

MULTI-OBJECT STATISTICAL *POSE+SHAPE* MODELS

M. N. Bossa and S. Olmos

GTC, I3A, Zaragoza University, Spain

ABSTRACT

Region of interest (ROI) analysis is a very common procedure for morphometry studies of brain structures, where each structure is usually isolated from the rest of the brain and aligned to a reference shape. In the alignment process all pose information is disregarded. However, considering the brain as a multi-object system formed by several structures, the relative pose among different structures may provide clinically relevant information. A methodology to build multi-object statistical *pose+shape* models is given in this work. The pose features for each structure are given by the parameters of a similarity transformation and the shape features are given by the coordinates of corresponding landmarks on the boundary. As pose and shape features do not live in an Euclidean vector space but in a Riemannian manifold, the methodology is based on performing standard multivariate statistical tools (such as PCA) on the tangent space. Experimental results are performed on brain structures such as the subcortical nuclei (caudate nucleus, hippocampus, amygdala, thalamus, putamen, pallidum) and lateral ventricles.

1. INTRODUCTION

Statistical shape analysis is being increasingly used to characterize brain anatomy either in control subjects as well as abnormalities of brain structures in patients with neuropsychiatric disorders [1]. To name a few, volumetry and morphometry studies of hippocampus and amygdala in Alzheimer's disease patients [2], volumetry and morphometry analysis of the thalamus in schizophrenic patients [3].

Statistical shape analysis is an emerging field with many applications on medical imaging and computer vision. Principal Component Analysis (PCA) is one of the most common procedures for rank reduction and statistical analysis. This technique should be only applied to multivariate data lying on an Euclidean space. However, most of pose and shape features live in a Riemannian manifold, and PCA should be replaced by its counterpart Principal Geodesic Analysis (PGA).

The main difficulty to be addressed in the analysis of human brain is that several 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 after alignment to a common coordinate system. In this way, geometrical relations among neighbor structures can not be properly analyzed. In order to overcome this limitation, multi-object shape models have been recently proposed in the literature. The first and straightforward proposal to build a joint statistical shape model of a set of objects was to concatenate shape features in a long vector, and to apply standard multivariate statistical tools. Some applications have been done in heart modeling using point sets as shape features [4]. In a recent work, the shape of several subcortical nuclei was characterized by m-reps and PGA was performed on the multi-object feature vector [5]. Later, a methodology to build multi-object pose models (MOPM) was introduced in [6], which can be applied in combination with any shape characterization. Very recently, a joint *pose+shape* model was proposed in [7], where m-reps was used as shape descriptors and pose was characterized in a different way to the proposal given in [6].

The aim of this work is to build multiobject *pose+shape* models using boundary point sets as shape features. This work is an extension of the multiobject pose model introduced in [6] with shape features. In this work shape is described by means of point distribution models.

2. JOINT *POSE+SHAPE* MODEL

Our proposal for joint modeling *pose+shape* consists on concatenating pose and shape features in a joint feature vector $\mathbf{f} = [w^p \mathbf{f}^p \quad w^s \mathbf{f}^s]^T$, where \mathbf{f}^p and \mathbf{f}^s denote pose and shape feature vectors respectively and $w^p, w^s \in \mathbb{R}^+$ are weighting factors that balance the relative importance between shape and pose within the model. Very often these features live in a Riemannian manifold, G , and their description at the tangent space is required in order to be able to compute statistics.

2.1. Principal Geodesic Analysis (PGA)

PGA is based on computing statistics on the tangent space at the mean, $T_\mu G$, which is an Euclidean space of the same dimension as the manifold, G , and it provides a local approximation to the manifold. The log operator provides the map-

This work was partially funded by research grants TEC2005-07801-C03-02, TEC2006-13966-C03-02, FIS PI04/1795 from Spain. M. Bossa work was funded by DGA under the FPI grant B097/2004.

ping between the manifold and the tangent space with two important properties: distances and angles on the manifold can be computed on the tangent space. The distance between any two elements $x, y \in G$ is:

$$d(x, y) = \|\log_x(y)\|. \quad (1)$$

In a similar way, the angle between geodesics in the manifold passing through the mean can be computed as the angle between their corresponding initial vectors in $T_\mu G$. The exponential mapping is the inverse operator that takes Euclidean points from the tangent space to the manifold.

The Fréchet mean of a set of elements $\{x_i\}_{i=1}^N \in G$, can be computed iteratively as [8]:

$$\mu_k = \exp_{\mu_{k-1}} \left(\frac{1}{N} \sum_i \log_{\mu_{k-1}}(x_i) \right). \quad (2)$$

A set of orthonormal principal vectors \mathbf{v}_k are obtained by standard Principal Component Analysis of the logarithm $\mathbf{u}_i = \log_\mu(x_i)$. The corresponding principal geodesics are given by $\gamma_k(t) = \exp_\mu(t \mathbf{v}_k)$, which are also orthogonal at the mean.

2.2. Pose features \mathbf{f}^p

A more detailed description of pose features was done in [6] and it is only reviewed here for brevity reasons.

Given the configuration matrices (landmark coordinates) of an object \mathbf{S}_i and a reference shape \mathbf{R} with known correspondence, pose is defined by the similarity transformation that minimizes the distance

$$T_i = \arg \min_{T \in Sim(3)} D^2(\mathbf{S}_i, T(\mathbf{R})), \quad (3)$$

where D^2 is the sum of square distances between corresponding points and $Sim(3)$ is the Lie group of similarity transformations.

A general similarity transformation T of a 3D point \mathbf{x} is:

$$\mathbf{x}' = T(\mathbf{x}) = s\mathbf{R}\mathbf{x} + \mathbf{d}, \quad (4)$$

where $\mathbf{R} \in SO(3)$ (3×3 orthogonal matrix with determinant one), $s \in \mathbb{R}^+$, and $\mathbf{x}', \mathbf{x}, \mathbf{d} \in \mathbb{R}^3$. A matrix representation of $Sim(3)$ is given by

$$\mathbf{T} = \begin{bmatrix} s\mathbf{R} & \mathbf{d} \\ \mathbf{0}^T & 1 \end{bmatrix}. \quad (5)$$

T operates on a 3D point in the following way $[\mathbf{x}'; 1] = \mathbf{T}[\mathbf{x}; 1]$. The log and exp mappings are computed as the standard log and exp functions of matrices. The tangent space

representation at the identity has the following form¹

$$\log \mathbf{T} = \begin{bmatrix} l & -r_z & r_y & x \\ r_z & l & -r_x & y \\ -r_y & -r_x & l & z \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad (6)$$

Therefore, the pose of a given object is characterized by the pose feature vector \mathbf{f}^p formed by the seven free parameters in $\log \mathbf{T}$, i.e., $\mathbf{f}^p = [x \ y \ z \ r_x \ r_y \ r_z \ l] \in \mathbb{R}^7$, where $l = \log s$. $\exp(\mathbf{f}^p)$ will denote the matrix obtained by exponentiating the matrix representation of \mathbf{f}^p described in the right hand side of (6), and $\exp_M(\mathbf{f}^p) = \mathbf{M} \exp(\mathbf{f}^p)$.

The relative magnitude of the different pose parameters (rotation, translation and scaling) is sensitive to the norm of the reference shape \mathbf{R} . This value, was selected in order to get commensurable units and more details can be found in [6].

Given a data set with I instances of complex systems formed by J objects $\{\mathbf{S}_{i,j}\}_{i=1, j=1}^{I, J}$, the pose transformation $T_{i,j}$ that takes the j -th reference object \mathbf{R}_j to the j -th object from the i -th instance, $\mathbf{S}_{i,j}$, is given by

$$T_{i,j} = \arg \min_{T \in Sim(3)} D^2(\mathbf{S}_{i,j}, T(\mathbf{R}_j)). \quad (7)$$

As the similarity transformations commute among different objects, the Lie group that describes the whole set of transformations is the direct product of the J similarity groups:

$$Sim(3)^J = \prod_{i=1}^J Sim(3) = Sim(3) \times Sim(3) \times \dots \times Sim(3).$$

The feature vector $\mathbf{f}_{i,j}^p$ is extracted from the logarithm of the residual transformations, $\log((\mathbf{M}_j^p)^{-1} \mathbf{T}_{i,j})$ and the mean transformation \mathbf{M}_j^p is obtained using (2). The multi-object pose feature vector is $\mathbf{f}_i^p = [\alpha_1 \mathbf{f}_{i,1}^p \ \alpha_2 \mathbf{f}_{i,2}^p \ \dots \ \alpha_J \mathbf{f}_{i,J}^p]^T \in \mathbb{R}^{7J}$, where α_j are scaling factors that take into account the relative size of each object and are computed as the geometric mean of the s_{ij} running for all instances.

2.3. Shape features \mathbf{f}^s

In this work shape is characterized by configuration matrices $\mathbf{S} \in \mathbb{R}^{3,n}$, or equivalently by its corresponding configuration vector $\mathbf{s} = \text{vec}(\mathbf{S})$. The set of configuration matrices that are Procrustes aligned to a reference shape \mathbf{r} , $\tilde{\mathbf{s}} = T^{-1}(\mathbf{s})$ forms the Procrustes shape space, denoted as Σ_n^3 , being n the number of landmarks. It has been shown in [9] that Σ_n^3 behaves locally as a Riemannian manifold and its exponential mapping is given by²

$$\tilde{\mathbf{s}} = \exp_{\mathbf{r}}(\mathbf{h}) = \cos\left(\frac{\|\mathbf{h}\|}{\|\mathbf{r}\|}\right) \mathbf{r} + \sin\left(\frac{\|\mathbf{h}\|}{\|\mathbf{r}\|}\right) \frac{\|\mathbf{r}\|}{\|\mathbf{h}\|} \mathbf{h} \quad (8)$$

¹Note that $[x \ y \ z]^T$ in the 4-th column in the r.h.s of (6) is different from \mathbf{d} in (4); actually, $[x \ y \ z]^T$ depends on \mathbf{R} , s and \mathbf{d} .

²In this work Σ_n^3 is a sphere of radius $\|\mathbf{r}\|$, while a unity radius was used in [9] without loss of generality.

where \mathbf{h} lies in the tangent space $T_{\mathbf{r}}\Sigma_n^3$ and $\|\tilde{\mathbf{s}}\| = \|\mathbf{r}\|$. Accordingly, \mathbf{h} is a $3n \times 1$ vector with seven linear constraints that only depend on the reference shape \mathbf{r} : $\mathbf{h}^T \mathbf{r} = 0$ and six extra linear constraints that takes into account translation and rotation invariance of the shape. Even though the dimension of $T_{\mathbf{r}}\Sigma_n^3$ is $3n - 7$, the number of instances I is usually the limiting factor of the number of degrees of freedom and a smaller dimension subspace can be computed from the observed data after pose alignment $\mathbf{h}_i = \log_{\mathbf{r}}(\tilde{\mathbf{s}}_i)$. $\mathbf{P} \in \mathbb{R}^{\min(3n-7, I), 3n}$ will denote an orthonormal basis of the subspace spanned by $\{\mathbf{h}_i\}$. In order to compute statistics, the reference \mathbf{r} is often selected as the mean shape, \mathbf{m}^s , which is iteratively computed using (2) and (3). The shape feature vector of a single object will be

$$\mathbf{f}^s = \mathbf{h}^T \mathbf{P}^T \in \mathbb{R}^{\min(3n-7, I)}. \quad (9)$$

The multi-object residual shape feature vector is given by $\mathbf{f}_i^s = [\alpha_1 \mathbf{f}_{i,1}^s \ \alpha_2 \mathbf{f}_{i,2}^s \ \cdots \ \alpha_J \mathbf{f}_{i,J}^s]^T$ where the scaling factors α_j are the same as for the pose feature vector. Note that each object may have a different number of landmarks n_j and their corresponding $\mathbf{f}_{i,j}$ might have different dimension in the case of $n_j < I$.

2.4. Joint statistical pose+shape model

After extracting pose and shape features from all instances and objects in the training set, any multivariate statistical technique can be used on the set of features

$$\mathbf{F} = [w^p [\mathbf{f}_1^p \ \mathbf{f}_2^p \ \cdots \ \mathbf{f}_I^p]; \quad w^s [\mathbf{f}_1^s \ \mathbf{f}_2^s \ \cdots \ \mathbf{f}_I^s]].$$

The weighting factor $w^p(w^s)$ was chosen such that the energy of pose(shape) features is equal to the corresponding pose(shape)-driven boundary displacement energy.

The most common multivariate statistical technique is PCA. The principal vectors of \mathbf{F} are denoted as \mathbf{v}_k , and can be object-wise splitted as

$$\mathbf{v}_k^T = \left[\left(\mathbf{v}_{k,1}^p \right)^T \cdots \left(\mathbf{v}_{k,J}^p \right)^T \quad \left(\mathbf{v}_{k,1}^s \right)^T \cdots \left(\mathbf{v}_{k,J}^s \right)^T \right]. \quad (10)$$

The principal geodesics are also splitted into the J objects. The j -th object generated by the k -th principal geodesic is given by

$$\Phi_{k,j} \left(\text{vec}^{-1} \left(\exp_{\mathbf{m}_j^s} \left(t \mathbf{P}_j^T \mathbf{v}_{k,j}^s / \alpha_j / w^s \right) \right); t \right), \quad (11)$$

where $\Phi_{k,j}(\cdot; t)$ denotes the similarity transformation defined by the matrix $\mathbf{M}_j^p \exp \left(t \mathbf{v}_{k,j}^p / \alpha_j / w^p \right)$ and $t \in \mathbb{R}$. Shape instances are first generated by means of the exponential map of the shape principal vectors and later on transformed by the corresponding principal pose.

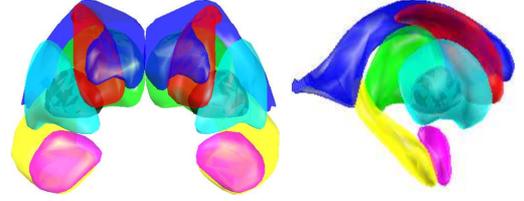


Fig. 1. Brain structures: Left) front view; Right) lateral view.

3. RESULTS

3.1. Data set and preprocessing

Experiments were performed on a data set of $I=18$ brain MRI studies from normal subjects from Internet Brain Segmentation Repository [10]. Seven brain structures from both hemispheres ($J=14$ objects in total) were selected for this study: lateral ventricle, thalamus, caudate nucleus, putamen, pallidum, hippocampus and amygdala. The total number of landmarks, $\sum n_j$, was 10459.

The preprocessing (global alignment and correspondence estimation) was done as in [6]. In a nutshell, Procrustes alignment and correspondence estimation is performed independently for each object by means of an iterative non-rigid registration procedure applied to a template structure based on the Robust Point Matching (RPM) algorithm [11]. The mean shape of each structure with its corresponding mean pose is illustrated in Fig. 1.

3.2. Experiments

The first experiment was to compute the relative importance of pose and shape within our training set. The total energy (TE) is defined as the variance of corresponding points after global alignment. The residual variance when each structure is independently aligned to its mean shape at the mean pose is named shape energy (SE). The pose energy is defined as the difference $PE = TE - SE$. Fig. 2 shows the ratio PE/TE for each structure.

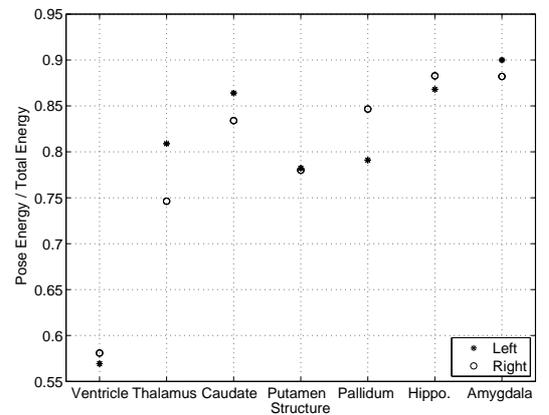


Fig. 2. Relative pose energy in the training set

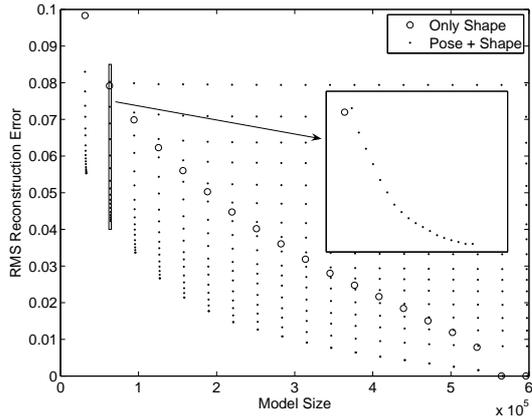


Fig. 3. Reconstruction error versus model size.

The second experiment was to compare reconstruction error performance between *pose+shape* versus only shape models. Reconstruction error was measured as the RMS distance between original and reconstructed instances at corresponding points. Model size is defined as the amount of data required to reconstruct instances from the model for a given number of modes. In the case of only shape model, the size is $(q^{os} + 1)(3 \sum n_j)$, with q^{os} the number of modes and n_j the number of landmarks of object j . In the case of *pose+shape* model, two data reduction strategies can be used: to reconstruct instances using the first q^{ps} modes, $k = 1 \dots q^{ps}$ in equation (10); and to perform data reduction of the matrix P in (9) (by means of SVD and selecting the first q^f modes). The model size in this case is $((7J) + q^f J)(q^{ps} + 1) + (q^f + 1)(3 \sum n_j)$. Remember that the number of landmarks, in our case 10459, is much larger than the total number of pose degrees of freedom and the number of instances.

Fig. 3 illustrates the reconstruction error vs. model size obtained by running the parameters q^f , q^{ps} and q^{os} . Increasing only the value of q^{ps} produces a very small increase of model size while a relevant decrease of reconstruction error. This is seen as nearly vertical dotted lines. A zoom of one of these lines is plotted at the right part of the figure.

4. DISCUSSION AND CONCLUSIONS

According to Fig. 2 most of the anatomical variability of subcortical nuclei (after global alignment) can be explained by pose transformations, which can be described by a multi-object pose model with a very small set of parameters. This model provides a coarse and compact representation of a the total anatomical variability, modeling about 70% of the total energy in most of the structures. A more accurate model of the shape boundary is obtained by adding shape features in the joint statistical model. These shape features characterize the residual anatomical variability after pose modeling.

In contrast, only shape models need a much larger model

size than *pose+shape* models in order to characterize the anatomical variability for the same reconstruction error.

The proposed model splits the anatomical variability of a set of objects into pose and shape parameters. Pose parameters are very compact and have a very natural understanding. In our vision, these models can provide valuable *a priori* information to segmentation and registration algorithms as initializations or regularizers.

5. REFERENCES

- [1] J Ashburner, JG Csernansky, et al., “Computer-assisted imaging to assess brain structure in healthy and diseased brains,” *The Lancet Neurology*, vol. 2, pp. 79–88, 2003.
- [2] AT Du, N Schuff, et al., “Atrophy rates of entorhinal cortex in AD and normal aging,” *Neurology*, vol. 60, pp. 481–486, 2003.
- [3] JG Csernansky, MK Schindler, et al., “Abnormalities of thalamic volume and shape in schizophrenia,” *Am J Psychiatry*, vol. 161, pp. 896–902, 2004.
- [4] S Ordas, L Boisrobert, et al., “Grid-enabled automatic construction of a two-chamber cardiac pdm from a large database of dynamic 3d shapes,” in *ISBI*, 2004, pp. 416–419.
- [5] G Gerig, S Joshi, et al., “Statistics of populations of images and its embedded objects: Driving applications in neuroimaging,” in *IEEE ISBI*, Apr 2006, pp. 1120–1123.
- [6] MN Bossa and S Olmos, “Statistical model of similarity transformations: building a multi-object pose model of brain structures,” in *IEEE CVPR-MMBIA*, New York, June 2006.
- [7] M Styner, K Gorczowski, et al., “Statistics of pose and shape in multi-object complexes using principal geodesic analysis,” in *Proc. MIAR conference*, 2006.
- [8] X. Pennec, “Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements,” in *Proc. of NSIP 1999*, June 20–23, Antalya, Turkey, 1999, IEEE-EURASIP, vol. 1, pp. 194–198.
- [9] M. N. Bossa and S. Olmos, “Statistical linear models in procrustes shape space,” in *Math. Foundations Computational Anatomy 2006*, Copenhagen, Oct 2006, pp. 102–111.
- [10] “IBSR 2004, Internet Brain Segmentation Repository (IBSR),” <http://www.cma.mgh.harvard.edu/ibsr/>, 2005.
- [11] Haili Chui and Anand Rangarajan, “A new point matching algorithm for non-rigid registration,” *Comput. Vis. Image Underst.*, vol. 89, no. 2-3, pp. 114–141, 2003.