Free-Form B-spline Deformation Model for Groupwise Registration

Published in final edited form as:
Med Image Comput Comput Assist Interv. 2007 ; 10(WS): 23–30. doi:10.1901/jaba.2007.10-23.

Free-Form B-spline Deformation Model for Groupwise Registration

Serdar K. Balci[1], **Polina Golland**[1], **Martha Shenton**[2], and **William M. Wells**[2]
[1]CSAIL, MIT, Cambridge, MA, USA

[2]Brigham & Women's Hospital, Harvard Medical School, Cambridge, MA, USA

Abstract

In this work, we extend a previously demonstrated entropy based groupwise registration method to include a free-form deformation model based on B-splines. We provide an efficient implementation using stochastic gradient descents in a multi-resolution setting. We demonstrate the method in application to a set of 50 MRI brain scans and compare the results to a pairwise approach using segmentation labels to evaluate the quality of alignment. Our results indicate that increasing the complexity of the deformation model improves registration accuracy significantly, especially at cortical regions.

1 Introduction

Groupwise registration is an important tool in medical image analysis for establishing anatomical correspondences among subjects in a population [1,2]. A groupwise registration scheme can be used to characterize anatomical shape differences within and across populations [3].

A common approach is to register every subject in the population to a reference subject using pairwise registrations. However, due to anatomical variabilities among human brains, a registration scheme based on a single subject introduces a bias towards the shape of the selected subject[3].

Seghers et al. [4] refrain from choosing an anatomical reference by considering all pairwise registrations between subjects in the population. They generate an atlas with a mean morphology by composing the deformations for each subject into a mean deformation. Park et al. [5] also consider all combinations of pairs in the population and use multi dimensional scaling to select the reference subject closest to the mean of the population. Both methods improve the pairwise approach by considering all combinations of pairs in the population; however, computational complexity can be high as the number of required pairwise registrations grows quadratically with the number of images.

More efficient groupwise registration methods can be developed by making use of a template constructed from the statistics of the population. Joshi et al. [6] register all subjects in the population to a mean intensity image in a large deformation diffeomorphic setting. Christensen et al. [7] consider deformations from the average intensity image to the population and minimize correspondence errors by using inverse consistent transforms. Bhatia et al. [8] make use of the statistics in the population by computing normalized mutual information between a reference frame and the joint histogram of all pairs of intensities. They describe a deformation model based on B-splines and define a common coordinate frame by constraining the average transformation in the population to be the identity transform. Twining et al. [9] combine groupwise registration and construction of a template in a minimum description length

framework using clamped-plate splines as a deformation model. In a recent work, Blezek and Miller [10] point out that a population can have a multi-modal distribution. They consider pairwise distances between subjects and construct templates from sub-sets of subjects by using mean shift algorithm to search for the modes in the population. Their results indicate that a registration method based on a single template may not account for all the variability in a population.

Selection of a template can be avoided if the population is considered as a whole and the deformations leading to the common coordinate frame are found simultaneously for all subjects. Studholme et al. [11] extend concepts from multi-modal image alignment and use the joint entropy of intensity distributions as an objective function. Their objective function measures probabilistic dependencies between intensity values in images and is robust both to global and local intensity variations. However, computing the entropy of a high dimensional distribution is a computationally expensive task.

Miller et al. [12] introduce an efficient groupwise registration method in which they consider sum of univariate entropies along pixel stacks as a joint alignment criterion. Their method provides a template-free approach to groupwise registration by simultaneously driving all subjects to the common tendency of population. Zollei et al. [13] successfully applied this method to groupwise registration of medical images using affine transforms and used stochastic gradient descent for optimization.

In our work, we extend Miller et al.'s [12] method to include free-form deformations based on B-splines in 3D. We describe an efficient implementation using stochastic gradient descent in a multi-resolution setting. We validate the method on a set of 50 MR images using segmentation label alignment as an independent measure of registration quality. To compare the results to a pairwise approach we use a template based method similar to Joshi et al. [6] where we register every subject to the mean image using sum of squared differences.

This paper is organized as follows. In the next section, we describe the stack entropy cost function, introduce B-spline based deformation model and discuss implementation details. In Section 3, we compare groupwise registration to the pairwise method and evaluate both methods using label prediction values. We conclude by discussing groupwise and pairwise registration approaches.

2 Methods

Given a set of images $\{I_1,\ldots, I_N\}$, each described by intensity values $I_n(\mathrm{x}_n)$ across the image space $\mathrm{x}_n \in \chi_n$, we can define a common reference frame $\mathrm{x}_R \in \chi_R$ and a set of transforms $\mathcal{T}$ which maps points from the reference frame to points in the image space

$$\mathcal{T}=\{T_n:\mathrm{x}_n=T_n(\mathrm{x}_R),\ n=1,\ldots,N\} \tag{1}$$

Mapped intensities in the reference frame x_R can be given by

$$I'_n(\mathrm{x}_R)=I_n\,(T_n\,(\mathrm{x}_R)) \tag{2}$$

In the following section, we describe the stack entropy cost function as introduced by Miller et al. [12] and applied to groupwise registration of medical images using affine transforms by Zollei et al. [13]. We continue by describing our transformation model and extend Miller et al.'s [12] method to include free-form deformations based on B-splines.

2.1 Stack Entropy Cost Function

In order to align all subjects in the population, we consider sum of pixelwise entropies as a joint alignment criterion. The justification for this approach is that if the images are aligned properly, intensity values at corresponding coordinate locations from all the images will form a low entropy distribution. This approach does not require the use of a reference subject; all subjects are simultenously driven to the common tendency of the population.

We let $x_v \in \chi_R$ be a sample from a spatial location in the reference frame and $H(I(T(x_v)))$ be the univariate entropy of the stack of pixel intensities $\{I_1(T_1(x_v)), \ldots, I_N(T_N(x_v))\}$ at spatial location x_v. The objective function for the groupwise registration can be given as follows

$$f = \sum_{v=1}^{V} H(I(T(x_v))). \tag{3}$$

To derive this objective function, the intensity samples over the space are assumed to be independent and are allowed to come from different distributions. At each particular spatial location x_v, pixel intensities along the image stack are assumed to be i.i.d samples, therefore allowing the evaluation of univariate entropies.

We employ a Parzen window based density estimation scheme to estimate univariate entropies [14]:

$$f = -\sum_{v=1}^{V} \frac{1}{N} \sum_{i=1}^{N} \log \frac{1}{N} \sum_{j=1}^{N} G_\sigma(d_{ij}(x_v)) \tag{4}$$

where $d_{ij}(x) = I_i(T_i(x)) - I_j(T_j(x))$ is the distance between intensity values of a pair of images evaluated at a point in the reference frame and G_σ is a Gaussian kernel with variance σ^2. The objective function achieves its minimum when the intensity differences are small. Using the entropy measure we obtain a better treatment of transitions between different tissue types, such as gray matter-white matter transitions in the cortical regions where intensity distributions can be multi-modal. Parzen window density estimator allows us to obtain analytical expressions for the derivative of the objective function with respect to transform parameters

$$\frac{\partial f}{\partial T_n} = \sum_{v=1}^{V} \frac{1}{\sigma^2 N} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{G_\sigma(d_{ij}(x_v)) d_{ij}(x_v)}{\sum_{k=1}^{N} G_\sigma(d_{ij}(x_v))} \frac{\partial d_{ij}(x_v)}{\partial T_n}. \tag{5}$$

2.2 Free-Form B-spline Deformation Model

For the nonrigid deformation model, we define a combined transformation consisting of a global and a local component

$$T(\mathrm{x}) = T_{local}(T_{global}(\mathrm{x})) \tag{6}$$

where T_{global} is a twelve parameter affine transform and T_{local} is a deformation model based on B-splines.

Following Rueckert et al.'s formulation [15], we let **Φ** denote an $n_x \times n_y \times n_z$ grid of control points $\varphi_{i,j,k}$ with uniform spacing. The free form deformation can be written as the 3-D tensor product of 1-D cubic B-splines.

$$T_{local}(\mathrm{x})=\mathrm{x}+\sum_{l=0}^{3}\sum_{m=0}^{3}\sum_{n=0}^{3}B_l(u)B_m(v)B_n(w)\phi_{i+l,j+m,k+n} \tag{7}$$

where $\mathbf{x} = (x, y, z)$, $i = \lfloor x/n_x \rfloor - 1$, $j = \lfloor y/n_y \rfloor - 1$, $k = \lfloor z/n_z \rfloor - 1$, $u = x/n_x - \lfloor x/n_x \rfloor$, $v = y/n_y - \lfloor y/n_y \rfloor$, $w = z/n_z - \lfloor z/n_z \rfloor$ and where B_l is l'th cubic B-spline basis function. Using the same expressions for u, v and w as above, the derivative of the deformation field with respect to B-spline coefficients can be given by

$$\frac{\partial T_{local}(x,y,z)}{\partial \phi_{i,j,k}}=B_l(u)B_m(v)B_n(w) \tag{8}$$

where $l = i - \lfloor x/n_x \rfloor + 1$, $m = j - \lfloor y/n_y \rfloor + 1$ and $n = k - \lfloor z/n_z \rfloor + 1$. We consider $B_l(u) = 0$ for $l < 0$ and $l > 3$. The derivative terms are nonzero only in the neighborhood of a given point. Therefore, optimization of the objective function using gradient descent can be implemented efficiently.

As none of the images are chosen as an anatomical reference, it is necessary to add a geometric constraint to define the reference coordinate frame. Similar to Bhatia et al. [8], we define the reference frame by constraining the average deformation to be the identity transform:

$$\frac{1}{N}\sum_{n=1}^{N}T_n(\mathrm{x})=\mathrm{x} \tag{9}$$

This constraint assures that the reference frame lies in the center of the population. In the case of B-splines, the constraint can be satisfied by constraining the sum of B-spline coefficients across images to be zero. In the gradient descent optimization scheme, the constraint can be forced by subtracting the mean from each update vector [8].

2.3 Implementation Details

We provide an efficient optimization scheme by using line search with the gradient descent algorithm. For computational efficiency, we employ a stochastic subsampling procedure [16]. In each iteration of the algorithm, a random subset is drawn from all samples and the objective function is evaluated only on this sample set. The number of samples to be used in this method depends on the number of the parameters of the deformation field. Using the number of samples on the order of the number of variables works well in practice.

To obtain a dense deformation field capturing anatomical variations at different scales, we gradually increase the complexity of the deformation field by refining the grid of B-spline control points. First, we perform a global registration with affine transforms. Then, we use affine transforms to initialize a low resolution deformation field at a grid spacing around 32 voxels. We increase the resolution of the deformation field to 8 voxels by using registration results at coarser grids to initialize finer grids.

As in every iterative search algorithm, local minima pose a significant problem. To avoid local minima we use a multi-resolution optimization scheme for each resolution level of the

deformation field. The registration is first performed at a coarse scale by downsampling the input. Results from coarser scales are used to initialize optimization at finer scales. For each resolution level of the deformation field we used a multi-resolution scheme of three image resolution levels.

We implemented our groupwise registration method in a multi-threaded fashion using Insight Toolkit (ITK) and made the implementation publicly available [17]. We run experiments using a dataset of 50 MR images with 256 × 256 × 128 voxels on a workstation with four CPUs and 8GB of memory. The running time of the algorithm is about 20 minutes for affine registration and six hours for non-rigid registration of the entire dataset. The memory requirement of the algorithm depends linearly on the number of input images and was around 3GB for our dataset.

3 Evaluation

We tested the groupwise registration algorithm on a MR brain dataset and compared the results to a pairwise approach [6]. The dataset consists of 50 MR brain images of three subgroups: schizophrenics, affected disorder and normal control patients. MR images are T1 scans with 256 × 256 × 128 voxels and 0:9375 × 0:9375 × 1:5 mm^3 spacing. MR images are preprocessed by skull stripping. To account for global intensity variations, images are normalized to have zero mean and unit variance. For each image in the dataset, an automatic tissue classification [18] was performed, yielding gray matter (GM), white matter (WM) and cerebro-spinal uid (CSF) labels. In addition, manual segmentations of four subcortical regions (left and right hippocampus and amygdala) and four cortical regions (left and right superior temporal gyrus and para-hippocampus) were available for each MR image.

We compare our groupwise algorithm to a pairwise method where we register each subject to the mean intensity using sum of square differences. The objective function for pairwise registration to the mean can be described as follows

$$f_{pair} = \sum_{n=1}^{N} (I_n(T_n(x)) - \mu(x))^2 \tag{10}$$

where μ is defined as the mean of the intensities $\mu(x) = \frac{1}{N} \sum_{n=1}^{N} I_n(T_n(x))$. During each iteration we consider the mean image as a reference image and register every subject to the mean image using sum of squared differences. After each iteration the mean image is updated and pairwise registrations are performed until convergence.

The images in Figure 1 show central slices of 3D images after registration. Visually, mean images get sharper and variance images becomes darker, especially around central ventricles and cortical regions. We can observe that anatomical variability at cortical regions causes significant blur for GM, WM and CSF structures using affine registration. Finer scales of B-spline deformation fields capture a significant part of this anatomical variability and the tissue label overlap images get sharper.

We evaluate registration results by measuring label prediction accuracy in a leave-one-out fashion for the two different sets of segmentation labels. To predict the segmentation labels of a subject, we use majority voting for the labels in the rest of the population. We compute the Dice measure between the predicted and the true labels and average over the whole population.

The prediction accuracy reported in Table 1 is lower than what is typically achieved by segmentation methods. This is to be expected as our relatively simple label prediction method

only considers voxelwise majority of the labels in the population and does not use the novel image intensity to predict the labels. Effectively, Table 1 reports the accuracy of the spatial prior (atlas) in predicting the labels before the local intensity is used to refine the segmentation. Increasing the complexity of the deformation model improves the accuracy of prediction. An interesting open problem is automatically identifying the appropriate deformation complexity before the registration overfits and the accuracy of prediction goes down. We also note that the alignment of the subcortical structures is much better than that of the cortical regions. It is not surprising as the registration algorithm does not use the information about geometry of the cortex to optimize the alignment of the cortex. In addition, it has been often observed that the cortical structures exhibit higher variability across subjects when considered in the 3D volume rather than modelled on the surface.

Our experiments highlight the need for further research in developing evaluation criteria for image alignment. We used the standard Dice measure, but it is not clear that this measurement captures all the nuances of the resulting alignment.

Comparing the groupwise registration to the pairwise approach, we observe that the sharpness of the mean images and the tissue overlaps in Figure 1 look visually similar. From Table 1, we note that groupwise registration performs slightly better than the pairwise setting in most of the cases, especially as we increase the complexity of the warp. This suggests that considering the population as a whole and registering subjects jointly brings the population into better alignment than matching each subject to a mean template image. However, the advantage shown here is only slight; more comparative studies are needed of the two approaches.

4 Conclusion

In this work, we used a free-form deformation model based on B-splines to extend a previously demonstrated entropy based groupwise registration method. We described an efficient implementation of this method by using stochastic gradient descent in a multi-resolution setting. We applied the method to a set of 50 MRI brain scans and compared the results to a pairwise approach. We evaluated registration accuracy using label agreement in the group to predict the labels for new subjects and observed that a groupwise registration algorithm can achieve better accuracy than a pairwise approach by robustly handling multi-modal distributions. Our results indicate that nonrigid registration improves registration accuracy significantly, especially at cortical regions.

Acknowledgments

We are thankful to Lilla Zollei at Massachusetts General Hospital for insightful discussion. This work was in part supported by the NIH NIBIB NAMIC U54-EB005149, NAC P41-RR13218 and mBIRN U24-RR021382 grants, as well as NIH NINDS R01-NS051826 grant and the NSF CAREER 0642971 grant.

References

1. Zitova B, Flusser J. Image registration methods: A survey. Image and Vision Computing 2003;21(977–1000)
2. Crum WR, Hartkens T, Hill DLG. Non-rigid image registration: Theory and practice. The British Journal of Radiology 2004;77(140–153)
3. Toga A, Thompson P. The role of image registration in brain mapping. Image and Vision Computing 2001;19(3–24)
4. Seghers D, et al. Construction of a brain template from mr images using state-of-the-art registration and segmentation techniques. MICCAI 2004:696–703.
5. Park H, Bland PH, Hero AO, Meyer CR. Least biased target selection in probabilistic atlas construction. MICCAI 2005:419–426. [PubMed: 16685987]

6. Joshi S, Davis B, Jomier M, Gerig G. Unbiased diffeomorphic atlas construction for computational anatomy. NeuroImage 2004;23:151–160.
7. Christensen GE, et al. Synthesizing average 3d anatomical shapes using deformable templates. Medical Imaging 1999;Volume 3661:574–582.
8. Bhatia KK, Hajnal JV, Puri BK, Edwards AD, Rueckert D. Consistent groupwise non-rigid registration for atlas construction. IEEE ISBI. 2004
9. Twining C, et al. A unified information-theoretic approach to groupwise non-rigid registration and model building. IPMI. 2005
10. Blezek DJ, Miller JV. Atlas stratification. MICCAI 2006:712–719. [PubMed: 17354953]
11. Studholme C, Cardenas V. A template free approach to volumetric spatial normalization of brain anatomy. Pattern Recogn. Lett 2004;25(10):1191–1202.
12. Miller E, Matsakis N, Viola P. Learning from one example through shared densities on transforms. IEEE CVPR 2000:464–471.
13. Zollei L, et al. Efficient population registration of 3d data. Computer Vision for Biomedical Image Applications, ICCV 2005:291–301.
14. Duda, R.; Hart, P. John Wiley and Sons; 1973. Pattern Classification and Scene Analysis.
15. Rueckert D, et al. Nonrigid registration using free-form deformations: Application to breast mr images. IEEE TMI 2003;22:120–128.
16. Klein S, Staring M, Pluim JP. A comparison of acceleration techniques for nonrigid medical image registration. WBIR 2006:151–159.
17. 2007. ”http://www.na-mic.org/svn/NAMICSandBox/trunk/MultiImageRegistration/”:”na-mic sandbox svn repository”
18. Pohl K, et al. Incorporating non-rigid registration into expectation maximization algorithm to segment mr images. MICCAI 2002:564–572.

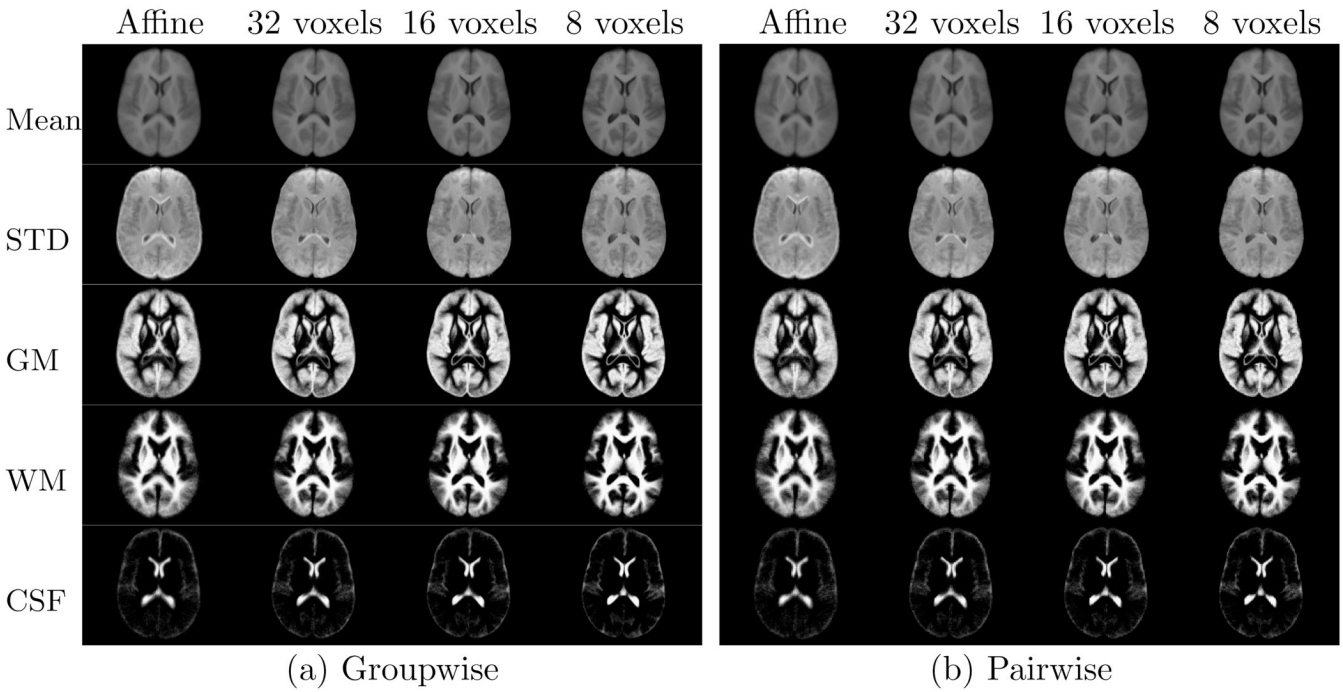

Fig. 1.

Central slices of 3D volumes for groupwise and pairwise approaches. Rows show mean and standard deviation images followed by label overlap images for GM, WM and CSF labels. The intensity of standard deviation images is scaled by 5 for visualization only. Columns display the results for affine and B-splines with grid spacing of 32 voxels, 16 voxels and 8 voxels, respectively.

Table 1

Average percent overlap (Dice) between the predicted and true labels for automatically segmented tissue types: GM,WM and manual segmentations for cortical and subcortical structures. Standard deviation across population are shown. Methods yielding higher overlap values are highlighted in boldface.

	GM		WM		Subcortical		Cortical	
	group	**pair**	**group**	**pair**	**group**	**pair**	**group**	**pair**
Affine	51 ± 2	51 ± 2	55 ± 2	55 ± 2	**55 ± 8**	53 ± 9	**18 ± 5**	16 ± 4
Bspline, 32 voxels	55 ± 2	55 ± 2	58 ± 1	58 ± 1	**57 ± 8**	56 ± 8	**28 ± 5**	26 ± 5
Bspline, 16 voxels	**57 ± 2**	56 ± 2	**61 ± 3**	60 ± 2	**56 ± 9**	53 ± 10	**32 ± 6**	29 ± 7
Bspline, 8 voxels	**60 ± 2**	59 ± 2	**65 ± 4**	63 ± 3	**53 ± 10**	51 ± 10	**38 ± 8**	34 ± 8