STATISTICAL SHAPE ANALYSIS VIA PRINCIPAL FACTOR ANALYSIS

Mauricio Reyes Aguirre (1),Marius George Linguraru (2), Kostas Marias (3), Nicholas Ayache (4), Lutz-Peter Nolte (1), Miguel Ángel González Ballester (1)

(1) Institute for Surgical Technology and Biomechanics, MEM Research Center, University of Bern, Stauffacherstrasse 78, CH-3014 Bern, Switzerland.
(2) Department of Engineering and Applied Sciences, Harvard University, 60 Oxford Street, Cambridge MA, 02138, USA.
(3) Institute for Computer Science, Foundation for Research and Technology – Hellas, Vassilika Vouton, PO Box 1385, GR-71110 Heraklion, Crete, Greece.
(4) Asclepios Research Project, INRIA, 2004 Route des Lucioles BP 93, F-06902 Sophia Antipolis Cedex, France.

ABSTRACT

Statistical shape analysis techniques commonly employed in the medical imaging community, such as Active Shape Models or Active Appearance Models, rely on Principal Component Analysis (PCA) to decompose shape variability into a reduced set of interpretable components. In this paper we propose Principal Factor Analysis (PFA) as an alternative and complementary tool to PCA providing a decomposition into modes of variation that can be more easily interpretable, while still being a linear efficient technique that performs dimensionality reduction (as opposed to Independent Component Analysis, ICA). The key difference between PFA and PCA is that PFA models covariance between variables, rather than the total variance in the data. The added value of PFA is illustrated on 2D landmark data of corpora callosa outlines. Then, a study of the 3D shape variability of the human left femur is performed. Finally, we report results on vector-valued 3D deformation fields resulting from non-rigid registration of ventricles in MRI of the brain.

Index Terms— Image shape analysis, Biomedical image processing, principal factor analysis, principal component analysis, morphometry

1. INTRODUCTION

The analysis of shape variability of anatomical structures is of key importance in a number of clinical disciplines, as abnormality in shape can be related to certain diseases. Statistical shape analysis techniques enjoy a remarkable popularity within the medical image analysis community. Its flagship, the Active Shape Model (ASM), proposed by Cootes et al.[1], provides a method to study the structure of point data sets or meshes. This technique was later extended to intensity information, and thus image data, as the Active Appearance Model (AAM) [2].

Nearly all existing statistical shape analysis methods rely on Principal Component Analysis (PCA) to build a compact model of principal 'modes of variation' from a training set. PCA belongs to a family of methods for multivariate analysis commonly known as Factor Analysis (FA). Reviews and comparative studies of FA techniques can be found elsewhere [3][4]. Such techniques can be classified into linear and non-linear, reflecting whether the shape variation can be expressed as a linear combination of basic deformation primitives. We contend that a factorial decomposition of shape variability, if it is to be easily interpretable, must follow a linear model, where each mode of variation has a scalar weight. PCA is a linear technique. Exceptionally, another linear FA method, Independent Component Analysis (ICA) or equivalently Maximum Autocorrelation Function (MAF) PCA [5], has been proposed. However, ICA does not provide a compact representation of shape variability, as one obtains as many independent components as there are variables [3].

A common difficulty encountered when using PCA for shape analysis is that of correlating the resulting modes of variation with intuitive shape descriptions employed by clinical partners. Thus, these components are often described as combinations of several localised shears, twists, rotations, etc., but these are most often simplistic approximations to complex deformation fields.

In this paper we compare the performance of PCA with another linear factor analysis technique, known as Principal Factor Analysis (PFA). According to the FA literature, PCA is recommended for dimensionality reduction, whereas PFA is adapted for the study of structure in the data [6]. In particular, Widaman [7] states that "principal component analysis should not be used if a researcher wishes to obtain parameters reflecting latent constructs or factors."

This paper will show that, PFA provides a better decomposition of shape variation in terms of interpretability.
Sections 2 and 3 describe PCA and PFA, respectively. Section 4 compares results of both PCA and PFA on 2D landmark data sets, 3D surfaces and 3D vector fields, all coming from medical imaging data. Discussion and conclusions are provided in section 5.

2. PRINCIPAL COMPONENT ANALYSIS

PCA is a projection model for FA aiming to find a low-dimensional manifold in the space of the data, such that the distance between the data and their projection on the manifold is small. PCA is the best, in the mean-square error sense, linear dimension reduction technique [3].
Given a set of training data $\{t_1,\ldots,t_n\}$ in a given orthonormal basis of $\Re^D$, PCA finds a new orthonormal basis $\{\mathbf{u_1},\ldots,\mathbf{u_d}\}$ with its axes ordered from highest to lowest variance. The principal components are the set of new ordered basis vectors.
To find the principal components is to compute the sample covariance matrix S of the data set and then find its eigenstructure $\mathbf{SU}=\mathbf{U\Lambda}$. $\mathbf{U}$ is a DxD matrix which has the unit length eigenvectors $\mathbf{u_1},\ldots,\mathbf{u_d}$ as its columns, and $\mathbf{\Lambda}$ is a diagonal matrix with the corresponding eigenvalues $\lambda_1,\ldots,\lambda_D$. The eigenvectors are the principal components and the eigenvalues their corresponding projected variances.

3. PRINCIPAL FACTOR ANALYSIS

In opposition to PCA, which is a projection model, PFA can be considered as a generative model for FA. Generative models try to model the density function that is assumed to have generated the data, under constraints that restricts the set of possible models to those with low intrinsic dimensionality. The following description is mainly based on Carreira-Perpiñán's paper [4].
PFA represents an observed D-dimensional continuous variable $\mathbf{t}$, as a linear function $\mathbf{f}$ of an L-dimensional (L<D) continuous latent variable $\mathbf{x}$ and an independent Gaussian noise process:

$$\mathbf{t} = \mathbf{\Lambda x} + \boldsymbol{\mu} + \mathbf{e}$$

Here $\mathbf{\Lambda}$ is the DxL 'factor loading matrix' defining the linear function $\mathbf{f}$, $\boldsymbol{\mu}$ is a D-dimensional vector representing the mean of the distribution of $\mathbf{t}$, and $\mathbf{e}$ is a D-dimensional vector representing the noise or individual variability associated with each of the D observed variables.
PFA assumes a Gaussian distributed prior and noise model, and a linear mapping from data space to latent space.
The columns of the matrix $\mathbf{\Lambda}$ are referred to as 'factor loadings'. The data space noise model $\mathbf{e}$ is normal with diagonal covariance matrix $\mathbf{\Psi}$:

$$p(\mathbf{t}|\mathbf{x}) \sim N(\mathbf{f}(\mathbf{x}), \mathbf{\Psi})$$

The D diagonal elements of $\mathbf{\Psi}$ are referred to as 'uniqueness'.
Following the Bayes rule, the posterior in latent space is also normal:

$$p(\mathbf{x}|\mathbf{t}) \sim N\left(\mathbf{A}(\mathbf{t}-\boldsymbol{\mu}), (\mathbf{I}+\mathbf{\Lambda}\mathbf{\Psi}^{-1}\mathbf{\Lambda})^{-1}\right)$$
$$\mathbf{A} = \mathbf{\Lambda}^t(\mathbf{\Lambda}\mathbf{\Lambda}^t + \mathbf{\Psi})^{-1} = (\mathbf{I}+\mathbf{\Lambda}^t\mathbf{\Psi}^{-1}\mathbf{\Lambda})^{-1}\mathbf{\Lambda}^t\mathbf{\Psi}^{-1}$$

The parameters of the PFA model may be estimated using the EM (Expectation-Maximization) algorithm:
- E step: This requires computing the moments. For each data point $\mathbf{t_n}$ given the current parameter values $\mathbf{\Lambda}^{(\tau)}$ and $\mathbf{\Psi}^{(\tau)}$:

$$E\{\mathbf{x}|\mathbf{t_n}\} = \mathbf{A}^{(\tau)}(\mathbf{t_n}-\boldsymbol{\mu})$$
$$E\{\mathbf{x}\mathbf{x}^t|\mathbf{t_n}\} = \mathbf{I} - \mathbf{A}^{(\tau)}\mathbf{\Lambda}^{(\tau)} + \mathbf{A}^{(\tau)}(\mathbf{t_n}-\boldsymbol{\mu})(\mathbf{t_n}-\boldsymbol{\mu})^t(\mathbf{A}^{(\tau)})^t$$

- M step: This results in the following updated equations for the factor loadings $\mathbf{\Lambda}$ and uniqueness $\mathbf{\Psi}$:

$$\mathbf{\Lambda}^{(\tau+1)} = \left(\sum_{n=1}^{N}\mathbf{t_n}E\{\mathbf{x}|\mathbf{t_n}\}^t\right)\left(\sum_{n=1}^{N}E\{\mathbf{x}\mathbf{x}^t|\mathbf{t_n}\}\right)^{-1}$$
$$\mathbf{\Psi}^{(\tau+1)} = \frac{1}{N}\mathrm{diag}\left(\sum_{n=1}^{N}\mathbf{t_n}\mathbf{t_n}^t - \mathbf{\Lambda}^{(\tau+1)}E\{\mathbf{x}|\mathbf{t_n}\}\mathbf{t_n}^t\right)$$

where the updated moments are used and the diag operator sets all off-diagonal elements of a matrix to zero.
Note that as opposed to PCA, the factor decomposition is not unique, since an orthogonal rotation of the factors ($\mathbf{\Lambda}^t = \mathbf{\Lambda R}$, where $\mathbf{R}$ is an orthogonal matrix) does not alter the distribution in data space. Thus, from all factor loading matrices $\mathbf{\Lambda}$, we are free to choose that which is easiest to interpret according to some criterion. We employ Varimax rotation, which finds an orthogonal rotation of the factors and maximize the sparseness of the retained modes. The resulting rotated matrix $\mathbf{\Lambda'}$ has many values clamped to (almost) 0. Thus, each factor involves only a few orthogonal variables. This simplifies the factor interpretation.

4. RESULTS

4.1. Corpus callosum - 2D landmark data

In this first experiment we attempt to compare the performance of PFA vs. PCA in the analysis of the shape variability of the corpus callosum. Our training data consists of 9 sets of 2D landmarks delimiting the corpora callosa of 9 patients as evidenced on MRI. Each data set has

63 landmarks. The delineations were performed manually by an expert.
Figure 1 shows the 3 principal components (top) and 3 principal factors (bottom) resulting from PCA and PFA, respectively. PCA does a good job at modelling the variance of the data in a few principal modes of variation, but such modes can be interpreted as a combination of several basic displacements, which are complex in most cases. On the contrary, PFA decomposes the variation into simple factors that have one or two defined directions of displacement.

4.2. Morphological analysis of human left femurs

We compare the results of PCA and PFA on a training set of 32 surface models extracted from CT data. These models represent the left femur, and are used in on-going research at our institute for computer-assisted surgery. Correspondences across data sets were established via a semiautomatic process in which a set of analogous anatomical landmarks (points, lines or surfaces) are identified in all data, and the remaining points are evenly spread to cover the surface of the object. These correspondences are further optimized via a Minimum Description Length optimization.
Figure 2 shows results obtained for PCA and PFA for the second, third and fourth mode (for the first mode, both, PCA and PFA similarly describe the change in length of the femur). The magnitude of deformation has been mapped in colors while its direction is represented with streamlines. For each mode it can be seen that PFA produces a more regular and homogeneous deformation and it also focuses more on specific areas of the bone (e.g. PFA mode 2 and 3, describes the deformation in the lower and upper part of the image, respectively, while PCA presents a mixture of scale and shape variations at different locations), which is more relevant for medical purposes.

4.3. Results on vector-valued displacement fields

The following example analyses the 3D vector-valued displacement fields resulting from non-rigid registration. In particular, 6 brain MRI data sets where registered affinely to a probabilistic brain atlas [8] for spatial normalization.
The aim of this study was to determine predominant patterns in ventricular shape variability. To this end, the images were first non-rigidly registered, resulting in a set of 3D deformation fields. PCA and PFA were then performed on the set of 6 ventricle displacement fields. Figure 3 shows the vector norm of the first principal component (left) and principal factor (right). It can be observed that the effect of the principal factor is more localized (towards the back of the ventricles), while the principal component affects most of the ventricular surface.

5. DISCUSSION AND CONCLUSIONS

The key difference between PFA and PCA is that PFA models covariance between variables, rather than the total variance in the data. PCA determines the factors that account for the total (unique and common) variance in the set of variables; PFA determines the least number of factors that can account for the common variance (correlation).
PFA provides better interpretability than PCA in terms of decomposition of shape variability, as evidenced by the results provided. The superiority of PFA over PCA has also been highlighted recently for fMRI analysis [9]. On the other hand, PCA is better suited to describe total variance and is optimal in reducing dimensionality, so PFA decomposition may require more factors to reconstruct the original data up to a certain threshold of accuracy.
For example, in a clinical setting, size and orientation are two important separate factors that must be easily distinguishable for understanding the geometrical and mechanical properties of organs and tissue. PCA, using global variability analysis, intermingles the variation components. Although it reduces dimensionality, it does not always simplify the interpretable data load or understanding of functionality by means of shape analysis. PFA, through its modes of variation, offers a better and simpler analysis of shape variability by intuitively distinguishable factors.
Alternative factor analysis techniques may be employed [3]. An exhaustive comparative study of such techniques for statistical shape analysis is in progress.
We proposed PFA as a viable alternative to PCA. Our results agree with those found in [6]; while still being a linear technique for dimensionality reduction, PFA models covariance between variables and offers decomposition into easily interpretable and distinguishable modes of variation.

6. REFERENCES

[1] Cootes, T. F., Taylor, C. J., Cooper, D. H., Graham, J.: Active Shape Models - Their Training and Applications. Computer Vision and Image Understanding, 61(2), 1995

[2] Cootes, T. F., Edwards, G. J., Taylor, C. J.: Active Appearance Models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(6):681-685, 2001

[3] Fodor, I.K.: A Survey of Dimension Reduction Techniques. Lawrence Livermore National Laboratory Technical Report no. UCRL-ID-148494, 2002

[4] Carreira-Perpiñán, M. A., Renals, S.: Dimensionality Reduction of Electropalatographic Data Using Latent Variable Models. Speech Communication, 26(4):259-282, 1998

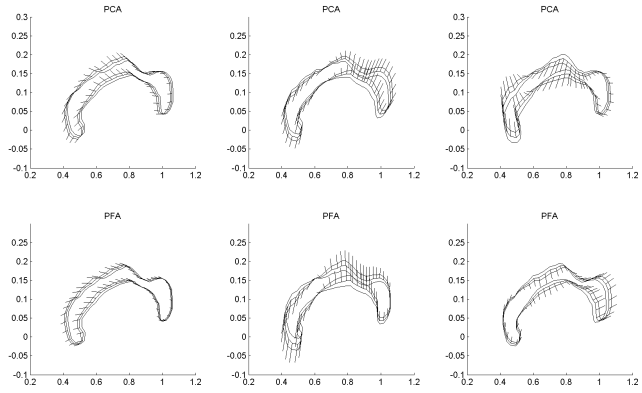

Figure 1. PCA and PFA results on 2D landmark data sets representing corpora callosa. Top: First 3 principal components (ordered from left to right according to the variance explained), and bottom: first 3 principal factors after Varimax rotation.

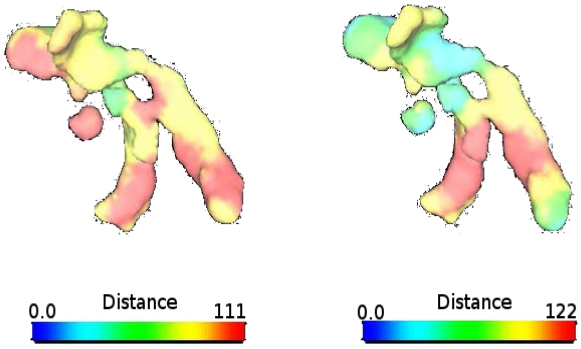

Figure 3. 3D view of the norm (mapped as colors on the surface of the mean shape ventricle) of the first principal component (left) and principal factor (right) extracted from a training set of 3D displacement vector fields on the surface of brain ventricles.

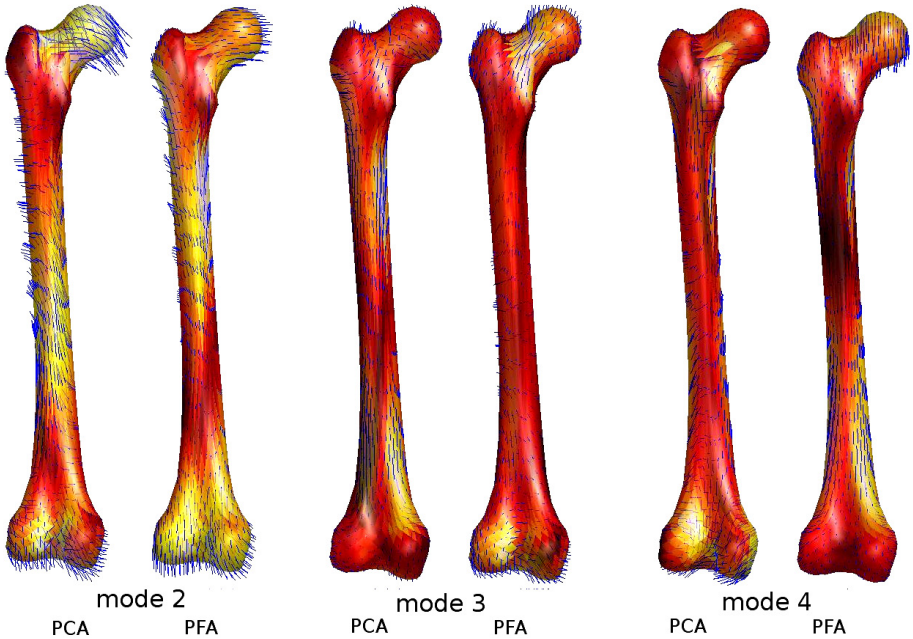

Figure 2. Shape analysis of human left femurs. Second, third and fourth mode of variation for PCA and PFA are represented with colormaps for the magnitude of the deformation (brighter regions correspond to higher magnitudes) and with lines for its direction. PFA represents shape variation with more regular and homogeneous deformations and it is more localized than PCA.

[5] Üzümcü, M., Frangi, A. F., Sonka, M., Reiber, J. H. C., Lelieveldt, B. P. F.: ICA vs. PCA Active Appearance Models: Application to Cardiac MR Segmentation. Procs. of MICCAI'2003, Lecture Notes in Computer Science, 2878:451-458, 2003

[6] Charles, D., Fyfe, C.: Kernel factor analysis with Varimax rotation. Neural Networks, 2000. IJCNN 2000, Proceedings
of the IEEE-INNS-ENNS Int. Joint Conf. on, vol.3, no.pp. 381-386

[7] Widaman, K.F.: Common Factor Analysis versus Principal Components Analysis: Differential Bias in Representing Model Parameters? Multivariate Behavioral Research, 28:263-311, 1993

[8] Collins, D.L. et al.: Design and Construction of a Realistic Digital Brain Phantom. IEEE Transactions on Medical Imaging, 17(3):463-468, 1998

[9] Huang, C.-C., Liou, M., Cheng, P. E., Chen, C.-C.: Noise Reduction in Functional MR Images by Common Factor Models. Procs. of International Conference on Neural Information Processing (ICONIP), 2001