Statistical model of similarity transformations: building a multi-object pose model of brain structures

Matías N. Bossa, Salvador Olmos
{bossa,olmos}@unizar.es
Communications Technology Group (GTC), I3A, University of Zaragoza
María de Luna 1, 3rd Floor, 50018, Zaragoza, Spain

Abstract

In most of computational anatomy studies, pose is disregarded because pose information mainly depends on non relevant external factors. However, the relative pose among different objects belonging to a complex multi-object system may be very useful for diagnosis, prognosis and monitoring. In this work a methodology to build statistical multi-object pose models (MOPM) is described. The methodology is based on Principal Geodesic Analysis because the space of similarity transformations does not form a vector space. Methods to compute statistics, namely averages and variation modes are described in detail. Experimental results are performed on neuroanatomical structures such as the subcortical nuclei (caudate nucleus, hippocampus, amygdala, thalamus, putamen, pallidum and lateral ventricles). We expect that multi-object pose models will be useful because they provide a valuable a priori information about relative location, orientation and scale, very useful as a coarse initialization for segmentation algorithms, or regularization of segmentation and registration algorithms.

1. Introduction

Computational Anatomy (CA) is being increasingly used to characterize brain anatomy either in control subjects as well as abnormalities of brain structures in patients with neuropsychiatric disorders [6, 2]. In the early 90's digital brain atlases were proposed mainly for corregistration of images from different modalities and volumetric studies [8]. However, the characterization of subtle anatomical abnormalities requires methodologies that provide local shape parameters of structures within the brain, such as the subcortical nuclei.

Many brain anatomy studies have been performed in order to search neuroanatomical markers of neuropsychiatric disorders. To name a few, volumetry and morphometry studies of the hippocampus and the amygdala in Alzheimer's disease (AD) patients [7], volumetry and morphometry analysis of the thalamus in schizophrenic patients [3, 5].

The main difficulty to be addressed in the analysis of human brain is that anatomical structures form a highly complex system, with a large shape variation among individuals. Many morphometry studies isolate a single structure from the rest of the brain and perform statistical inferences of patient groups regarding clinical categories.

In shape analysis theory, shape is often defined as all the geometrical information of an object living in an N dimensional space, which is invariant to location, orientation and very often size. Pose is usually defined as the similarity transformation that takes a shape from a reference frame to its actual real world coordinates.

Shape is usually described as the residual of a pose removal procedure. Accordingly, shape description depends on the particular pose definition. In many actual problems, pose is disregarded because mainly provides non useful information, as it depends on external factors (*e.g.* position and orientation of the patient within the scanner). However, the relative pose of the different objects belonging to a multi-object system may provide a very useful and compact information. In neuroanatomy, while overall size, position and orientation depend on external factors, the relative pose among different anatomical structures (*e.g.* caudate nucleus, hippocampus, thalamus, etc) may provide valuable information for diagnosis, prognosis, disease and therapy monitoring of several neuropsychiatric disorders. For example, hemispheric pose asymmetry could be characterized by a very small set of pose parameters.

Statistical shape analysis is an emerging field with many applications on medical imaging and computer vision. Principal Component Analysis (PCA) is the preferred procedure for rank reduction and statistical shape analysis. This technique should be only applied to multivariate data lying on an Euclidean space. However, similarity transformations form a Lie group, which is a group and a Riemannian manifold.

To overcome this limitation, Principal Geodesic Analysis

(PGA) has been recently proposed as an extension of PCA, where the concept of straight lines is replaced by geodesic curves.

Pennec [16, 15] defined uniform and Gaussian distributions on general manifolds, and later tensor metrics were studied in detail, including interpolation of Diffusion Tensor images [17]. Fletcher *et al*. used Lie group framework to define PGA for statistical analysis of 3D boundary representations based on medial atoms [10, 11, 12] and analysis of diffusion tensor images [9, 13]. A more detailed review of previous works on this field can be obtained in [13].

The aim of this paper is twofold. Firstly, to define the methodology for building statistical models of similarity transformations. Secondly, to apply this methodology to the case of a multi-object pose model (MOPM). Experiments are performed on anatomical structures within the human brain.

2. Statistics of similarity transformations

Moments are concise but often used and useful statistical descriptors. The nth moment, and central moment of a probability density function $f(x)$, are defined as: $\mu'_n = E(X^n) = \int_{-\infty}^{\infty} x^n f(x)dx$, and $\mu_n = E((X-\mu'_1)^n) = \int_{-\infty}^{\infty}(x-\mu'_1)^n f(x)dx$ respectively, where $E(.)$ denotes the mathematical expectation. The simplest and most used way to describe a data set is in terms of the first moment μ'_1 and the second central moment μ_2, *i.e*. mean and variance. If the probability distribution is assumed Gaussian, these two values provide a complete description of the whole distribution.

In order to compute moments of geometrical transformations, a mathematical characterization is required (see sec. 2.3) as well as the concept of distance between any pair of transformations. Similarity transformations form a Lie group, where the concept of distance is well defined. In fact, it is known how to compute mean and variance, and the analog of principal components, in any Lie group (see sec. 2.2).

2.1. Lie groups

Lie groups are groups that are also differentiable manifolds. A group G is defined as a set of elements together with a binary operation (not necessarily commutative, when this happens the group is called Abelian) or multiplication $x \bullet y = xy = z, x, y, z \in G$. The operation must fulfill the following properties: associativity ($(xy)z = x(yz)$); existence of an identity element $e \in G$ such that $ex = xe = x$; and $\forall x \in G$ exists $x^{-1} \in G$ such that $xx^{-1} = x^{-1}x = e$.

Roughly, a differentiable manifold is a space that locally looks like an Euclidean space, but globally may be curved in some way (like a smooth surface). The tangent space to G at the point $g \in G$, denoted as T_gG, is an Euclidean space of the same dimension of G. When an inner product $\mathbf{a} \cdot \mathbf{b} \in \mathbb{R}$, $\mathbf{a}, \mathbf{b} \in T_gG$ is defined in the tangent space of G at each point g, such that it varies smoothly from point to point, the manifold is called Riemannian. Distances or angles are well defined in Riemannian manifolds.

The tangent space of a Lie group G at the identity e, T_eG, is called Lie algebra $\mathfrak{g}$.

The length of a curve $\gamma(t) \in G$, $t \in \mathbb{R}$, between two points $\gamma(a)$ and $\gamma(b)$ is $L_a^b(\gamma) = \int_a^b \|\gamma'(t)\|dt$, where the derivative $\gamma'(t) \in \mathfrak{g}$ and the norm is given by $\sqrt{\gamma'(t) \cdot \gamma'(t)}$, and $\cdot$ is de Euclidean inner product. The distance between two elements $x, y \in G$ is defined as the length over the geodesic (shortest path between two elements of G) that passes through both elements: $d(x, y) = \min_\gamma(L_x^y(\gamma))$. This metric is called intrinsic or Riemannian. The extrinsic metric is the Euclidean distance between points when the manifold is embedded in a larger Euclidean vector space.

For every $\mathbf{v} \in \mathfrak{g}$, there is a unique geodesic $\gamma(t) \in G$ that starts at the identity $\gamma(0) = e$ and has initial velocity $\mathbf{v}$: $\gamma'(0) = \mathbf{v}$. This geodesic is defined by means of the so called exponential map $\mathfrak{g} : \exp(\mathbf{v}) \rightarrow G$ ($\gamma(t) = \exp(t\mathbf{v})$ is a geodesic with $\gamma(0)' = \mathbf{v}$ and $\gamma(0) = e$). Straight lines in $\mathfrak{g}$ that pass through the origin are mapped to geodesics in G passing through the identity e. Also, the angle at the identity between two geodesics intersecting with velocities $\mathbf{u}$ and $\mathbf{v}$ is $\theta = \cos^{-1}(\mathbf{u} \cdot \mathbf{v})$. The log map $G : \log(x) \rightarrow \mathfrak{g}$ is the inverse of the exponential map.

The distance from the identity e and any element $x \in G$ is equal to the Euclidean distance, in the tangent space, from the origin to $\mathbf{v} = \log(x)$:

$$d(e, x) = \|\log(x)\|, \tag{1}$$

and the distance between any two elements x, y of G is:

$$d(x, y) = d(e, x^{-1}y) = \|\log(x^{-1}y)\|. \tag{2}$$

Note that because x and y don't commute in general, expressions like $\log(xy) = \log(x) + \log(y)$ or $\exp(\mathbf{u} + \mathbf{v}) = \exp(\mathbf{u})\exp(\mathbf{v})$ are not valid.

Also note that the distances $\|\log(x^{-1}y)\| = \|\log(yx^{-1})\| = \|\log(y^{-1}x)\| = \|\log(xy^{-1})\|$ are all equal, but the vectors $\log(x^{-1}y)$, $\log(yx^{-1})$, $\log(y^{-1}x)$ and $\log(xy^{-1})$, in general are different. The vector in $\mathfrak{g}$ that is the log of the geodesic segment $x \rightarrow y$ is $\mathbf{dl}_{x\rightarrow y} = \log(yx^{-1})$. This is called left invariant metric, because it is invariant under left multiplication: $\mathbf{dl}_{zx\rightarrow zy} = \log((zx)^{-1}zy) = \log(x^{-1}z^{-1}zy) = \log(x^{-1}y) = \mathbf{dl}_{x\rightarrow y}$.

2.2. Principal Geodesic Analysis (PGA)

Once distance and angles are well defined in the manifold, mean and variance can be computed.

The Fréchet mean of a set of elements $\{x_i\}_{i=1}^N \in G$, is defined as the point $\mu \in G$ which minimizes the sum of squared distances $d^2(\mu, x_i)$:

$$\mu = \arg\min_{\mu \in G} \sum_{i=1}^{N} d^2(\mu, x_i). \tag{3}$$

This definition is inspired in the Euclidean mean definition. Unlike in Euclidean space, the mean in Riemannian manifolds may not exist, or be unique. Luckily, for localized enough data the Fréchet mean exists, is unique and can be computed iteratively as [15]:

$$\mu_k = \mu_{k-1} \exp\left(\frac{1}{N}\sum_{i=1}^{N} \log(\mu_{k-1}^{-1} x_i)\right). \tag{4}$$

The starting point was $\mu_0 = e$ in this work, but alternative initialization can be used [15].

Analogously, the Fréchet variance of a set of points is also defined analogously to the Euclidean case:

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} d^2(\mu, x_i). \tag{5}$$

Fletcher [11] defined Principal Geodesics (PG) as successive geodesics $\gamma_l(t) \in G$, $l = 1, 2, ...$, that passes through the mean ($\gamma_l(0) = \mu$), for which the data projected on them presents maximal variance, and are orthogonal to all the previous ones. In other words, PG are a set of orthogonal ordered geodesics such that the data projected on the first P ones maximize the variance for any P. The projection on a geodesic $\gamma(t) = \mu \exp(t\mathbf{v})$ is defined as

$$p_\gamma x = \arg\min_{y \in \gamma} d^2(x, y)^{[1]}. \tag{6}$$

The projection operator can not be computed with a closed formula, nevertheless an approximation can be done when x is in a small neighbourhood of the mean [11]. The approximation is the following: $\log(x^{-1}y) = \log(\tilde{x}^{-1}\tilde{y}) \approx \log(\tilde{y}) - \log(\tilde{x})$, where $\tilde{x}$ denote the residual $\tilde{x} = \mu^{-1}x$, $\forall x \in G$. With this approximation, and in terms of the residuals, the projection operator (6) becomes:

$$p_{\tilde{\gamma}}\tilde{x} \approx \arg\min_{\tilde{y} \in \tilde{\gamma}} \| \log(\tilde{x}) - \log(\tilde{y}) \|^2, \tag{7}$$

As y lies in the geodesic, it can be written as $y = \gamma(r) = \mu \exp(r\mathbf{v})$, for some r. And so, $\log(\tilde{y}) = \log(\mu^{-1}y) = r\mathbf{v}$, where $r \in \mathbb{R}$ and $\mathbf{v} \cdot \mathbf{v} = 1$. Finding the minimum in equation (7) is the same that finding $\arg\min_{r \in \mathbb{R}} \|\mathbf{u} - r\mathbf{v}\|^2$, with $\mathbf{u} = \log(\tilde{x})$. The minimum is obtained when $r = \mathbf{u} \cdot \mathbf{v}$, and finally equation (7) becomes:

$$p_{\tilde{\gamma}}\tilde{x} \approx \tilde{y} = \exp(r\mathbf{v}) = \exp\left((\mathbf{u} \cdot \mathbf{v})\mathbf{v}\right). \tag{8}$$

[1] Again, this may not exist, or be unique, if elements are not close enough to the mean.

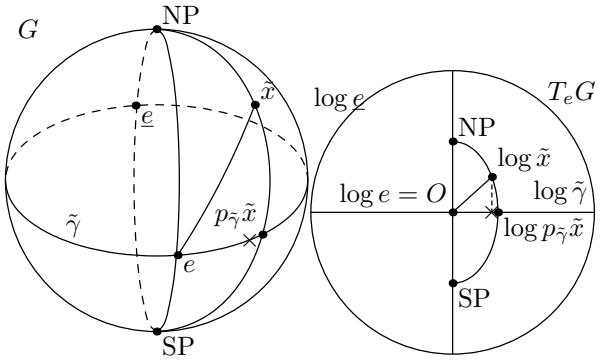

Figure 1. Schematic view of a group G and its tangent space at identity T_eG. The antipodal point of the identity, indicated as $\underline{e}$, is called cut locus of e, and it is the point where all the geodesics passing through e meet. All the points in T_eG for which their images by exponentiation are $\underline{e}$ are the circle of radius π ($\log \underline{e}$). Inside this circle, the exp function is a bijection. The points NP and SP are the north and south poles respectively, indicated only for clarity. $P_{\tilde{\gamma}}\tilde{x}$ is the projection of $\tilde{x}$ in the geodesic $\tilde{\gamma}$ (equator in the figure), and the approximation (8) is indicated with a $\times$.

Figure 1 shows schematically the error done with the approximation in (7) when the manifold is a 2-sphere in 3D. In general, the error tends to zero for data close enough to the identity, and for small angles (when $\tilde{x} \approx p_{\tilde{\gamma}}\tilde{x}$).

The amount of variance projected on $\tilde{\gamma}$ is

$$\sigma_{\tilde{\gamma}}^2 = \frac{1}{N}\sum_{i=1}^{N} d^2(e, p_{\tilde{\gamma}}\tilde{x}_i) = \frac{1}{N}\sum_{i=1}^{N} (\mathbf{u}_i \cdot \mathbf{v})^2, \tag{9}$$

with $\mathbf{u}_i = \log(\mu^{-1}x_i)$. Therefore finding the geodesics $\gamma_k(t)$ that maximizes projected variance is the same as finding vectors in $\mathfrak{g}$ that maximizes Euclidean projected variance. By computing usual Principal Component Analysis of the logarithm of residuals $\mathbf{u}_i$, a set of orthonormal principal components $\mathbf{v}_l \in \mathfrak{g}$ are obtained. The corresponding PG are given by $\gamma_l(t) = \mu \exp(t\mathbf{v}_l)$, and are also orthogonal at the mean. The projected variance on each geodesic is equal to the variance of the PCA modes.

2.3. Similarity group

The similarity group $Sim(3)$ is the set of similarity transformations in 3D. A similarity transformation $T(\cdot)$ applied to a 3D point is:

$$\mathbf{x}' = T(\mathbf{x}) = s\mathbf{R}\mathbf{x} + \mathbf{d}, \tag{10}$$

where $\mathbf{R} \in \mathrm{SO}(3)$ (3×3 orthogonal matrix with determinant one), $s \in \mathbb{R}^+$, and $\mathbf{x}', \mathbf{x}, \mathbf{d} \in \mathbb{R}^3$. The composition of two similarity transformations T_1 and T_2 is given by $T_3(\mathbf{x}) = T_2(T_1(\mathbf{x}))$, with the following parameters: $s_3 = s_2 s_1$, $\mathbf{R}_3 = \mathbf{R}_2\mathbf{R}_1$ and $\mathbf{d}_3 = s_2\mathbf{R}_2\mathbf{d}_1 + \mathbf{d}_2$.

A matrix representation of similarity transformations can be achieved by using projective geometry:

$$\mathbf{X} = (\mathbf{x}, 1) = (x, y, z, 1)^T. \tag{11}$$

The transformation is then written as $T(\mathbf{X}) = \mathbf{T}\mathbf{X}$, with:

$$\mathbf{T} = \begin{bmatrix} s\mathbf{R} & \mathbf{d} \\ \mathbf{0}^T & 1 \end{bmatrix}. \tag{12}$$

The set of matrices of the form (12) form a Lie group with matrix multiplication. With this matrix representation of the group, the composition of similarity transformations can be written as matrix multiplication ($\mathbf{T}_3 = \mathbf{T}_2\mathbf{T}_1$). Exponential and logarithm mappings are computed as standard matrix exponential and logarithm.

It can be easily checked that the following exponentials give the group elements that span $Sim(3)$:

$$\exp \begin{bmatrix} 0 & 0 & 0 & x \\ 0 & 0 & 0 & y \\ 0 & 0 & 0 & z \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & x \\ 0 & 1 & 0 & y \\ 0 & 0 & 1 & z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{13}$$

$$\exp \begin{bmatrix} \log s\mathbf{I}_3 & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} = \begin{bmatrix} s\mathbf{I}_3 & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{14}$$

$$\exp \begin{bmatrix} 0 & -r_z & r_y & 0 \\ r_z & 0 & -r_x & 0 \\ -r_y & r_x & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \tag{15}$$

where

$$\mathbf{R} = \exp(\mathbf{A}) = \begin{cases} \mathbf{I}, & if\ a = 0, \\ \mathbf{I} + \frac{sin(a)}{a}\mathbf{A} + \frac{1-cos(a)}{a^2}\mathbf{A}^2, & if\ a \neq 0, \end{cases} \tag{16}$$

$a = \sqrt{r_x^2 + r_y^2 + r_z^2}$ and $\mathbf{A}$ is the 3×3 upper left submatrix of left matrix in (15). The rotation parameters r_x, r_y and r_z (components of the skew symmetric matrix $\mathbf{A}$) define the rotation axis by means of $\mathbf{r} = (r_x, r_y, r_z)$, and the rotation angle as $\phi = \sqrt{\mathbf{r} \cdot \mathbf{r}}$.

The complete set of parameters that defines a similarity transformation in the tangent space $\mathfrak{sim}(3)$ is

$$\mathbf{b} = [r_x\ r_y\ r_z\ x\ y\ z\ l]^T, \tag{17}$$

with $l = \log s$.

The set of 4×4 matrices $\{\mathbf{B}_p\}_{p=1}^7$, with components:

$$\begin{aligned} [\mathbf{B}_1]_{3,2} = [\mathbf{B}_2]_{1,3} = [\mathbf{B}_3]_{2,1} &= 1, \\ [\mathbf{B}_1]_{2,3} = [\mathbf{B}_2]_{3,1} = [\mathbf{B}_3]_{1,2} &= -1, \\ [\mathbf{B}_4]_{1,4} = [\mathbf{B}_5]_{2,4} = [\mathbf{B}_6]_{3,4} &= 1, \\ [\mathbf{B}_7]_{1,1} = [\mathbf{B}_7]_{2,2} = [\mathbf{B}_7]_{3,3} &= 1 \\ \text{otherwise } 0, & \end{aligned} \tag{18}$$

form an orthogonal basis of the Lie algebra $\mathfrak{sim}(3)$, with inner product $(\mathbf{C} \cdot \mathbf{D}) = \mathrm{tr}(\mathbf{C}\mathbf{D}^T)$. $[\mathbf{A}]_{i,j}$ denotes the ith and jth component of matrix $\mathbf{A}$.

Any similarity transformation can be written as:

$$\mathbf{T}(\mathbf{b}) = \exp(\sum_p [\mathbf{b}]_p \mathbf{B}_p), \tag{19}$$

and any geodesic as $\mathbf{T}(t\mathbf{b})$, for $\mathbf{b}$ fixed, and t running in $\mathbb{R}$.

2.4. Pose model

The similarity transformation that takes an object from a reference frame pose A to a pose B is well defined and usually easy to compute from the object representation. But if the objects in A and B are different, we need to define when two object has the same pose.

If correspondence between the points on the surface of different objects are given, we define the similarity transformation that takes an object S_i to the same pose of a different object S_j as:

$$T_{i \to j} = \arg \min_{T \in Sim(3)} D^2(S_j, T(S_i)), \tag{20}$$

where D^2 is the sum (integral in the case of continuous shape representation) of square distances from corresponding points. For point distribution models, transformation (20) can be computed with a closed form formula [14].

In order to build a statistical pose model, a template shape is required. The template shape was selected as the mean shape, defined as:

$$R = \arg \min_{S \in \text{Shape space}} \sum_{i=1}^{N} D_p^2(S_i, S), \tag{21}$$

where D_p^2 is the squared Procrustes distance, *i.e.* the distance D^2 after Procrustes alignment:

$$D_p^2(S_i, S) = \min_{T \in Sim(3)} D^2(S_i, T(S)). \tag{22}$$

The pose of each shape S_i is characterized by the similarity transformation T_i that best fit mean shape R into S_i:

$$T_i = \arg \min_{T \in Sim(3)} D^2(S_i, T(R)), \tag{23}$$

as shown in Figure 2. Finally, the pose model is obtained by applying PGA defined before to the set of T_i.

It is worthy to note that according to (21) the pose of the reference shape R can be chosen arbitrarily. The orientation and location of R have no consequences on PGA, but the size of R will control the relative magnitude between translation and rotation of residuals. It is well known that it is important to normalize multivariate data in order to obtain all parameters in commensurable magnitudes before computing SVD. By setting the mean radius of R to unity, boundary displacements driven by rotations of R will be equal to rotation angle in radian units.

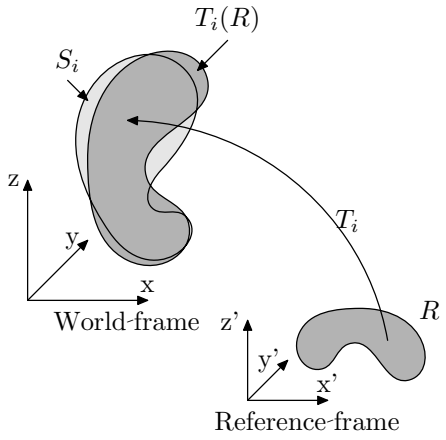

Figure 2. Transformation T_i takes the mean shape R from the reference frame to shape S_i in world frame.

3. Multi-object pose model (MOPM)

Given a data set consisting of N instances of complex systems with L objects for each instance $S_{i,l}$, the set of NL pose transformations that takes the L reference objects R_l to the instances $S_{i,l}$ are given by

$$T_{i,l} = \arg \min_{T \in Sim(3)} D^2(S_{i,l}, T(R_l)). \quad (24)$$

$\mathbf{T}_{i,l}$ denotes the matrix representation of $T_{i,l}$.

As the similarity transformations commute among different objects, the group that describes the whole set of transformations $\mathbf{T}_{i,l}$ is the direct product of the L similarity groups: $Sim(3)^L = \prod_{i=1}^{L} Sim(3) = Sim(3) \times Sim(3) \times ... \times Sim(3)$. Its Lie algebra is given by the $7L$ dimensional Euclidean space consisting of L copies of $\mathfrak{sim}(3)$.

The logarithm of an element $T_i = (\mathbf{T}_{i,1}, \mathbf{T}_{i,2}, .., \mathbf{T}_{i,L}) \in Sim(3)^L$ is $\log T_i = (\log \mathbf{T}_{i,1}, \log \mathbf{T}_{i,2}, \cdots, \log \mathbf{T}_{i,L})$. The exponential of an element $U_i = (\mathbf{U}_{i,1}, \mathbf{U}_{i,2}, \cdots, \mathbf{U}_{i,L}) \in \mathfrak{sim}(3)^L$ is $\exp U_i = (\exp \mathbf{U}_{i,1}, \exp \mathbf{U}_{i,2}, \cdots, \exp \mathbf{U}_{i,L})$. Similarly, products and inversions are performed component wise.

PGA in $Sim(3)^L$ is computed in an analogous way to PGA in $Sim(3)$: the multi-object pose mean is $M = (\mathbf{M}_1, \mathbf{M}_2, .., \mathbf{M}_L)$, being $\mathbf{M}_l$ the mean of $\mathbf{T}_{i,l}$ (averaging in $i = 1, .., N$); the residuals U_i are given by

$$U_i = \log(M^{-1} T_i). \quad (25)$$

In terms of the orthonormal basis $\mathbf{B}_p$ in (18), residuals can be written as the $7L$ dimensional vector $\mathbf{u}_i = (\mathbf{u}_{i,1}^T, \mathbf{u}_{i,2}^T \cdots \mathbf{u}_{i,L}^T)^T$ where

$$\mathbf{u}_{i,l} = (\mathbf{U}_{i,l} \cdot \mathbf{B}_1, \mathbf{U}_{i,l} \cdot \mathbf{B}_2, \cdots, \mathbf{U}_{i,l} \cdot \mathbf{B}_7)^T. \quad (26)$$

The residuals of each structure $\mathbf{u}_{i,l}$ can be weighted with positive scalars α_l in order to emphasize the relative importance of some objects.

SVD analysis of the residual matrix formed by $\mathbf{u}_i$ provides the direction of the principal geodesics $\mathbf{v}_k = (\mathbf{v}_{k,l}^T, \mathbf{v}_{k,2}^T, \cdots, \mathbf{v}_{k,L}^T)^T$ for each mode k as well as the corresponding singular values σ_k.

The principal geodesics are given by:

$$G_k(t) = (\mathbf{M}_1 \exp(t\mathbf{V}_{k,1}/\alpha_1), \mathbf{M}_2 \exp(t\mathbf{V}_{k,2}/\alpha_2), ... \mathbf{M}_L \exp(t\mathbf{V}_{k,L}/\alpha_L)), \quad (27)$$

being $\mathbf{V}_{k,l} = \sum_p [\mathbf{v}_{k,l}]_p \mathbf{B}_p$

4. Experiments and results

4.1. Data set and preprocessing

Experiments were performed on a data set of N=18 brain MRI studies from normal subjects available at Internet Brain Segmentation Repository [1] from Massachusetts General Hospital. Seven brain structures from both hemispheres (L=14 objects in total) were selected for this study: lateral ventricle, thalamus proper, caudate nucleus, putamen, pallidum, Hhppocampus and amygdala.

For each object, a 3D point distribution model (PDM) with point correspondence was built from the binary segmentation images available at IBSR according to the following procedure. A dense mesh surface representation was obtained by thresholding an isosurface representation of the binary image. Point correspondence was estimated by means of an iterative non-rigid registration procedure applied to a template structure based on the Robust Point Matching (RPM) algorithm [4]. The initial template was a cloud of points uniformly distributed on a sphere, where the number of points were proportional to the mean surface area of each object.

After the PDM representation, the mean shape of each object was obtained as an iterative process of Procrustes alignment (22) and point averaging. This procedure provide the pose transformation for each instance and object (24). Fig. 3 illustrates the mean of the PDM for each structure.

4.2. Multi object pose model of brain structures

Similarity transformations defined by the MOPM consists of a mean pose and a set of pose variation modes. The mean pose of each object l was iteratively computed following (4), where the generic group element x_i was replaced by T_{il}. The mean pose is illustrated in Fig. 3.

Pose modes were computed by means of SVD of the residual matrix formed by the set of vectors in (25), where the scaling factors α_l were selected proportional to the mean volume. The pose modes are illustrated in Fig. 4-6 in the interval $(-3\,\sigma, 3\,\sigma)$. This range is a bit exaggerated in order to

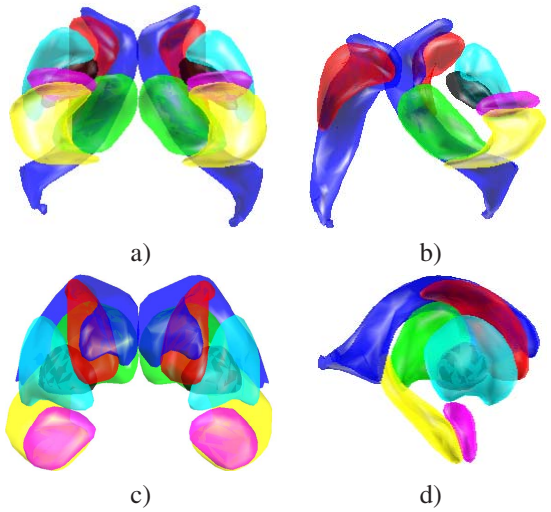

Figure 3. Brain structures: lateral ventricle; thalamus proper; caudate; putamen; pallidum; hippocampus; amygdala. a) bottom view; c) front view; b) oblique view (only left hemisphere structures plus right ventricle and caudate nucleus are shown); d) lateral view.

best visualize variations. Brain structures have been packed into three groups for a better visualization: lateral ventricle and caudate nucleus in group A; thalamus, putamen and pallidum in group B; and hippocampus and amygdala in group C.

Panels a), c) and d) in Figs. 4-6 illustrate the effect of pose variations spanned by the first mode on the mean shape. Solid surfaces represent mean shape transformed by $G_1(3\sigma_1)$ in (27) and transparent surfaces by $G_1(-3\sigma_1)$ The color (RGB components) of the surface for each object represents the relative importance of rotation (red), translation (green) and scaling (blue) parameters.

Panels b) show the values of the 7 pose parameters corresponding to the arguments of the exponential functions in (27), *i.e.* $3\sigma_1 \mathbf{v}_{1,l}/\alpha_l$ for each l-th object. Solid lines represent left hemisphere objects, while dashed lines refer to right hemisphere objects. The coordinate axis were oriented as follows: x axis was set to right-left orientation; y axis from front to back and z axis as down-up orientation. In this plot, rotation parameters in y and z direction, and translation parameter in x direction, were sign reversed for right objects. In such way, the paired lines in panel b) corresponding to transformations with hemispherical symmetry will be similar.

An additional representation of the seven pose parameters is illustrated by means of a set of spheres, vectors and segments located at the center of mass of each structure. The magnitude of scale parameter $|l| = |\log s|$ is represented as the sphere radius and the sign as sphere color: yellow for positive scalings or dilations and red color for contractions. The remaining 6 pose parameters are repre-

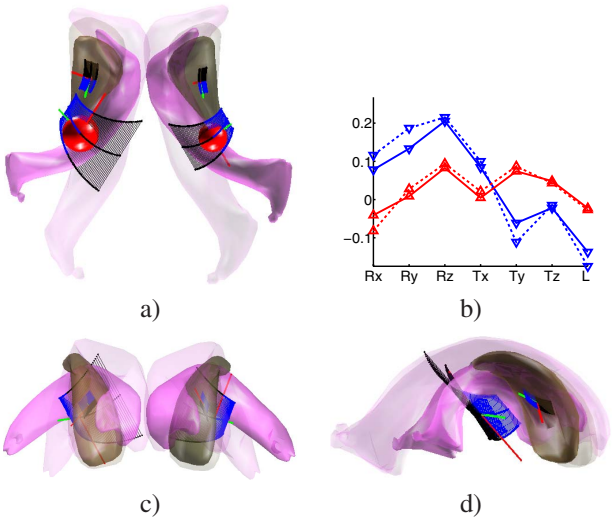

Figure 4. First pose mode of structures in group A: lateral ventricle and caudate nucleus. a) Bottom view. b) Pose parameters. c) Front view. d) Lateral view. More details are given in the text.

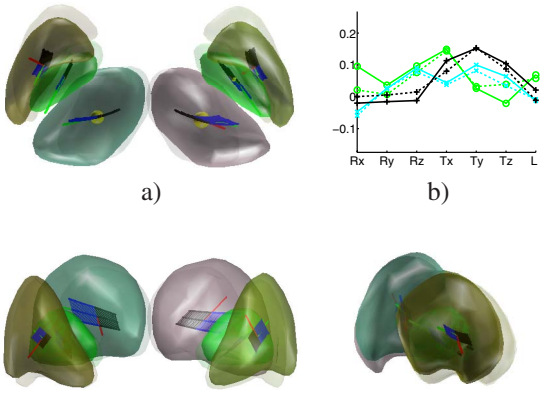

Figure 5. First pose mode of structures in group B: thalamus proper,putamen, pallidum. a) Bottom view. b) Pose parameters. c) Front view. d) Lateral view. More details are given in the text.

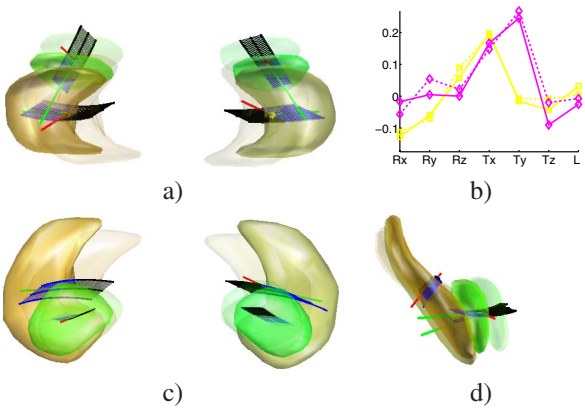

Figure 6. First pose mode of structures in group C:hippocampus; amygdala. a) Bottom view. b) Pose parameters. c) Front view. d) Lateral view. More details are given in the text.

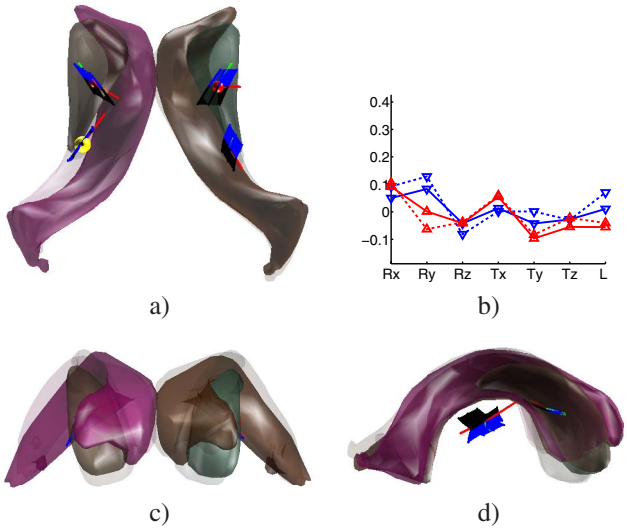

Figure 7. Second pose mode of structures in group A: lateral Ventricle and caudate nucleus. a) Bottom view. b) Pose parameters. c) Front view. d) lateral view. More details are given in the text.

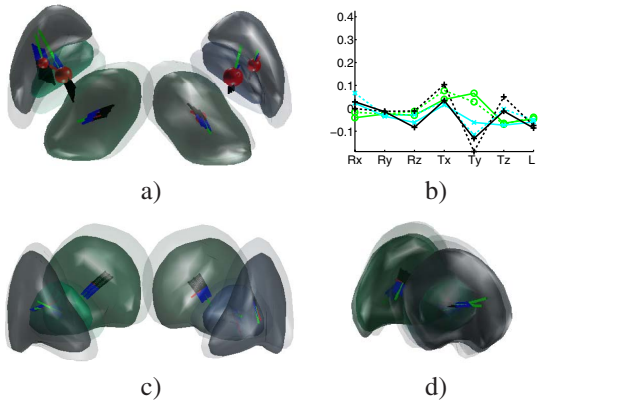

Figure 8. Second pose mode of structures in group B: thalamus proper, putamen, pallidum. a) Bottom view. b) Pose parameters. c) Front view. d) Lateral view. More details are given in the text.

sented by vectors. Red color vectors are use for rotation parameters, while green vectors for translation. Vector direction denotes either rotation axis or translation direction. Vector magnitude denote either rotation angle or translation magnitude. Additionally a blue and black band is generated by applying the geodesic transformation to a segment located at the center of mass. The segment is oriented along the rotation axis and its length is equal to the rotation magnitude. The geodesic runs from -6σ (black color) up to 6σ (blue color) for illustration purposes.

It is worthy to note that the orientation of the pose mode is just a matter of convention, and therefore what is relevant is the relative signs and magnitude among different pose parameters.

The second mode is illustrated in Fig. 7– 9.

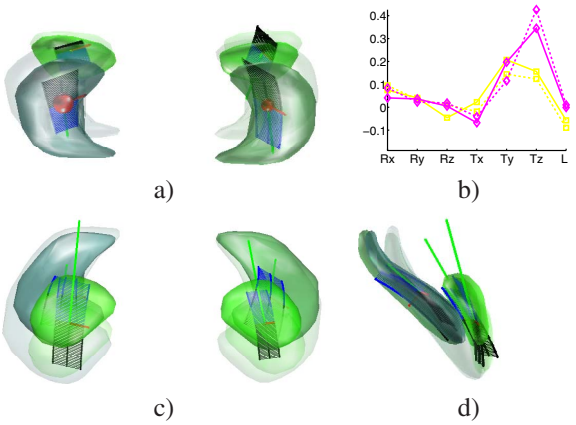

Figure 9. Second pose mode of structures in group C:hippocampus; amygdala. a) Bottom view. b) Pose parameters. c) Front view. d) Lateral view. More details are given in the text.

4.3. Discussion

The dominant modes of the MOPM explain most of the pose variance within the training set. The first mode shows an important hemispheric symmetry on most of the subcortical nuclei and lateral ventricles. This is clearly seen on panel b) of Fig. 4-6, where paired solid and dashed lines are quite similar. Accordingly, surface colors from paired objects in panels a) c) and d) are also similar.

The objects with the largest displacements are the lateral ventricles (see displacement of solid and transparent surfaces in panels a), c) and d) as well as the magnitude of lateral ventricles pose parameters in panel b) of Fig. 4). Although the pose parameters of the amygdala (magenta lines in panel b) of Fig. 6) have similarly large values, the amygdala displacement is relatively small (see panels a) c) and d) in Fig. 6) because pose parameters in panel b) must be weighted by the amygdala volume α_l.

Lateral ventricles also experience an important contraction while the rest of the structures remain mainly unscaled. It can be also seen that some structures have a correlated displacement (such as the hippocampi tails, posterior part of the thalami and lateral ventricles tails), consisting of a strong lateral displacement getting farther away from midsagittal plane, and a slight forward and upward displacement.

In addition, the frontal part of most nuclei (head of hippocampi, head of caudate nuclei, front part of putamen and pallidum) and lateral ventricles experience a backward displacement.

The second mode explains a much more slight displacement of the lateral ventricles than in the first mode. This time the displacement is larger in the right ventricle, pro-

viding an important asymmetry component. In contrast, the remaining structures have a symmetric behavior.

The body of the right lateral ventricle moves to the right lateral side at the same time that a slight dilation.

On the other hand, the structures in Group B moves symmetrically to the lateral sides. Both thalami move backwards while putamen and pallidum forward. The whole displacement can be explained as a radial movement outwards at the same time as a structure contraction. Therefore, the surrounding white matter enlarges.

The tail of the hippocampi move backwards, in a similar way than thalami, while the head of the hippocampi contracts. The amygdalae follow the movement of the heads of the hippocampi and remain unscaled.

It should be kept in mind that all the pose information provided by the MOPM will depend on the accuracy of the similarity transformations estimated from the training set. The two most relevant error sources are segmentation errors and correspondence estimation. While the former will be significantly larger in some structures with low contrast such as hippocampi, the latter will mainly affect to structures with a large anatomical variability, such as the lateral ventricles.

5. Conclusion

In this work a methodology to build statistical models of similarity transformations was proposed. As similarity transformations does not live in an Euclidean space, distances, angles, and statistics (namely averages and modes of variation) were defined within the Lie group framework.

This methodology provides the required tools in order to introduce multi-object pose models (MOPM). These statistical models learn from a training set of data previously segmented, mean and variation modes of the relative pose among different objects belonging to a multi-object system.

We expect that MOPM will provide a valuable *a priori* information about relative location, orientation and scale of a set of objects, which may be of great help in medical image segmentation and registration. In addition pose parameters may be useful as neuroanatomical markers of neurological diseases.

References

[1] IBSR 2004, Internet Brain Segmentation Repository (IBSR). http://www.cma.mgh.harvard.edu/ibsr/, 2005. 5

[2] J. Ashburner, J. G. Csernansk, C. Davatzikos, N. C. Fox, G. B. Frisoni, and P. M. Thompson. Computer-assisted imaging to assess brain structure in healthy and diseased brains. *The Lancet Neurology*, 2(2):79–88, Feb. 2003. 1

[3] M. Buchsbaum, T. Someya, C. Teng, L. Abel, S. Chin, A. Najafi, R. Haier, J. Wu, and J. Bunney, WE. PET and MRI of the thalamus in never-medicated patients with schizophrenia. *Am J Psychiatry*, 153(2):191–199, Feb. 1996. 1

[4] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Comput. Vis. Image Underst.*, 89(2-3):114–141, 2003. 5

[5] J. G. Csernansky, M. K. Schindler, N. R. Splinter, L. Wang, M. Gado, L. D. Selemon, D. Rastogi-Cruz, J. A. Posener, P. A. Thompson, and M. I. Miller. Abnormalities of thalamic volume and shape in schizophrenia. *Am J Psychiatry*, 161(5):896–902, 2004. 1

[6] J. G. Csernansky, L. Wang, S. C. Joshi, J. Tilak Ratnanather, and M. I. Miller. Computational anatomy and neuropsychiatric disease: probabilistic assessment of variation and statistical inference of group difference, hemispheric asymmetry, and time-dependent change. *NeuroImage*, 23(Supplement 1):S56–S68, 2004. 1

[7] A. T. Du, N. Schuff, X. P. Zhu, W. J. Jagust, B. L. Miller, B. R. Reed, J. H. Kramer, D. Mungas, K. Yaffe, H. C. Chui, and M. W. Weiner. Atrophy rates of entorhinal cortex in AD and normal aging. *Neurology*, 60(3):481–486, 2003. 1

[8] A. C. Evans, S. Marrett, J. Torrescorzo, S. Ku, and L. Collins. MRI-PET correlation in three dimensions using a volume-of-interest (VOI) atlas. *J Cerebr Blood F Met*, 11(2):A69–A78, Mar 1991. 1

[9] P. T. Fletcher and S. C. Joshi. Principal geodesic analysis on symmetric spaces: Statistics of diffusion tensors. In M. Sonka, I. A. Kakadiaris, and J. Kybic, editors, *ECCV Workshops CVAMIA and MMBIA*, volume 3117 of *Lect Notes Comput Sc*, pages 87–98. Springer, 2004. 2

[10] P. T. Fletcher, S. C. Joshi, C. Lu, and S. M. Pizer. Gaussian distributions on Lie groups and their application to statistical shape analysis. In C. J. Taylor and J. A. Noble, editors, *IPMI*, volume 2732 of *Lect Notes Comput Sc*, pages 450–462. Springer, 2003. 2

[11] P. T. Fletcher, C. Lu, and S. C. Joshi. Statistics of shape via principal geodesic analysis on Lie groups. In *CVPR (1)*, pages 95–101. IEEE Computer Society, 2003. 2, 3

[12] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Trans. Med. Imaging*, 23(8):995–1005, 2004. 2

[13] T. Fletcher. *Statistical Variability in Nonlinear Spaces: Application to Shape Analysis and DT-MRI*. PhD thesis, Department of Computer Science, University of North Carolina, Aug 2004. 2

[14] B. K. P. Horn. Closed-form solution of absolute orentation using unit quaternions. *JOSA*, 4(4):629–642, Apr. 1987. 4

[15] X. Pennec. Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements. In A. Cetin, L. Akarun, A. Ertuzun, M. Gurcan, and Y. Yardimci, editors, *Proc. of NSIP 1999*, volume 1, pages 194–198, June 20-23, Antalya, Turkey, 1999. IEEE-EURASIP. 2, 3

[16] X. Pennec and N. Ayache. Uniform distribution, distance and expectation problems for geometric features processing. *J of Mathl Imaging and Vis*, 9(1):49–67, July 1998. 2

[17] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *Int J Comput Vis*, 66(1):41–66, January 2006. 2