Computational Anatomy to Assess Longitudinal Trajectory of Brain Growth

G. Gerig[1,3], B. Davis[1,2], P. Lorenzen[1,2], Shun Xu[1], M. Jomier[3], J. Piven[3] and S. Joshi[1,2]

Departments of [1]Computer Science, [2]Radiation Oncology and [3]Psychiatry

University of North Carolina

Chapel Hill, NC 27599

Abstract

This paper addresses the challenging problem of statistics on images by describing average and variability. We describe computational anatomy tools for building 3-D and spatio-temporal 4-D atlases of volumetric image data. The method is based on the previously published concept of unbiased atlas building, calculating the nonlinear average image of a population of images by simultaneous nonlinear deformable registration. Unlike linear averaging, the resulting center average image is sharp and encodes the average structure and geometry of the whole population. Variability is encoded in the set of deformation maps. As a new extension, longitudinal change is assessed by quantifying local deformation between atlases taken at consecutive time points. Morphological differences between groups are analyzed by the same concept but comparing group-specific atlases. Preliminary tests demonstrate that the atlas building shows excellent robustness and a very good convergence, i.e. atlases start to stabilize after 5 images only and do not show significant changes when including more than 10 volumetric images taken from the same population.

1. Introduction

Statistical modeling is concerned with the construction of a compact and stable description of the population mean and variability. Applied to a population of images taken from a common domain and only differing by natural variability, we are facing the fundamental question of the definition of an *average image*. It is obvious that averaging images after linear alignment of pose, e.g. by affine registration, cannot be sufficient and results in a blurred result (see for example Fig. 2 top). Variability among the set of images requires alignment of much higher order, to ensure that features are in correspondence. This geometric variability of the anatomy cannot be represented by elements of a flat space since the space of transformations is not a vector space but rather the infinite dimensional group of diffeomorphisms of the underlying coordinate system. A fundamental difficulty for the development of a processing scheme is the high dimensionality of the set of features given the relatively small sample size, typically 256^3 features versus 20 to 50 sample images in volumetric neuroimaging studies as presented in this paper.

Mapping of populations of 3-D images into a common coordinate space is a central topic of the brain mapping neuroimaging community. The mapping establishes correspondence between sets of images for statistical testing of group differences and for warping subject images into a reference frame with known anatomical coordinates (e.g. Talairach atlas [1]). Whereas linear global alignment was sufficient in early analysis of coarse resolution positron emission tomography brain imaging (PET) and to some extent in functional MRI (fMRI), todays spatial resolution of scanning technology has significantly improved and shows excellent levels of detail. This improvement in spatial resolution requires appropriate new analysis methods.

Most digital brain atlases so far are based on a single subject's anatomy [1, 2, 3, 4]. Although these atlases provide a standard coordinate system, they are limited because a single anatomy cannot faithfully represent the complex structural variability between individuals. Extending this framework to be less dependent on a single template, Rohlfing et al. [5] showed automatic segmentation of images by transforming sets of labeled templates to a new unknown image, encoding variability of the atlas.

A major focus of computational anatomy has been the development of image mapping algorithms [6, 7, 8, 9] that can map and transform a single brain atlas on to a population. In this paradigm the atlas serves as a deformable template[10]. The deformable template can project detailed atlas data such as structural, biochemical, functional as well as vascular information on to the individual or an entire population of brain images. The transformations encode the variability of the population under study. Statistical analysis of the transformations can be used to characterize different populations [11, 12, 13, 14]. For a detailed review of deformable atlas mapping and the general framework for computational anatomy see [15, 16].

Most other previous work [17, 18] in atlas formation has focused on the small deformation setting in which arithmetic averaging of displacement fields is well defined. Guimond et. al develop an iterative averaging algorithm to re-

duce the bias [17]. In the latest work of [18], explicit constraints requiring that the sum of the displacement fields add to zero is enforced in the proposed atlas construction methodology. These small deformation approaches are based on the assumption that a transformations of the form $h(x) = x + u(x)$, parameterized via a displacement field, $u(x)$, are close enough to the identity transformation such that composition of two transformations can be approximated via the addition of their displacement fields:

$$h_1 \circ h_2(x) \approx x + u_1(x) + u_2(x) .$$

In more recent and related work Avants and Gee [19, 20] developed an algorithm in the large deformation diffeomorphic setting for template estimation by averaging velocity fields.

One of the fundamental limitations of using a single anatomy as a template is the introduction of a bias based on the arbitrary choice of the template anatomy. The focus of this paper is to build on methodology developed by Avant et al. [20] and Joshi et al. [21, 22] which describe contruction of atlases by simultaneously estimating the set of transformations and the unbiased template in the large deformation setting. The unbiased template is a result of the processing, and is representing a new image centered in the population of images. We will discuss stability, convergence and validation, and will extend this framework to 4-D by describing deformation between atlases.

2 Methodology

This section reviews existing methods and provides analysis of of stability and convergence. In later sections, we will extend the concept to longitudinal and cross-sectional group difference analysis and present preliminary results from an ongoing clinical research study.

2.1 Unbiased atlas building

Construction of atlases is a key procedure in population-based medical image analysis. A simple averaging of images after a linear transformation, most often affine, is know to result in a blurred image. Nonlinear registration to a template requires the choice of a template that is close to the expected average, but the result is biased by the choice of a template. Both problems can be overcome by nonlinear processing via large deformation registration and population-based simultaneous nonlinear averaging of sets of images [20, 21].

We follow the notation as presented in [21]. As discussed in the introduction, the geometric variability of anatomy is not a vector space. For representations in which the underlying geometry is parametrized as a Euclidean vector space, training data can be represented as a set of vectors $x_1, \cdots, x_N$ in a vector space V. In a vector space, with addition and scalar multiplication well defined, an average representation of the training set can be computed as the linear average

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i .$$

In the group of diffeomorphisms, the addition of two diffeomorphisms is not generally a diffeomorphism and, hence, a template based on linear averaging of transformations is not well defined.

Frechet [23] extended the notation of averaging to general metric spaces. For a general metric space M, with a distance $d : M \times M \to R$, the *intrinsic mean* for a collection of data points x_i can be defined as the minimizer of the sum-of-squared distances to each of the data points. That is

$$\mu = \operatorname*{argmin}_{x \in M} \sum_{i=1}^{N} d(x, x_i)^2 .$$

This approach, combined with the mathematical metric theory of diffeomorphisms developed by Miller and Younes [24], presents the core of the unbiased atlas methodology. A detailed description is not the topic of this paper, and a reader might consult the literature for more details.

Applied to sets of images, we need to solve the following estimation problem. Given a metric on a group of transformations, the template construction problem can be stated as that of estimating an image $\hat{I}$ that requires the minimum amount of deformation to transform into every population image I_i. More precisely, given a transformation group $\mathcal{S}$ with associated metric $D : \mathcal{S} \times \mathcal{S} \to R$, along with an image dissimilarity metric $E(I_1, I_2)$, we wish to find the image $\hat{I}$ such that

$$\{\hat{h}_i, \hat{I}\} = \operatorname*{argmin}_{h_i \in \mathcal{S}, I} \sum_{i=1}^{N} E(I_i \circ h_i, I)^2 + D(e, h_i)^2$$

where e is the identity transformation and h_i the resulting deformation maps.

Results presented in this paper are based on a greedy implementation of fluid flow algorithm. We are currently working on implementing the full space time optimization based on the Euler-Lagrange equations derived in [24].

Figure 1 illustrates the construction of an atlas of a population of 14 3-D MRI images of pediatric subjects. A qualitative check shows that the resulting atlas is still sharp and that its anatomical objects seem to represent the expected average shape geometry. The resulting deformation maps guaranteee diffeomorphism, i.e. transformations are invertible and do not show eventual overfolding of space. Invertible transformations are of significant advantage as users

can transform images to an atlas but also atlas information back to the whole set of images, e.g. for automatic labeling or segmentation.
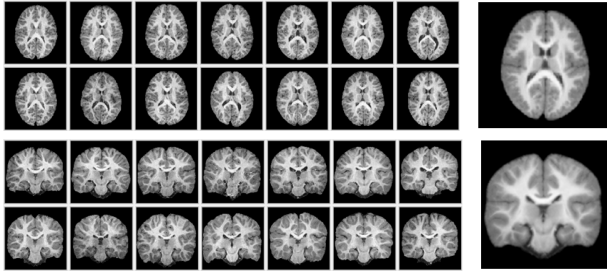

Figure 1: Atlas building by simultaneous nonlinear deformation of a population of images. A set of 14 3D MRI datasets of a clinical pediatric database is processed. Top and bottom row show axial and coronal sections of the set of images (left) and the resulting atlas (right).
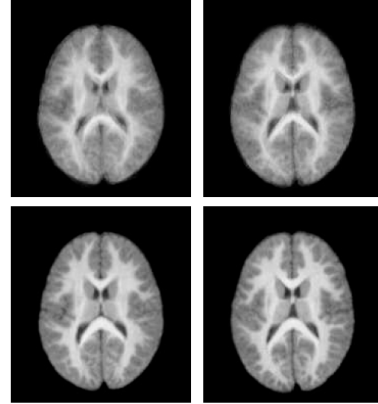

Figure 2: Test of stability by creating atlases from two mutually exclusive sets of 7 images (left and right column). Top and bottom rows illustrate linear averaging and deformable registration.

2.2. Stability and convergence

The stability and convergence rate of atlas building is discussed by Lorenzen et al. [25]. To qualitatively test stability, atlases were built from two sets of mutually exclusive sets of 7 images each (Fig. 2). The anatomy represented by two atlases becomes very similar. A comparison of top and bottom rows clearly demonstrates the improved quality and sharpness of the deformable registration scheme over linear processing. Convergence was tested by creating atlases from a different number of images selected from the same population. Figure 3 top illustrates three orthogonal slices of atlases created from 9, 10, 13 and 14 images. Entropy as a quantitative measure of convergence has been proposed by Lorenzen et al. [25]. This metric is based on the assumption that high-quality sharp atlases need less number of bits to represent the distribution of image intensities. Entropy is defined as $H(p) = -\sum_{i=1}^{n} p(i) \log p(i)$, where $[p(i), \ i = 1 \cdots n]$ represents the image intensity histogram with n bins. In addition to entropy, we currently explore the use variance across the stack of co-registered images (see Fig. 3 third and sixth rows) and the use of gradient histogram information, both measures related to image sharpness.

3. Motivation and clinical study

Imaging studies of early brain development get increasing attention as improved modeling of the pattern of normal development might lead to a better understanding of origin, timing and nature of morphologic differences in neurodevelopmental disorders. Quantitative MR imaging studies face the challenge that cross-sectional inter-individual variability is very large in relation to longitudinal change, which underscores the critical importance of a longitudinal design of such studies. It is our goal to model the trajectory of early brain development, primarily focusing on the most challenging group of very young children in the age range from birth to 6 years, as a 4-dimensional atlas that is represented by a time series of 3-D images and quantitative description of local growth. In addition, the same technique will be applied to generate representative atlases for various groups, e.g. group-specific atlases for female/male populations and for healthy controls and patients.

This project is driven by the needs of several clinical pediatric studies at UNC Chapel Hill. This includes an autism study (51 autistic (AUT) and 25 control individuals (14 typically developing (TYP), 11 developmentally delayed (DD)) with baseline scans at 2 years and follow-up at 4 years. So far, the new method has been applied to a subset of subjects from this autism study. We have selected 5 subjects each from the TYP and AUT groups. For eight of these subjects, we had longitudinal data and could include MRIs at 2yrs and 4 years of age. We applied the unbiased atlas building procedure to the two groups at both time points and computed image deformations longitudinally (comparing the 2yrs and 4yrs atlases for AUT and TYP) and cross-sectionally (TYP versus AUT at 2yrs and 4yrs).

4. Analysis of growth trajectories

Longitudinal analysis is known to have increased power over cross-sectional group tests as it compares the rate of change and not absolute differences. In neuroimaging applications, cross-sectional variability is usually much higher

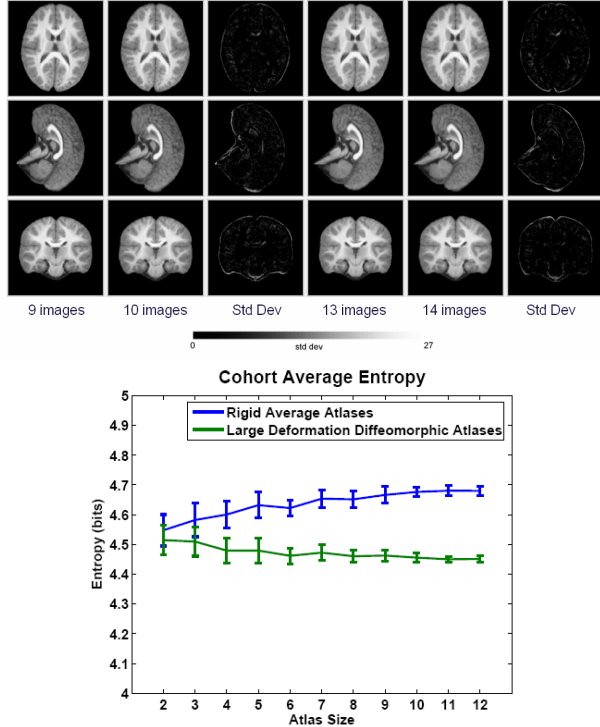| 9 images | 10 images | Std Dev | 13 images | 14 images | Std Dev |

Figure 3: Convergence demonstrated by building atlases from different numbers of images (top). The third and sixth columns represent intensity variability of the set of images for the atlases with 10 and 14 images. The bottom image shows entropy calculated for linear and nonlinear atlas building, with error estimates obtained by permutation tests.

than the expected change over time, even in pediatric populations (see Fig. 4 top). We propose to use unbiased atlas building to create average images of populations at different time points and then compare deformation between the resulting atlases via quantitative analysis of the deformation field.

4.1. Invididual growth versus growth trajectory between group atlases

The existing set of sample images allows a validation of the deformation between atlases. Since we have subjects with baseline and follow-up scans (see Fig. 4), we can compare the growth pattern obtained from the atlas with the growth patterns of the invidual cases. Deformable registration was calculated between scans of each subject at age 2 and 4. The log of the determinant of the Jacobian was then mapped into the common atlas space for point to point comparison. These five deformation maps can be compared to the deformation map between the pair of atlases. Fig. 5 illustrates the five images and atlases at both time points and the resulting
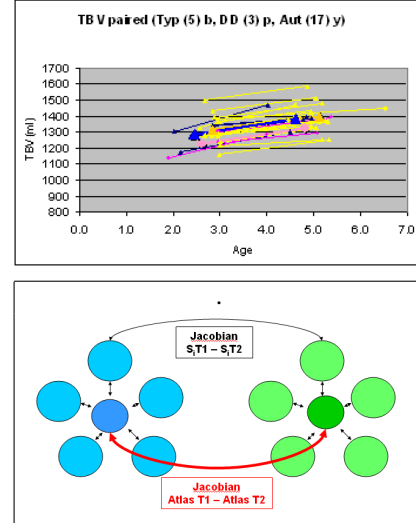

Figure 4: Design of our experimental study. Top: Age versus total brain volume (TBV) plotted for autistic (yellow), typically developing (blue) and developmentally delayed (pink) subjects. Connecting lines represent individuals measured at two different times. The growth rate is significantly smaller than the cross-sectional variability. Bottom: Diagram illustrating the concept of 4-D growth modeling. The deformation map between the two atlases describes local volumetric change. The change pattern between the atlases can be validated by calculating deformable registrations between individual subjects and map these into the common atlas space.

deformation maps. Qualitatively, each of the individual images represents the asymmetry pattern as also observed in the atlas. The concentration of growth in cortical gray matter shown in the atlas space is also represented in each individual subject. Beyond this purely qualitative analysis, we are currently developing a testing scheme for quantitative comparison of the Jacobian maps.

4.2. Extension to growth trajectory analysis between groups

Following the processing outlined before, we can extend the scheme by calculating unbiased atlases for different groups at different time points. In a preliminary experiment, we selected 5 children from a typically developing (TYP) group and 5 from an autistic (AUT) group. For 8 of these subjects, we had baseline scans at year 2 with follow-up at year 4. Two more subjects without follow-up were added to each group based on optimal match of gender and age. We built atlases for each of the four groups, TYP and AUT at age 2 and 4 years. Figure 7 illustrates the four atlases and the four deformation maps color-coded as differences in the mag-
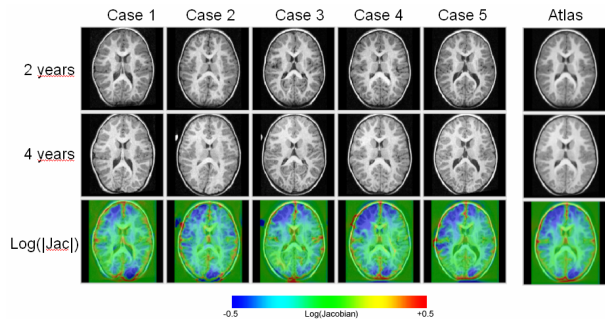
Figure 5: Validation of local growth pattern between the five individual subjects and the resulting atlases. Top and middle row represent images and atlas at age 2 and 4 years. Bottom row shows the color-coded $\log(|Jacobian|)$, which blue and red representing local growth and atropy.
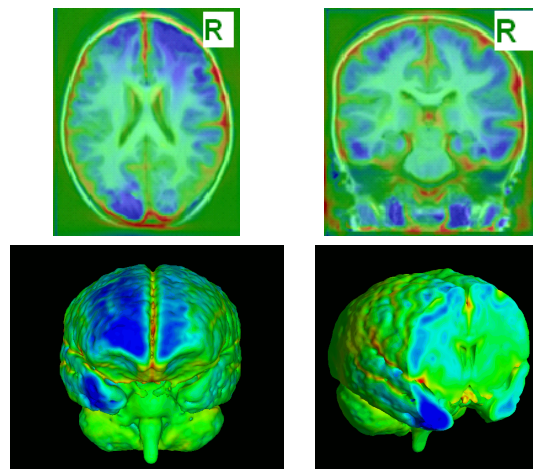

Figure 6: 2-D and 3-D illustrations of local volumetric growth between populations at age 2 to 4 years. Blue indicates local growth, green regions of no change, and red local atrophy. Please note the significant asymmetry of growth of the right frontal and right temporal regions. Also, growth regions are mostly located at the rim of cortical gray matter.

nitude of the deformation fields. Again, the similarity of the four atlases is striking, especially between the TYP and AUT groups. We calculated local growth maps between groups and across age (see Fig. 8. The growth maps for both groups could even be compared via another processing step, a scheme that would represent a group analysis between growth patterns. Such analysis is highly relevant for studying early development in patient groups to explore differences in location and timing of brain growth, which can be associated with differences in maturation of specific brain functions. Entropies for the four atlases are calculated as 3.70 and 4.06 for TYP at age 2 and 4, and 3.65 and 3.87 for AUT at age 2 and 4. The observed smaller value for AUT at the second time point, also reflected in the entropies of each invididal image (not shown here), needs further explanation.

5. Results

We have presented a computational anatomy methodology to build nonlinear averages of population images and to study differences between atlases. All the results are preliminary feasibility tests on a very small set of sample images, and ongoing work is extending the analysis to the full set of images in this study. Stability and convergence is demonstrated by qualitative and quantitative comparison. Our preliminary tests demonstrate that the atlas building shows excellent robustness and a very good convergence, i.e. atlases start to stabilize with 5 images only and do not show significant changes when including more than 10 volumetric images, which is a surprisingly low number given the complex appearance of the 3-D volume data. Currently, we are developing improved metrics to measure the quality of atlases and to systematically explore stability and convergence in a quantitative study. Our processing uses a diffeo-

morphic registration scheme, assuming that there is a one-to-one correspondence between features in sets of images. The fluid warping further makes the assumption that there is a continuous flow along geodesic paths between these images. In studies of brain images, we see that these assumptions are applicable to the type of changes under analysis. A generalization to other regions of the body or images with pathology, however, is not straightforward since topology changes cannot be described by the proposed technique.

So far, we have developed a scheme to generate average images. Variability is encoded in the set of deformation fields which maps each image to the unbiased atlas. We will develop a new method for analysis of this variability, which is necessary to for a full group hypothesis testing scheme. In parallel, we are investigating the properties of anatomical shapes embedded in the volumetric images [26, 27]. Although the atlases look correct, they do not provide explicit information about the preservation of shapes and the relationship of resulting average shapes to the set of original shapes. Explicit shape statistics of parametrized objects will be compared with point-to-point correspondence provided by the unbiased atlas scheme.

The most striking result of the longitudinal growth analysis between 2 and 4 years is the apparent cerebral asymmetry and brain torque. There is a consistent right frontal > left frontal and a left posterior parietal/occipital > right posterior parietal/occipital pattern, commonly called torque or brain torsion. This growth trajectory finding is consistent for both the TYP and AUT groups. Gender differences
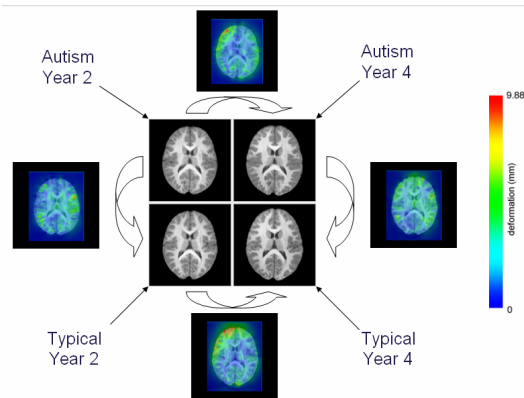
Figure 7: Atlas-building applied to a longitudinal clinical study with autistic subjects and typically developing children scanned at age 2 and 4 years. Studying deformation between atlases allows cross-sectional (groups at same age) and longitudinal (each group across time) analysis. Longitudinal growth maps for both groups can then be compared, representing an analysis scheme for growth analysis between groups. The four atlases look very similar, again demonstrating the excellent stability of the atlas building.

could not yet be explored due to the small sample size. The temporal lobes show a similar pattern as the frontal lobes, with right temporal > left temporal growth. Local growth is mostly evident in cortical gray, which seems to account for the major brain growth during this age period. Lateral ventricles are stable, but the third ventricle illustrates a significant width reduction, along with a closing of the ascending ramus of the Sylvian fissure. Group tests between TYP and AUT subjects reveal a strong size difference of the cerebellum, which is much more pronounced at age 2 and lessens towards age 4. We are currently confirming these exploratory findings with independent samples.

Our preliminary findings indicate that the new methodology shows excellent potential to explore longitudinal change, difference between groups, and differences between growth trajectories between groups. The simultaneous analysis of the whole volumetric brain is a major strength, as it will reveal morphometric changes of structures with embedding context, e.g. studying cortical growth in relationship to adjacent white matter, and examining groups of subcortical structures and even whole circuits.
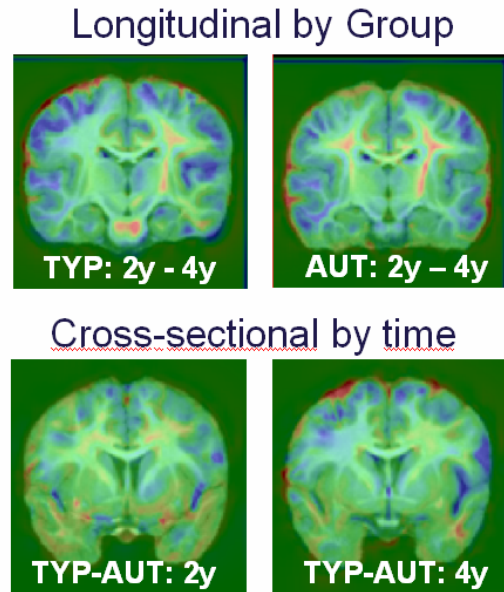
Acknowledgements

Figure 8: Results of longitudinal (top) and cross-sectional (bottom) group differences. Both groups show similar local growth patterns, concentrated in cortical gray. Group differences show a lateralized difference in the Sylvian fissue, in particular at age 4.

References

[1] J. Talairach and P. Tournoux, *Co-Planar Stereotaxis Atlas of the Human Brain*, Thieme Medical Publishers, 1988.

[2] K.H. Höhne, M. Bomans, M. Riemer, U. Tiede, R. Schubert, and W. Lierse, "A 3d anatomical atlas based on a volume model," *IEEE Comput. Graph. Appl.*, pp. 72–78, Dec 1992.

[3] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, and A. C. Evans, "BrainWeb: Online interface to a 3D MRI simulated brain database," *NeuroImage*, vol. 5, no. 4, 1997.

[4] D. L. Collins, A. P. Zijdenbos, V. Kollokian, J. G. Sled, N. J. Kabani, C. J. Holmes, and A. C. Evans, "Design and construction of a realistic digital brain phantom," *IEEE TMI*, vol. 17, no. 3, pp. 463–468, June 1998.

[5] Torstens Rohlfing, Daniel B. Russakoff, and Calvin.R. Maurer, "Extraction and application of expert priors to combine multiple segmentations of human brain tissue," in *Proceedings Medical Image Computing and*

Computer Assisted Intervention MICCAI, R.E. Ellis and T.M. Peters, Eds., 2003, vol. 2879 of *Springer LNCS*, pp. 578–585.

[6] A. W. Toga, *Brain Warping*, Academic Press, 1999.

[7] Michael I. Miller, Sarang C. Joshi, and Gary E. Christensen, "Large deformation fluid diffeomorphisms for landmark and image matching," in *Brain Warping*, Arthur W. Toga, Ed., chapter 7. Academic Press, 1999.

[8] J. C. Gee M. Reivich R. Bajcsy, "Elastically deforming an atlas to match anatomical brain images.," *J. COmput. Assis. Tomogr.*, vol. 17, pp. 225–236, 1993.

[9] Dinggang Shen and Christos Davatzikos, "Hammer: Hierarchical attribute matching mechanism for elastic registration," *IEE TMI*, vol. 21, no. 11, pp. 1421–1439, November 2002.

[10] U. Grenander, *General Pattern Theory*, Oxford Univ. Press., 1994.

[11] J. C. Csernansky S. Joshi L. Wang M. Gado J. P. Miller U. Grenander M. I. Miller, "Hippocampal morphometry in schizophrenia by high dimensional brain mapping," *Proceedings of the National Academy of Science*, vol. 95, pp. 11406–11411, September 1998.

[12] P. M. Thompson J. Moussai S. Zohoori A. Goldkorn A. A. Khan M. S. Mega G. W. Small J. L. Cummings A. W. Toga, "Cortical variability and asymmetry in normal aging and alzheimer's disease," *Cerebral Cortex*, vol. 8, no. 6, pp. 492–509, September 1998.

[13] S. Joshi, U. Grenander, and M.I. Miller, "On the geometry and shape of brain sub-manifolds," *International Journal of Pattern Recognition and Artificial Intelligence: Special Issue on Processing of MR Images of the Human*, vol. 11, no. 8, pp. 1317–1343, 1997.

[14] J. Ashburner and K.J. Friston, "High-dimensional image warping," in *Human brain function, 2nd edn.*, R. Frackowiak, Ed., pp. 673–694. Academic Press, 2004.

[15] A. W. Toga P. M. Thompson, "A framework for computational anatomy," *Computing and Visualization in Science*, , no. 5, pp. 13–34, 2002.

[16] U. Grenander M. I. Miller, "Computational anatomy: An emerging discipline," *Quarterly of Applied Mathematics*, vol. 56, pp. 617–694, 1998.

[17] A. Guimond, J. Meunier, and J.-P. Thirion, "Average brain models: a convergence study," *Comput. Vis. Image Underst.*, vol. 77, no. 2, pp. 192–210, 2000.

[18] K.K. Bhatia, J.V. Hajnal, B.K. Puri, A.D. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *IEEE International Symposium on Biomedical Imaging*, 2004.

[19] B. Avants J.C. Gee, "Symmetric geodesic shape averaging and shape interpolation," in *Computer Vision Approaches to Medical Image Analysis (CVAMIA) and Mathematical Methods in Biomedical Image Analysis (MMBIA) Workshop 2004 in conjunction with the 8th European Conference on Computer Vision, Prague, CZ.*, 2004.

[20] B. Avants and J.C. Gee, "Geodesic estimation for large deformation anatomical shape averaging and interpolation," *Neuroimage*, vol. 23, pp. 139–150, 2004.

[21] S. Joshi B. Davis M. Jomier G. Gerig, "Unbiased diffeomorphic atlas construction for computational anatomy," in *NeuroImage*, 2004, vol. 23, pp. S151–S160.

[22] B. Davis P Lorenzen S. Joshi, "Large deformation minimum mean squared error template estimation for computational anatomy," in *ISBI*, 2004, pp. 173–176.

[23] Maurice Fréchet, "Les elements aleatoires de nature quelconque dans un espace distancie," in *Ann. Inst. Henri Poincare*, 1948, number 10, pp. 215–310.

[24] M.I. Miller, A. Trouve, and L. Younes, "On the metrics and eulerlagrange equations of computational anatomy," *Annu. Rev. Biomed. Eng.*, vol. 4, pp. 375–405, 2002.

[25] Peter Lorenzen, Brad Davis, and Sarang Joshi, "Unbiased atlas formation via large deformations metric mapping," in *Medical Image Computing and Computer Assisted Intervention MICCAI*. Oct 2005, vol. 3750 of *Lecture Notes in Computer Science LNCS*, pp. 411–418, Springer Verlag.

[26] Shun Xu, Martin Styner, Brad Davis, Sarang Joshi, and Guido Gerig, "Group mean differences of voxel and surface objects via nonlinear averaging," in *Proc. International Symposium on Biomedical Imaging (ISBI'06), Macro to Nano*, April 2006.

[27] Guido Gerig, Sarang Joshi, Tom Fletcher, Kevin Gorczowsky, Shun Xu, Stephen M. Pizer, and Martin Styner, "Statistics of populations of images and its embedded objects: Driving applications in neuroimaging," in *Proc. International Symposium on Biomedical Imaging (ISBI'06), Macro to Nano*, April 2006.