Automatic Sub–volume Registration by Probabilistic Random Search

Jingfeng Han[a], Min Qiao[a], Joachim Hornegger[a],
Torsten Kuwert[b], Werner Bautz[c], Wolfgang Römer[b]

[a]Institute of Pattern Recognition, University of Erlangen–Nürnberg, Erlangen, Germany;
[b]Clinic of Nuclear Medicine, University of Erlangen–Nürnberg, Erlangen, Germany;
[c]Institute of Radiology, University of Erlangen–Nürnberg, Erlangen, Germany

ABSTRACT

Registration of an individual's image data set to an anatomical atlas provides valuable information to the physician. In many cases, the individual image data sets are partial data, which may be mapped to one part or one organ of the entire atlas data. Most of the existing intensity based image registration approaches are designed to align images of the entire view. When they are applied to the registration with partial data, a manual pre–registration is usually required. This paper proposes a fully automatic approach to the registration of the incomplete image data to an anatomical atlas. The spatial transformations are modelled as any parametric functions. The proposed method is built upon a random search mechanism, which allows to find the optimal transformation randomly and globally even when the initialization is not ideal. It works more reliably than the existing methods for the partial data registration because it successfully overcomes the local optimum problem. With appropriate similarity measures, this framework is applicable to both mono–modal and multi–modal registration problems with partial data. The contribution of this work is the description of the mathematical framework of the proposed algorithm and the implementation of the related software. The medical evaluation on the MRI data and the comparison of the proposed method with different existing registration methods show the feasibility and superiority of the proposed method.

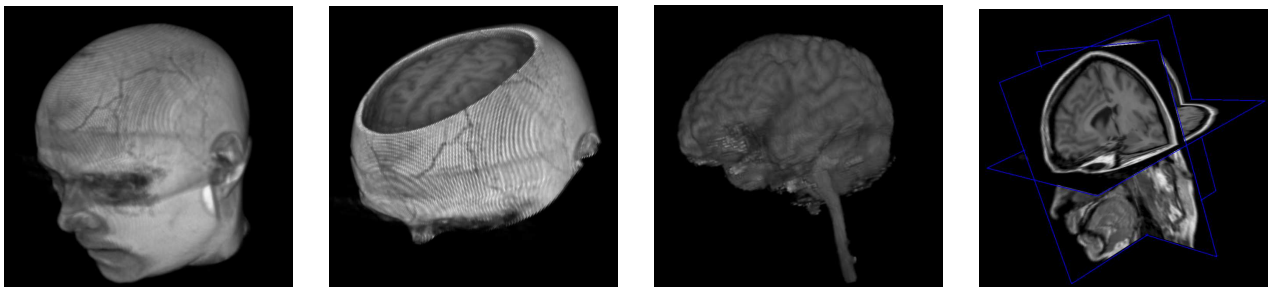
Keywords: Image registration, Partial data registration, Parameter estimation, Probabilistic optimization, Adaptive random searching

1. INTRODUCTION

Recently, atlas-to-subject image registration becomes more and more important for many medical applications, especially for brain tumor radiotherapy. The atlas, for example the computerized Human Brain Atlas, is constructed out of many high–resolution magnetic resonance images (MRI) of normal subjects. The atlas provides the labelling, the mean shape and position of anatomical structures.[1,2] The fundamental task of atlas-to-subject registration is to build the mapping between a specific patient's image with the atlas image, so that physicians can make use of the atlas annotations to study the individual image, for example, to segment several structures simultaneously.

In order to deform the atlas to the patient image, the registration is usually performed in two steps. The first step is a rigid matching between the atlas and the patient image. The second step is the refinement of the transformation by a non–rigid registration. One challenge of such two–step scheme is that the non–rigid registration is generally difficult for multi–subject images. The normal registration using uniform regularization cannot find the ideal transformation that should be smooth in some places while containing fine details in others. Recently, some studies have been conducted on the registration with inhomogeneous regularization. The idea is to accommodate the regularization according to the deformability, so that the regularization is strong where the local deformability is low.[2–5] Another open challenge is how to automatically compute the rigid transformation between the partial patient image and a full atlas. In many cases the patient images are just partial images,

Further author information: (Send correspondence to Jingfeng Han.)
E-mail: jingfeng@informatik.uni-erlangen.de, Telephone: +49 9131 85 27826



(a) (b) (c) (d)

Figure 1. The examples of partial data registration problem. (**a**): The full atlas volume. The patient image data could be a sub volume (**b**), segmented organs (**c**) or MPR 2D slices (**d**).

which may map to one part or one organ in the entire atlas. An illustrative example is given in Figure 1. Most existing rigid registration methods, like iterative closest point (ICP) algorithm[6] or the correlation based algorithms, are designed to align the images of entire view. In practice the user needs to pre–register the partial image data to the atlas manually.

In this article, we focus on the development of an automatic approach for registration of the atlas image and partial patient image without any close initialization. We firstly use an adaptive random search[7] to find the reasonable candidates of parameter configuration. A stochastic sampling scheme simultaneously searches the multiple optimal configurations in a large parameter space and effectively avoids the problem of local convergence. Then we use the result of random search as the starting point of common local search to accurately find the optimal rigid transformation. Finally the non–rigid registration refines the transformation. The experiments show that the random search automatically generates reasonable initialization for the following registrations and significantly improves the reliability of registration process.

2. METHOD

Let scalar–valued functions $f : \Omega_1 \to \mathbb{R}$ and $g : \Omega_2 \to \mathbb{R}$ denote the intensity functions of the atlas volume and the patient volume. The sets $\Omega_1 \subset \mathbb{R}^n$ and $\Omega_2 \subset \mathbb{R}^n$ are the domains of the atlas volume and the patient volume respectively. The goal of the atlas-to-subject registration is to find a spatial transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ such that the transformed atlas is close to the patient volume. It can be reflected in the following minimization problem:

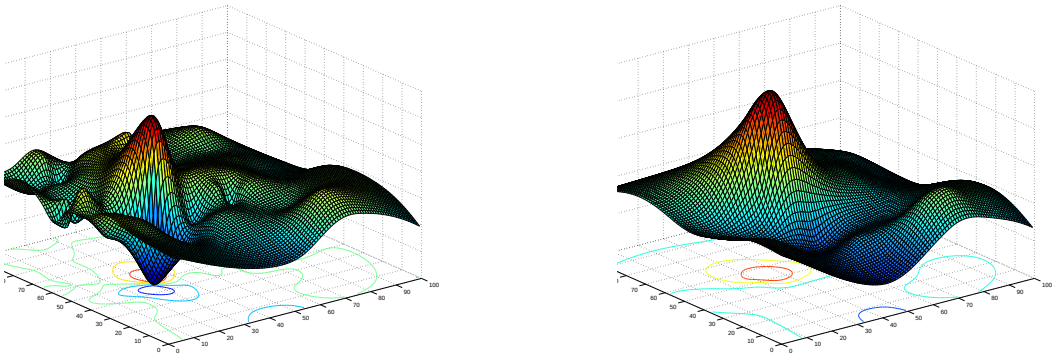
$$T = \text{argmin}_T D^{[f,g]}(T). \tag{1}$$

Here the distance measure $D^{[f,g]}$ indicates the dissimilarity between two data sets. Some survey studies[8,9] have discussed and compared different distance measures for the image matching. Without losing generality, we use the sum of square difference (SSD) as the distance measure in this work. It is defined as

$$D^{[f,g]}(T) := \int_{T(\Omega_1) \bigcap \Omega_2} (f(T(x)) - g(x))^2 \, dx. \tag{2}$$

In this work we assume that the patient volume is a sub–volume of the atlas volume, i.e., $\Omega_2 \subset T(\Omega_1)$. Therefore, the domain of the integration in (2) is simplified to be the domain of patient volume Ω_2 and the minimization problem is reduced to:

$$T = \text{argmin}_T \int_{\Omega_2} (f(T(x)) - g(x))^2 \, dx. \tag{3}$$



(a) (b)

Figure 2. The rough energy landscape computed from a sparse number of uniformly distributed sampling point may smear out the true minimum. (**a**): The true energy landscape. (**b**): The rough energy landscape is computed by B-Spline interpolation of the scatter sampling points.

2.1. Global probabilistic rigid registration

For the rigid registration, the spatial transformation T is defined by the translation parameters $\boldsymbol{t} = (t_i)_{i=1,...,n}$ and the rotation parameters $\boldsymbol{\theta}$. In our implementation, the rotation parameters were treated differently in 2D and 3D. In the 2D case, $\boldsymbol{\theta}$ is the common Euler angle in respect of the center of images. In the 3D case, we used the quaternion representation of rotation, e.g. $\boldsymbol{\theta} = (q_i)_{i=1,2,3,4}$. Let $\boldsymbol{\phi} = (\boldsymbol{t}|\boldsymbol{\theta})^T$ denote the m-dimensional parameter vector of rigid transformation. Since the patient volume is a sub–volume of the atlas volume, the search space of $\Omega_{\boldsymbol{\phi}}$ is clearly a bounded domain.

The global probabilistic rigid registration is actually a process of stochastic sampling in the parameter space $\Omega_{\boldsymbol{\phi}}$. Each sampling point in the $\Omega_{\boldsymbol{\phi}}$ corresponds to one configuration of transformation $\boldsymbol{\phi}$. Since we have no clue about a reasonable initialization of $\boldsymbol{\phi}$, in the first sampling we distribute N_0 sampling points uniformly in the entire $\Omega_{\boldsymbol{\phi}}$. Then the distance measure $D^{[f,g]}(\boldsymbol{\phi})$ is evaluated for each sampling point $\boldsymbol{\phi}_i (i = 1, ..., N_0)$. In this way a rough landscape of $D^{[f,g]}(\boldsymbol{\phi})$ in $\Omega_{\boldsymbol{\phi}}$ can be generated. However, the true landscape of distance measure may be much more complicated, especially when the patient volume g is a sub–volume of atlas f. The rough landscape based on the limited number of sampling points needs to be refined, otherwise the true global minimum may be smeared out. See the example in the Figure 2.

Algorithm 1 Global Probabilistic Rigid Registration Algorithm[7]

---

1. Input the full-volume f and the sub-volume g.
2. Define a covariance matrix Σ, a contraction factor $\gamma < 1$ and maximum iteration number K.
3. Distribute N_0 points in the search space uniformly.
4. Select the best N_1 ($N_1 < N_0$) points in the search space with lowest distance measure $D^{[f,g]}(\boldsymbol{\phi})$ and sort these N_1 points into a list.
5. The points in the list are used to generate the other N_1 new random points in the following way. Build up N_1 normal distributions, where the mean vectors correspond to the computed list elements and the covariances are given by Σ. Compute one new random point using each normal distribution.
6. Decrease the covariance matrix Σ by multiplication with the contraction factor γ.
7. The steps 4–6 are repeated until iteration number exceeds.

The refinement is evolved as following process. From the N_0 initial sampling points we select $N_1 (N_1 < N_0)$

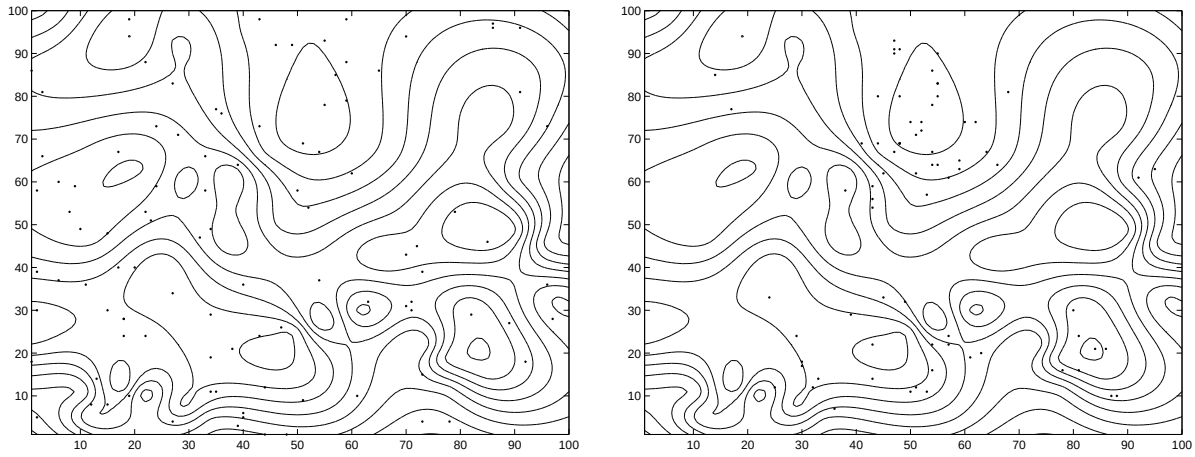


Figure 3. A example of random global optimization of 2D function. The contour–map of distance function with the uniformly distributed initial points (**a**) and the sampling points generated by probabilistic search (**b**).

"best" ones that have lowest distance measure. Let $\Phi = (\boldsymbol{\phi}_i)_{i=1,\ldots N_1}$ denotes the collection of these best points. Given a covariance matrix $\Sigma = \mathrm{diag}(\sigma_i)_{i=1,\ldots,m}$ we construct a group of m-dimensional Gaussian distributions $G = (G_i)_{i=1,\ldots,N_1}$ such that the mean of the G_i is ϕ_i and the covariance is Σ. Then N_1 new sampling points $\Phi^{'} = (\boldsymbol{\phi}_i^{'})_{i=1,\ldots N_1}$ are generated according to the group of distribution G. Now we have $2N_1$ sampling points together with the original ones Φ and the new generated ones $\Phi^{'}$. From these $2N_1$ sampling points we only reserve N_1 best ones according to the associated distance measure $D^{[f,g]}(\boldsymbol{\phi})$. In the next round of sampling, we update the N_1 sampling points again in the same way except that the system of random generation is annealed. The annealing is expressed by decreasing the covariance of Gaussian distributions with a contraction factor $\gamma \in]0,1[$, i.e., $\Sigma^{(k+1)} = \gamma\Sigma^{(k)}$ where k is the number of iteration. For the clarity of presentation, we summarize the global probabilistic rigid registration in the Algorithm 1. For the detail of the random search algorithm, we refer to.[7]

The most striking property of the probabilistic searching is that the sampling is controlled by an annealing process. At the beginning, these Gaussian distributions are given large variances. The generated sampling points are allowed to deviate from the previous sampling points within large ranges. The parameter searching is performed more globally and more randomly in the early stage to avoid that optimization is trapped by the local optimums. In the following iterations, the variances are decreased by the contraction factor. Thus the new test points are more likely to be confined in the neighboring area of the previous ones. With the annealed sampling, the parameters are searched more locally and more deterministically in order to converge to the global optimum. An illustrative example is given in Figure 3, where the objective function has simply two unknown parameters.

2.2. Local deterministic rigid registration

Local deterministic optimization seeks to more accurately search the optimal parameter configuration. Since the global probabilistic random search provides good initial guesses, the deterministic optimization is usually performed in the finest resolution and in very few iterations. Many local optimization techniques have been proposed for the rigid registration. In this work, we choose the steepest descent optimization, in which the parameter vector is updated in the following way:

$$\boldsymbol{\phi}^{(k+1)} = \boldsymbol{\phi}^{(k)} - \lambda^{(k)} \cdot \nabla D^{[f,g]}(\boldsymbol{\phi}^{(k)}), \tag{4}$$

where $\nabla D^{[f,g]}(\boldsymbol{\phi}^{(k)})$ is the gradient of distance measure at the position $\boldsymbol{\phi}^{(k)}$ and $\lambda^{(k)}$ is the learning rate. The denotation "steepest" targets the way how the $\lambda^{(k)}$ are computed: a linear search along the descent direction is

performed to get the best possible learning rate within a given accuracy. For the details of the steepest descent optimization we refer to.[10]

2.3. Elastic non–rigid registration

The purpose of the non–rigid registration is to further refine the transformation T such that the multi–subject volumes are optimally aligned. The transformation T is modelled by a non–parametric function, the so called displacement field $u : \mathbb{R}^n \to \mathbb{R}^n$, i.e.,

$$T := x - u(x). \tag{5}$$

For the non–rigid registration, the objective function usually contains a regularizer $S[u] : \mathbb{R}^n \to \mathbb{R}$ to remedy the arbitrary irregularity of the displacement field, i.e.,

$$u := \text{argmin}_u(D^{[f,g]}[u] + \alpha S[u]), \tag{6}$$

where α controls the importance of the regularizer S. The typical choices of regularizers for non–rigid image registration have been discussed in.[11] In this work, we implemented the standard elastic regularization.[12] The optimal displacement field u is characterized by the following Euler–Lagrange equation:

$$\mu \triangle u + (\lambda + \mu) \nabla \text{div } u = h \tag{7}$$

where λ and μ are the Lamé constants. The vector field $h : \Omega_2 \to \mathbb{R}^n$ denotes the first variation of $D^{[f,g]}[u]$ with respect to u, which is computed as

$$h(x) := (g(x) - f(x - u(x)))\nabla f(x - u(x)) \tag{8}$$

For the numerical implementation of elastic registration, we refer to.[11]

3. EXPERIMENTS

In the 2D experiments, we compared the proposed method, which was initialized by probabilistic random search (PRS), with the standard multi-resolution gradient descent method[13] (MGD) which had arbitrary initializations. The experiment was designed as following. The MRI slice of brain in Figure 5a was the moving full image. We had two partial images in Figure 5b and Figure 5c which were segments from the corresponding slice of different acquisitions of the same patient. In the following we were going to seek to transform the full image to the partial image in Figure 5b (denoted as experiment $\mathbf{a} \to \mathbf{b}$) and the partial image in Figure 5c (denoted as experiment $\mathbf{a} \to \mathbf{c}$). In order to simulate the arbitrary initialization, the initial parameters were generated via a uniform distribution in the parameter space. We treated the results of the manual registration as the ground truth. If the error of the translation is less than 1 pixel and the error of the rotation is less than 1.5°, we defined that the registration experiment succeeded. Otherwise it failed.

We performed the proposed PRS method for 1000 times for each partial data registration experiment. Both of experiments reached quite high success rate. ($\mathbf{a} \to \mathbf{b}$: 99.8% and $\mathbf{a} \to \mathbf{c}$: 99.6%), because the probabilistic pre-registration provided ideal initializations for the following deterministic registration in most cases (See Figure 4). Meanwhile we performed also the standard MGD method on the same experiments. For the experiment $\mathbf{a} \to \mathbf{b}$ we carefully select parameter setting of the MGD algorithm, so that the resulting success rate achieved quite high (99.9%). In the experiment $\mathbf{a} \to \mathbf{c}$ we applied same algorithm parameter setting as in the experiment $\mathbf{a} \to \mathbf{b}$, however, the deterministic MGD method frequently failed (27.3%). In this article we did not intend to underrate the standard multi-resolution gradient descent registration method. We also assured that a careful adjustment of the algorithm parameter setting for each initialization could achieve a correct match in most cases. However, the global probabilistic registration appeared more robust in respect to the poor initialization. One should note that all the experiments with the probabilistic pre-registration were performed with a constant parameter setting.

We compared the computational performance of the successful experiments of the proposed PRS and the standard MGD method. The average computational time included the time of elastic registrations as well. The comparison was summarized in Table 1. The setting of the experimental environment was given in Table 2. In Figure 5, we presented one successful $\mathbf{a} \to \mathbf{c}$ experiment with a probabilistic initialization (pre–registration).

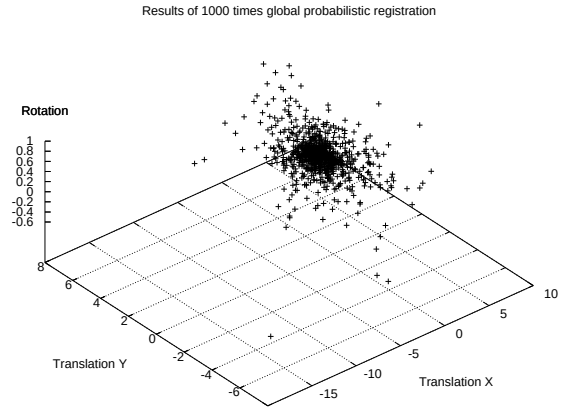


Figure 4. The results of 1000 times global probabilistic registrations in the experiment $\mathbf{a} \rightarrow \mathbf{c}$. In most cases, the probabilistic approach successfully pre-registered the images and provided ideal initializations for the following deterministic registration.

Exp.	$\mathbf{a} \rightarrow \mathbf{b}$		$\mathbf{a} \rightarrow \mathbf{c}$	
Method	Succ.	Time (s)	Succ.	Time (s)
PRS	99.8%	102.5	99.6%	101.8
MGD	99.9%	98.3	27.3%	57.2

Table 1. Success rates (Succ.) and average computational time (Time) of the experiment $\mathbf{a} \rightarrow \mathbf{b}$ and $\mathbf{a} \rightarrow \mathbf{c}$ with two approaches: the multi-resolution gradient decent registration (PRS) and the proposed method with probabilistic pre-registration (MGD). The PRS method reached quite high success rate in both experiments with a constant algorithm parameter setting. The MGD method had relatively high success rate in experiment $\mathbf{a} \rightarrow \mathbf{b}$ with the optimal algorithm parameter setting, however, frequently failed (Succ.27.3%) in the experiment $\mathbf{a} \rightarrow \mathbf{c}$ with the same algorithm parameter setting.

CPU	:	AMD Athlon 2.20GHz
RAM	:	2.00GB
OS	:	WindowXP
Compiler	:	VC++ 6.0 (Debug)
Image	:	128×128

Table 2. Test environment setting.

In Figure 6, we plotted the parameter evolution pathes of this experiment, in which the probabilistic search appeared more random and more global while the following deterministic search more local. By the way, the probabilistic optimization was actually a multi-path search which found multiple configurations in each iteration. In Figure 6 we plotted only the best configuration with lowest distance measure.

We also verified the proposed method with 3D data sets, in which a MRI volume of the full brain (in Figure 1a) was registered to a sub volume (in Figure 1b) and a segmented brain organ (in Figure 1c). The fusion images of two 3D experiments are presented in Figure 7 and Figure 8 respectively.

4. SUMMARY AND CONCLUSION

In this article we introduced and evaluated a new framework to register a sub-volume data to a full volume data. For this partial data registration problem, a key challenge is to find an appropriate way to avoid the local convergence. The presented probabilistic pre-registration makes use of random search technique to globally search the transformation parameter and successfully overcomes the problem of local optimums. The comparison of the proposed algorithm to the well–established algorithms shows that our approach turns out to be more reliable in partial data registration than the standard deterministic approaches. Consequently, the registration system with such probabilistic pre-registration mechanism requires minimum user interaction. Obviously, the presented probabilistic pre-registration methods can also be applied to map images using any parametric transforms, like affine transform. With different distance measures, for example, mutual information, the presented framework can be extended to multi-modal registration problem, like PET/CT registration.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of Deutsche Forschungsgemeinschaft (DFG) under the grant SFB 603, TP, C10. The authors are also thankful to HipGraphics for providing the volume rendering software (InSpace).

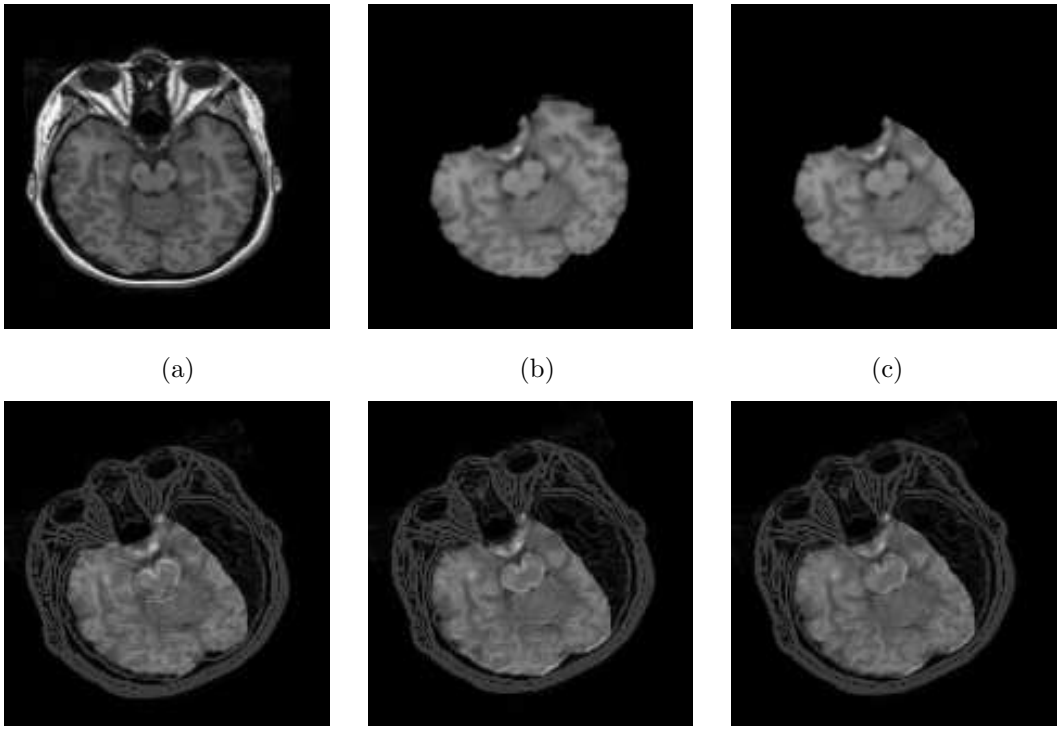


Figure 5. The experiment on 2D MRI slices. Top: The moving full image (a), two target partial images (b) and (c). Bottom: Overlap the partial image (c) with the edge set of the transformed image after global probabilistic registration (d), after local deterministic registration (e) and after the non-rigid elastic registration (f).

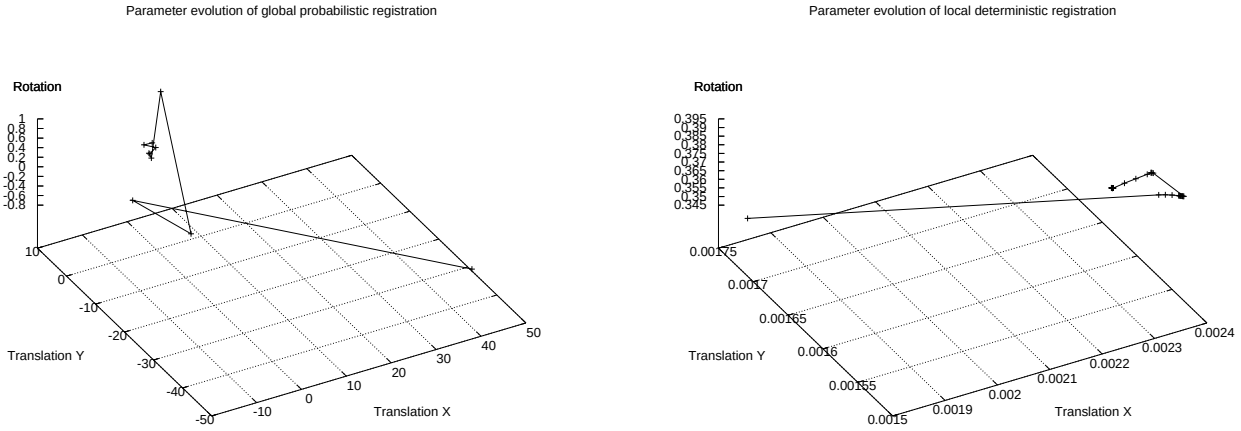


Figure 6. An example of evolution path in the parameter space. Left: global probabilistic registration. Right: local deterministic registration. The probabilistic registration searched the transformation parameters randomly and globally in order to avoid to be trapped by the local optimums, while the following deterministic registration searched more locally in order to converge to the local optimums. Note the coordinate units of two plots were different.

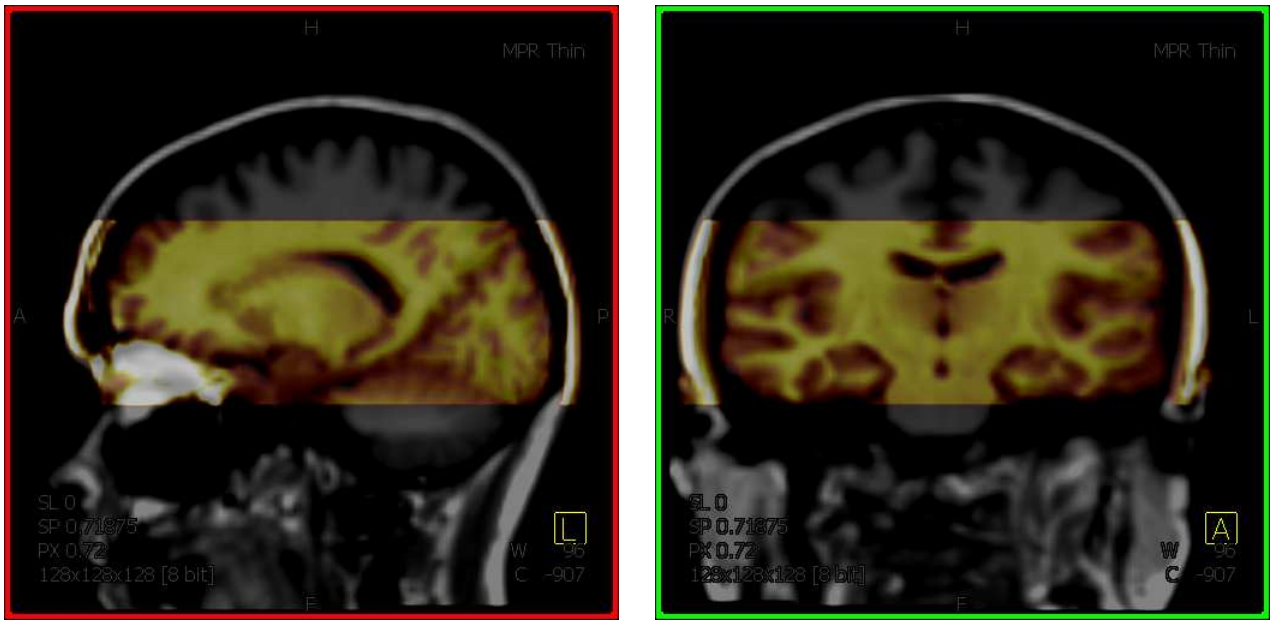


Figure 7. The results of the experiment on 3D MRI data set: fusion of the transformed full volume and the sub volume. The moving full volume is the volume in Figure 1(a). The fixed sub volume is the one appearing in Figure 1(b)

Figure 8. The results of the experiment on 3D MRI data set: fusion of the transformed full volume and the segmented brain organ. The moving full volume is the volume in Figure 1(a). The brain organ is the one appearing in Figure 1(c)

REFERENCES

1. P. Y. Bondiau, G. Malandain, S. Chanalet, P. Y. Marcy, J. L. Habrand, F. Fauchon, P. Paquis, A. Courdi, O. Commowick, I. Rutten, and N. Ayache, "Atlas-based automatic segmentation of mr images: validation study on the brainstem in radiotherapy context," *International journal of radiation oncology, biology, physics.* **61**(1), pp. 289–98, 2005.
2. E. Bardinet, D. Dormont, G. Malandain, M. Bhattacharjee, B. Pidoux, C. Saleh, P. Cornu, N. Ayache, Y. Agid, and J. Yelnik, "Retrospective cross-evaluation of an histological and deformable 3d atlas of the basal ganglia on series of parkinsonian patients treated by deep brain stimulation (dbs)," in *Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005*, J. Duncan and G. Gerig, eds., *LNCS* **3750**, pp. 385–393, Springer Verlag, (Palm Springs, CA, USA, October 26-29), 2005.
3. O. Commowick, R. Stefanescu, P. Fillard, V. Arsigny, N. Ayache, X. Pennec, and G. Malandain, "Incorporating statistical measures of anatomical variability in atlas-to-subject registration for conformal brain radiotherapy," in *Proceedings of the 8th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005, Part II*, J. Duncan and G. Gerig, eds., *LNCS* **3750**, pp. 927–934, Springer Verlag, (Palm Springs, CA, USA, October 26-29), 2005.
4. H. Lester, S. R. Arridge, K. M. Jansons, L. Lemieux, J. V. Hajnal, and A. Oatridge, "Non-linear registration with the variable viscosity fluid algorithm," in *IPMI '99: Proceedings of the 16th International Conference on Information Processing in Medical Imaging*, pp. 238–251, Springer-Verlag, (London, UK), 1999.
5. R. Stefanescu, X. Pennec, and N. Ayache, "Grid powered nonlinear image registration with locally adaptive regularization," *Medical Image Analysis* **8**(3), pp. 325–342, 2004. MICCAI 2003 Special Issue.
6. P. J. Besl and N. D. McKay, "A method for registration of 3-d shapes," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(2), pp. 239–256, 1992.
7. J. Hornegger and H. Niemann, "Probabilistic modeling and recognition 3-d objects," *Internaltional Journal of Computer Vision* **39**(3), pp. 229–251, 2000.
8. B. Zitova and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing* **21**(11), pp. 977–1000, 2003.
9. J. B. A. Maintz and M. A. Viergever, "A survey of medical image registration," *Medical Image Analysis* **2**(1).
10. P. Venkataraman, *Applied Optimization with Matlab Programming*, ch. Numerical Techniques for Unconstrained Optimization. John Wiley and Sons Inc., 2002.
11. J. Modersitzki, *Numerical Methods for Image Registration*, Oxford University Press, New York, 2004.
12. C. Broit, *Optimal Registration of Deformed Images.* Ph.d. thesis, Computer and Information Science, Univeristy of Pennsylvania, 1981.
13. D. Hahn, J. Hornegger, W. Bautz, T. Kuwert, and R. W., "Unbiased rigid registration using transfer function," in *Proc. SPIE Medical Imaging*, J. M. Fitzpatrick and J. M. Reinhardt, eds., **5747**, pp. 151–162, (San Diego, USA), 2005.