Multi-Atlas Tensor-Based Morphometry and its Application to a Genetic Study of 92 Twins

Natasha Leporé, Caroline Brun, Yi-Yu Chou, Agatha Lee, Marina Barysheva, Greig I. De Zubicaray, Matthew Meredith, Katie Macmahon, Margaret Wright, Arthur Toga, et al.

Multi-Atlas Tensor-Based Morphometry and its Application to a Genetic Study of 92 Twins

Natasha Leporé[1], Caroline Brun[1], Yi-Yu Chou[1], Agatha D. Lee[1], Marina Barysheva[1], Greig I. de Zubicaray[2], Matthew Meredith[2], Katie L. McMahon[2], Margaret J. Wright[3], Arthur W. Toga[1], and Paul M. Thompson[1]

[1] Laboratory of Neuro Imaging, UCLA, Los Angeles, CA 90095, USA
[2] Centre for Magnetic Resonance, University of Queensland, Brisbane, Australia
[3] Genetic Epidemiology Lab, Queensland Institute of Medical Research, Brisbane, Australia

Abstract. Here we develop a multi-template analysis for tensor-based morphometry (TBM) which aims to reduce error from the registration step. In conventional TBM, a single template is nonlinearly registered to all images in the study, and the deformation field statistics are computed from the transformations. Using an MRI dataset from 23 monozygotic and 23 dizygotic twin pairs, we instead registered each individual twin image to 9 additional brain templates using a Riemannian fluid algorithm [3]. Average deformation tensors from multiple registrations were computed within each image, using a log-Euclidean framework [1]. To quantify improvements as the number of registration templates increased from 1 to 9, sequential t-tests assessed the significance of any error reduction, as each new template was added. For each number of templates, we also computed two tensor-derived metrics, and maps of the intraclass correlation of local volume differences, to evaluate any power advantages of multi-atlas TBM.

1 Introduction

Template selection is an important step in group analyses of brain MR images. In particular, in tensor-based morphometry (TBM), a set of images is non-linearly registered to a common reference image, and a statistical analysis is performed on the deformation tensors $S = \sqrt{J^T J}$, where J represents the Jacobian matrices derived from the deformation fields.

Detection power depends on several factors, and key among these is the quality of the non-linear registration, which also depends on the common target to which all images are mapped. Typically, nonlinear registration is performed either to one of the controls, or to an average of them [8, 7, 13]. In theory the construction of a mean anatomical atlas can reduce the bias induced by registering images to an individual control subject. In practice however, anatomical boundaries and image gradients are often blurrier in the average image, which may reduce the accuracy of the registration. In [5] for instance, the ICBM53 average brain was compared to a single individual template in a TBM study of

HIV/AIDS patients. Greater effect sizes per voxel were found with the single template when compared with those found using the ICBM53 brain as a registration target. Even so, individual variability may cause a bias in the registration when a single template is used, as the template may have a shape and intensity distribution that is closer to some subjects than to others. For instance, in [15], statistical power for TBM with single template registration was found to depend on which particular individual was selected as reference. Templates that were most *average-* in the sense of inducing the smallest deformation tensors when registered to other brains in the study- tended to generate more powerful statistics.

Our aim here was to design a new method that eliminates the dependence on individual variability, while retaining the sharp features associated with registration to a single target. One way to combine both of these requirements consists of moving the averaging step until after the non-linear registration. The gist of our averaging method consists of the following steps: starting from a set of brain MR images and a set of templates, we non-linearly register all images to all templates individually. We then compute the deformation tensors for each image as an average of those generated from the registration to individual templates. Other solutions to the template selection problem include 'targetless' normalization as in [22], [19], [21], [10], and the selection of the optimal individual target [9, 15].

In the standard version of TBM [19, 2], statistics are performed on the determinants of the Jacobian matrices, $detJ$, or equivalently on the determinants of the deformation tensors generated from the deformation field. In [14], multivariate statistics were computed instead on the full deformation tensors. As deformation tensors do not form a vector space under standard matrix addition and scalar multiplication, computations were performed in the log-Euclidean framework [1]. Since statistics are computed on the deformation tensors or a function of its components, for example the determinant, here we also perform a log-Euclidean averaging on those tensors.

Our analysis was performed on a dataset of MR images from 23 monozygotic (MZ) and 23 dizygotic (DZ) twin pairs, as well as 10 template brain MR images from identically scanned healthy subjects who did not belong to those pairs. We used the intraclass correlation of the $detJ$ as a statistic to characterize the influence of shared genes on local brain volume. Voxelwise permutation statistics were computed to assess the significance of the results.

2 Method

2.1 Data

3D T1-weighted images were acquired from 23 pairs of monozygotic twins (MZ; 20 males/26 females; 25.11.5SD years old) and 23 pairs of same-sex dizygotic twins (DZ; all same-sex pairs; 20 males/26 females; 23.52.2 years), as well as 10 individuals of comparable age, scanned identically. All MR images were collected

using a 4 Tesla Bruker Medspec whole body scanner (Bruker Medical, Ettingen, Germany) at the Center for Magnetic Resonance (University of Queensland, Australia). Three-dimensional T1-weighted images were acquired with an inversion recovery rapid gradient echo (MP-RAGE) sequence to resolve anatomy at high resolution. Acquisition parameters were as follows: inversion time (TI)/repetition time (TR)/echo time (TE) = 1500/2500/3.83 msec; flip angle =15 degrees; slice thickness = 0.9 mm, with an acquisition matrix of 256 x 256 x 256. The study was approved by the Institutional Review Boards at the University of Queensland and at UCLA; all subjects gave informed consent.

Non-brain tissues were removed from all images using the Brain Surface Extraction tool (BSE) of BrainSuite [18]. The masked image was spatially normalized to the Colin27 standard brain by a 9-parameter (3 translations, 3 rotations, 3 scales) transformation, using the FLIRT software [11].

2.2 Template averaging

Each individual image I was non-linearly registered to each of 9 templates using a fluid (large-deformation) version of a Riemannian registration algorithm [3, 17, 6], which guarantees diffeomorphic mappings. In order to compute averages in a common space, all templates were also registered to a tenth template. The deformation fields resulting from the first registration step were concatenated with those of the second registration to obtain 9 sets of deformation fields in the common space for each image I. The deformation tensors $S^1(x, I)$ for the single template case were then computed at each voxel x.

In the log-Euclidean framework, the deformation tensors are projected to the tangent plane of the manifold on which they are defined, via their matrix logarithm. Tensor addition is performed in this space and the result is projected back to the original manifold. Thus, the n-templates average deformation tensor $S^n(x, I)$ at voxel x on image I is given by:

$$S^n(x, I) = \exp \frac{1}{n} \Sigma_{j=1}^{n} \log S_j^1(x, I) \tag{1}$$

2.3 Selecting the number of templates

Two measures were used to determine any improvements from increasing the number of templates. We describe both procedures in this section. Briefly, our first comparison method consists of comparing the total magnitude of deformation tensors before and after the addition of a new template. In effect, this amounts to determining how 'average' the effective template would have to be to generate the given set of deformation tensors. The second method we use begins with the assumption that the registration error diminishes with the number of templates used. Thus we wish to compare the distance between the deformation tensors obtained from a given number of templates to that which would have been found had we had access to an infinite number of templates. In practice, we will use values found with 9 templates as our gold standard.

In the log-Euclidean framework, the distance between two deformation tensors S_1 and S_2 is

$$d(S_1, S_2) = ||\log S_1 - \log S_2||,$$

where $||.||$ denotes a norm, and log is the matrix logarithm. Following [1], here we use

$$d(S_1, S_2) = (\text{Trace}(\log S_1 - \log S_2)^2)^{1/2}. \quad (2)$$

A measure of the size E_i^n of the tensors $S^n(., i)$ for n templates integrated over image i is given in this framework by [13]:

$$E^n(i) = \int ||\log S^n(x, i)||^2 d^3x = \int Tr(\log(S^n(x, i))^2) d^3x \quad (3)$$

For our first test, the $E^n(i)$ are computed for each image and a t-test is performed between the sets of $E^{n-1}(i)$'s and the $E^9(i)$'s to assess the significance of the results.

For the second test, we start with the following hypothesis

$$\log S^n(x, i) = \log S^\infty(x, i) + e^n(x, i) \quad (4)$$

where $e^n(., i)$ is the error in the logarithm of the deformation tensors in image i from using n templates, and $S^\infty(., i)$ is the hypothetical deformation tensor field that would be obtained from averaging with an infinite number of templates. Integrating over the image volume, we obtain

$$||e^n||_i = \int Tr(\log(S^n(x, i) - S^\infty(x, i))^2) d^3x \quad (5)$$

Here a t-test is again performed, in this case between the errors $||e^{n+1}||_i$ and $||e^n||_i$. In practice, as we do not have the value of S^∞, we compare all tensors to the ones found using all 9 templates.

2.4 Twin statistics

The determinants $detJ$ were computed at each voxel, to assess local tissue volume differences between individuals, after global brain scale differences across subjects were discounted using 9-parameter registration. $DetJ > 1$ and $detJ < 1$ respectively represent larger and smaller local volumes in the subject studied, with respect to the reference (template) image. We use the intraclass correlation (ICC) as a statistic to assess the influence of genes on these parameters. The ICC measures the correlation between unordered pairs and is defined as:

$$ICC = \sigma_b^2/(\sigma_b^2 + \sigma_w^2). \quad (6)$$

Here σ_b^2 is the pooled variance between pairs, while σ_w^2 is the variance within pairs. The voxelwise intraclass correlation is computed over the whole brain volume, and the significance assessed by comparing the ICC values to a permutation distribution [16], for which we randomly reassign subject labels (5000 permutations).

3 Results

In **Fig. 1a**, $\Sigma_{i=1}^{I} E^n(i)$ is plotted against n. The p-values in Table 1 give the significance of the error reduction from adding each successive template, i.e., as n increases. For both the MZs and DZs, the averaging converges for 4 templates, in the sense that no statistically significant energy reduction is detected. The registration error summed over all images, $\Sigma_{i=1}^{I} ||e^n||_i$, is shown in **Fig.1b**. P-values for MZs are below the 0.05 threshold for significance for the first 3 templates, while the DZs converge after the second averaging. In all cases, the size of the improvement decreases with an increasing number of templates.

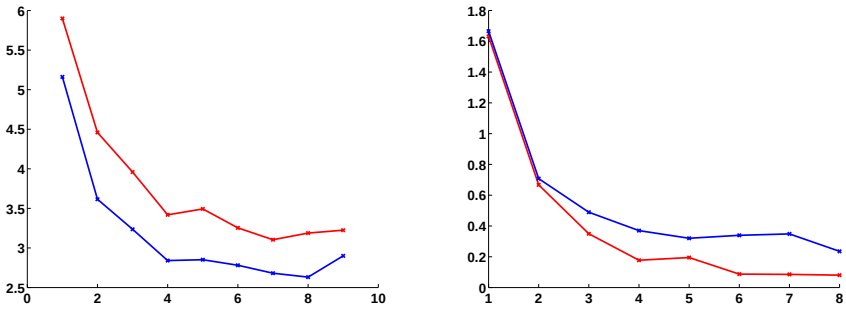

Fig. 1. Left: Size of the deformation tensors for the MZ (red line) and the DZ (blue line) groups. The x-axis shows the number of templates n included in the averaging, while the y-axis represents $\Sigma_{i=1:I} E^n(i)$, the size of the deformation tensors summed over all images in either the MZ or DZ dataset. Here I is the total number of subjects in each of the two groups. For each value of n, a Students t-test was performed between E_i^n and E_i^9 to assess the significance of adding one more image to the sample, when compared to the result of using all 9 templates. The p-values derived from these tests are shown in **Table 1**. **Right:** Magnitude of the registration error for the MZ (red line) and the DZ (blue line) groups. The x-axis is the number of templates n, and the error $\Sigma_{i=1:I} ||e^n||_i$ is plotted on the y-axis. A t-test was again performed for each successive value of n, this time between $||e^n||_i$ and $||e^{n+1}||_i$ to assess the significance of adding one more reference image. The p-values are shown in **Table 1**.

In order to verify our results in a TBM analysis, we computed the intra-class correlation for the MZs and the DZs for each number of templates. These analyses would be expected to show higher correlations in cases where less registration error was present, if other factors were equal. **Fig.2** shows the p-values that were found using the optimal number of templates, which we determined to be 4 from **Table 1**. The p-values indicate the statistical significance of the correlation ($p < 0.05$ are shown in red). As expected, the p-values are generally lower in the MZ twins (denoting higher intraclass correlations), as MZ twins share all of their genes, and regional brain volumes are known to be under genetic

no of templates	2	3	4	5	6	7	8	9
$E^n(i)$ p-values for the MZs	$>$ **0.0001**	**0.053**	**0.014**	0.72	0.23	0.42	0.64	0.85
$E^n(i)$ p-values for the DZs	$>$ **0.0001**	**0.0065**	**0.0032**	0.92	0.60	0.49	0.71	0.26
$\|\|e^n\|\|_i$ p-values for the MZs	$>$ **0.0001**	**0.0027**	$>$ **0.0001**	0.77	0.082	0.92	0.91	N/A
$\|\|e^n\|\|_i$ p-values for the DZs	**0.0016**	0.46	0.69	0.87	0.94	0.98	0.71	N/A

Table 1. The first two rows show the p-values from perfoming a t-test comparing the E_i^n's to the $E_i^n + 1$ for all n's. P-values from a t-test comparing the error $||e^n||_i$ to $||e^n||_i$.

control. This is indeed the case.We also computed the cumulative distribution functions of the p-values for maps derived using 1,4 and 9 templates, for each of the two groups. For a null distribution, the cumulative distribution function is expected to fall along the $x = y$ line. Larger upward deviations from the diagonal are generally indicative of greater statistical power and larger effect sizes [14]. However, none ofthe distributions differed significantly from each other, indicating thus indicating that statistical powerwas not detectably different for the 3 maps using progressively higher number of templates. Despite this, any detected differences may have greater validity as the analysis removes the potential bias resulting from selecting an individual brain as a registration target, which may affect the accurate quantification of regional brain volumes.

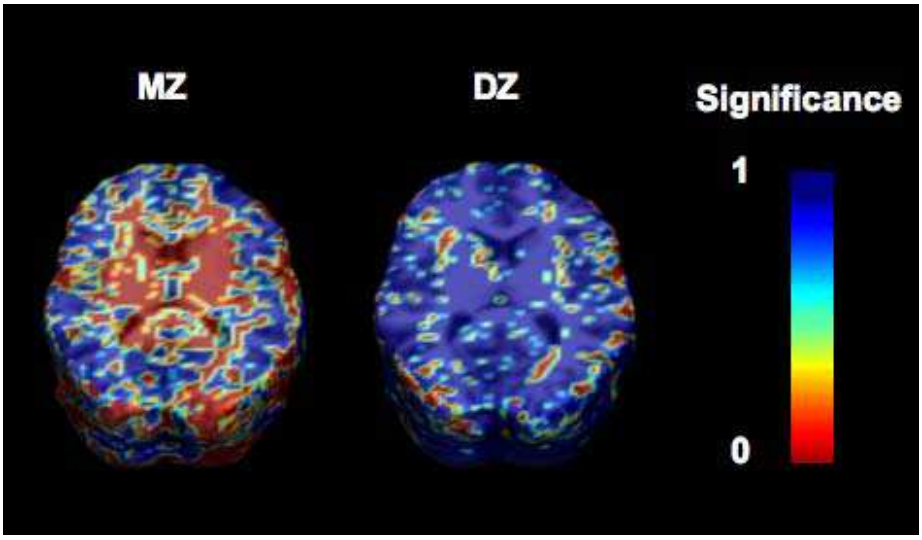

Fig. 2. Maps of p-values for the intraclass correlation, shown here in logarithmic scales. **Left:** MZ twins. **Right**: DZ twins. Red p-values indicate ICCs significant at a level of $p = 0.05$.

4 Discussion

Here we presented a new multi-atlas version of Tensor-Based Morphometry. A log-Euclidean averaging procedure was performed on the deformation tensors generated from the multiple registrations, resulting in the reduction of registration error and improved statistical power. At least for the case presented here, the error was significantly reduced up to an averaging of 4 templates. Future studies will assess how much multi-template TBM boosts power relative to other influential factors, including the sample sizes used, scan quality or field strength, the regularization model and data fidelity term [5, 3], and the tensor statistics used [14].

References

1. Arsigny V et al., *Log-Euclidean metrics for fast and simple calculus on diffusion tensors*, Mag. Res. in Med. 56, (2006) 411–421.
2. Ashburner J, *A fast diffeomorphic image registration algorithm*, Neuroimage 38, (2007) 95–113.
3. Brun C et al., *Comparison of standard and Riemannian elasticity for Tensor-Based Morphometry in HIV/AIDS*, MICCAI workshop on Statistical Registration: Pairwise and Group-wise Alignment and Atlas Formation (2007).
4. Chou YY et al., *Automated Ventricular Mapping with Multi-Atlas Fluid Image Alignment Reveals Genetic Effects in Alzheimer's Disease*, accepted for publication in Neuroimage (2007).
5. Chiang MC et al., *3D pattern of brain atrophy in HIV/AIDS visualized using tensor-based morphometry*, Neuroimage 34, (2007) 44–60.
6. Christensen GE et al., *Deformable templates using large deformation kinematics*, IEEE-Trans. Imag. Process. 5, (1996) 1435–1447.
7. Guimond et al. *Average brain models: a convergence study*, Comp. Vis. and Im. Understanding 77, (1999) 192–210.
8. Kochunov P et al., *Regional spatial normalization: toward an optimal target*, J. Comp. Assist. Tomogr. 25, (2001) 805–816.
9. Kochunov P et al. *An optimized individual target brain in the Talairach coordinate system*, Neuroimage 17, (2003) 922–927.
10. Kochunov P et al., *Mapping structural differences of the corpus callosum in individuals with 18q deletions using targetless regional spatial normalization*, Hum. Brain Map. 24, (2005) 325–331.
11. Jenkinson M and Smith S, *A global optimisation method for robust affine registration of brain images*, Med. Imag. Anal. 5, (2001) 143–56.
12. Leow AD et al., *Statistical properties of Jacobian maps and inverse-consistent deformations in non- linear image registration*, IEEE-Trans. Med. Imag. 26, (2007) 822-832
13. Leporé N et al., *Mean template for Tensor-Based Morphometry using deformation tensors*, MICCAI (2007) 826–833.
14. Leporé N et al., *Multivariate statistics of the Jacobian matrices in Tensor-Based Morphometry and their application to HIV/AIDS*, IEEE-Trans. Med. Imag., in press (2008).
15. Leporé N et al., *Best individual template selection from deformation tensor minimization*, ISBI (2008).

16. Nichols TE, Holmes AP, *Non parametric permutation tests for functional neuroimaging: a primer with examples*, Hum. Brain Mapp. 15, (2001) 1–25.
17. Pennec X et al., *Riemannian elasticity: A statistical regularization framework for non-linear registration*, MICCAI, (2005) 943–950.
18. Shattuck DW and Leahy RM, *BrainSuite: an automated cortical surface identification tool*, Med. Imag. Anal. 8, (2002) 129–141.
19. Studholme C et al., *Deformation tensor morphometry of semantic dementia with quantitative validation*, Neuroimage 21, (2004) 1387-1398
20. Studholme C, Cardenas V, *A template free approach to volumetric spatial normalization of brain anatomy*, Patt. Recogn. Lett. 25, (2004) 1191–1202.
21. Twining CJ, *A unified information-theoretic approach to groupwise non-rigid registration and model building* MICCAI, (2005) 190–193.
22. Thompson PM et al., *Mathematical/Computational Challenges in Creating Population-Based Brain Atlases*, Hum. Brain Mapp. 9, (2000) 81–92.