SHAPE MODELLING USING MAXIMUM AUTOCORRELATION FACTORS

Rasmus Larsen

Department of Mathematical Modelling, Technical University of Denmark,
Building 321, DK-2800 Kgs. Lyngby, Denmark
rl@imm.dtu.dk

ABSTRACT

This paper addresses the problems of generating a low dimensional representation of the shape variation present in a training set after alignment using Procrustes analysis and projection into shape tangent space. We will extend the use of principal components analysis in the original formulation of Active Shape Models by Timothy Cootes and Christopher Taylor by building new information into the model. This new information consists of two types of prior knowledge. First, in many situation we will be given an ordering of the shapes of the training set. This situation occurs when the shapes of the training set are in reality a time series, e.g. snapshots of a beating heart during the cardiac cycle or when the shapes are slices of a 3D structure, e.g. the spinal cord. Second, in almost all applications a natural order of the landmark points along the contour of the shape is introduced. Both these types of knowledge may be used to defined Shape Maximum Autocorrelation Factors. The resulting point distribution models are compared to ordinary principal components analysis using leave-one-out validation.

1. INTRODUCTION

The Active Shape Model (ASM) and Active Appearance Model (AAM) were proposed by Cootes & Taylor in [1, 2]. The ASM is based on learning shape variation for a given class from a series of training shapes represented by a set of corresponding points (landmarks). After alignment wrt. translation, rotation, and scale (e.g. bmo. Procrustes analysis [3]), and projection into shape tangent space the shape coordinates are analysed by a principal components analysis (PCA). A low dimensional representation is then obtained by only retaining the first t of the eigen modes. The number t can be determined so that the retained modes account for a given proportion of the variation (e.g. 98 %), or so that an independent validation set is approximated sufficiently well.

The AAM is an extension to ASM that simultaneously accounts for shape and texture variation. Again a low dimensional representation is obtained by PCA.

We propose an extension to ASM and AAM using the maximum autocorrelation factor (MAF) analysis instead of PCA. The MAF analysis was originally proposed as an alternative transformation of multivariate spatial imagery to the celebrated PCA transform by Paul Switzer [4]. In the MAF analysis we seek a transformation that maximizes the autocorrelation between neighbouring observations (i.e. pixels). The basic assumption of the MAF analysis is that the interesting signal exhibits high autocorrelation, whereas the noise exhibits low autocorrelation. By building the additional information of the structure of the observations into the model application examples (cf. [5, 6, 7, 8]) show a more satisfying ordering and compression of the data. This is particularly the case when some noise components have higher variance than some signal components. In this case the principal components will fail to give an intuitive order of image quality. The MAF analysis requires knowledge of or estimation of the variance-covariance matrix of the data as well as the variance-covariance matrix of the difference between the original data and the a spatially shifted version of the data. This may be formulated as a canonical correlation analysis problem [9]. A similar approach to filter design is shown in [10].

We propose to apply this technique to ASM/AAM. This can be achieved in two ways. First, in situations where a natural ordering of the training shapes is given we may apply the algorithm directly. Such situations can arise when the training data a sampled as a time series (e.g. echocardiograms of the heart chambers recorded during the cardiac cycle [1]), or when the training samples are 2-D slices sampled along a 3-D structure. In this way we build in new information in the form 'neighbouring observations are similar' into the transformation. We will then seek information that is present in neighbours simultaneously. Information that is not similar between neighbours is considered noise.

Alternatively, we may also utilise that the shape representation usually consists of landmark points with a natural ordering (i.e. along contours). This approach is based on a reformulation of the problem using the Eckart-Young theorem [11]. By representing the shape coordinates of each training observation as the rows of a data matrix, the PCA of the original ASM is an R-mode analysis of this matrix. In the statistical sense - in this analysis - the variables are the point coordinates and the observations are the training shapes. The estimated eigenvectors are used to deform the

mean shape.

In we instead make a Q-mode analysis then the variables are the (unordered) training shapes and the observations are the point coordinates. In this case the deformation of the mean shape is introduced not by the eigenvectors but by the transformed variables. Q-mode and R-mode analysis are equivalent in this case. However, we can solve the MAF eigenproblem in the Q-mode fashion as well. Then we arrive at new (and different) transformed variables.

Other transformations used to describe the contour are Fourier models [12] and wavelets [13].

2. DATA

We will illustrate the proposed methods on a dataset that consists of annotations of the contour of 24 metacarpals, i.e. a bone in the human hand. An example is shown in Figure 1.

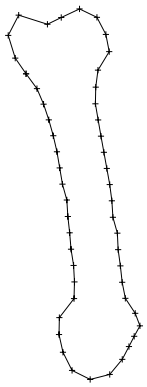

Fig. 1. A metacarpal annotated using 50 landmarks. The proximal end is up and the distal end is down.

The annotations are based on 2-D wrist radiographs of human hands. The annotations are prone to errors in the proximal and distal ends due to the bones being overlaid in the projection of the radiograph and thus difficult to discern.

3. METHODS

Let there be given p training examples for a given shape class, and let each example be represented by a set of n landmark points (u_{ij}, v_{ij}), $i = 1, \ldots, p$ and $j = 1, \ldots, n$. Then each example is given by a $2n$ vector

$$\boldsymbol{u}_i = (u_{i1}, \ldots, u_{in}, v_{i1}, \ldots, v_{in})^T. \tag{1}$$

The p training examples are aligned to a common mean using a full generalised Procrustes analysis (e.g. see [14]).

The tangent space coordinates are the projections of the full Procrustes coordinates into the tangent plane to the shape space at the full Procrustes mean. In the vicinity of this pole the Euclidean distances of the tangent space are good approximations to the Procrustes distances. Let the tangent space coordinates with the origin placed at the pole of the tangent space for each training example be

$$\boldsymbol{x}_i = (x_{i1}, \ldots, x_{in}, y_{i1}, \ldots, y_{in})^T. \tag{2}$$

Note, that $\sum_{i=0}^{p} \boldsymbol{x}_i = \boldsymbol{0}$. We organise the tangent space coordinates of the training examples in a $p \times 2n$ data matrix

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{x}_1^T \\ \boldsymbol{x}_2^T \\ \vdots \\ \boldsymbol{x}_p^T \end{bmatrix} \tag{3}$$

This matrix may then be decomposed using Eckart-Young's theorem [11]

$$\boldsymbol{X} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{U}^T. \tag{4}$$

Where $\boldsymbol{U}$ $(2n \times r)$ and $\boldsymbol{V}$ $(p \times r)$ are orthogonal matrices, and $\boldsymbol{\Lambda}$ $(r \times r)$ is a diagonal matrix with positive diagonal elements. The diagonal elements of $\boldsymbol{\Lambda}$ are called the singular values of $\boldsymbol{X}$. This decomposition is also called the singular value decomposition (SVD).

By direct calculation using Eckart-Young's theorem we have the following two eigenvalue decompositions

$$\boldsymbol{X}\boldsymbol{X}^T = \boldsymbol{V}\boldsymbol{\Lambda}^2\boldsymbol{V}^T \tag{5}$$

$$\boldsymbol{X}^T\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Lambda}^2\boldsymbol{U}^T \tag{6}$$

The diagonal elements of $\boldsymbol{\Lambda}^2$ are the squared diagonal elements of $\boldsymbol{\Lambda}$ and these are the positive eigenvalues of $\boldsymbol{X}\boldsymbol{X}^T$ and $\boldsymbol{X}^T\boldsymbol{X}$. The analysis of $\boldsymbol{X}^T\boldsymbol{X}$ is called a R-mode analysis, and the analysis of $\boldsymbol{X}\boldsymbol{X}^T$ is called a Q-mode analysis. The relation between the eigenvectors corresponding to the positive eigenvalues for the two problems are given by

$$\boldsymbol{V} = \boldsymbol{X}\boldsymbol{U}\boldsymbol{\Lambda}^{-1} \tag{7}$$

$$\boldsymbol{U} = \boldsymbol{X}^T\boldsymbol{V}\boldsymbol{\Lambda}^{-1} \tag{8}$$

The estimated variance-covariance matrix of the tangent space coordinates of the training examples in Eq. (2) is

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{p-1}\boldsymbol{X}^T\boldsymbol{X} \tag{9}$$

the eigenvectors (i.e. the principal components) of which are given by the columns of $\boldsymbol{U}$.

The ASM model then consists of retaining the $t \leq r$ first principal components. Deviations from the Procrustes mean (in tangent space) can then be modelled by

$$\boldsymbol{x} = \boldsymbol{U}'\boldsymbol{b} \tag{10}$$

where $\boldsymbol{U}'$ is a matrix consisting of the first t columns of $\boldsymbol{U}$, and $\boldsymbol{b}$ defines a set of t parameters of the deformable model.

However, from Eq. (8) we see that by solving the problem in Q-mode, i.e. solve for $\boldsymbol{V}$ we could generate the same ASM by

$$\boldsymbol{x} = \boldsymbol{X}^T\boldsymbol{V}'\boldsymbol{b} \tag{11}$$

where $\boldsymbol{V}'$ is a matrix consisting of the first t columns of $\boldsymbol{V}$.

Solving the problem in Q-mode corresponds to an eigenvalue decomposition of a matrix consisting of sums-of-squares of the deviations of a stochastic variable, $\boldsymbol{Z}$, from $\mathbf{0}$, examples of which are given by the coordinates of each point across the shape training examples, i.e.

$$\begin{aligned} \boldsymbol{z}_j &= (x_{1j}, \ldots, x_{pj}), \quad \text{for } j = 1, \ldots, n \\ \boldsymbol{z}_j &= (y_{1j}, \ldots, y_{pj}), \quad \text{for } j = n+1, \ldots, 2n. \end{aligned} \tag{12}$$

This matrix is given by

$$\hat{\boldsymbol{\Pi}} = \frac{1}{2n-2}\boldsymbol{X}\boldsymbol{X}^T \tag{13}$$

The eigenvectors (i.e. the principal components) of this matrix corresponding to the positive eigenvalues are given by the columns of $\boldsymbol{V}$.

3.1. Maximum Autocorrelation Factors

In this section we will describe two approaches to using the maximum autocorrelation factors (MAF) [4]
transform instead of principal components for formulation of an ASM. The first approach is based on a prior ordering of the training shapes. The second approach assumes an ordering of the landmarks.

We will begin by reviewing the MAF transform. Consider the spatial covariance function

$$\boldsymbol{\Gamma}(\Delta) = \text{Cov}\{\boldsymbol{Z}_k, \boldsymbol{Z}_{k+\Delta}\} \tag{14}$$

where Δ denotes a spatial shift. Evidently $\boldsymbol{\Gamma}^T(\Delta) = \boldsymbol{\Gamma}(-\Delta)$. Defining the variance-covariance matrix $\boldsymbol{\Sigma}_\Delta = D\{\boldsymbol{Z}_k - \boldsymbol{Z}_{k+\Delta}\}$, we find

$$\boldsymbol{\Sigma}_\Delta = 2\boldsymbol{\Sigma} - \boldsymbol{\Gamma}(\Delta) - \boldsymbol{\Gamma}(-\Delta) \tag{15}$$

We are now able to compute the covariance between the original variables an the shifted variables

$$\begin{aligned} &\text{Cov}\{\boldsymbol{w}_i^T\boldsymbol{Z}_k, \boldsymbol{w}_i^T\boldsymbol{Z}_{k+\Delta}\} \\ &= \boldsymbol{w}_i^T\boldsymbol{\Gamma}(\Delta)\boldsymbol{w}_i = \boldsymbol{w}_i^T\boldsymbol{\Gamma}^T(\Delta)\boldsymbol{w}_i \\ &= \frac{1}{2}\boldsymbol{w}_i^T(\boldsymbol{\Gamma}(\Delta) + \boldsymbol{\Gamma}(-\Delta))\boldsymbol{w}_i \\ &= \boldsymbol{w}_i^T(\boldsymbol{\Sigma} - \frac{1}{2}\boldsymbol{\Sigma}_\Delta)\boldsymbol{w}_i, \end{aligned} \tag{16}$$

which results in the following correlation coefficient

$$\text{Corr}\{\boldsymbol{w}_i^T\boldsymbol{Z}_k, \boldsymbol{w}_i^T\boldsymbol{Z}_{k+\Delta}\} = 1 - \frac{1}{2}\frac{\boldsymbol{w}_i^T\boldsymbol{\Sigma}_\Delta\boldsymbol{w}_i}{\boldsymbol{w}_i^T\boldsymbol{\Sigma}\boldsymbol{w}_i}. \tag{17}$$

In order to minimize that correlation we must maximize the Rayleigh coefficient

$$R(\boldsymbol{a}) = \frac{\boldsymbol{w}^T\boldsymbol{\Sigma}_\Delta\boldsymbol{w}}{\boldsymbol{w}^T\boldsymbol{\Sigma}\boldsymbol{w}}. \tag{18}$$

Let $\kappa_1 \leq \cdots \leq \kappa_m$ be the eigenvalues and $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ corresponding conjugate eigenvectors of $\boldsymbol{\Sigma}_\Delta$ with respect to $\boldsymbol{\Sigma}$. Let

$$\boldsymbol{W} = \begin{bmatrix} \boldsymbol{w}_1 & \boldsymbol{w}_2 & \cdots & \boldsymbol{w}_m \end{bmatrix}. \tag{19}$$

Then the minimum/maximum autocorrelation factor transform is given by

$$\boldsymbol{Y}_k = \boldsymbol{W}^T\boldsymbol{Z}_k. \tag{20}$$

Note that ordering is defined so that the first MAF has maximum autocorrelation. An additional problem will arise when the number of training examples is less than the dimensionality of the problem. Then the variance-covariance matrix in the nominator of Eq. 18 is not positive definite. In this case the optimization must be carried out in the subspace spanned by the eigenvectors corresponding to nonzero eigenvalues of this matrix.

3.2. Shape R-MAF

This is the most straightforward application of the MAF transform to ASMs. We now assume that the numbering of the training examples constitute a natural order. This order may come from the shapes being sampled on a time axis or on a spatial axis (e.g. slicing a 3D structure). This is a generalisation of the ordinary R-mode analysis of ASMs. This generalisation works also for the AAMs, where shape information is extended with texture information.

An estimate of the variance-covariance matrix is given by Eq. (9). An estimate of the difference variance-covariance matrix is given by

$$\hat{\boldsymbol{\Sigma}}_\Delta = \frac{1}{p-1}\boldsymbol{D}^T\boldsymbol{D} \tag{21}$$

where

$$\boldsymbol{D} = \begin{bmatrix} (\boldsymbol{x}_1 - \boldsymbol{x}_2)^T \\ (\boldsymbol{x}_2 - \boldsymbol{x}_3)^T \\ \vdots \\ (\boldsymbol{x}_{p-1} - \boldsymbol{x}_p) \\ (\boldsymbol{x}_p - \boldsymbol{x}_1)T \end{bmatrix}. \tag{22}$$

The MAF ASM is then built by retaining the $t \leq r$ first maximum autocorrelation factors. Deviations from the Procrustes mean (in tangent space) is then be modelled by

$$\boldsymbol{x} = \boldsymbol{W}'\boldsymbol{b} \tag{23}$$

where $\boldsymbol{W}'$ is a matrix consisting of the first t columns of $\boldsymbol{W}$ (Eq. (19)), and $\boldsymbol{b}$ defines the set of t parameters of the deformable model.

In ordinary ASMs limits are applied to the parameters $\boldsymbol{b}$ according to the variance of the parameters across the training set. In the ASM case the variance of the i^{th} parameter equals the i^{th} eigenvalue of the principal component analysis. Here the result does not simplify as well. The variance of the i^{th} parameter across the training set is given by

$$V\{b_i\} = \frac{\boldsymbol{w}_i^T\boldsymbol{\Sigma}\boldsymbol{w}_i}{\boldsymbol{w}_i^T\boldsymbol{w}_i}. \tag{24}$$

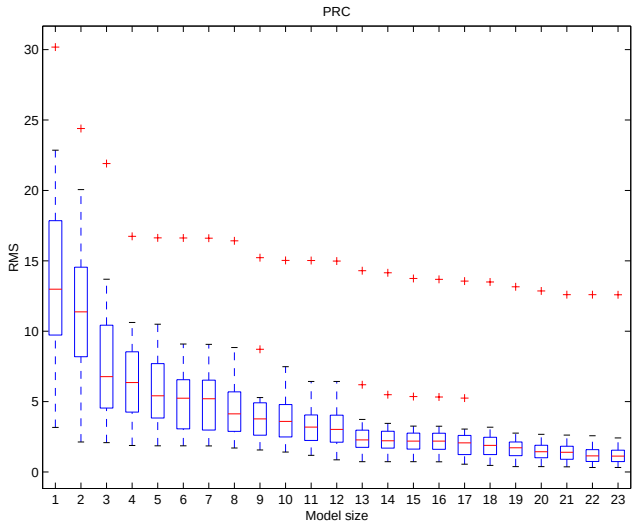

Fig. 2. Boxplots of the leave-one-out RMS error for the principal components analysis (PCA) analysis as a function of model size (model 1 is using the mean shape only). The box defines the interquartile range (IQR), the whiskers show the extend of the data, their length is at maximum 1.5 IQR. Data outside the whiskers are denoted outliers and drawn individually.

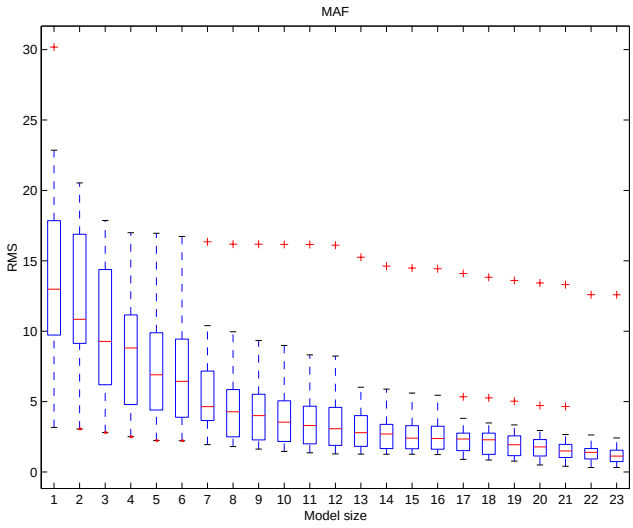

Fig. 3. Boxplots of the leave-one-out RMS error for the Q-mode maximum autocorrelation factor (MAF) analysis as a function of model size (model 1 is using the mean shape only). The box defines the interquartil range (IQR), the whiskers show the extend of the data, their length is at maximum 1.5 IQR. Data outside the whiskers are denoted outliers and drawn individually.

3.3. Shape Q-MAF

Another extension of the ASM uses the MAF transform based on the Q-mode analysis. It is therefore not as straightforward as the previous method. In general, it should be more applicable, since a natural ordering of the training examples is seldom present, but an ordering of the landmark points is almost always present.

In the MAF analysis in Section 3.1 we substitute the matrix $\hat{\mathbf{\Pi}}$ from Eq. (13) for the variance-covariance matrix $\mathbf{\Sigma}$. The estimate of the difference variance-covariance matrix is given by

$$\hat{\mathbf{\Sigma}}_{\Delta} = \frac{1}{2n-2}\boldsymbol{E}\boldsymbol{E}^T \tag{25}$$

where

$$\boldsymbol{E}^T = \begin{bmatrix} (\boldsymbol{z}_1 - \boldsymbol{z}_2)^T \\ \vdots \\ (\boldsymbol{z}_{n-1} - \boldsymbol{z}_n)^T \\ (\boldsymbol{z}_n - \boldsymbol{z}_1)^T \\ (\boldsymbol{z}_{n+1} - \boldsymbol{z}_{n+2})^T \\ \vdots \\ (\boldsymbol{z}_{2n-1} - \boldsymbol{z}_{2n})^T \\ (\boldsymbol{z}_{2n} - \boldsymbol{z}_{n+1})^T \end{bmatrix} \tag{26}$$

where $\boldsymbol{z}_j$ is defined in Eq (12).

The MAF ASM in this case is built by retaining the $t \leq r$ first maximum autocorrelation factors. From Eq. (11) deviations from the Procrustes mean (in tangent space) is then be modelled by

$$\boldsymbol{x} = \boldsymbol{X}^T\boldsymbol{W}'\boldsymbol{b} \tag{27}$$

where $\boldsymbol{W}'$ is a matrix consisting of the first t columns of $\boldsymbol{W}$ (Eq. (19)), and $\boldsymbol{b}$ defines the set of t parameters of the deformable model.

4. RESULTS AND DISCUSSION

The Q-MAF transformation is performed on the metacarpal dataset and compared to ordinary PCA. The resulting eigen modes are shown in Fig. 4. In the figure all eigen modes for the two transformations are shown. The eigen modes are visualised as deviations from the (full Procrustes) mean shape. The mean shape and the deviations from the mean shape is shown for each eigen mode as ± 5 standard deviations across the training, respectively. Note that if we assume that the variation from the mean value across the training set can be modelled by a Gaussian distribution, then we would expect (almost) all deviations to be within ± 3 standard deviation.

From the plots in Fig. 4 we see that the Q-MAF transformation results in different eigen modes from the PCA transformation. The Q-MAF modes constitute a decomposition of (localized) spatial frequency along the contour with frequency increasing with mode number. Furthermore, the first two modes are easily interpreted as thickness of the cortical bone, mode three as bending, and mode four as thickness of the proximal (top) end. In the high order number modes

variations composed of neighbouring points deforming in opposite directions are concentrated.

The PCA eigen modes are less easily interpreted and it seems that many low number modes are devoted to descriptions of variations of the proximal end. These are variations that may well stem from annotation arbitrariness.

In order to truly evaluate which method is better at retaining the real variations of the specimens of the training set and excluding annotation arbitrariness and unique factors we would have to evaluate the repeatability/reproducibility of the annotation procedure for this class of shape. Such data have not been available.

A leave-one-out evaluation where the root-mean-square error as a function of model size (retained eigen modes) for the two methods has been carried out. Boxplots of the root-mean-square error are shown in Figs. 2 and 3. Those data are inconclusive with respect to disqualifying one of the methods.

5. CONCLUSION

In this paper we have described a new alternative to using principal components analysis in active shape models. This method is based on the maximum autocorrelation factors transform and may be applied in a Q-mode or a R-mode type of analysis depending on the nature of the shape data.

In Q-mode we seek deformation modes that exhibit similar deformation of neighbouring points of the shape. It is assumed that this is a natural choice. On an example we have demonstrated that the Q-MAF transformation leads to a decomposition of spatial frequency along the contour and that the resulting modes are easily interpreted.

ACKNOWLEDGMENTS

The annotations of the metacarpal dataset, an example of which is shown in Fig. 1, were carried out by Hans Henrik Thodberg, Ph.D., manager of the image processing group of Pronosco A/S. The author is grateful for being allowed to use this dataset.

REFERENCES

[1] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models – their training and application," *Computer Vision, Graphics and Image Processing*, vol. 61, no. 1, pp. 38–59, Jan. 1995.

[2] T. F. Cootes and C. J. Taylor, "Statistical models of appearance for computer vision," Tech. Rep., Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering, University of Manchester, Manchester M13 9PT, United Kingdom, 2000, 99 pp.

[3] C.Goodall, "Procrustes methods in the statistical analysis of shape," *Journal of the Royal Statistical Society, Series B*, vol. 53, no. 2, pp. 285–339, 1991.

[4] Paul Switzer, "Min/max autocorrelation factors for multivariate spatial imagery," in *Computer Science and Statistics*, L. Billard, Ed. 1985, pp. 13–16, Elsevier Science Publishers B.V. (North Holland).

[5] Knut Conradsen, Bjarne Kjær Nielsen, and Tage Thyrsted, "A comparison of min/max autocorrelation factor analysis and ordinary factor analysis," in *Proceedings from Symposium in Applied Statistics*, Lyngby, Jan. 1985, pp. 47–56.

[6] Andrew A. Green, Mark Berman, Paul Switzer, and Maurice D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, Jan. 1988.

[7] Allan Aasbjerg Nielsen and Rasmus Larsen, "Restoration of GERIS data using the maximum noise fractions transform," in *Proceedings of the First International Airborne Remote Sensing Conference and Exhibition*, Strasbourg, France, 1994, vol. 2, pp. 557–568.

[8] Allan Aasbjerg Nielsen, Knut Conradsen, and James J Simpson, "Multivariate alteration detection (MAD) and MAF post-processing in multispectral, bitemporal image data: New approaches to change detection studies," *Remote Sensing of Environment*, vol. 19, pp. 1–19, 1998.

[9] Harold Hotelling, "Relations between two sets of variables," *Biometrika*, vol. 28, pp. 321–377, 1936.

[10] Hans Knutsson, Mats Andersson, Magnus Borga, and Johan Wiklund, "Automated generations of representations in vision," in *Proceedings of the 15th International Conference on Pattern Recognition*, A. Sanfeliu, J. J. Villanueva, M. Vanrell, R. Alquézar, T. Huang, and J. Serra, Eds., Los Alamitos, California, USA, 2000, vol. 3, pp. 63–70, IEEE Computer Society.

[11] R. M. Johnson, "On a theorem stated by Eckart and Young," *Psychometrika*, vol. 28, pp. 259–263, 1963.

[12] Lawrence H. Staib and James S. Duncan, "Boundary finding with parametrically deformable models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 11, pp. 1061–1075, 1992.

[13] A. Neuman and C. Lorenz, "Statistical shape model based segmentation of medical images," *Computerized Medical Imaging and Graphics*, vol. 22, pp. 133–143, 1998.

[14] Ian Dryden and Kanti Mardia, *Statistical Shape Analysis*, John Wiley & Sons, Chichester, 1998, xx + 347 pp.

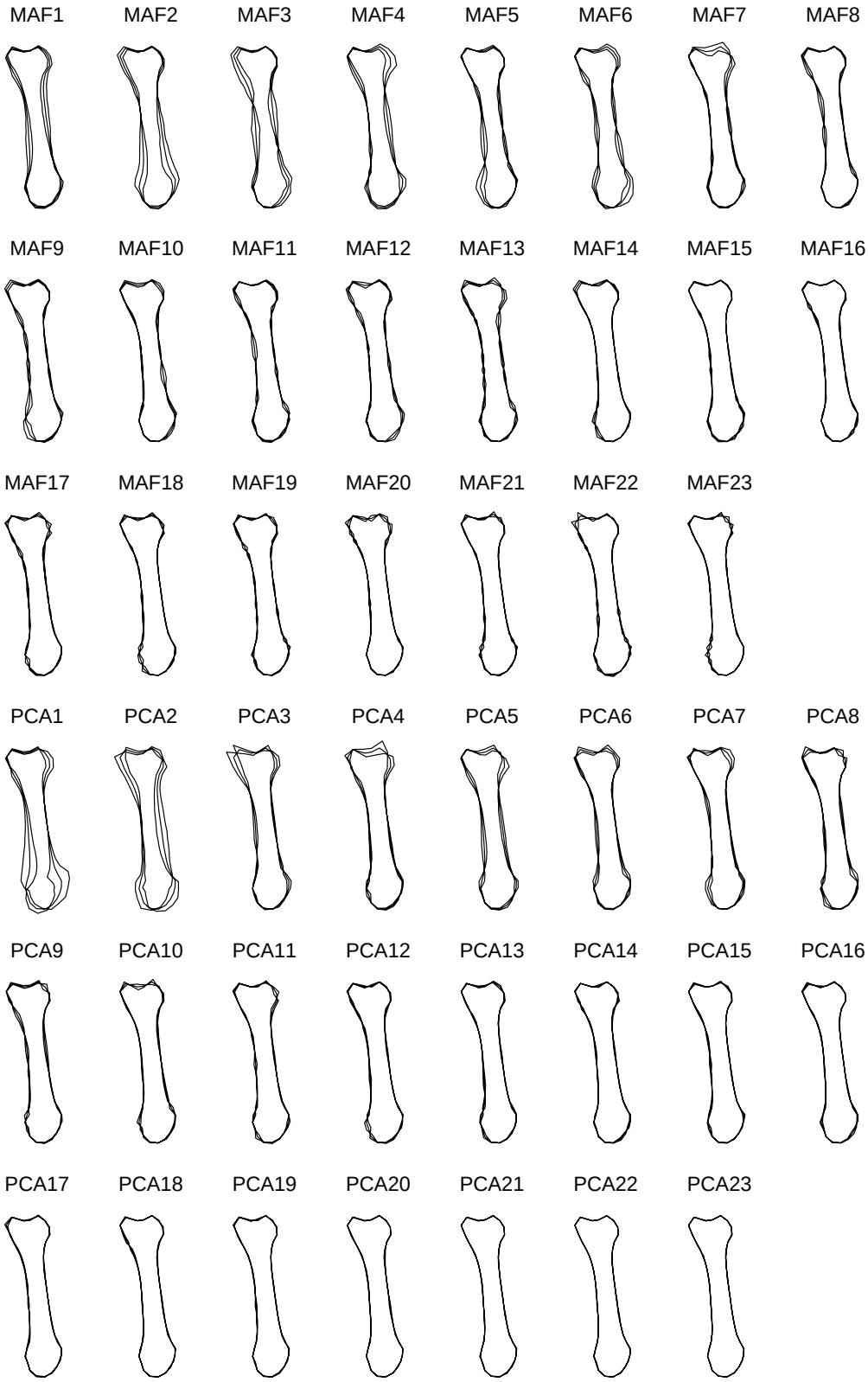

Fig. 4. The eigen modes for the shape Q-MAF and the PCA transformation are visualised as deviations from the (full Procrustes) mean shape. The mean shape and the deviations from the mean shape is shown for each eigen mode as ± 5 standard deviations across the training, respectively.