# Groupwise Morphometric Analysis based on Morphological Appearance Manifold

Nai-Xiang Lian* Christos Davatzikos
Department of Radiology,
University of Pennsylvania, Philadelphia, PA
liannaixiang@gmail.com, christos@rad.upenn.edu

## Abstract

*The field of computational anatomy has developed rigorous frameworks for analyzing anatomical shape, based on diffeomorphic transformations of a template. However, differences in algorithms used for template warping, in regularization parameters, and in the template itself, lead to different representations of the same anatomy. Variations of these parameters are considered as confounding factors as they give rise to non-unique representation. Recently, extensions of the conventional computational anatomy framework to account for such confounding variations have shown that learning the equivalence class derived from the multitude of representations can lead to improved and more stable morphological descriptors. Herein, we follow that approach, estimating the morphological appearance manifold obtained by varying parameters of the template warping procedure. Our approach parallels work in the computer vision field, in which variations lighting, pose and other parameters leads to image appearance manifolds representing the exact same figure in different ways.*

*The proposed framework is then used for groupwise registration and statistical analysis of biomedical images, by employing a minimum variance criterion on selected complete morphological descriptor to perform manifold-constrained optimization, i.e. to traverse each individual's morphological appearance manifold until group variance is minimal. Effectively, this process removes the aforementioned confounding effects and potentially leads to morphological representations reflecting purely biological variations, instead of variations introduced by modeling assumptions and parameter settings. The nonlinearity of a morphological appearance manifold is treated via local approximations of the manifold via PCA.*

*This work was supported by the NIH-funded Grant R01-CA104976.

## 1. Introduction

Computational anatomy provides a powerful tool for characterizing differences between normal and pathologic anatomies by analyzing their variations relative to a common template. Diffeomorphic shape transformations [1, 2, 3] are first estimated to warp all anatomies to a template or vice versa; various descriptors are then derived to quantify their morphological characteristics.

Template transformations are often derived from image similarity measures, in the intensity-driven methods [1, 4], either by employing intensity differences or via mutual information [5, 3]. Topology is maintained by imposing smoothness constraints either via physical models [1, 6] or directly on the deformation field [7]. Other approaches, such as [5, 3], ensure biological correspondence through feature-based approaches by introducing biologically, anatomically and geometrically significance attributes in shape morphological representations.

Various approaches have been presented in the literature, under the umbrella of computational anatomy: *large deformation diffeomorphic metric mapping* (LDDMM) [8], *deformation based morphometry* (DBM) [9, 10], *voxel based morphometry* (VBM) [11, 12], *tensor based morphometry* (TBM) [13], depending on the aspects of the template transformation being measured. DBM, for instance, establishes group differences based on local deformation of anatomical structures through the Jacobian of the diffeomorphism. VBM, on the other hand, factors out global differences via diffeomorphism, before analyzing anatomical differences. VBM is, therefore, considered as complementary to DBM, since the former utilizes the information not represented by the transformation.

However, the inherent complexity of the problem poses a major challenge to these approaches. First, anatomical correspondence may not be uniquely determined from intensity-based image attributes, which drive template warping algorithms. Second, exact anatomical correspondence may not exist at all due to anatomical variability

across subjects. As a result, the choice of the template plays an important role in the accuracy of analysis. The aforementioned challenges lead to residual information that the transformation fails to capture.

To remedy this problem, some approaches have been proposed to use average anatomies as templates [14], to facilitate the template matching procedure. In most practical cases, considerable differences still persist between samples and the average brain. A very promising approach in this situation is groupwise registration [15, 16], which solves the problem to a certain extent, in the sense that instead of minimizing individual dissimilarity it minimizes combined cost. Bhatia et al. [15], for instance, implicitly find the common coordinate system by constraining the sum of all deformations from itself to each subject to be zero. Davis et al. [14] compute the most representative template image through a combined cost functional on the group of diffeomorphisms. Such groupwise registration based representations is, therefore, more consistent across the samples, even though they might still fail to eliminate residuals.

In the approach presented herein, we follow the work of [17], which uses a complete morphological descriptor of the form [Transformation, Residual]; any morphological information not captured by the transformation is captured by the residual, hence no morphological characteristic is discarded. An entire class of many anatomically equivalent descriptors is generated by varying parameters of the template transformation, as well as the template itself, all representing the same anatomy. The resultant anatomical equivalence class (AEC) maps the underlying anatomy to a manifold embedded in a high dimensional space, which we call a *morphological appearance manifold* (MAM). Although one limitation of such representation is that such manifolds can be nowhere differentiable [18], they can become differentiable by smoothing the images. Instead of directly smoothing the images prior to estimating a MAM, herein we follow a different approach: we estimate its local structure by fitting a hyperplane that approximates its tangent plane, as described below. In a group-wise registration framework, among the members of an AEC of each individual, one is selected according to the criterion that the variance across individuals is minimized. This problem is solved via a manifold-constrained optimization approach, which locally estimates the structure of the MAM via local PCA, and moves along the manifold to minimize the variance across individuals. Intermediate steps of projection onto the manifold are necessary to guarantee that each individual's representation remains on its respective manifold. Standard voxel-based analysis methods are then used to compare individuals and groups, however they are applied to the optimal morphological signatures (OMS) obtained through this optimization. Unlike the work in [17], which used a globally linear approximation, by using a local approximation of the MAM we follow the nonlinearities of the underlying AEC.

The other contribution of this paper is to investigate features used for optimization and find a good one. In [17], minimizing the distance between manifolds was used as optimization criterion, and two complete morphological descriptors, combined feature [Jacobian determinant of Transformation, Residual] and tissue density maps [12] (TDM), were used to estimate group difference and longitudinal atrophy, respectively. However, the paper did not tell the difference between using combined feature and TDM, and also did not discuss the weights of two components when computing distance metrics. Because the minimum distance of Jacobian is achievable for infinite regularization while minimum distance of residuals is achievable via the most aggressive registration, different weight of two components will yield a different optimal solution. In TDM, the weight of the two components is fixed. Our experimental results show that the distance of TDM is highly correlated with distance of residuals, implying that the weight of residuals in TDM is much larger than that of the Jacobian. Minimizing the distance of TDM is therefore close to minimizing the distance of residuals, and not necessarily on optimal distance. See more detail discussion in Section 3.

The approach described herein is akin to methods on image appearance manifolds that has been used in computer vision [19] to model variations in images that are caused by measurement parameters, such as illumination and pose. Such variations are confounding factors when, for example, one is interested in face recognition, as the same person appears different for different parameters. As in our work, learning such variations is important for determining robust parameters to be used for analysis and recognition.

## 2. Anatomical Equivalence Class Manifold

### 2.1. Anatomical Equivalence Class

Computational anatomy involves characterizing anatomical differences between a subject $S$ and a template $T$ by mapping the template space $\Omega_T$ to the subject space $\Omega_S$ through a diffeomorphism $h : \Omega_T \rightarrow \Omega_S, \mathbf{x} \mapsto h(\mathbf{x})$ by maximizing some similarity criterion between $T$ and normalized subject $S_T$. A zero residual mapping is usually very difficult if possible at all, resulting in a residual:

$$R_h(\mathbf{x}) := T(\mathbf{x}) - S(h(\mathbf{x})), \mathbf{x} \in \Omega_T \qquad (1)$$

Our approach herein is to use a template warping algorithm that captures some of the morphological differences between the template and the target anatomies, in combination with the respective residual, which captures everything else. There exist two ways to combine the two types of information. In [17], the transformation and the residual were

combined into a concatenated descriptor:

$$\mathcal{Q}_h(\mathbf{x}) := (J_h, R_h), \tag{2}$$

where $J_h$ is the Jacobian determinant of the transformation $h$. However, it is not intuitive how the transformation and the residual should be weighted relative to each other, when calculating distances. In [20], a slightly different way of combining $h(.)$ and $R_h(.)$ was used, for quantifying patterns of local variations of sizes of biological tissues, as, for example, brain atrophy. Such local volumetric measurements can be quantified by TDM [11, 12], defined as follows:

$$\mathcal{Q}_h(\mathbf{x}) := J_h(\mathbf{x})[T(\mathbf{x}) - R_h(\mathbf{x})], \mathbf{x} \in \Omega_T, \tag{3}$$

For brevity, we refer to the morphological representation obtained via $\mathcal{Q}_h(.)$ as a complete morphological descriptor (CMD); the term "complete" stems from the fact that the residual completes the representation obtained via the transformation $h(.)$ so that no morphological information is discarded.

$\mathcal{Q}_h$, depends not only on the underlying anatomy but also on transformation parameters. An entire family of anatomically equivalent CMDs - they are equivalent to reconstruct the same anatomy, and form an (anatomical) equivalence class (AEC) - may be generated by varying $h$. In practice, it is impossible to sample all possible transformations and residuals to characterize an AEC, therefore we concentrate on two important parameters for variations of $h$: $h_{\lambda,\tau}$, where $\lambda$ denotes the amount of regularization and $\tau$ denotes an intermediate template. Since analysis eventually has to be carried out in a common space, we ultimately bring all anatomies to the template space $\Omega_T$, however intermediate templates capture the variation we expect to see when the same anatomy is seen "via different templates". Variations of the regularization parameter $\lambda$ reflect variations observed by varying the degree of conformality of the transformation, with large values of $\lambda$ indicating very smooth transformations and large residuals, and $\lambda = 0$ indicating the most conforming transformation of the template under the assumptions of a respective deformable registration algorithm that seeks to minimize the residual. For tractability and notational simplicity, we combine confounding factors together to represent $\boldsymbol{\theta} := (\tau, \lambda)$, and the corresponding definition of AEC becomes

$$\mathcal{A}(S) = \{\{(\mathcal{Q}_{h_{\boldsymbol{\theta}}}(\mathbf{x})) : S(h_{\boldsymbol{\theta}}(\mathbf{x})) = T(\mathbf{x}) - R_{h_{\boldsymbol{\theta}}}(\mathbf{x}), \forall \mathbf{x} \in \Omega_T\}, \forall \boldsymbol{\theta} \in \boldsymbol{\Theta}\} \tag{4}$$

All AECs construct a manifold, that we called morphological appearance manifold (MAM).

## 2.2. Nonlinear Manifold Representation

Practical considerations limit us to sampling the MAM at a discrete number of locations. Figure 1 shows an example of the manifolds obtained using three intermediate templates (shown in Figure 2): each curve represents one intermediate template and multiple $\lambda$'s. Since these MAMs are embedded in a very high-dimensional space, for visualization purposes we calculated the first three principal components and projected onto them, so that 3D plots can be shown.

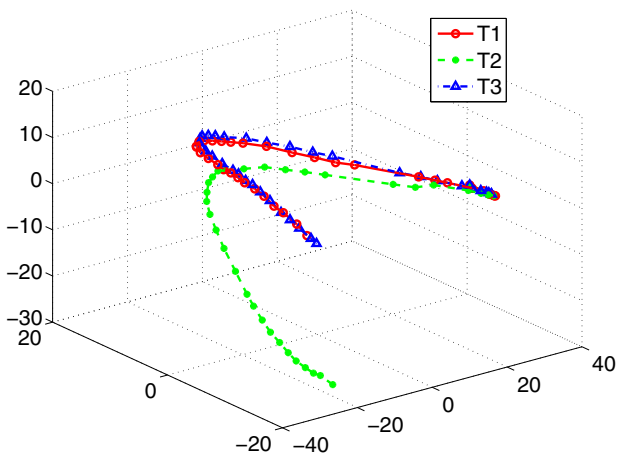



Figure 1. Manifold for one subjects using three intermediate templates.

(a) (b) (c)

Figure 2. Templates (a) T1, (b) T2 and (c) T3.

In [17], the AEC manifold $\mathcal{Q}_{h_{\boldsymbol{\theta}_i}}$ was approximated via its principle directions, computed via principle component analysis (PCA).

$$\mathcal{Q}_{h_{\boldsymbol{\theta}_i}} = \hat{\mathcal{Q}}_{h_{\boldsymbol{\theta}_i}} + \sum_{j=1}^{n} \alpha_{ij} V_{ij} \tag{5}$$

where $\hat{\mathcal{Q}}_{h_{\boldsymbol{\theta}_i}}$ is the population mean, and $\alpha_{ij} \in [\alpha_{ij}^{min} \alpha_{ij}^{max}]$ is the variation in the principle direction $V_{ij}$, which is originally represented by $\boldsymbol{\theta}_i$. However, the AEC manifold is nonlinear (Figure 1) and PCA-based manifold approximation may include some samples that are not on the manifold, resulting in the optimal morphological signature (OMS) that may not belong to the manifold. To achieve a better representation, we assume linearity in a local neighborhood. The same assumption is used for the popular nonlinear approximation methods ISOMAP [21] and LLE [22]. Under this assumption, we apply PCA to represent the local neigh-

borhood of the AEC manifold:

$$\mathcal{Q}_{h_{\boldsymbol{\theta}_i}}^{(\delta\boldsymbol{\theta}_i^{center})}(\boldsymbol{\theta}_i) = \hat{\mathcal{Q}}_{h_{\boldsymbol{\theta}_i}}^{(\delta\boldsymbol{\theta}_i^{center})} + \sum_{j=1}^{n} \alpha_{ij} V_{ij}^{(\delta\boldsymbol{\theta}_i^{center})} \quad (6)$$

where $\delta\boldsymbol{\theta}_i^{center}$ is the local neighborhood space around the center sample at $\boldsymbol{\theta}_i^{center}$. This optimization problem is solved iteratively. During each iteration, the local linear approximation of the MAM is recomputed, centered on the current optimal point $\mathcal{Q}_{h_{\boldsymbol{\theta}_i^*}}$ on the manifold:

$$\boldsymbol{\theta}_i^{center} = \boldsymbol{\theta}_i^* \quad (7)$$

We refer to this optimization process as manifold-constrained optimization.

However, setting the current point as center for local MAM approximation could make the optimization vulnerable to local minima nearest to initial point. Then the optimization performance will depend on selection of initial point. Increasing the neighborhood size used may solve the problem, but also lead to poor global linear approximation of a highly nonlinear manifold. To make the optimization more robust of selection of initial point without increasing neighborhood size, we set the center sample of each individual anatomy as the one that has largest possibility to be next optimal point. We find the point by optimizing the $l$-norm distance criterion (9),

$$\boldsymbol{\theta}_i^{center} = \arg\min_k \sum_j \left\| \mathcal{Q}_{h_{\boldsymbol{\theta}_{ik}}}, \mathcal{Q}_{h_{\boldsymbol{\theta}_j^*}} \right\|_l^l \quad (8)$$

where $\mathcal{Q}_{h_{\boldsymbol{\theta}_{ik}}}$ is the $k$th available samples on the manifold.

Finding optimal morphologic signatures (OMS) on these manifolds needs a suitable CMD feature for optimization among MAMs, which is the topic in the following section.

## 3. Optimization Criterion and Feature Selection

Remember that varying $\boldsymbol{\Theta} = [\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_N]$ effectively allows every individual representation to slide along its own manifold, thereby leading to multiple possible ways of representing each individual. The goal of the optimization procedure presented in this section is to find a single optimal solution for each individual, i.e. to find a unique morphological representation. The optimality criterion is minimal group variance: each individual slides along its own manifold until the group variance is minimized, which is assumed to occur when the confounding effects of parameters and templates are removed, thereby allowing us to examine biological differences instead of modeling differences; see Figure 3. Specifically,

$$\boldsymbol{\Theta}^* = \arg\min_{\boldsymbol{\Theta}} \sum_{i,j} \left\| \mathcal{Q}_{h_{\boldsymbol{\theta}_i}}, \mathcal{Q}_{h_{\boldsymbol{\theta}_j}} \right\|_l^l \quad (9)$$

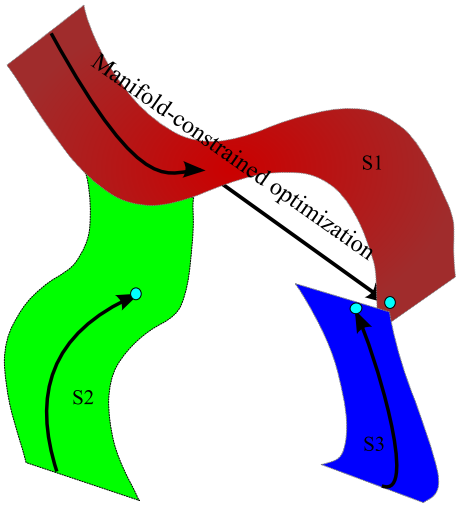



Figure 3. Manifold-constrained optimization: the AEC of each individual resembles a manifold embedded in $R^N$, where $N$ is the dimensionality of the measurement space. An optimal member of each person's AEC is found by locally approximating the structure of this manifold with PCA, and iteratively traversing the manifold until a certain criterion of minimum variance is met. This procedure removes variations that are introduced by confounding variables during the calculation of a template transformation.

where $\|\cdot\|_l$ is the $l$-norm distance criterion, and $\mathcal{Q}_{h_{\boldsymbol{\theta}_i}}$ is CMD, either of combined feature (2) or TDM (3), to trade-off between the Jacobian and the residual.

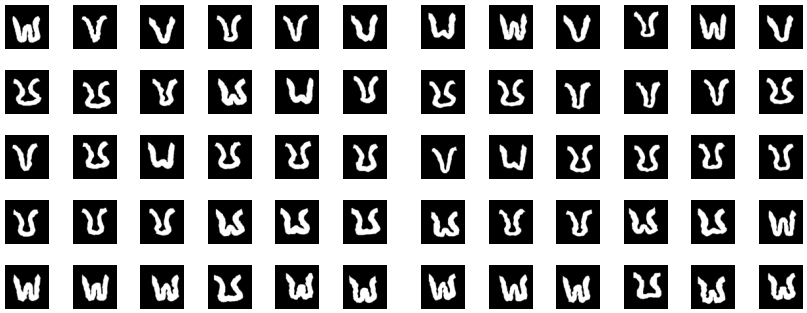
However, TDM is not necessarily a best feature for the group difference estimation. Such a TDM is determined from pre-defined multiplicative relationship between the Jacobian and the residual. Minimizing the distance of TDM may not optimally leverage upon the trade-off between the Jacobian and the residual, and might fail to find a solution that best reflects shape differences.

In general, we seek to find an optimal value for $\mu$ in

$$\mathcal{Q}_{h_{\boldsymbol{\theta}_i}|\mu} = (R_{h_{\boldsymbol{\theta}_i}}, \mu J_{h_{\boldsymbol{\theta}_i}}) \quad (10)$$

and then apply it in the criterion (9) to find a solution best reflecting shape deference. Setting $\mu = 0$ is equivalent to only minimizing distance of residual and leads to aggressive registration. Setting $\mu = \infty$ is equivalent to only minimizing distance of Jacobian and leads to rigid transformations. The best value of $\mu$ tends to be between $0 \sim \infty$.

We performed an experiment to investigate the relationship between the value of $\mu$ and the ability to detect group shape difference. A group of simulated 2D subjects were used; see Figure 4. These subjects simulate different biological shapes of gray matter of brain, and any shape in right column is obtained by shrinking the middle third of each ribbons of left column. We used two groups of subjects: one is centered - most shapes are roughly aligned



(a) (b)

Figure 4. Simulated 2D subjects (a) without and (b) with thinning in the middle third of each ribbon

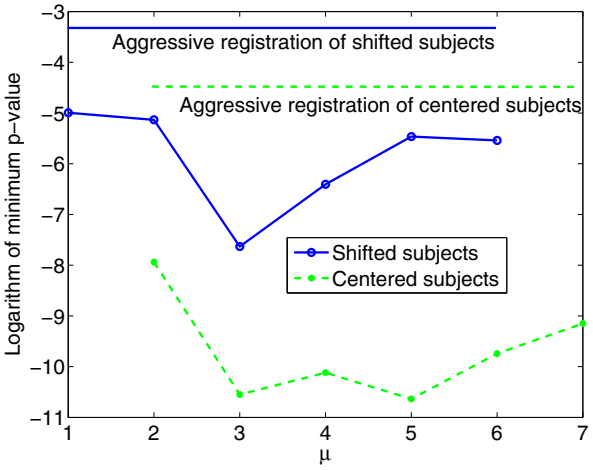



Figure 5. The minimum p-value *vs.* the value of $\mu$ for simulated 2D subjects

around center location and their group difference is mainly due to the shapes; the other is shifted - most shapes are misaligned by applying a random shift. These subjects were registered to 12 templates, and 37 values of $\lambda$ were used. We apply Hotelling $T^2$ test on smoothed (Gaussian filter with standard deviation $\sigma = 7$) optimized samples, and take logarithm of minimum p-value as an indicator of significant group difference determined for a given $\mu$.

Figure 5 shows the significance levels *vs.* the value of $\mu$. Because the group difference in centered subjects is mainly due to the shape thinning in their middle portion, a large value, $\mu = 5$, yields best significance level. On the other hand, shifting the shapes will increase the importance of alignment and then require a smaller value, $\mu = 3$. However, the significance levels are not very different for $\mu \in \{2 \sim 6\}$. For example, the maximum of logarithm of minimum p-value for centered subjects is -8, still a good one. Therefore, we decide to choose one value for all experiments: $\mu = 3$.

In [17], combined features were normalized via conversion to z-scores. Because the Jacobian has smaller variance than the Residual, this normalization puts more weight on the Jacobian. Experimental investigation shows that normalization on centered shapes roughly corresponds to $\mu = 8$ and normalization on shifted shapes roughly corresponds to $\mu = 6$. Hence, an optimal $\mu$ value obtained better p-values for the same shapes.

## 4. Evaluation

To evaluate our optimization method, we apply it on two datasets. The first is a group of simulated 2D subjects, which are used in Section 3. The second is real 3D MR brain images. We apply Hotelling $T^2$ test on the smoothed unoptimized/optimized combined features for p-maps.

### 4.1. Simulated 2D shapes

First we compare the significance of p-maps based on various values of the smoothing parameter - $\sigma$, the standard deviation of the Gaussian filter used to smooth the Residual prior to the statistical test. We plot significance levels - minimum logarithm of p-maps as the function of $\sigma$ in Figure 6. The traditional method - using the Jacobians - is one of the worst of all methods in terms of detecting the group difference. So is the combined features with small regularization, because it resembles the Jacobian. Larger regularization can bring much better results for centered subjects, because it can better reflect shape difference. On the other hand, larger regularization did not yield the better results on shifted subjects due to its mis-alignment of subjects. Our optimization finds an optimal regularization, which was found to improve the estimated significance for both groups regardless of the smoothness levels - $\sigma$. The improvement of significance could be from $10^{-9.2}$ to $10^{-10.9}$ (when $\sigma = 7$) for centered subjects, and from $10^{-4.1}$ to $10^{-8.3}$ (when $\sigma = 5$) for shifted subjects.

Figure 6 also compares global PCA and local PCA manifold approximation. Local PCA can better follow the non-

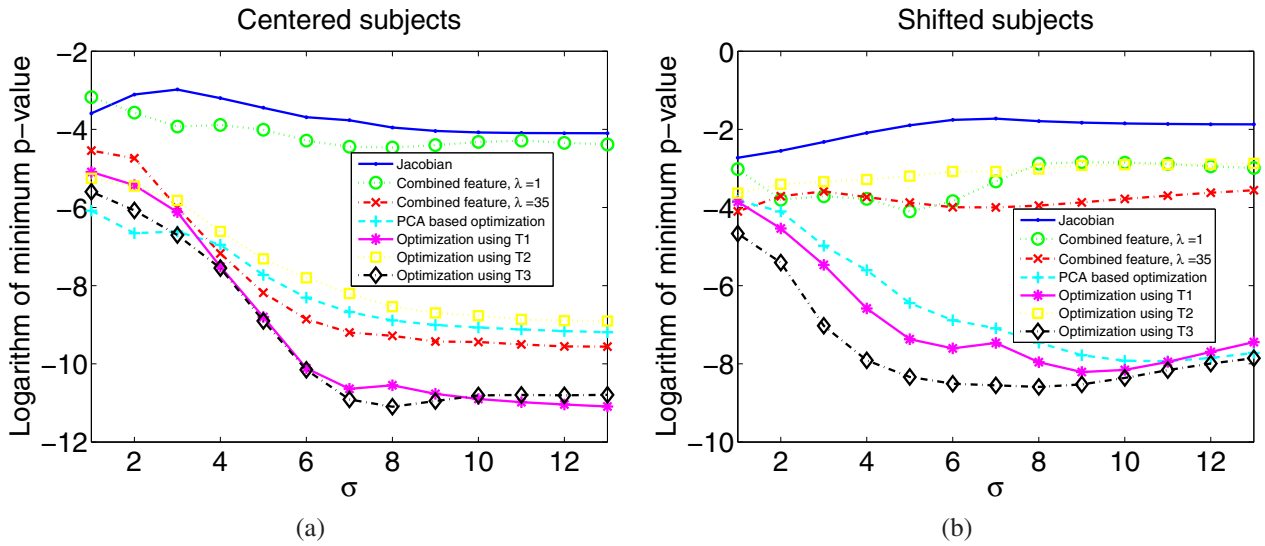



Figure 6. Compare curves of minimum p-values of unoptimized and optimized combined features on (a) centered subjects and (b) shifted subjects. if approximation or global template is not specified, local PCA and T1 are used.

linearity of MAM, and then help reach a better minima in optimization. For centered subjects, the resultant improvement is from $10^{-9.0}$ to $10^{-10.8}$ (when $\sigma = 9$), and for shifted subjects, from $10^{-7.5}$ to $10^{-8.5}$ (when $\sigma = 7$).

We used three different global templates shown in Figure 2. Figure 6 shows that using a more convoluted global template (T3) can achieve better estimation. Our optimization is more robust when using more-convoluted global templates, T1 and T3. In general, these results show that selection of the final template can be important.

Results of p-maps from group comparisons are shown in Figures 7 and 8. Using the traditional Jacobian cannot find the group difference region. For centered subjects, the region can be detected by using combined features with large regularization, because they can better reflect the shape difference. But for shifted subjects, the detection is worse due to misalignment for large $\lambda$. Our optimization can further improve the detection by obtaining more significance p-value, especially when using local PCA for a better nonlinear approximation of MAM.

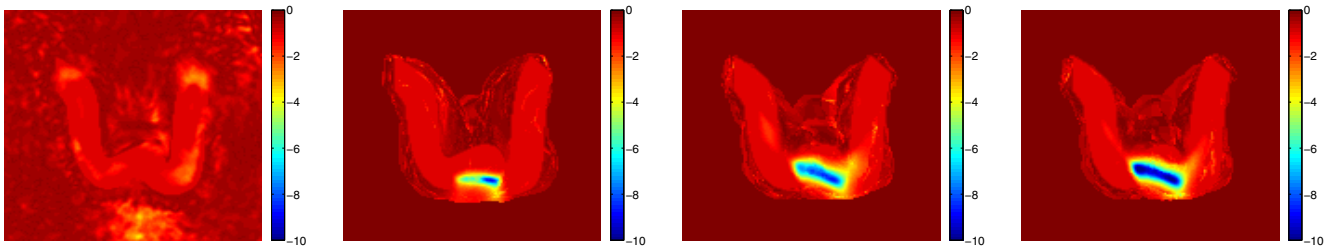
### 4.2. Real 3D data

We used MR brain images from 30 different persons, 15 of them were diagnosed as mild cognitive impairment (MCI) and the other 15 were normal elderly. Six templates and nine regularizations were used to sample the MAMs. For optimization, we used $\mu = 3$, determined from the simulated images as discussed in Section 3.

We did not use minimum p-value, that may come from noise, but directly computed the p-map of smoothed unoptimized/optimized combined features to evaluate the estimated group differences. Gaussian filters ($\sigma = 5$) are used to smooth the Jacobian and Residual prior to the statistical test. The Jacobian and the TDMs give patches of significance. The combined features tend to outline the hippocampus, which is actually the area in which we anticipated to find anatomical differences in these early MCI patients. Moreover, the combined features using local PCA tend to give a result that seems to emphasize the hippocampus most. We can only evaluate these results visually, as there is no ground truth for these datasets.

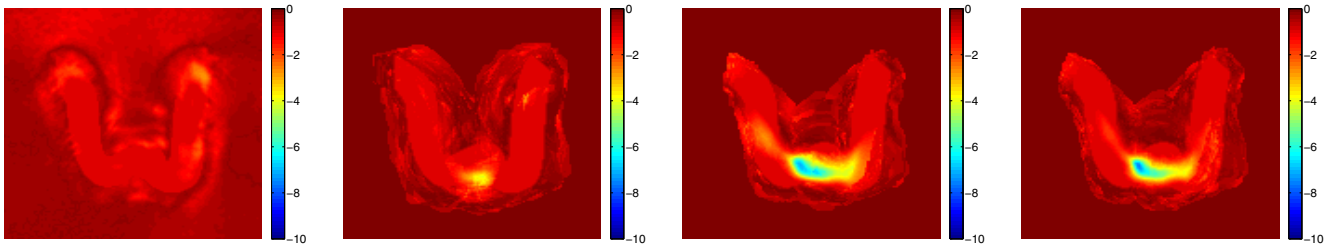
## 5. Conclusions

In this paper, we proposed a framework for morphological analysis of medical images, and for group-wise registration, by building upon the work in [17]. The transformation from the template to an individual anatomy was combined with residual for a complete representation of anatomy, and each anatomy was represented through an equivalence class of descriptors, by varying transformation regularization and intermediate templates and constructing a morphological appearance manifold. We used local linear approximations of the manifold to follow the nonlinearity of a MAM. We employed a minimum variance criterion on selected complete morphological descriptor and performed manifold-constrained optimization, i.e. traversed each individual's MAM, to find the points on all MAMs that brought the group of morphological descriptors closest to each other. The proposed method can reduce the confounding variation in each MAM from noise, parameter and template selection, and preserve the true variations that relate to true underlying morphological change. This is confirmed in our experimental results, where our approach improves the performance




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

Figure 7. The logarithm of P-maps of simulated shifted subjects using (a) Jacobian , (b) Combined features with $\lambda = 35$, and optimized combined features based (c) global and (d) local PCA approximation and L2-norm distance criterion, and T1 global template.




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

Figure 8. The logarithm of P-maps of simulated shifted subjects using (a) Jacobian , (b) Combined features with $\lambda = 35$, and optimized combined features based (c) global and (d) local PCA approximation and L2-norm distance criterion, and T1 global template.

of estimating group difference for simulated and real volumetric datasets.

## References


[1] M. Miller, G. Christensen, Y. Amit, and U. Grenander, "Mathematical textbook of deformable neuroanatomies," *Proc. the National Acad. of Sciences* **90**(24), pp. 11944–11948, 1993.

[2] M. Miller, A. Banerjee, G. Christensen, S. Joshi, and et al., "Statistical methods in computational anatomy," *Statis. Methods in Med. Research* **6**, pp. 267–299, 1997.

[3] D. Shen and C. Davatzikos, "HAMMER: Hierarchical attribute matching mechanism for elastic registration," *IEEE Trans. Med. Imag.* **21**(11), pp. 1421–1439, 2002.

[4] J. Ashburner and K. Friston, "Nonlinear spatial normalization using basis functions," *Hum. Brain Mapp.* **7**(4), pp. 254–266, 1999.

[5] P. Thompson and A. Toga, "A surface-based technique for warping three-dimensional images of the brain," *IEEE Trans. Med. Imag.* **15**, pp. 402–417, 1996.

[6] H. Johnson and G. Christensen, "Consistent landmark and intensity-based image registration," *IEEE Trans. Med. Imag.* **21**(5), pp. 450–461, 2002.

[7] B. Karacali and C. Davatzikos, "Estimating topology preserving and smooth displacement fields," *IEEE Trans. Med. Imag.* **23**(7), 2004.

[8] M. Miller, A. Trouvé, and L. Younes, "On the metrics and euler-lagrange equations of computational anatomy," *Annual Review of Biomed. Engin.* **4**, pp. 375–405, 2002.

[9] C. Davatzikos, M. Vaillant, S. Resnick, J. Prince, S. Letovsky, and R. Bryan, "A computerized approach for morphological analysis of the corpus callosum," *J. Comp. Assis. Tomogr.* **20**(1), pp. 88–97, 1996.

[10] J. Ashburner, C. Hutton, R. Frackowiak, I. Johnsrude, C. Price, and K. Friston, "Identifying global anatomical differences: deformation-based morphometry," *Hum. Brain Mapp.* **6**(6), pp. 348–357, 1998.

[11] J. Ashburner and K. J. Friston, "Voxel-based morphometry – the methods," *NeuroImage* **11**(6), pp. 805–821, 2000.

[12] C. Davatzikos, A. Genc, D. Xu, and S. Resnick, "Voxel-based morphometry using RAVENS maps: methods and validation using simulated longitudinal atrophy," *Neuroimage* **14**, pp. 1361–1369, 2001.

[13] P. Thompson, J. Giedd, R. Woods, D. MacDonald, A. Evans, and A. Toga, "Growth patterns in the

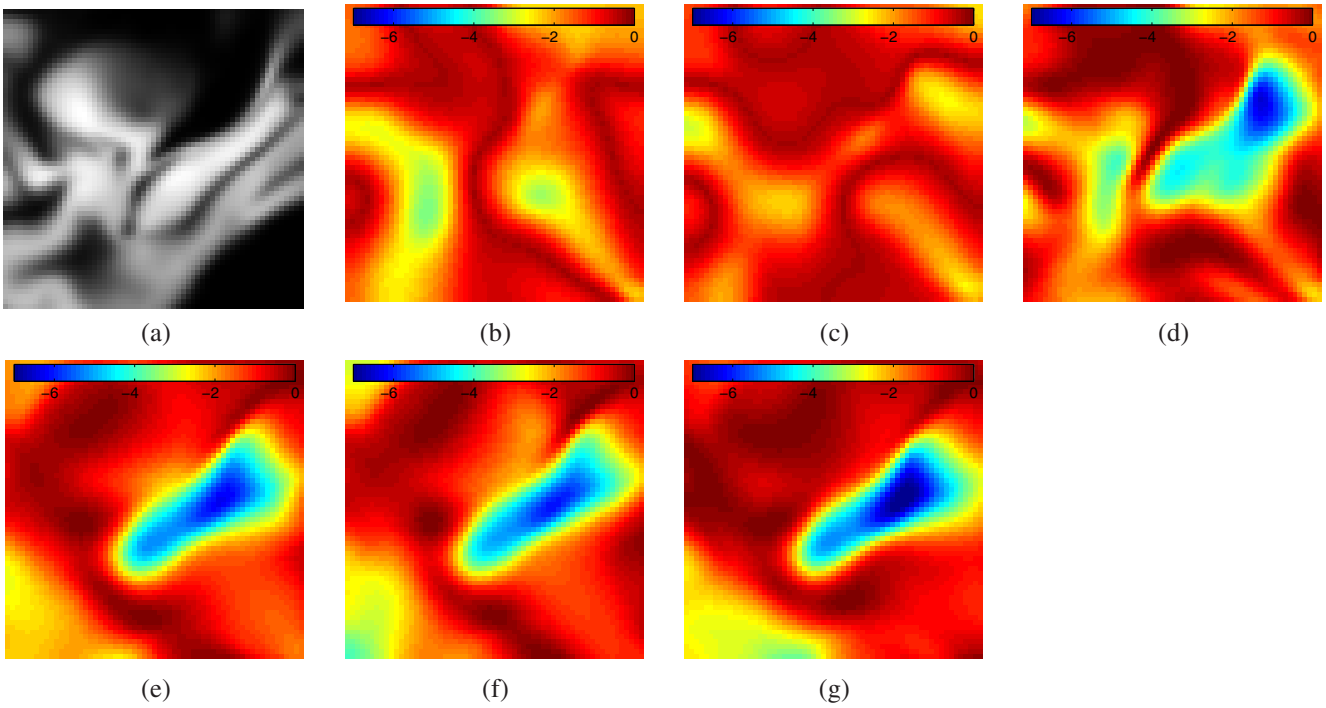



Figure 9. (a) Averaged TDM and corresponding logarithm of P-maps of 3D MR brain images using (b) Jacobian , (c) TDM with $\lambda = 1$, Combined features with (d) $\lambda = 1$, (e) $\lambda = 6$, and optimized combined features with (f) global and (g) local PCA approximation.

developing human brain detected using continuum-mechanical tensor mapping," *Nature* **404**(6774), pp. 190–193, 2000.

[14] B. Davis, P. Lorenzen, and S. Joshi, "Large deformation minimum mean squared error template estimation for computational anatomy," in *ISBI'04*, pp. 173–176, 2004.

[15] K. Bhatia, J. Hajnal, B. Puri, A. Edwards, and D. Rueckert, "Consistent groupwise non-rigid registration for atlas construction," in *ISBI'04*, pp. 908–911, 2004.

[16] C. Twining, T. Cootes, S. Marsland, V. Petrovic, R. Schestowitz, and C. Taylor, "A unified information-theoretic approach to groupwise non-rigid registration and model building," in *IPMI'05*, pp. 1–14, 2005.

[17] S. Baloch, R. Verma, and C. Davatzikos, "An anatomical equivalence class based joint transformation-residual descriptor for morphological analysis," *IPMI, 2007* **4584**, pp. 594–606, 2007.

[18] M. B. Wakin, D. L. Donoho, H. Choi, and R. G. Baraniuk, "The multiscale structure of non-differentiable image manifolds," in *SPIE Wavelets XI, San Diego, Californi*, pp. 1–14, July, 2005.

[19] J. Ham and D. D. Lee, "Separating pose and expression in face images: A manifold learning approach," *Neural Inform. Process. C Letters and Reviews* **11**(4-6), pp. 91–100, April-June 2007.

[20] N.-X. Lian and C. Davatzikos, "Morphological appearance manifolds in computational anatomy: Groupwise registration and morphological analysis," *IJ - 2008 MICCAI Workshop - Manifolds in Medical Imaging: Metrics, Learning and Beyond* , 2008.

[21] J. Tenenbaum, V. de Silva, and J. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science* **290**, pp. 2319–2323, 2000.

[22] S. Roweis and L. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science* **290**, pp. 2323–2326, 2000.