

Multichannel Image Registration using Gabor Wavelet Transform

Jisha.M.R

Department of Electronics & Communication, Govt.College of Engineering/CUSAT, India

Abstract: This paper proposes a multichannel image registration method which combines information from different modalities/channels to produce a unified joint registration. Multichannel images are created by co-registering multimodality images of the same subject to utilize information across modalities comprehensively. The existing methods combine the information at the image/intensity level. But the proposed method uses feature-level information fusion method to combine the complementary information from different modalities that characterize different tissue types. This is through Gabor wavelets transformation and Independent Component Analysis (ICA), to produce an inter-subject registration. Experiments on both simulated and real multichannel images illustrate the applicability and robustness of the proposed registration method that combines information across modalities.

Keywords: Deformable registration, diffusion tensor imaging (DTI), independent component analysis (ICA), information fusion, Gabor filter, multichannel image registration.

I. Introduction

The advancements in technology made more and more imaging modalities to be available for clinical/research studies such as X-ray, computed tomography (CT), positron emission tomography (PET) and magnetic resonance imaging (MRI). In MRI, different protocols (T1, T2, FLAIR, mPRAGE and diffusion tensor imaging (DTI), etc.) can also be viewed as different modalities. Each of these modalities provides some unique and often complementary characterization of the underlying anatomy and tissue microstructure. For example, structural T1 images provide contrast between the gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). On the other hand, DTI captures the directional micro-structural information within WM as a tensor, which complements the missing orientation information in structural images. We define a multichannel image for each subject as the co-registered collection of all single modality images that represent the same anatomy. Analysis of such images will be referred to as multichannel image analysis (MIA), which has become an active area of research with the growth in the number of imaging techniques, that provide different characterization for the same anatomy.

Specifically, using a true inter-subject multichannel registration algorithm has several benefits. 1) By combining complementary information from different modalities which characterize different aspects of the anatomy, the multichannel image registration is expected to be more robust and accurate. 2) The deformation field obtained is common to all the modalities, therefore the spatially normalized images of all modalities can be jointly statistically analyzed, in contrast to using separate registrations for each of the modalities which usually generates inconsistent deformation fields. 3) Although different modalities can provide complementary information, a considerable portion of the information could be overlapping and hence “redundant” or less reliable (such as the WM information from T1 in comparison to that from DTI). Therefore, by controlling the redundancy appropriately, the multichannel image registration can be more efficient and reliable. Multichannel inter-subject registration is rendered challenging because there could be competing information from different modalities. I propose a multichannel inter subject registration that combines information across modalities in such a way as to alleviate these challenges and provide an effective joint registration.

More formally, we define a multichannel image as a set of images from different modalities or scanning parameters, which depict the same subject: $\mathbf{I}=[I_1, \dots, I_n]$, where n is the total number of modalities in the multichannel image. We assume that, all the image channels I_1, \dots, I_n , have been co-registered to each other through mutual information (MI)-based intra-subject multimodal registration method, such as [1]. Although there are several established methods for single modality deformable registration, to the best of my knowledge, methods for multichannel image registration remain scarce in the literature, mainly owing to the challenge of combining information effectively. In [2], Park *et al.* proposed a deformable multichannel image registration method using multichannel demons [3], in which, the multichannel image consists of a T2-weighted image and other scalar images derived from DTI. After these scalar images $I(x)$ and $J(x)$ were co-registered within each subject, two multichannel images and were registered as multidimensional vector images using the metric, $D(\mathbf{I}(\mathbf{x}_1), \mathbf{J}(\mathbf{x}_2)) = \sqrt{\sum_{i=1}^n d(I_i(\mathbf{x}_1) - J_i(\mathbf{x}_2))^2}$ where d is the scalar image metric. In this method, every image channel is assigned equal importance in the image matching at every voxel. More recently, in [4],

multiple channels consisting of T1, DTI and cerebral blood flow (CBF) image were registered by defining the multichannel similarity metric as a weighted summation of each modality's similarity metric: $D(I(x_1), J(x_2)) = \sqrt{\sum_{i=1}^n w_i \cdot d_i (I_i(x_1) - J_i(x_2))^2}$ where w_i and d_i are the weights and similarity metric for the i th image modality.

The main challenge associated with multichannel image registration is how to best fuse the multimodal information to guide the registration, so that competing and redundant information is minimized, and complementary information is emphasized. The meaning of "fuse" is two-fold here. Firstly, on different tissue types, it means to adaptively select the features that best characterize the underlying anatomical structure. In this way, the complementary information from different modalities are merged and, thereby, making the registration more accurate. Secondly, it aims at removing the redundant/less-reliable information. For a particular tissue type, usually, there is one or several modalities that give relatively more reliable information. From another point of view, removing redundancy may also be understood as alleviating the possible "contradictoriness" between different modalities. In such cases, the most reliable modality is selected, omitting the less reliable. For example, any GM voxel will be better characterized by T1 than by fractional anisotropy (FA) measures obtained from DTI. Adding all the FA information to characterize GM may confuse the registration with CSF or fiber crossings. So losing the redundant part of FA information for this voxel may actually enhance the matching ability of this voxel. Therefore, to improve the efficiency and reliability of the multichannel image registration, for each tissue type, it is beneficial to only keep the most pertinent information from all the modalities, and discard the part that overlaps or less reliable. Such information will be referred to as "redundant" in the paper.

In contrast to the existing methods that are trying to solve the challenging information fusion problem in multichannel image registration at image/intensity level, in this paper, I propose to fuse the multimodal information using Gabor Wavelets transformation. The main drawback of dealing with the information fusion problem at image/intensity level is that, it is difficult to extract and distinguish between complementary and redundant information, since for a specific tissue type, a higher intensity value does not necessarily indicate a stronger tissue characterization ability which reflects how a modality can discriminate one tissue type from another. Here an example is specified to explain this. Fig. 1 shows the co-registered PD, T2 and T1 images of a single subject and the intensities have been normalized into the same range a, b and c are three corresponding points on WM from the three different modalities, respectively.

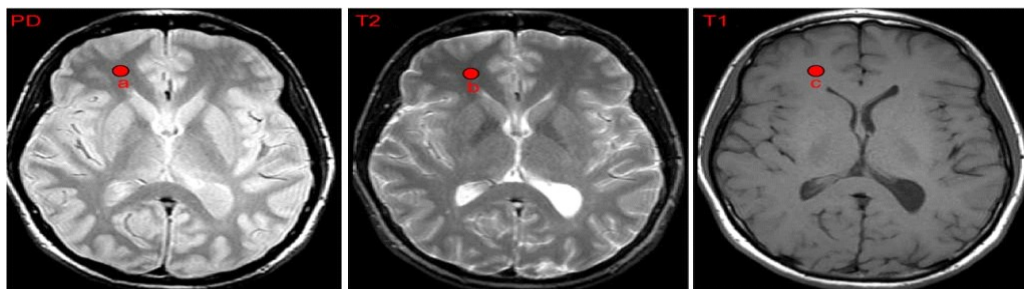


Fig. 1. Characterization ability of modalities. Images from three different modalities of the same subject are co-registered, in which, (a), (b), and (c) show the corresponding location on WM. The intensity values of the three points are related as: $b < a < c$. However, the PD image gives better characterization of the underlying WM structure, because it gives better contrast between WM and the surrounding tissues, although a has neither the lowest, nor the highest intensity value.

To measure the characterization ability, multiresolution analysis methods, such as, discrete wavelets transform (DWT) or Gabor wavelets transform have been shown to be appropriate tools as the strength of the output signal from the high-pass (wavelets) or band-pass (Gabor) filter, which reflects the contrast level, is a reliable indicator of the characterization ability at different scales and locations. Based on this idea of measuring the characterization ability, multiscale decomposition (MSD)-based information fusion scheme was proposed and has been successfully used in image fusion. In this paper, by adopting a similar information fusion scheme based on Gabor wavelets transform, a novel multichannel image registration algorithm is proposed. The results show that the proposed multichannel image registration method can integrate and enhance complementary information, while eliminating the less reliable/redundant information from different channels and leads to a more accurate and robust inter-subject registration at low computational cost.

II. Proposed Multichannel Image Registration Method

Every feature-based registration method requires feature extraction and a definition of a metric that will help match points and regions between the two images, based on these features. This is then incorporated into an optimization framework with appropriate smoothing terms. Before going into the details, a schematic diagram of the entire framework is provided in Fig. 2 and explained as the following:

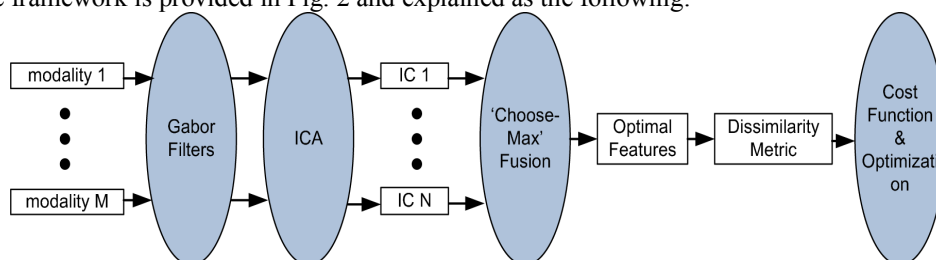


Fig.2 Flowchart of the proposed multichannel information fusion scheme and the registration framework

2.1 Feature Extraction and the Measure of Tissue Characterization Ability:

For the multiscale decomposition method, we use Gabor wavelets [5], as Gabor wavelet transformation has been shown to be optimal in the sense of minimizing the joint uncertainty in space and frequency, and has been widely used for feature extraction [6], and hence, more appropriate for the purpose of matching/registration. At the same time, just like coefficients of DWT, the magnitude of the coefficients of Gabor wavelets is also a reliable measure of the characterization ability of each modality. Thus Gabor wavelets are a better choice for both feature extraction/matching and tissue characterization ability measurement, for our proposed information fusion-based multichannel image registration.

Although the registration is for 3D images, to alleviate the computational cost, we use 3 perpendicular (axial, coronal, and sagittal) 2D Gabor filter banks to extract the features. A 2D Gabor filter can be viewed as a sinusoidal plane of particular frequency and orientation, modulated by a Gaussian envelope: $G(x,y)=s(x,y).g(x,y)$, where $s(x,y)$ is a complex sinusoid: $s(x,y)=\exp[-j2\pi(u_0x+v_0y)]$ and $g(x,y)$ is a 2D Gaussian envelope:

$$g(x, y) = \frac{1}{\sqrt{2\pi\sigma_x\sigma_y}} \exp \left[-\frac{1}{2} \left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2} \right) \right]$$

Where σ_x and σ_y characterize the spatial extent and bandwidth of g along the respective axes, u_0 and v_0 are the shifting frequency parameters in the frequency domain. Using $G(x,y)$ as the mother wavelet a class of self-similar functions can be obtained by appropriate dilations and rotations of $G(x,y)$ through:

$G_{m,n}(x, y) = a^{-m}G'(x', y')$, Given an image $I(x,y)$, the Gabor transform with orientation n and scale m can be computed as:

$$F_{m,n}(x, y) = \int I(x_1, y_1)G_{m,n}^*(x - x_1, y - y_1)dx_1dy_1$$

Where * indicates the complex conjugate.

2.2 Tissue-Pertinent Information Extraction Through Independent Component Analysis

The “Choose-Max” criteria is directly applied between different modalities so that, the residual useful information (usually a small amount) from the other “non-max” modality will be discarded. In some cases, if information about certain tissue type is quite evenly distributed among different modalities, the cost of this simple “Choose-Max” scheme would be high. To solve this problem and facilitate the subsequent information fusion, for each tissue type, it would be very beneficial to merge all the pertinent information of this tissue type from all modalities. In the proposed framework, independent component analysis (ICA) is adopted to achieve this goal. By applying ICA to the Gabor features extracted from different modalities, each of the resultant independent components (ICs) combines all the information regarding a particular tissue type in the feature space. In other words, after ICA, the information pertaining to different tissue types is “concentrated” into different ICs, and the input information is no longer separated by their sources (modalities), but by their targets (different tissue types). This information “repartitioning” is important to the subsequent “Choose-Max” step.

Specifically, to apply ICA to the analysis of multichannel images, there are two steps: training and decomposition. In the training step, a number of voxels are randomly selected from a few sample multichannel

images and input to the ICA algorithm. After the above iterative optimization, when the algorithm converges and the error is below a specified threshold, the obtained decomposition matrix can be applied to decompose the new multichannel image into independent component images, each of which highlights one of the tissue types. This decomposition process can be viewed as an information repartitioning process. Before decomposition, each image channel contains information about a combination of many tissue types. However, after decomposition, each independent component image is “specialized” for capturing one tissue type and contains all the relevant information (taken from all the modalities) pertaining to that tissue type. Because of this “purity,” the decomposed independent component image provides higher contrast and better characterization than the original image. It is worthwhile to clarify that although each tissue type will be highlighted in one IC, using ICA here is not expected to produce tissue segmentation, but to merge all the pertinent information regarding one tissue type into an IC. Therefore, the correspondence between each IC and the segmentation label, such as WM, GM, or CSF is not needed in our work, as the subsequence “choose-max” algorithm dose not needs this information. In this work, instead of applying ICA to the multichannel image, we apply ICA to the multivariate Gabor features, since our goal is to separate the information in Gabor feature space according to different tissue types. For each scale and orientation, we perform an ICA on all the Gabor features obtained from different modalities.

2.3 Information Fusion

As the underlying tissue type could be different for different voxels, different independent components (IC) are needed to acquire the best characterization. Therefore, “Choose-Max” scheme [7] is adopted to select the optimal IC according to the underlying tissue type. In this spatially-adaptive manner, the complementary information is kept and enhanced while less reliable/redundancy is reduced, via information fusion.

After the ICA step, each IC of Gabor features is “specialized” in depicting one particular tissue type. Therefore, by using the Choose-Max scheme on each voxel, we can choose the optimal Gabor features from the corresponding IC to characterize the underlying structure. Suppose we have two multichannel images: $I=[I_1, \dots, I_M]$ and $J=[J_1, \dots, J_M]$, in which I_1, \dots, I_M and J_1, \dots, J_M are two collections of co-registered images generated by different modalities. For each of the image channels, the raw Gabor features are extracted on ‘S’ scales and ‘O’ orientations. For orientation ‘n’ and scale ‘m’, after ICA, the i th IC of the Gabor features of I_i and J_i are denoted as

$F_{mn}^{I_i}$ and $F_{mn}^{J_i}$ respectively. Using the choose-max scheme, the dissimilarity between two voxels $I(x_1)$ and $J(x_2)$ can be measured as:

$$D_1(I(x_1), J(x_2)) = \sqrt{\sum_{m=1}^S \sum_{n=1}^O \sum_{i=1}^M w_{m,n,i} \left(\|F_{m,n}^{I_i}(x_1)\| - \|F_{m,n}^{J_i}(x_2)\| \right)^2}$$

where

$$w_{m,n,i} = \begin{cases} 1, & \text{if } i = \operatorname{argmax}_{j=1 \dots M} \|F_{m,n}^{I_j}(x_1)\| \\ 0, & \text{otherwise} \end{cases}$$

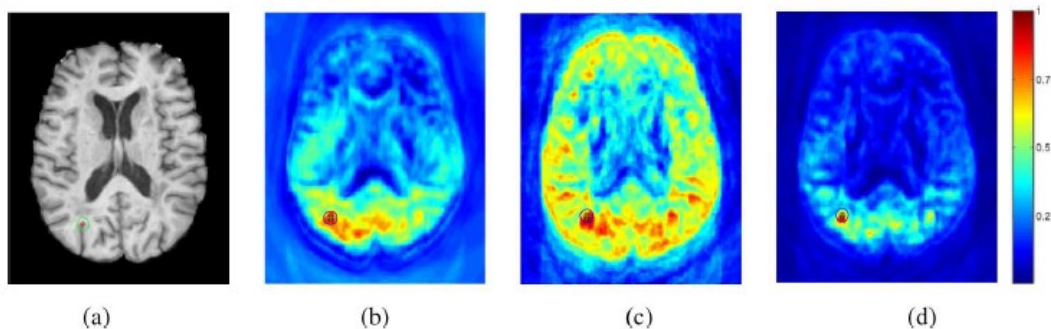


Fig.3 Similarity map of different metric

- a) It shows the reference point (denoted by a red cross) on T1.
- b) It is the similarity map generated by D1
- c) It is the similarity map generated by metric (normalized to [0 1]).

d) It is the similarity map generated by metric D
 Red indicates high similarity. Metric D1 gives more specific similarity map than the coarse metric L . By combining the two metrics in a coarse-to-fine manner, metric D gives the most specific similarity map which indicates the highest discriminative ability.

2.4 Dis-Similarity Metric

For each voxel, based on the optimal IC obtained through ICA and “Choose-Max” scheme, a dissimilarity metric is defined to find the correspondence between two multichannel images.

Let $P_{m,n}^I(\mathbf{x})$ denote the label of the IC which generates the strongest Gabor feature at scale m and orientation n on $I(\mathbf{x})$. Its then arranged in matrix form: $\mathbf{P}^I(\mathbf{x}) = [p_{1,1}^I(\mathbf{x}), \dots, p_{S,O}^I(\mathbf{x})]$. Let $L(I(\mathbf{x}_1), J(\mathbf{x}_2))$ be the number of equal labels between $\mathbf{P}^I(\mathbf{x}_1)$ and $\mathbf{P}^J(\mathbf{x}_2)$. L is a similarity metric with increasing resolution as S and O increase. The three perpendicular 2D Gabor filter banks each have four scales and six orientations. Therefore, L has the ability to differentiate the brain tissue into $4*6*3=72$ classes. By combining, our metric for multichannel image matching is finally defined as:

$$D(I(\mathbf{x}_1), J(\mathbf{x}_2)) = \left[1 - \frac{L(I(\mathbf{x}_1), J(\mathbf{x}_2))}{3SO + 1} \right] \cdot D_1(I(\mathbf{x}_1), J(\mathbf{x}_2))$$

Where D is defined as a coarse-fine metric , L gives a coarse/discrete similarity measure , D_1 refines the measure.

2.5 Cost Function and Optimization

Finally, based on the dissimilarity metric, the image matching problem is formulated as a cost function. By optimizing it,[8] the deformation field of the registration is obtained.. We will illustrate each step with an example of registering T1+ DTI multichannel images.

III. Comparisons

For a clearer comparison, the errors on GM and WM are shown separately

- (a) The proposed method (labeled as T1+ DTI) gives the smallest registration error on GM amongst all the 6 methods. This result shows that fusing information from DTI with T1 improves the registration even in the GM
- (b) The result on WM. When compared with the method using Gabor features only from T1 or DTI, it is found that, by using the complementary information from the T1-weighted image, the registration accuracy is improved on WM as well.

DROID gives the best accuracy on WM, as the features that DROID uses are WM specific and it utilizes the orientation information.The proposed method with the fused information gives better accuracy than most of the methods (except DROID which is slightly better) using single channel

Fig 4. Comparison of the registration accuracy on simulated images

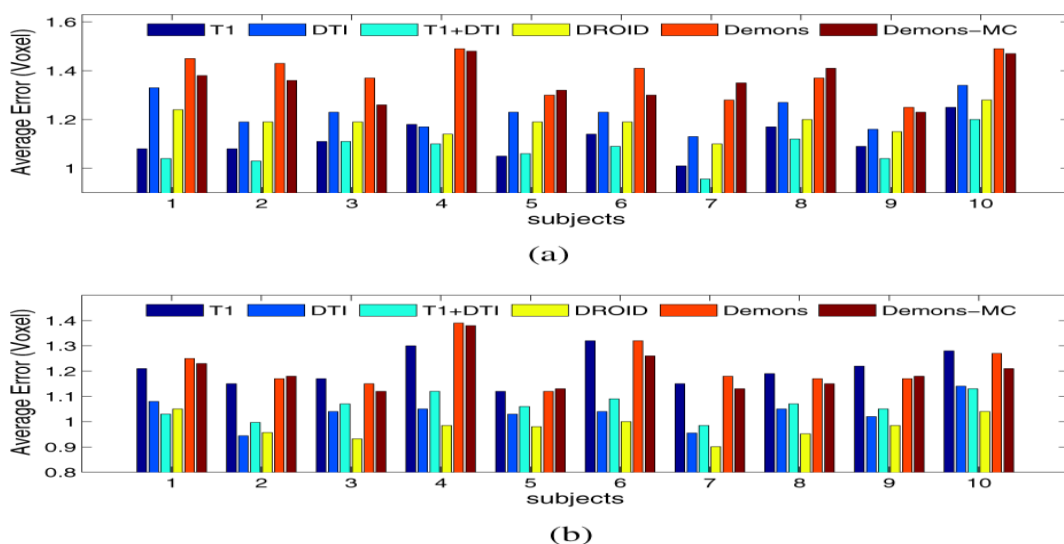


Table 1

SUMMARIZATION OF THE EXPERIMENT RESULTS ON SIMULATED IMAGES

Methods	T1	DTI	T1+DTI	DROID	Demons	Demons-MC
mean \pm std (voxel)	1.18 \pm 0.07	1.09 \pm 0.06	1.05 \pm 0.05	1.05 \pm 0.04	1.27 \pm 0.08	1.20 \pm 0.13
min (voxel)	1.09	1.01	0.97	0.97	1.16	1.10
max (voxel)	1.27	1.20	1.15	1.12	1.51	1.47

IV. Conclusion

This paper, proposes a multichannel inter-subject image registration framework that combines information from different modalities based on feature-level information fusion. The registration produces spatially normalized images of all the modalities acquired in the study. Thus the statistical analysis done on these jointly spatially normalized images is more comparable, as a unified registration scheme has been used to register them. I expect the proposed registration method to be very useful in large population clinical studies that acquire several modalities and a unified spatial normalization is needed for subsequent statistical analysis. Although the method is general and is applicable to any number of modalities on which features can be computed, here it is applied to the joint registration of T1 and DTI images, which are routinely acquired in all clinical studies. Experiments on both simulated and real multichannel images DTI+ T1 illustrate that the proposed method can effectively fuse the information from different modalities and result in a more accurate and robust registration.

Acknowledgement

The author would like to acknowledge the help from the work of Yang Li and Ragini Verma

References

- [1] M. Jenkinson, P. Bannister, J. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, 2002.
- [2] H.-J. Park, M. Kubicki, and M. Shenton, "Spatial normalization of diffusion tensor MRI using multiple channels," *Neuroimage*, vol. 20, pp. 1995–2009, 2003.
- [3] J. P. Thirion, "Non-rigid matching using demons," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 1996, pp. 245–251.
- [4] B. Avants, J. Duda, H. Zhang, and J. Gee, "Multivariate normalization with symmetric diffeomorphisms for multivariate studies," in *Med. Image Computing Computer-Assisted Intervent. (MICCAI)*, 2007, pp. 359–366.
- [5] J. Daugman, "Complete discrete 2D Gabor transforms by neural networks for image analysis and compression," *IEEE Trans. Acoustics*, vol. 36, no. 7, pp. 1169–1179, Jul. 1988.
- [6] W. Ma and B. Manjunathi, "Texture features for browsing and retrieval of image data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 8, pp. 837–842, Aug. 1996.
- [7] A. Hyvarinen, J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001
- [8] D. Shen, "Image registration by local histogram matching," *Pattern Recognit.*, vol. 40, pp. 1161–1171, 2007.

AUTHOR



JISHA.M.R received B.Tech in Biomedical Engineering from Sahrdaya College of Engineering & Technology under Calicut University, Thrissur, Kerala, India. She has got 1 year of teaching experience as Lecturer in the department of Electronics & Communication at K.R.Gouramma College of Engineering for women, Cherthala under Kerala University. Presently she is pursuing M.Tech in Signal Processing (EC) at Government College of Engineering, Cherthala under CUSAT.