term for landmarks. Different to our previous work [14], a correspondence term $\sigma \|c - l\|^2$ is added, so that the landmark energy can be optimized together with the update field u on the Lie group to ensure the diffeomorphism:

$$\begin{aligned} E(c,s) =& Sim(F, M \circ c) + \sigma \|s - c\|^2 + \sigma \|c - l\|^2 \\ &+ \sigma\gamma \|s - l\|^2 + \sigma\lambda Reg(s), \end{aligned} \quad (8)$$

where l is the estimated transformation according to the landmark. Given n corresponding landmarks $\mathbf{q}_i$ and $\mathbf{r}_i$, $i = 1, \ldots, n$ that have been localized in F and M. The energy term with respect to the landmarks can be written as:

$$\sigma\gamma \|s - l\|^2 = \sum_{i=1}^{n} (\mathbf{r}_i - \mathbf{q}_i \circ s)^2. \quad (9)$$

The optimization of this modified energy function with respect to c, l, s leads to the following steps:

1. Minimizing $\sigma\gamma \|s - l\|^2$ with respect to l. This is easily done by guiding the moving landmarks towards the reference landmarks.

2. Find the correspondence c of the dense field by minimizing $Sim(F, M \circ c) + \sigma \|s - c\|^2 + \sigma \|c - l\|^2$ with respect to c. c turns out to be a combination of update field u and splines of Gaussian radial basis functions, which is projected on the Lie group through the exponential map:

$$c(x) = s(x) \circ \exp\left(u(x) + \sum_{i=1}^{n} \alpha_i G(x - \mathbf{q}_i)\right), \quad (10)$$

   where G is the Gaussian kernel, α is a scalar value and α_i is a vector of weighting parameters.

3. Find the estimated transformation s by minimizing $\sigma \|s - c\|^2 + \sigma\lambda Reg(s)$ with respect to s.

3. RESULTS

To validate our distortion correction approach, we applied the proposed algorithm to six clinical DWI datasets. Each dataset consists of one T2w image with $\mathrm{b} = 0\ \mathrm{s/mm}^2$ and six T1w images with $\mathrm{b} = 500\mathrm{s/mm}^2$ with different gradient orientations. The resolution of the images is $384 \times 384 \times 44$ with voxel spacing of $0.5365 \times 0.5365 \times 3\ \mathrm{mm}^3$. The evaluation of our hybrid approach is decoupled into two parts:

Landmark Correspondence Accuracy

We first examined the accuracy of landmark correspondences since it is crucial for the whole registration algorithm. Typically, about 70 landmarks were detected through all the slices in both images. For quantitative validation, we manually annotated the correspondencing landmarks in the target image as the "bronze standard" for the list of detected landmarks in reference image. Table 1 shows the results between correspondences detected by our matching algorithm and the bronze standard. The mean error is around $1\ \mathrm{mm}$, while the maximum error reaches more than $4\ \mathrm{mm}$ due to the missing correspondences in some regions.

Table 1: Mean Error Between Detected Correspondences and Bronze Standard [mm]

Dataset	1	2	3	4	5	6
Mean	0.86	0.96	0.84	0.94	1.06	1.31
Max	2.91	3.20	2.65	2.74	3.77	4.80

Registration Accuracy

A common difficulty in validating results from non-rigid image registration algorithms, is the lack of ground truth. Therefore, we are limited to using indirect measures to establish the reliability of our variational approach. In this case, we first performed visual inspections of the registration results by superimposing the contour of the B0 image onto the other images. The eddy current distortion is observed in the occipital lobe in Figure 2(b), while Figure 2(c) shows that the image was well corrected as the contours mostly correspond with the B0 image.

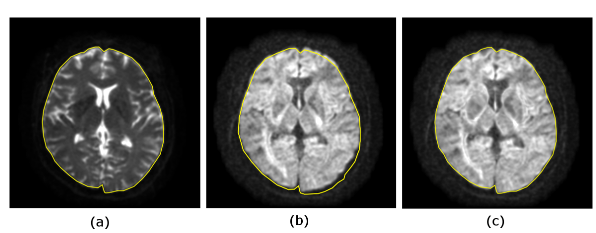

Fig. 2: T2w B0 image contour (yellow curve) overlays on (a) B0 image (b) initial distorted image and (c) registered image

For quantitative analysis, 15 pairs of corresponding anatomical landmarks were also identified in both images to evaluate the geometric registration error. All the dataset were tested using the hybrid approach as well asthe intensity-based method. From Figure 3, one can see that the automatic hybrid approach outperforms the intensity-based in terms of mean error and maximum error thanks to the additional information provided by the landmarks.

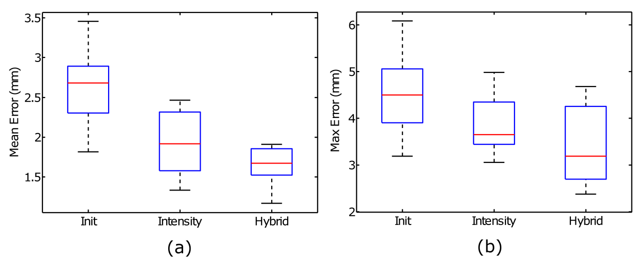

Fig. 3: (a) Mean error and (b) maximum error between the "bronze standard" and different registration results.

4. DISCUSSION

In this paper we proposed a novel approach for DWI distortion correction using hybrid multimodal registration. The algorithm couples the automatically extracted landmarks and the multimodal demons registration in a diffeomorphic space. Due to the richer information used in the landmark detection, voxels in salient regions can establish more reliable correspondences than just using point-wise mutual information, therefore potential local-minima can be avoided. Besides, it is always flexible to incorporate different information in landmark detection for task-specific registration problems.

Experiments on clinical data demonstrate that the distortion is better corrected using the proposed hybrid approach than using a pure intensity-based method. On the other hand, due to the fact that the accuracy of the registrationsignificantly depends on the landmark detection, the applicability of this approach could be restrained for tackling problems like missing correspondences. Therefore, future works will be devoted to improve the landmark detection in terms of accuracy but more importantly robustness. Moreover, the successful DWI distortion correction opens a wide range of applications.

5. REFERENCES

[1] G. Rohde, A. Barnett, P. Basser, S. Marenco, and C. Pierpaoli, "Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI," *Magn Res Med*, vol. 51, 2004.

[2] X. Tao and J. Miller, "A method for registering diffusion weighted magnetic resonance images," in *Proc. of MICCAI'06*, vol. 4191, pp. 594–602.

[3] J-F. Mangin, C Poupon, C Clark, D Le Bihan, Bihan D Le, and I Bloch, "Distortion correction and robust tensor estimation for MR diffusion imaging," *Med Imag Anal*, vol. 6, no. 3, pp. 191–8, 2002.

[4] H. Lu, M. Reyes, A. Serifovic, S. Weber, Y. Sakurai, H. Yamagata, and Ph. C. Cattin, "Multi-modal diffeomorphic demons registration based on point-wise mutual information," in *In Proc. of ISBI'10*, pp. 372–375.

[5] P. Cachier, J-F. Mangin, X. Pennec, D. Riviere, D. Papadopoulos, J. Regis, and N. Ayache, "Multisubject non-rigid registration of brain MRI using intensity and geometric features," in *In Proc. of MICCAI'01*, pp. 734–742.

[6] H. J. Johnson and G. E. Christensen, "Consistent landmark and intensity-based image registration," *MedImg*, vol. 21, no. 5, pp. 450–461, May 2002.

[7] K. Rohr, P. Cathier, and S. Worz, "Elastic registration of electrophoresis images using intensity information and point landmarks," *Pattern Recognition*, vol. 37, no. 5, pp. 1035 – 1048, 2004.

[8] A. Azar, C. Xu, X. Pennec, and N. Ayache, "An interactive intensity- and feature-based non-rigid registration framework for 3d medical images," in *Proc. of ISBI'06*, pp. 824–827.

[9] A. Biesdorf, S. Wörz, H-J. Kaiser, C. Stippich, and K. Rohr, "Hybrid spline-based multimodal registration using local measures for joint entropy and mutual information," in *Proc. of MICCAI'09*, pp. 607–615.

[10] T. Hartkens and A. J. Martin, "Using points and surfaces to improve voxel-based non-rigid registration," in *Proc. of MICCAI'02*, 2002, vol. 2489, pp. 565–572.

[11] A. Sotiras, Y. Ou, B. Glocker, C. Davatzikos, and N. Paragios, "Simultaneous geometric - iconic registration," in *Proc. of MICCAI'10*, pp. 676–683.

[12] T. Kadir and M. Brady, "Saliency, scale and image description," *Int J of Computer Vision*, vol. 45, no. 2, pp. 83–105, 2001.

[13] S. Ourselin, A. Roche, S. Prima, and N. Ayache, "Block matching: A general framework to improve robustness of rigid registration of medical images," in *Proc. of MICCAI'00*, pp. 373–373.

[14] H. Lu, P. C. Cattin, and M. Reyes, "A hybrid multimodal non-rigid registration of mr images based on diffeomorphic demons," in *Proc. of EMBC'10*.