Subject-Matched Templates for Spatial Normalization

Torsten Rohlfing¹, Edith V. Sullivan², and Adolf Pfefferbaum^{1,2}

¹ Neuroscience Program, SRI International, Menlo Park, CA, USA
torsten@synapse.sri.com, dolf@synapse.sri.com

² Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford CA, USA
edie@stanford.edu

Abstract. Spatial normalization of images from multiple subjects is a common problem in group comparison studies, such as voxel-based and deformation-based morphometric analyses. Use of a study-specific template for normalization may improve normalization accuracy over a study-independent standard template (Good et al., *NeuroImage*, 14(1):21-36, 2001). Here, we develop this approach further by introducing the concept of subject-matched templates. Rather than using a single template for the entire population, a different template is used for every subject, with the template matched to the subject in terms of age, sex, and potentially other parameters (e.g., disease). All subject-matched templates are created from a single generative regression model of atlas appearance, thus providing a priori template-to-template correspondence without registration. We demonstrate that such an approach is technically feasible and significantly improves spatial normalization accuracy over using a single template.

1 Introduction

A essential task in group comparison studies is the normalization of all individual anatomies in a subject population to a joint template, which provides the reference coordinate system for statistical analysis. Hundreds such studies haven been published in the fields of voxel-based morphometry [1] and deformation-based morphometry [2] alone, and others come from functional and diffusion tensor magnetic resonance (MR) imaging studies.

As noted by Good *et al.* [3], it is typically advantageous to use a template generated from the population specific to a study itself, rather than a generic template shared by many different studies. Despite many years of research, the image registration algorithms used for normalization are still imperfect. In particular, these algorithms tend to produce less accurate coordinate correspondences when the images to register are very different, be it in terms of morphology or image intensities. Thus, it is not surprising that a study comparing the brains of younger adults would achieve better normalization using a template representing a younger brain than one representing an older brain.

This example can be considered an application of a minimum deformation template (MDT) [4]. The fundamental principle of MDT is to construct a template that minimizes, on average, the deformation applied to all individual images when they are deformed to that template, thus increasing the average registration accuracy as deformation is minimized.

Fig. 1. Illustration of the difference between a minimum-deformation template and subject-matched templates from a single regression appearance model. (a) Each subject image (numbered 1 through 7) is registered to the Minimum Deformation Template. (b) Each subject image is registered to its own subject-matched template. All subject-matched templates relate to the model mean template via known transformations (dashed arrows), which do not need to be computed by registration.

One problem with study-specific templates is that their unique quality poses a serious challenge for comparison of results across studies. We have recently suggested, albeit not experimentally demonstrated, a technique [5], to address this problem by generating study-appropriate templates, rather than study-specific ones, from a single study-independent appearance model of brain anatomy, so that all different templates remain compatible via *a priori* dense correspondences between them. The appearance model itself is created by a regression-based modeling framework, which is similar to the shape-regression framework by Davis *et al.* [6].

Our contribution in this paper is to take the idea of study-specific or study-appropriate templates further by using subject-matched individual templates. In other words, we propose to use a different template for normalization of each subject's images, such that the subject demographic parameters (e.g., age, sex) match the corresponding template parameters. As all templates are generated from the same regression appearance model, they all relate to a common “mean model” via *a priori* coordinate transformations. The mean model provides a natural reference system for all studies that use instances of the same model as their templates.

The concept of using subject-matched instances of an appearance model as templates is similar to registration using an active appearance model [7], wherein registration is performed by varying model parameters until the generated model instance optimally matches the target image. Our approach is substantially different, however, in that it decouples the model parameter determination, which we achieve directly by using a regression-based rather than PCA-based model, from the residual nonrigid registration. The ultimate normalization transformation for each subject is thus a concatenation of a transformation from mean template to subject-matched template, which is defined by the appearance model, with a second transformation from subject-matched template to subject image, which is determined via registration. This is illustrated in Fig. 1.

The goal of our proposed method is to improve registration accuracy and achieve better normalization of all subjects by reducing the residual deformations between subject-matched templates and subject images. We demonstrate the effect in this paper by normalizing images from 64 normal subjects to subject-matched as well as per-study templates and by comparing the overlaps of tissue segmentations as a measure of spatial normalization accuracy.

2 Methods

2.1 Test Subjects

To test spatial normalization, MR images from 64 normal controls, 30 men and 34 women, age range 22.4 to 79.2 years ($\text{mean} \pm \text{std.dev.} = 50.5 \pm 15.2$ years) were acquired as part of an ongoing study in our laboratory. For each subject, a T_1 -weighted three-dimensional image with $256 \times 256 \times 124$ pixels (pixel size $0.9375 \times 0.9375 \times 1.25$) was acquired on a 3T GE clinical scanner using a SPoiled Gradient Recalled echo (SPGR) sequence. All images were bias-field corrected [8] using in-house software, skull stripped using FSL BET [9], and segmented into three tissue classes (gray matter: GM, white matter: WM, cerebrospinal fluid: CSF) using FSL FAST [10].

2.2 Template Model Generation

The template model was created from MR images of 36 normal subjects (none of them part of the above test set) scanned at 3T. Skull-stripped SPGR images from all subjects were aligned using a simultaneous groupwise nonrigid registration algorithm [11]. A regression appearance model was then created that relates the two independent variables age and sex to anatomical shape and image intensities. Details of the regression appearance modeling procedure and the input data used here can be found in Rohlfing *et al.* [5].

In short, all input images were first aligned using template-free, unbiased groupwise nonrigid image registration. Template shape was then modeled analogous to the active deformation model [12], but using generalized multi-linear regression instead of principal components analysis. The resulting model relates each subject's image space to a common template space, where the latter depends on the independent variables of the model. To create actual atlas images in template space, image intensity at each pixel was also modeled via the same regression model, rather than simple averaging [13]. The result of the modeling procedure is an appearance model that can be instantiated for arbitrary values of the independent variables. For the model used here, each such instance represents a brain of given age and sex, but all instances are related to each other via known coordinate transformations defined by the regression model.

2.3 Experimental Procedure

The SPGR image of each of the 64 test subjects was registered independently to: a) each of seven mixed-sex template, instantiated for ages 20, 30, 40, 50, 60, 70, and 80 years, b) a template that matched subject age and sex, c) a mixed-sex template that matched subject

Fig. 2. Templates created from the continuous atlas regression model for ages 20 through 80 years in 10 year increments. *Top row:* SPGR channel; *bottom row:* three-class tissue segmentation.

Fig. 3. Scatter plots and linear regression fits of tissue overlaps for (a) age-matched vs. age- and sex-matched templates, and (b) age-matched vs. decade-matched templates.

age, and d) a mixed-sex template that matched the age of the subject by closest decade¹. All registrations were computed via a variant of the algorithm by Rueckert *et al.* [14] with a multi-resolution deformation strategy and 2.5 mm final control point spacing.

For each template that was needed for registration, SPGR and tissue segmentation images were created at 1 mm isotropic resolution (shown in Fig. 2 for the decade atlases). The SPGR channel was used for registration to the subject SPGR images, the tissue segmentation channel was used for computing the overlaps (in terms of fraction of matching pixels) between the template tissue image and the subject tissue images,

¹ Registration to the template with the closest age by decade was already covered by the registrations mentioned in a), but for the decade-matched template evaluation we grouped the results differently.

Fig. 4. Tissue overlaps plotted vs. subject-template age difference. (a)–(g), results for mixed-sex templates instantiated for ages 20, 30, 40, 50, 60, 70, and 80 years. See text for details.

after reformatting the latter into template space. To make the results comparable across templates, tissue images of all subjects were actually reformatted, via concatenation of transformations, into the space of the model mean template (mixed-sex, age=52.2 years) and overlaps computed in that space.

3 Results

The different subject matching techniques (age matched, age and sex matched, decade matched) are compared via scatter plots in Fig. 3. These results suggest that all three matching strategies work comparably well. Because including sex and matching the exact subject age each increase the number of templates that need to be generated from the atlas appearance model, we limit further consideration to the decade-matched templates as the more computationally efficient option.

The influence of age difference between subject and template on tissue overlap is illustrated in Fig. 4. In each plot, the mean of overlaps for the decade-matched templates is shown for comparison as a dashed horizontal line, and the range of ± 1 standard deviation as a gray box. Each plot also shows a second-order polynomial regression line fitted to the individual overlaps via nonlinear least squares. These plots, in particular the incremental tilt of the regression lines, suggest that indeed normalization accuracy decreases with increasing age difference, although it appears that younger subjects are more easily registered to an older template than the other way around.

As Fig. 5 shows, the average performance of middle-aged templates comes close to the performance of the decade-matched templates. Going back to Fig. 4, however, it is clear that the situation would be more favorable for the decade-matched templates if we excluded the middle-aged subjects from the test set and compared only the very young and the very old.

Fig. 5. Comparison of tissue overlaps over all subjects vs. template age. *Left:* mean $\pm$ standard deviation of overlap vs. template age. For comparison, the dashed line and gray box show mean $\pm$ standard deviation for decade-matched templates. *Right:* Actual numerical values represented by the plot on the left, including results of two-sided, paired t-tests of overlap values by subject for each of the single atlases vs. decade-matched subject-specific atlases.

4 Discussion

This paper has introduced the concept of using individually age-matched templates for spatial normalization of neuroimage data. Because all matched templates were generated from a single, regression-based appearance model, they were all related to each other through known spatial correspondences without the need for template-to-template registration.

Using test data from 64 subjects and a model from a separate set of 36 subjects we have demonstrated that matched templates can slightly, but significantly, increase the accuracy of spatial normalization. By excluding middle-aged subjects from the test set, we could have made the results appear even more favorable, but chose not to do so because it is important to investigate the performance of our method even in less-than-ideal circumstances.

Likewise, using a weaker nonrigid (or even an affine) registration algorithm for testing would have amplified the superiority of our method, because clearly a perfect registration algorithm would deliver perfect label overlaps regardless of the template used. Instead, we still observed an improvement in label overlap using a registration algorithm [14] that has been found to be at least on par with all other currently available algorithms (“IRTK”, see [15]).

Using matched templates for normalization incurs only moderate additional computational complexity once the appearance model for template generation has been created. It is particularly encouraging to note that we achieved essentially the same results using templates matched to subjects only by age rounded to the nearest decade, which greatly reduces the number of templates needed to cover a subject population.

Fundamental problems with our approach could arise in situations where, for example, the actual ages of subjects do not match their “brain ages,” e.g., due to neurodegenerative disorders. The obvious solution to this problem would be to create subject-matched templates using a more appropriate regression model. Using a model that includes disease factors as additional independent variables would then allow subjects to be matched to templates in terms of these variables as well, in addition to age and sex.

Acknowledgments. This work was supported under Grants AG019717, AA017923, AA005965, and AA017347. Source code of the software tools used for registration and modeling is available in the Computational Morphometry Toolkit (CMTK), <http://nitrc.org/projects/cmtk/>, supported by the NIBIB under Grant EB008381.

References

1. Ashburner, J., Friston, K.J.: Voxel-based morphometry — the methods. *NeuroImage* 11(6), 805–821 (2000)
2. Ashburner, J., Hutton, C., Frackowiak, R., Johnsrude, I., Price, C., Friston, K.: Identifying global anatomical differences: Deformation-based morphometry. *Hum. Brain Map.* 6(5-6), 348–357 (1998)
3. Good, C.D., Johnsrude, I.S., Ashburner, J., Henson, R.N.A., Friston, K.J., Frackowiak, R.S.J.: A voxel-based morphometric study of ageing in 465 normal adult human brains. *NeuroImage* 14(1), 21–36 (2001)

4. Kochunov, P., Lancaster, J.L., Thompson, P., Woods, R., Mazziotta, J., Hardies, J., Fox, P.: Regional spatial normalization: toward an optimal target. *J. Comput. Assist. Tomogr.* 25(5), 805–816 (2001)
5. Rohlfing, T., Sullivan, E.V., Pfefferbaum, A.: Regression models of atlas appearance. In: Prince, J.L., Pham, D.L., Myers, K.J. (eds.) *IPMI 2009*. LNCS, vol. 5636, pp. 151–162. Springer, Heidelberg (2009)
6. Davis, B.C., Fletcher, P.T., Bullitt, E., Joshi, S.: Population shape regression from random design data. In: *IEEE 11th International Conference on Computer Vision, ICCV*, October 2007, pp. 1–7 (2007)
7. Cootes, T.F., Beeston, C.J., Edwards, G.J., Taylor, C.J.: A unified framework for atlas matching using active appearance models. In: Kuba, A., Sámal, M., Todd-Pokropek, A. (eds.) *IPMI 1999*. LNCS, vol. 1613, pp. 322–333 Springer, Heidelberg (1999)
8. Likar, B., Viergever, M.A., Pernus, F.: Retrospective correction of MR intensity inhomogeneity by information minimization. *IEEE Trans. Med. Imag.* 20(12), 1398–1410 (2001)
9. Battaglini, M., Smith, S.M., Brogi, S., De Stefano, N.: Enhanced brain extraction improves the accuracy of brain atrophy estimation. *NeuroImage* 40(2), 583–589 (2008)
10. Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imag.* 20(1), 45–57 (2001)
11. Balci, S.K., Golland, P., Shenton, M., Wells, W.M.: Free-form B-spline deformation model for groupwise registration. In: *MICCAI 2007 Workshop Statistical Registration: Pair-wise and Group-wise Alignment and Atlas Formation*, pp. 23–30 (2007)
12. Rueckert, D., Frangi, A.F., Schnabel, J.A.: Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. *IEEE Trans. Med. Imag.* 22(8), 1014–1025 (2003)
13. Guimond, A., Meunier, J., Thirion, J.P.: Average brain models: A convergence study. *Comput. Vision Image Understanding* 77(2), 192–210 (2000)
14. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L.G., Leach, M.O., Hawkes, D.J.: Nonrigid registration using free-form deformations: Application to breast MR images. *IEEE Trans. Med. Imag.* 18(8), 712–721 (1999)
15. Klein, A., Andersson, J., Ardekani, B.A., Ashburner, J., Avants, B., Chiang, M.C., Christensen, G.E., Collins, L.D., Gee, J., Hellier, P., Song, J.H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R.P., Mann, J.J., Parsey, R.V.: Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage* 46(3), 786–802 (2009)