Morphological appearance manifolds in computational anatomy: Groupwise registration and morphological analysis

Sajjad Baloch *, Christos Davatzikos

University of Pennsylvania, Philadelphia, PA, USA

A R T I C L E   I N F O

Article history:
Received 15 October 2008
Accepted 15 October 2008
Available online 12 November 2008

Keywords:
Computational anatomy
Groupwise registration
Deformation based morphometry
Voxel-based morphometry
Morphological appearance manifolds

A B S T R A C T

Existing approaches to computational anatomy assume that a perfectly conforming diffeomorphism applied to an anatomy of interest captures its morphological characteristics relative to a template. However, the amount of biological variability in a groupwise analysis renders this task practically impossible, due to the nonexistence of a single template that matches all anatomies in an ensemble, even if such a template is constructed by group averaging procedures. Consequently, anatomical characteristics not captured by the transformation, and which are left out in the residual image, are lost permanently from subsequent analysis, if only properties of the transformation are examined.

This paper extends our recent work [Makrogiannis, S., Verma, R., Davatzikos, C., 2007. Anatomical equivalence class: a computational anatomy framework using a lossless shape descriptor. IEEE Trans. Biomed. Imag. 26(4), 619–631] on characterizing subtle morphological variations via a lossless morphological descriptor that takes the residual into account along with the transformation. Since there are infinitely many transformation, residual] pairs that reconstruct a given anatomy, we treat them as members of an Anatomical Equivalence Class (AEC), thereby forming a manifold embedded in the space spanned by [transformation, residual]. This paper develops a unique and optimal representation of each anatomy that determines the optimal template and transformation parameters for each individual anatomy, and eliminates respective confounding variation in the data. It, therefore, constitutes the second novelty, in that it represents a group-wise optimal registration strategy that individually adjusts the template and the smoothness of the transformation according to each anatomy. Experimental results support the superiority of our morphological analysis framework over conventional analysis, and demonstrate better diagnostic accuracy.

© 2008 Published by Elsevier Inc.

Introduction

Computational anatomy typically characterizes 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 diffeomorphic shape transformation $h: \Omega_T \rightarrow \Omega_S, x \rightarrow h(x)$. The resulting transformation $h$ carries information about morphological differences between a subject and a template. Various descriptors may then be derived from $h$ for quantifying these morphological characteristics (Davatzikos et al., 1996, 2001; Ashburner et al., 1998; Ashburner and Friston, 2000; Thompson et al., 2000; Chung et al., 2001; Chetelat and Desgranges, 2002; Leow et al., 2006).

Abbreviations: DBM, Deformation based morphometry; TBM, Tensor based morphometry; VBM, Voxel based morphometry; AEC, Anatomical equivalence class; CMD, Complete morphological descriptor; OMS, Optimal morphological signature; UJD, Log Jacobian determinant; TDM, Tissue density map.

* Corresponding author.
E-mail addresses: sajjad.baloch@gmail.com (S. Baloch), christos@rad.upenn.edu (C. Davatzikos).

Such approaches date back to D’Arcy Thompson (Thompson, 1917), who, in 1917, studied differences among species by measuring deformations of the coordinate grids from images of one species to those of another (Grenander, 1983) later capitalized on this idea to propose the deformable template based morphometrics, which were further extended by the landmark-based approach (Bookstein, 1989). Later work (Christensen et al., 1993; Miller et al., 1997) provided a foundation for numerous approaches (Grenander and Miller, 1998; Miller and Younes, 2001; Miller et al., 2002; Geng et al., 2005; Lorenzen et al., 2006) based on the diffeomorphic transformations. Deformation based morphometry (DBM), for instance, establishes group differences based on the local deformation, and displacement (Ashburner et al., 1998; Gaser et al., 1999; Joshi, 1998; Collins et al., 1998; Cao and Worsley, 1999), or the divergence of the displacement of various anatomical structures (Thirion and Calmon, 1999). A feature of particular interest in this case is the Jacobian determinant (JD), which identifies regional volumetric changes (Davatzikos et al., 1996; Ashburner and Friston, 1999; Chung et al., 2001). Tensor based morphometry (TBM) (Thompson et al., 2000; Leow et al., 2006; Studholme and Cardenas, 2007) utilizes the tensor information for
capturing local displacement. Variants of TBM include voxel compression mapping (Fox et al., 2001), which describes tissue loss rates over time.

Another class of methods known as voxel-based morphometry (VBM) (Ashburner and Friston, 2000; Davatzikos et al., 2001; Baron and Chetelat, 2001; Bookstein, 2001; Chetelat and Desgranges, 2002) factors out global differences via a relatively low dimensional transformation, before analyzing them for anatomical differences. VBM is, therefore, considered as complementary to DBM or TBM, since the former utilizes the information not represented by the transformation. RAVENS maps (Davatzikos et al., 2001) addressed some of the limitations of both approaches, in a mass preserving framework, by combining the corresponding complimentary features. An important characteristic of the resulting maps was their ability to retain the total amount of a tissue under a shape transformation in any arbitrary region by accordingly increasing or decreasing the tissue density.

Irrespective of the features employed in various types of analyses, their accuracy largely depends on the ability of finding a spatial transformation that allows perfect warping of anatomical structures. Such transformations are typically driven by image similarity measures, in the so-called intensity-driven methods (Christensen et al., 1993; Woods et al., 1993, 1999; Ashburner and Friston, 1999), either by directly employing intensity differences or via mutual information. Topological relationships among anatomical structures are maintained by imposing smoothness constraints either via physical models (Christensen et al., 1993; Beg et al., 2005) or directly on the deformation field. Although these methods have provided remarkable results, their ability to establish anatomically meaningful correspondences is not always certain, since image similarity does not necessarily imply anatomical correspondence. Alternative methods (Thompson and Toga, 1996; Wang et al., 2003) are based on feature identification and matching. Hybrid methods, such as (Shen and Davatzikos, 2002), incorporate geometrically significant features to identify anatomically more consistent transformations.

Despite remarkable success of the above methods, their fundamental shortcomings result from inherent complexity of the problem. First, anatomical correspondence may not be uniquely determined from intensity-based image attributes, which drive these algorithms. Second, exact anatomical correspondence may not exist at all due to anatomical variability across subjects. For instance, it may not be possible to perfectly warp a single-folded cingulate sulcus to a double-folded cingulate sulcus through a biologically meaningful transformation. As a result, the choice of a template plays an important role in the accuracy of the analysis. Anatomies closer to the template are well represented by a diffeomorphism. However, large differences between an individual and the template lead to the residual information that the transformation does not capture. In other words, a priori fixing a single template for all subjects biases the analysis (Bookstein, 2001; Blezek and Miller, 2006); this has been a fundamental assumption in computational anatomy. Although one may argue that this residual may be eliminated by a very aggressive registration, this may create biologically

Fig. 1. Effect of aggressive registration: (a) Subject; (b) Template; (c) Subject warped through a viscous fluid algorithm. Aggressively registering a bifolded sulcus to a singlefolded sulcus creates very thin needle like structures, which are atypical of a human brain.

Fig. 2. (a) Template; (b) Mean of 31 spatially normalized brains; (c) A representative subject; (d) Spatial normalization of (b) using HAMMER; (e) Corresponding residual. While crispness of the mean brain indicates reasonably good anatomical correspondence, there are still significant anatomical differences, which may be large enough to offset disease specific atrophy in a brain.
implausible correspondences, as illustrated in Fig. 1. Aggressive warping may also lead to noisy deformation fields that are unsuitable for subsequent statistical analysis. To avoid these situations, typically some level of regularity is always desirable in the transformation. Both parameters, i.e., the choice of a template as well as the level of regularity, are often arbitrarily or empirically chosen. More importantly, they are not optimized for each anatomy; individually this leads to unwanted confounding variation in the morphological measurements, which may reduce our ability to detect subtle abnormalities.

Fig. 2 shows a characteristic example of residuals that persist after registration of brain images. MRI scans of 31 human brains were spatially normalized to the template given in Fig. 2(a) using deformable registration of (Shen and Davatzikos, 2002), with the smoothness parameter adjusted to avoid the creation of biologically incorrect warping. The average of spatially normalized brains is given in Fig. 2(b), whose clarity indicates relatively good registration. A typical subject (Fig. 2(c)), on the other hand, when warped to the template (Fig. 2(d)) still exhibits significant residual as shown in Fig. 2(e). Depending on the template, this residual may be large enough to offset disease specific atrophy in a brain, and may easily confound subtle anatomical variations. The question under consideration is, therefore, how to carry out accurate statistical analysis in such situations.

Some approaches have recently been proposed that utilize the mean brain as a template (Davis et al., 2004; Avants et al., 2006). In most practical cases, considerable differences still persist between some samples and the mean brain, and consequently, the residual is never negligible. A very promising approach in this situation is the groupwise samples and the mean brain, and consequently, the residual is never practical cases, considerable differences still persist between some brain as a template (Davis et al., 2004; Avants et al., 2006). In most complete morphological descriptor

The concept of an appearance manifold has gained a great deal of attention in the past 5 years in the computer vision community (Lee et al., 2003; Christoudias et al., 2004; Shan et al., 2005; Lina et al., 2008), albeit in a different context. In particular, object recognition depends largely on lighting and pose, which are variable and are considered confounding parameters. Image appearance manifolds are often constructed by varying the lighting and pose parameters, and learning the resulting image variations. Morphological appearance manifolds herein are constructed in an analogous way: parameters such as regularization constants and templates are varied for each anatomy, thereby allowing us to construct an appearance manifold of that anatomy, and to factor out variations that do not reflect underlying biological characteristics.

The paper is organized as follows. We start with problem formulation and morphometric analysis framework given next. Proposed optimal descriptor will be presented in the later part of Morphological descriptor framework section. Experiments section presents and discusses the experimental results on synthetic 2D and real 3D volumetric datasets with simulated atrophy. We conclude in Conclusions section with a discussion of results and future directions.

Morphological descriptor framework

Motivation

The fundamental principle of computation anatomy is that differences between various individuals, $S_i, i=1,...,L$ are characterized relative to a template $T$. This requires mapping the template space $\Omega$ to individual subject space $\Omega_S$ through a diffeomorphism $h_i \in H_{\Omega_S}: \Omega_T \rightarrow \Omega_S, x \rightarrow h_i(x)$, where $H_{\Omega}$ is the set of all optimal diffeomorphic transformations that maximize some similarity criterion between $T$ and normalized subject $S^T$. The resulting

Fig. 3. Manifold structure and intersubject comparisons based on Euclidean distance.
transformation $h_i, i=1,..., L$ then carry morphological information of different individuals relative to the template (Davatzikos et al., 1996, 2001; Miller et al., 1997; Ashburner et al., 1998; Joshi, 1998; Ashburner and Friston, 2000; Thompson et al., 2000; Chung et al., 1996, 2001; Miller et al., 1997; Ashburner et al., 1998; Joshi, 1998; 

Fig. 4. OMS versus CMD: (a) Randomly selecting CMDs (random parameter selection) reduces inter-group separation. Dots marked with arrows represent group means; (b) OMS results in optimal separation between the two groups.

denoted by a vector $\theta$. An entire family of CMDs may be generated by varying $\theta \in \mathbb{D}_\theta$, where $\mathbb{D}_\theta$ represents the domain of $\theta$, thereby, defining a parametric manifold, should $\theta$ be a continuous variable. In short, $\mathcal{M}_\theta$ is generated by combining vectorial forms of $h$ and $R_h$. 

**Definition 1.** Two CMDs $(h, R_h) \in \mathcal{X}$ are anatomically equivalent ($\sim$) if for a given template $T$, they reconstruct the same anatomy $S$:

$$
\begin{align*}
(h, R_h) \sim (h, R_h) & \iff T(h, R_h(y)) = T(h, R_h(y)) \\
& = S(y), \forall y \in \Omega_S.
\end{align*}
$$

where $\mathcal{X}$ is the space spanned by all such CMDs, and $y \in \Omega_T$.

This non-uniqueness of representation will be removed in the next section, where we first develop the framework in a general setting. It will then be followed by a discussion on specific examples of $\theta$.

**Assumption 1.** For a given anatomy $S$, the class of anatomically equivalent CMDs, referred to as anatomically equivalent class (AEC), is generated by varying transformation parameters $\theta \in \mathbb{D}_\theta$:

$$
\mathcal{A}(S) = \left\{ (h, R_h(x)) : S(h, R_h(x)) = T(x) - R_h(x), \forall x \in \Omega_T \right\}, \forall \theta \in \mathbb{D}_\theta.
$$

forms a smooth manifold $\mathcal{Q}_S$ embedded in the subspace $\mathbb{X} \subset \mathbb{R}^n$ spanned by CMDs, $n$ being twice the cardinality of discretized $\Omega_T$.

Note that all CMDs, each represented as a set of discretized components $\{h, R_h(x) \in \mathbb{X} \}$ in Eq. (3), satisfy the constraint $S(h, R_h(x)) = T(x) - R_h(x), \forall x \in \Omega_T$, which according to Definition 1 makes them anatomically equivalent. It is known that image appearance manifolds constructed by varying such parameters are continuous but not differentiable, if the images contain edges (Wakin et al., 2005). They can, however, be approximated by smooth manifolds, either by fitting parameterized manifolds to them, or by smoothing the images from which they are constructed (Wakin, 2007). Herein, we take the former approach, by approximating the 

Fig. 5. Approximation of AEC manifolds with hyperplanes. $\tau$ and $\lambda$ may be two of the confounding factors invariance to which is sought, as explained later in Robustness to template and regularization parameters section.

Fig. 6. Constructing AECs: Each subject is normalized to $\Omega_T$ via intermediate templates at different smoothness levels of the warping transformations.

### Anatomical equivalence class framework

The CMD, $\mathcal{M}_\theta$, depends not only on the underlying anatomy $S$ but also on transformation parameters, which collectively are

Fig. 7. Intermediate templates aid registration: (a) Template of Fig. 2(a); (b) Direct warping of Fig. 2(c); (c) Warping via an intermediate template.
manifold by a hyperplane within which it is contained. Since the dependence of \( M^A_\theta \) on diffeomorphism \( h \) is through the transformation parameters \( \theta \), in the rest of the text we drop the subscript to represent CMDs as \( M^A(\theta) \).

**Anatomical comparisons based on AEC manifolds**

Although the resulting AEC represents the entire range of variability in \( R_\theta \), it is not always clear what \( \theta \) should be selected for analysis. For instance, a small residual does not necessarily correspond to the best CMD, as we will experimentally show later, although minimizing residuals is a common target in most of deformable registration algorithms. It is, therefore, imperative to find an appropriate \( \theta \) for each individual that is optimized to the underlying anatomy according to a certain criterion \( J \), leading to optimal parameters \( \Theta=(\theta_1,\ldots,\theta_L) \), where \( L \) is the number of subjects.

In order to define our criterion for optimality of \( \Theta \), we first consider the simplest case of two subjects. To find the distance between two anatomies represented by their respective manifolds, one may define \( J \) as the minimum separation between their corresponding manifolds. Physically this amounts to inter-orbital distance and is computed by moving along the manifolds such that the distance between corresponding points is minimized.

\[
\text{dist}(S_A, S_B) = \min \left\{ d \left( M^A_{\theta_0}, M^B_{\theta_0} \right) : \forall \theta_0 \in \mathcal{F}_{S_A}, \forall \theta_0 \in \mathcal{F}_{S_B} \right\},
\]

(4)

where \( d \) represents Euclidean distance defined on \( \chi \), and \( \mathcal{F}_{S_A} \) and \( \mathcal{F}_{S_B} \) respectively denote the spaces of diffeomorphisms for anatomies \( A \) and \( B \). While this works for two anatomies, comparing more than two anatomies becomes problematic as shown in Fig. 3, where the optimal representations for comparison between \( S_A \) and \( S_B \) differ from those for comparison between \( S_B \) and \( S_C \).

To generalize the cost functional, we notice that for two subjects, optimization allows sliding along the respective manifolds to find two representations that yield minimum distance. These representations best highlight differences between these anatomies, since together they eliminate confounding effects of \( \Theta \). For \( L \) anatomies, we allow their representations to slide along respective AECs to minimize the sum of pairwise squared distances of all individuals. The objective functional, therefore, becomes:

\[
J(\Theta) = \sum_{i=1}^{L} \sum_{j \neq i \leq L} d^2 \left( M^i(\theta_i), M^j(\theta_j) \right),
\]

(5)

and the optimization is constrained to respective manifolds such that we in effect find optimal parameters as:

\[
\Theta^* = \arg \min_{\Theta=(\theta_1,\ldots,\theta_L)} J(\Theta),
\]

(6)

where \( M^i(\theta_i) \) is the CMD of subject \( i \) for \( \Theta=(\theta_1,\ldots,\theta_L) \). \( \Theta^* \) represents the optimal selection of parameters, whose values provide the optimal \( \Theta \) for each individual, which differs, in general, across individuals.

It is trivial to show that the criterion of Eq. (6) minimizes the variance of morphological descriptors over entire ensemble with respect to confounding factors, leading to:

\[
\Theta^* = \arg \min_{\Theta=(\theta_1,\ldots,\theta_L)} \sum_{i=1}^{L} \frac{1}{L-1} d^2 \left( M^i(\theta_i), \overline{M}(\Theta) \right),
\]

(7)

where:

\[
\overline{M}(\Theta) = \frac{1}{L} \sum_{i=1}^{L} M^i(\theta_i),
\]

represents the mean descriptor. The resulting OMS, \( M^i(\Theta^*) \), corresponding to optimal parameters, \( \Theta^{*}_i \), for each individual, and, therefore, removes arbitrariness due to these parameters, as illustrated schematically in Fig. 4.

**Optimal morphological signature**

For simplicity and tractability, we approximate AEC manifolds with hyperplanes to solve the optimization problem of Eq. (7), as illustrated in Fig. 5.
Each manifold is first independently represented in terms of its principal directions, computed through principal component analysis (PCA). If $\{V_j^i\}_{j=1}^n$ represent principal directions of subspace $\mathcal{X}_i$, in which the manifold $\mathcal{Q}_i$ of subject $S_i$ is embedded, and $\hat{M}^i$ denotes corresponding subject mean, then the linear hyperplane approximating the corresponding AEC manifold is given by:

$$\hat{M}^i(\theta) = \hat{M}^i + \sum_{j=1}^n \alpha_j V_j^i,$$

where $\alpha_j \in [\alpha_j^{\text{min}}, \alpha_j^{\text{max}}], j=1, \ldots, n$ capture transformation dependent variability originally represented by $\theta$. Basically, $\alpha_j$ identifies the component of a complete descriptor $\hat{M}^i(\theta)$ along direction $V_j^i$, and by varying it in the interval $[\alpha_j^{\text{min}}, \alpha_j^{\text{max}}]$, we allow sliding along the manifold. The bounds $\alpha_j^{\text{min}}$ and $\alpha_j^{\text{max}}$ define the extents of the manifold, and may be computed from corresponding principal modes.

In the experiments of this paper, these bounds were selected so as to allow moving twice the standard deviation along each principal direction.

The objective function of Eq. (7), therefore, becomes:

$$A^* = \arg\min_{A = (a_1, \ldots, a_L)} \sum_{i=1}^L d^2 \left( \hat{M}^i + \sum_{j=1}^n \alpha_j V_j^i - \hat{M}(A) \right),$$

where

$$\hat{M}(A) = \frac{1}{L} \sum_{i=1}^L \left( \hat{M}^i + \sum_{j=1}^n \alpha_j V_j^i \right)$$

is the mean CMD across subjects for current correspondence $A^* = (\alpha_1, \ldots, \alpha_n)$.

Solution to this constrained problem is an algorithm that allows moving along individual hyperplanes, to minimize the objective function. As shown in Fig. 5, at each optimization iteration, an update of $M^i(\alpha_j)$ is computed, which yields the current floating mean $\hat{M}(A)$. The procedure is repeated until the minimum of Eq. (9) is attained. Analytically it leads to the following solution subject to constraints given above:

$$
\alpha_j^* = \frac{1}{\sum_{i=1}^L \left( \hat{M}^i + \sum_{j=1}^n \alpha_j V_j^i \right) - \hat{M}^i} \left[ \sum_{i=1}^L \hat{M}^i \right], j = 1, \ldots, L. \tag{10}
$$

When combined with Eq. (8), optimal $\alpha_j^*$ yields OMS, $A^*$, which is then used for subsequent analysis. It provides optimal combination of transformation and residual by finding optimal selection of transformation parameters $\theta$.

### Robustness to template and regularization parameters

In this section, we particularize the above formulation to the problem at hand, by noting that the residual is mainly a consequence of two parameters, namely the template and the regularization of the diffeomorphism, as mentioned in Introduction section. The AEC of a given anatomy, $S$, is generated by varying these two parameters $\lambda \in \mathbb{R}$, and $\tau \in T$ with $T$ representing the set of all possible templates. Since analysis eventually has to be carried out in a common space, we ultimately bring all warped anatomies to a common template space $\Omega_0$ as illustrated in Fig. 6. However, variations caused by the selection of different templates have already been represented by the intermediate templates $\tau$ (Fig. 6).

Suppose that for a given $\tau$, $\mathcal{F}_{S\tau}$ represents the set of all diffeomorphisms that warp $S$ to $\tau$ to yield $S_{\tau}$, whereas $\mathcal{G}_{S\tau}$ denotes the set of diffeomorphisms that warp $S$ to $T$:

$$\mathcal{F}_{S\tau} := \{ f_{S\tau} \in H_5 : (Sf_{S\tau}) = \tau \in T \}, \quad \mathcal{G}_{S\tau} := \{ g_{S\tau} \in H_5 : (Sg_{S\tau}) = \tau \in T \}. \tag{11}$$

Then, the set of transformations that take $S$ to $T$ is given by:

$$\epsilon_S := \{ h_{S\tau} = f_{S\tau} \ast g_{S\tau} \in H_5 : f_{S\tau} \in \mathcal{F}_{S\tau}, g_{S\tau} \in \mathcal{G}_{S\tau}, \forall \lambda \in \mathbb{R} \}, \forall \tau \in T.$$
Algorithm 1

Optimal morphological representation

**Input:** Individual Anatomies \(S_i, i = 1, \ldots, L\), Intermediate templates \(T\), Final template \(T^*\)

**Output:** Optimal Parameter Selections \(\alpha^*_i\), Optimal Morphological Signatures 

\[
\mathcal{M}_i(\alpha^*_i), i = 1, \ldots, L
\]

1. \(\text{begin}\)
2. \(\text{for } i = 1, \ldots, L \text{ do}\)
3. \(\text{for each } \lambda \in R, \text{ do}\)
4. \(\text{Estimate the diffeomorphism } f_{\lambda, \tau} \text{ by registering } S_i \text{ to } \tau \text{ to get } S_{\lambda, \tau}\)
5. \(\text{Estimate the diffeomorphism } g_{\omega, \tau} \text{ by registering } S_{\lambda, \tau} \text{ to } \tau \text{ to get } S_{\lambda, \tau}\)
6. \(\text{Determine the composite diffeomorphism } h_{\lambda, \omega} = f_{\lambda, \tau} \circ g_{\omega, \tau}\)
7. \(\text{Find the corresponding residual: } R_i = T^* - S_{\lambda, \tau}\)
8. \(\text{Construct a CMD } \mathcal{M}(h_{\lambda, \omega}) \text{ from } R_i \text{ and } h_{\lambda, \omega}, \text{ as described in Section 2.6}\)
9. \(\text{end}\)
10. \(\text{end}\)
11. \(\text{Construct AEC manifold } \mathcal{Q} = \cup_{\lambda, \omega, \tau} \mathcal{M}(h_{\lambda, \omega})\)
12. \(\mathcal{M}(\{V^0(\omega), 1, \ldots, n\}) = \mathcal{P}(\mathcal{Q}) \text{ according to Eq. (8)}\)
13. \(\text{end}\)
14. \(\text{Solve optimization problem of Eq. (9) to get } \alpha^*_i, i = 1, \ldots, L\)
15. \(\text{Compute OMSs } \mathcal{M}_i(\alpha^*_i), i = 1, \ldots, L \text{ according to Eq. (8)}\)
16. \(\text{end}\)

**Feature extraction**

The previously discussed framework was derived in a very general setting, with its applicability going beyond medical image analysis. In this paper, we leave out discussion on the application of this approach to other areas of computer vision, and instead focus on how to formulate features for application to biomedical image analysis.

So far we have used a general notation \(\mathcal{M}_k\) for CMDs, which is suitable for groupwise registration. However, several avenues open if one concentrates on the problem at hand. For instance, if one is interested in characterizing volumetric deficits for diagnosis of Alzheimer’s disease, Jacobian determinant (JD), \(J_h\), of the transformation \(h\) is a feature of interest. For modeling brain development patterns, one may utilize some measure of diffusivity, such as fractional anisotropy (Basser and Pierpaoli, 1996). For consistent labeling, intensity values provide a good feature of interest.

Weighted combination of features

One may readily suggest weighting individual features with their \(z\)-scores. In this paper, we adopt a more natural combination with their \(z\)-scores. For example, skull-stripped brain MR images are typically segmented with \(c = 3\) into grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Pre-processed cardiac MR images are also typically segmented with \(c = 3\) into muscle, blood, and fat. Based on segmented images one may readily extract tissues of interest. Suppose \(I_k(h(x))\) and \(K_k(x)\) respectively denote indicator functions for \(k = 1, \ldots, c\) tissues in the individual image \(S\) and the template \(T\). Then corresponding residuals are defined as:

\[
R^k_h(x) = K_k(x) - I_k(h(x)).
\] (12)

Since JD represents the degree of growth or atrophy, a similar quantity is derived by weighting tissue based residuals \(R_k^h\) to have values of \([-1,0,1]\). With such weighting +1 indicates unit increase in volume, and –1 represents a unit decrease. Total tissue loss or gain as reflected by the residual may be computed by integrating the residual over the region of interest, which is tantamount to counting up or down the number of non-zero voxels (assuming that a voxel has a unit positive or negative volume), where the sign indicates volumetric growth or deficit. As a result, if the \(J_h\) shows an atrophy at \(x \in \Omega\), actual volumetric deficit at \(x\) may be lower or higher depending on the corresponding \(R^k_h\).

Tissue density maps

\(h\) and \(R_k\) may also be combined to form a morphometric descriptor based on residual and shape-based features, referred to as tissue density map (TDM) (RAVENS maps in (Davatzikos et al., 2001)) that relates to the spatial distribution of anatomical tissues in the target anatomy:

**Definition 2.** A tissue density map (TDM), \(\mathcal{D}^k: \Omega_T \rightarrow \Omega_T\), for \(k = 1, \ldots, c\) tissues is defined as:

\[
\mathcal{D}^k(x) = \int_{\Omega_T} (K_k(x) - R^k_h(x)) \, dx.
\] (13)

TDMs may be computed in a mass-preserving framework of (Davatzikos et al., 2001) by warranting that the amount of tissues in \(S\) are preserved exactly under the transformation.

**Theorem 1.** TDM defined by Eq. (13) is a mass preserving map (Makrogiannis et al., 2007).

**Proof.** For any region \(R \subset \Omega_T\) defined in the template space,

\[
\int_R \mathcal{D}^k(x) \, dx = \int_R (K_k(x) - R^k_h(x)) \, dx
\]

\[
= \int_R (K_k(x) - R^k_h(x)) \int_{\Omega_T} (h(y)) \, dy
\]

\[
= \int_R (K_k(x) - R^k_h(x)) \, dh(x)
\]

As a consequence of this theorem, regions that contract under \(h^{-1}\) demonstrate an increase in density. Since a relatively larger anatomical region contracts to a smaller one, its density increases if the tissue mass is to be preserved, as illustrated in Fig. 8.

In further discussion, CMDs derived from these two types of features will be discussed.

**Experiments**

In this section, we present extensive experimental results to support the hypothesis that residual carries significant amount of information for identifying group differences and that OMS yields superior performance by maintaining group separation between normal and pathologic anatomies in a voxel-wise statistical comparison framework.

For comparison, we perform two types of tests on CMDs in addition to OMSs: (1) \(t\)-tests on individual log \(\log_{J_h} \) and \(R_k\), components, and (2) \(T^2\) test on \(\log_{J_h} \) and \(R_k\) descriptors, to compute \(p\) values. This assigns a level of significance to each feature in terms of differences between healthy and pathologic anatomies. For CMDs, we randomly selected intermediate templates for each subject before conducting tests for all smoothness levels \(\lambda\). Note that residual was smoothed with a Gaussian filter with various selections of smoothness.
parameter \( \sigma \) prior to statistical tests mainly due to two reasons. First, it ensures the Gaussianity of the smoothed residual. Second, since the residuals appear only on tissue boundaries, even if tissue atrophy is in the interior of the structure, smoothing produces a more spatially uniform residual. JD, on the other hand, was not smoothed for the \( T^2 \) test due to its inherent smoothness.

Three datasets are considered comprising 2D toy data and 3D simulated cross-sectional and longitudinal MRI data.

It should also be mentioned that the proposed approach does not depend on the choice of the registration algorithm. We, therefore, utilized viscous fluid based registration for 2D dataset, whereas HAMMER for the 3D case. As discussed later, both algorithms yielded consistent results. In order to simulate the effect of variation in the regularization \( \lambda \) in transformation (shown in Fig. 6), we first registered all individuals with “maximal” flexibility (lowest \( \lambda \)) to multiple templates. These highly aggressive diffeomorphisms were then gradually smoothed to generate transformations \( h_{\lambda \tau} \) with various levels of regularity. In addition, we smooth residual slightly (a Gaussian kernel of standard deviation of 0.5) before optimization in order to ensure regularity of the AEC manifolds.

2D Synthetic dataset

A 2D dataset of 60 shapes was generated by introducing random variability in 12 manually created templates given in Fig. 9. First, the shape of each template is represented by a number of control points, which are then randomly perturbed to simulate the variability across subjects resembling anatomical differences encountered in the gray matter folds of the human brain. Thinning was introduced (5%) in center one-third of the fold of 30 subjects to simulate atrophy (patient data). All subjects were spatially normalized to \( T_{12} \) via \( T_{p...T_{12}} \), for smoothness levels of \( \lambda=0,...,42 \) to construct individual AECs.

Effect of transformation parameters on registration accuracy is illustrated in Figs. 10 and 11. Note how registration accuracy...
deteriorates with inappropriate selection of $\lambda$ and $T$. Anatomies that are topologically similar to the template are represented well, whereas others require an intermediate level of regularity to avoid the creation of biologically incorrect structures.

Voxel-wise statistical tests

First we compare the level of significance of group differences based on CMDs and OMSs by computing $p$-value maps for various values of $\sigma$ (and $\lambda$ for tests on CMDs). Minimum of $p$-value maps is plotted in Fig. 12 as a function of $\sigma$ to indicate the best achievable performance for the two descriptors.

It may be observed from results based on CMDs that residual achieves considerably lower $p$ values as compared to JD. The significance of both log JD ($\text{LJD}$) and the residual increases with $\sigma$ and $\lambda$ up to a point after which it starts degrading. Similarly, $T^2$ test also shows best performance for intermediate values of $\lambda$ (Fig. 13(a)), which means that an overly aggressive transformation tends to contaminate statistical analysis. This agrees with (Cachier, 2001), which argued the need of moderate regularization in the transformation. These observations are also in accordance with our hypothesis that the residual carries anatomical information that is complementary to, and perhaps is more important than, the transformation. After optimization, the absolute minimum $p$ value improves from $10^{-9}$ at $\lambda=23$ for CMD to $10^{-10}$ for OMS (corresponding to $\sigma=13$) as shown in Fig. 13(a).

It should be noted that the minimum $p$ value plots given in Figs. 12 and 13(a) do not completely represent spatial distribution. For classification, it is essential to use a descriptor that leads to small $p$-values in the entire ROI. For OMS, $p$ values were consistently found to be small in the entire ROI, whereas for CMDs, it is typically an isolated voxel that yields low $p$-value. A comparison of mean $p$ values in regions with $p \leq 10^{-2}$ is given in Fig. 13(b), which demonstrates that OMS ($p=10^{-4.5}$) slightly outperforms the best possible CMD ($p=10^{-4}$, $\lambda=39$). It may, however, be immediately inferred from the figure that the OMS significantly outperforms the CMDs in general, i.e., over the entire range of $\lambda$ values.

$p$-significant ROIs ($p=10^{-2}$) were computed for OMS and CMD corresponding to the parameter selection that yielded best performance ($\lambda=23$ for CMD and $\sigma=13$) as shown in Fig. 14. Note how OMS helps in precisely localizing atrophy, which is in accordance with the objective function of Eq. (9). On the other hand, CMDs fail to accurately localize atrophy, with a considerably large number of false positives.

![Fig. 14. 2D synthetic dataset — $T^2$-test based log$_{10}$ $p$-value maps for 2D simulated data at $\sigma=13$ thresholded to $p \leq 10^{-2}$: (a) CMD $\lambda=23$; (b) OMS.](image1)

![Fig. 15. 2D synthetic dataset — Hotelling’s $T^2$ test on CMD and OMS corresponding to the best case of Fig. 13, i.e., $\sigma=13$, and $\lambda=23$. Note that $p$-value maps are thresholded to $p \leq 10^{-4}$. Results are shown on a log$_{10}$-scale: (a) CMD with $T_1$; (b) CMD with $T_2$; (c) CMD with $T_3$; (d) OMS with $T_1$; (e) OMS with $T_2$; (f) OMS with $T_3$.](image2)

![Fig. 16. 3D dataset: (a) A subject without atrophy; (b) With 10% simulated atrophy.](image3)
CMD were considered corresponding to their best parameter selections. As indicated by the p-value maps of Figs. 14 and 19, not all voxels in morphological descriptors are discriminating. We, therefore, employ Hotelling’s $T^2$ test to rank all features according to their associated p-values, and then select a subset of features with $p < 10^{-2}$. Note that the resulting feature vectors, in general, form a low dimensional embedding in a high dimensional space. It is, therefore, necessary to find the embedding or to carry out dimension reduction to avoid the curse of dimensionality (Hastie et al., 2001). Since our proposed method is distance based, we used isomap (Tenenbaum et al., 2000) to find the embedding thereby preserving the distance structure. The neighborhood parameter was varied from 4 to 15 (one fourth of the dataset), and the dimensionality of the embedding was consistently found to be 3 as indicated by the elbow of the residual variance curve. Consequently, a support vector machine based classifier is learned to partition the embedding through corresponding low dimensional feature vectors.

In order to account for nonlinearity of the class boundary, we utilized a radial basis function (RBF) kernel. SVM classifier was trained and tested for both CMD and OMS through 5-fold cross validation, where classification rates were found to increase from 73% for CMD to 91% for OMS. This improvement clearly demonstrates the superiority of the proposed method over traditional approaches.

**3D MRI sectional data with simulated atrophy**

The second dataset consisted of real MRI scans of 31 subjects. To simulate patient data, 10% atrophy was introduced in 15 randomly selected subjects in a spherical region as shown in Fig. 16 using the simulator of (Karacali and Davatzikos, 2006). Five intermediate templates were selected for spatial normalization to generate AECs for all subjects with smoothing levels of $\lambda = 0, \ldots, 7$.

Minimum of p-value maps were computed for all values of $\sigma$ (and $\lambda$ for tests on CMDs). Minimum log$_{10}$p plots given in Fig. 17 show that residual achieves considerably lower p values as compared to LJQ, again indicating the significance of residual for capturing group differences. Best performance is achieved at high $\lambda$ ($\lambda = 7$). The dependence of $(J_{\theta_{13}}, R_{\theta_{13}})$ on $\lambda$ is eliminated through optimization as indicated in Fig. 18 by OMSs. However, no improvement in minimum p value was achieved through optimization. In this particular example, CMDs perform slightly better ($p = 10^{-9.75}$) than OMSs ($p = 10^{-10.5}$). However, OMS appears superior to CMD due to its relative insensitivity to $\sigma$. The variation in p-values for OMS is in the range $\sigma = 2–6$ is $10^{-0.75}$ whereas that for CMDs is $10^{-5.5}$ which highlights that CMDs are much more affected by $\sigma$. Small variations in $\sigma$, therefore, may considerably degrade CMD-based analysis. OMS is, hence, not only more robust due to better dynamic range, but it also maintains the separation between the two groups as indicated by very

**Fig. 17.** 3D dataset – The effect of regularization parameter $\lambda$ and smoothing $\sigma$ on the performance of morphological descriptors without optimization, as depicted by minimum p-value plots based on the t-test for capturing significant group differences: (a) log$_{10}$ J$_{\theta_{13}}$; (b) R$_{\theta_{13}}$. For log$_{10}$ J$_{\theta_{13}}$, best performance is achieved for low regularization ($\lambda = 0$), whereas R$_{\theta_{13}}$ performs the best for high regularization ($\lambda = 7$).

**Fig. 18.** 3D dataset – Hotelling’s $T^2$-test based minimum p-value plots for CMDs and OMS. The performance of CMDs is highly dependent on $\lambda$. Optimization, on the other hand, removes this dependency as evident from the largely stable curve for OMS, which is also less sensitive to $\sigma$.

**Fig. 19.** Atrophy maps as captured by CMD and OMS for the 3D dataset – $T^2$ test based p-value maps corresponding to the best results for each descriptor thresholded to $p < 10^{-5}$: (a) CMD $\lambda = 7$, $\sigma = 4$; (b) OMS with $\sigma = 5$. 

Invariance to template selection

In order to evaluate the robustness of OMS-based statistical analysis, we vary the template $T$ to which all anatomies are finally warped (see Fig. 6). Three different choices of $T$ were considered ($T_1$, $T_2$, $T_3$) to set up three optimization problems. For each $T_0$, Hotelling’s $T^2$ test was conducted on optimized as well as unoptimized $(\log_{10} J_{\theta_{13}}, R_{\theta_{13}})$ descriptors. We specifically considered the best parametric selections for CMDs and OMS, which correspond to $\sigma = 13$ and $\lambda = 23$ as indicated by Fig. 13(b). p-value maps thus computed after appropriate thresholding ($p = 10^{-4}$) are given in Fig. 15. Figs. 15(a)–(c) indicate that statistical test on CMDs lead to somewhat different regions of significant differences for different template selections. On the other hand, statistical analysis based on OMS is robust to variations in the choice of the template, and better agrees with the true underlying atrophy. In addition, OMS results in low p-values throughout the region of atrophy ($p = 10^{-8} - 10^{-10}$), whereas those for CMDs are mostly in the range $p = 10^{-5} - 10^{-7}$.

Classification

In this section, we further test the performance of the CMD and OMS through pattern classification. For comparison, both OMS and
When \( p \) was thresholded to find regions with values \( 10^{-2} \), CMDs resulted in false negatives as shown in Fig. 19(a). On the other hand, OMS helped in precisely localizing atrophy Fig. 19(b), which is in accordance with the objective function of Eq. (9).

3D MRI simulated longitudinal data (serial scans)

The third set of experiments evaluated the concept of building AECs from a different perspective, by generating simulated longitudinal aging profile in an individual MRI scan. 50% atrophy was introduced in three different regions of the brain: (1) posterior cingulate; (2) hippocampus; and (3) superior temporal gyrus over 12 time points (simulating a period of 12 years). Such datasets are quite common in practical studies for modeling normal decay of GM versus disease specific atrophy. The reason for performing this experiment can be appreciated, if one considers Fig. 4, where each of the manifolds now correspond to the same anatomy measured at different time points. The true longitudinal change may then be better estimated if we measure distances across manifolds, rather than distances of individual measurements, which can artificially appear larger than true longitudinal variation.

GM TDMs were computed by warping these 12 images to a common template via multiple (five) “intermediate” templates to construct anatomical manifolds. Each manifold, thus, accounted for the morphological description of a particular time point, simulating the situation where an anatomy evolves through a series of manifolds as an individual progresses in age. Factorization of variability along these manifolds is, therefore, necessary to find optimal descriptors that retain only temporal variation.

We compare the proposed approach with traditional TDM based analysis (most aggressive registration with direct warping), by evaluating mean temporal profiles of TDMs in regions with atrophy (Fig. 20) as well as regions without atrophy (Fig. 21). Results highlight the limitations of traditional analysis, which suffers from random fluctuations in TDMs resulting from arbitrariness due to transformation parameters. Consequently, they fail to correctly characterize temporal profiles. For instance, in Fig. 21, a linear regression model highlights atrophy, where in reality no atrophy was present. OMS, on the other hand, helps in minimizing this arbitrariness, to accurately account for underlying atrophy.

In order to understand the effect of optimization, regression maps were computed. For instance, that for traditional descriptor (Fig. 22 (a)) indicates local tissue growth in the highlighted area, which is in fact a side effect of registration errors. It should be noted that a local
growth in such cases contaminates the amount of atrophy modeled by traditional descriptors, as observed in Figs. 20 and 21. These effects were minimized by OMS based analysis, as indicated by Fig. 22(b), which correctly highlights atrophy in GM.

**Conclusions**

In this paper, we proposed a fundamentally novel framework for morphological analysis with three major contributions. First, the transformation, normalizing an anatomy to a common template space, was combined with the residual for a complete representation of the anatomy. Second, each anatomy was represented through an AEC manifold, reflecting variability due to different templates and regularization parameters. Third, an unsupervised approach was consequently developed for factoring out this unwanted variation due to these parameters, thereby measuring the true inter-subject differences. Such an approach ensures that optimal descriptors do not depend on group associations. Moreover, it yields similar morphological descriptors if the underlying anatomies are similar, irrespective of their group membership.

In the experiments presented in the paper, we focussed on atrophy based analysis, and validated the proposed approach with 2D synthetic and 3D real datasets with simulated cross-sectional as well as longitudinal atrophy. Two features were, therefore, considered, namely (LJD, residual), and TDM. Several improvements on traditional descriptors were readily observed. Residual was consistently found to be highly significant for characterizing group differences, as indicated by considerably low \( p \)-values. \( T^2 \) tests on CMDs confirmed our hypothesis that best performance is achieved for not so aggressive registration. Decreasing the level of regularity from very aggressive registration improves performance up to a certain extent after which it starts deteriorating. This suggests the importance of optimal selection of transformation parameters, which actually depends on the individual anatomy.

Optimization was found to introduce several improvements in the performance. First, it resulted in significantly low \( p \)-values indicating its ability to detect significant group differences. Second, the \( p \)-values were found to be consistently low in the entire ROI of the true atrophy, whereas the CMD results in very low \( p \)-values for only a few of the voxels. Consequently, the CMDs were found less discriminating for classification purposes. This was confirmed by the classification of morphological descriptors into healthy and pathologic anatomies through nonlinear support vector machines. As a consequence of optimization, classification rate was found to improve by 18%.

One of the most important aspects of optimization is its ability to precisely highlight the regions of significant differences. The CMD, however, resulted in relatively poorer localization of such regions, with considerable false positives and false negatives, thus lowering the specificity and sensitivity of the analysis. Invariance to the choice of the template was also readily observed for the OMS, in contrast to the unoptimized representation. It should be noted that the best performance of the CMD in terms of the lowest \( p \)-value requires intermediate values of \( \lambda \), which are not exactly known a priori. On could potentially choose the \( \lambda \) that yields the most significant group differences. However, such an approach by construction amplifies the multiple comparison problem, which is prevalent in voxel-based statistical analysis of medical images. In addition, \( p \)-values were found to be less sensitive to \( \sigma \) after optimization, yielding a better dynamic range.

For the longitudinal dataset, the proposed approach was able to detect group differences where traditional analysis failed. Optimization helped in minimizing random fluctuations in the temporal profiles for more accurate characterization. Application of this approach for measuring longitudinal changes in serial scans is particularly interesting and encouraging. The problem of “jittery” measurements from serial scans is very important when evaluating subtle changes due to disease progression or response to treatment. These random variations that are unrelated to the true underlying morphological changes significantly reduce sensitivity and specificity of these measurements. Our approach effectively removed the confounding variations that are due to the template and smoothness parameter selection, and allowed us to obtain considerably more stable estimates of longitudinal change.

Future work includes application to real datasets for computer aided diagnosis as well as nonlinear modeling of AEC manifolds.

**Conflict of interest statement**

The authors declare that there are no conflicts of interest.

**Acknowledgment**

This work was supported in part by the National Institutes of Health under grant number R01AG14971.

**References**


