

What's in a Smile? Initial Results of Multilevel Principal Components Analysis of Facial Shape and Image Texture

DJJ Farnell^{1†}, J Galloway¹, A Zhurov¹, S Richmond¹, P Pirttiniemi²,
and Raija Lähdesmäki²

¹School of Dentistry, Cardiff University, Heath Park, Cardiff CF14 4XY

²Faculty Of Medicine, University Of Oulu, Oulu University Hospital MRC, Finland

FarnellD@cardiff.ac.uk (†corresponding author)
GallowayJL@cardiff.ac.uk
ZhurovAI@cardiff.ac.uk
RichmondS@cardiff.ac.uk
Pertti.Pirttiniemi@oulu.fi
raija.lahdesmaki@oulu.fi

Abstract. Multilevel principal components analysis (mPCA) has previously been shown to provide a simple and straightforward method of forming point distribution models that can be used in (active) shape models. Here we extend the mPCA approach to model image texture as well as shape. As a test case, we consider a set of (2D frontal) facial images from a group of 80 Finnish subjects (34 male; 46 female) with two different facial expressions (smiling and neutral) per subject. Shape (in terms of landmark points) and image texture are considered separately in this initial analysis. Three-level models are constructed that contain levels for biological sex, “within-subject” variation (i.e., facial expression), and “between-subject” variation (i.e., all other sources of variation). By considering eigenvalues, we find that the order of importance as sources of variation for facial shape is: facial expression (47.5%), between-subject variations (45.1%), and then biological sex (7.4%). By contrast, the order for image texture is: between-subject variations (55.5%), facial expression (37.1%), and then biological sex (7.4%). The major modes for the facial expression level of the mPCA models clearly reflect changes in increased mouth size and increased prominence of cheeks during smiling for both shape and texture. Even subtle effects such as changes to eyes and nose shape during smile are seen clearly. The major mode for the biological sex level of the mPCA models similarly relates clearly to changes between male and female. Model fits yield “scores” for each principal component that show strong clustering for both shape and texture by biological sex and facial expression at appropriate levels of the model. We conclude that mPCA correctly decomposes sources of variation due to biological sex and facial expression (etc.) and that it provides a reliable method of forming models of both shape and image texture.

Keywords: multilevel principal components analysis; shape and image texture; biological sex; facial expression

1 Introduction

Populations of subjects or images used in medical imaging often contain naturally occurring groups or clusters [1], e.g., by: manual mark-ups of the same set of images by different image experts (each subject forms its own cluster); ethnicity, biological sex, or age (etc.); multiple images or shapes from the same subjects (each subject again forms its own cluster) such as time series of images. One approach to dealing with such effects is effectively to ignore them; one analyzes all subjects or images irrespective of any naturally occurring groupings in the data. However, important effects might be omitted by this approach. If any such factors (e.g., by ethnicity, biological sex, or age) are modelled implicitly then we cannot ensure that these factor have not become “mixed together.” Another approach is to analyze results from each group separately. However, in this case it might be hard (if not impossible) to extrapolate the results for one group or cluster to another (perhaps new) group. Also, this approach is not statistically efficient because one is not using all of the data in a single model. Multilevel statistical methods [1] present a third way of dealing with such effects by modelling both the differences *between* groups and differences *within* groups separately at different levels of the model.

Multilevel principal components analysis (mPCA) has previously been shown [2-5] to provide a simple and straightforward method of forming point distribution models, which are of use in active shape models (ASMs) [6-10]. One such previous application of mPCA in ASMs related to the segmentation of the human spine [2]. The authors stated that their results showed that “such a modelization offers more flexibility and allows deformations that classical statistical models can simply not generate.” They also noted that “the idea is to decompose the data into a within-individual and a between-individual component.” This is an important point: we have shown that mPCA analysis [4,5] of landmark points relating to 3D facial shape obtained from Konica Minolta Vivid laser cameras allows us to determine the relative importance of biological sex, ethnicity, in facial shape by examination of eigenvalues. Modes of variation made sense because changes in shape were seen to correspond to biological sex and ethnicity at the correct levels of the model and no “mixing” of these effects was observed [4,5]. Principal component “scores” also showed strong clustering, which were again at the correct levels of the mPCA model. These results have great implication to the analysis and modelling (and detection in some cases) of subtle effects on facial and cranial shape such as orthognathic surgery, age, fetal alcohol syndrome, and Treacher Collins syndrome, (etc.).

Here we extend the mPCA approach to model “image texture” as well as shape based on landmark points. By “image texture” we mean the pattern of intensities (or colors) across an image (or image patch) as in active appearance models

(AAMs) [11-16]. As a test case, we consider a set of (frontal) facial images from a group of 80 Finnish subjects (34 male; 46 female) each for two different facial expressions (smiling and neutral). A three-level model illustrated by Fig. 1 is constructed that contains biological sex, facial expression, and “between-subject variation” at different levels of the model. We find the mPCA model provides sensible results for both shape and image texture. We are able to answer the question “what’s in a smile?” We then present our conclusions.

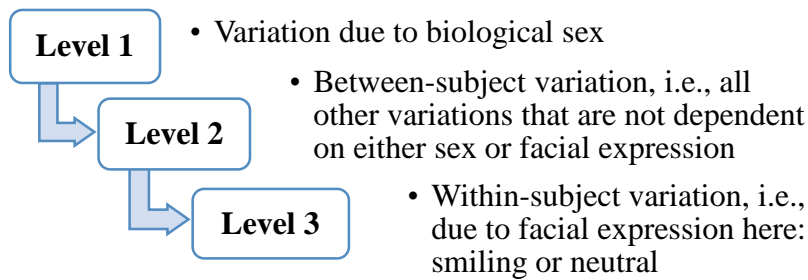


Fig. 1. Flowchart illustrating the multilevel model of facial shape

2 Methods

2.1 Multilevel Principal Components Analysis

Active shape models (ASMs) have previously been discussed extensively elsewhere [6-10]. However, it is worth noting that shapes are represented as vectors containing the components for a set of landmark points, z_i , after scaling, centering, and rotation (see below for details of preprocessing). Multilevel principal components analysis (mPCA) models variation in the sets of points due to different influences at the various levels of the model and separate covariance matrices are formed at each level. The implementation of mPCA to describe point distribution models used in ASMs is discussed in Refs. [2-5] and the interested reader is referred to these articles for more details about mathematical formalism. However, we note also here that one firstly carries out principal components analysis with respect to the covariance matrices at each level separately. This process allows us to isolate specific sources of variation in the point sets, which is encapsulated by the principal components at each level of the model. The eigenvalues tell us how much variation in the point sets is captured by each principal component: the larger the eigenvalue, the more important it is. Hence, eigenvalue plots yield direct quantitative evidence as to what the important factors are that affect facial shape. The l^{th} eigenvalue at level 1 (biological sex: male or female) is denoted by λ_l^1 with associated eigenvector u_l^1 , whereas the l^{th} eigenvalue at level 2 (between-subject variation) is denoted

by λ_l^2 with associated eigenvector is denoted by u_l^2 . Finally, the l^{th} eigenvalue at level 3 (within-subject variation) of the model is denoted λ_l^3 and its eigenvector is denoted by u_l^3 .

The maximum rank of a covariance matrix, which gives the maximum number of non-zero eigenvalues, is equal to $n - 1$, where n is the number of subjects or the number of groups at a given level for mPCA. Due to the homogeneity of the population (i.e., all Finnish and all 46 years old), the rank of the covariance matrix at level 1 of the models presented here is given exactly by one because it is evaluated with respect to only two groups only (male and female). Clearly, any restriction on the rank of the covariance matrix is a limitation of the mPCA model, although other multilevel methods (such as multilevel Bayesian approaches) ought not to be as strongly constrained as mPCA. An exploration of these topics will form the contents of future research.

We rank all of the eigenvalues into descending order at each level of the model separately, and then we retain the first l_1 , l_2 and l_3 eigenvectors of largest magnitude for the three levels, respectively. Any new shape is now modelled by

$$z = \bar{z} + \sum_{l=1}^{l_1} a_l^1 u_l^1 + \sum_{l=1}^{l_2} a_l^2 u_l^2 + \sum_{l=1}^{l_3} a_l^3 u_l^3 \quad , \quad (1)$$

where \bar{z} is the ‘‘grand mean.’’ The mPCA model is fitted to a shape by using all three levels in the model at the same time, which is achieved by finding the sets of weights $\{a_l^1\}, \{a_l^2\}, \{a_l^3\}$ via a global optimization procedure in MATLAB R2017 with respect to an overall cost / distance function between the new shape and the model of Eq. (1). These weights are also referred to as component ‘‘scores’’ here. Inspection of point-to-point errors suggests that the correct solution is identified, as errors tend to zero as the number of components / modes at each level is increased [3-5]. The component scores may finally be standardized to give ‘‘standard scores’’ by dividing the scores for each subject with respect each principal component by its standard deviation (i.e., the square root of its eigenvalue: $\{a_l^1 / \sqrt{\lambda_l^1}\}, \{a_l^2 / \sqrt{\lambda_l^2}\}, \{a_l^3 / \sqrt{\lambda_l^3}\}$) such that all scores can be plotted on the same (standardized) scale. The extension to image texture [11-16] was carried out straightforwardly. The vector z_i contained the grayscale values of 7339 pixels and the three-level mPCA model was constructed in exactly the same manner as for the analysis of shape. All calculations were carried out by using MATLAB R2017.

2.2 Image Capture, Preprocessing, and Subject Characteristics

Photographic images of the 80 adult Finnish subjects (34 female; 46 male) were obtained. Patients were from the Northern Finland Birth Cohort NFBC-66 and all patients were 46 years old when the images were collected. Ethical approval for this study was given by the Hospital District of Oulu Central Hospital. Images were captured at various distances from the lens and a small amount of variation in both head tilt and in illumination was present. As in previous studies [3-5], twenty one reliable facial landmarks¹ shown in Fig. 2 were found manually and these points are used here in the analysis of shape.

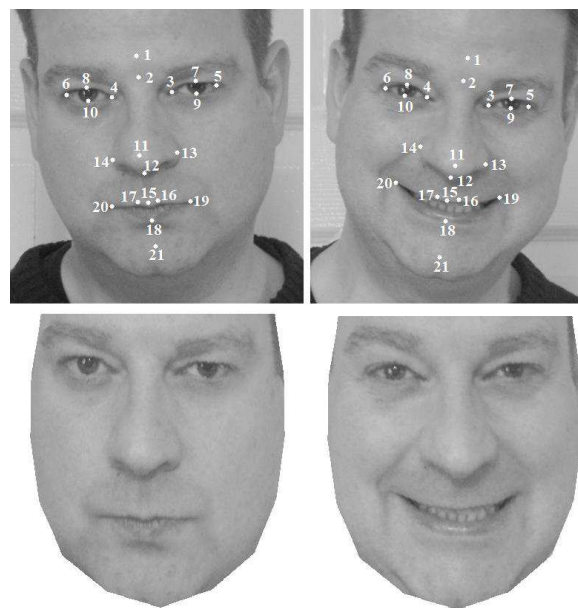


Fig. 2. Illustration of preprocessing for photographic images of one of the authors (DJF). (Top Left) Neutral facial expression; (Top Right) Smiling facial expression; (Bottom) images corrected for head tilt and overall illumination, and image texture region of interest (ROI) was defined.

The outline of the face was found also by placement of thirteen regularly spaced points. Color images were transformed into 16-bit grayscale images (Corel-Draw X5). The extension to colour images is straightforward, e.g., by having a

¹ (1) Glabella (g); (2) Nasion (n); (3) Endocanthion left (enl); (4) Endocanthion right (enr); (5) Exocanthion left (exl); (6) Exocanthion right (exr); (7) Palpebrale superius left (psl); (8) Palpebrale superius right (psr); (9) Palpebrale inferius left (pil); (10) Palpebrale inferius right (pir); (11) Pronasale (prn); (12) Subnasale (sn); (13) Alare left (all); (14) Alare right (alr); (15) Labiale superius (ls); (16) Crista philtri left (cphl); (17) Crista philtri right (cphr); (18) Labiale inferius (li); (19) Cheilion left (chl); (20) Cheilion right (chr); (21) Pogonion (pg).

vector, z , that is three times larger to represent the different RGB channels. In this initial study, we concentrate on gray-scale images only. The centerline of the face was identified from the set of landmark points, and this information was to rotate the shapes in order to correct for head tilt, shown in Fig. 2. These shapes were also scaled by the average distance to the centroid in order to correct for distance from the lens during image capture (length scales for the landmark points from the centroid of each shape were therefore of order 1) and these sets of points were finally translated so that their centroid was congruent with the origin. These shapes were then analyzed by using mPCA, as described above. A region of interest (ROI) / image patch based on the outline of the face using the thirteen landmark points was formed after rotation of the image data, again in order to correct for head tilt. All images were scaled in x - and y -directions independently such that the image size was 86×100 and the ROI (of total size 7339 pixels) was selected for each image, as illustrated by Fig. 2. Images were finally corrected for differences overall illumination level, also as illustrated by Fig. 2. Again, the analysis of image texture via mPCA was carried out separately to the analysis of shape in this initial study.

3 Results

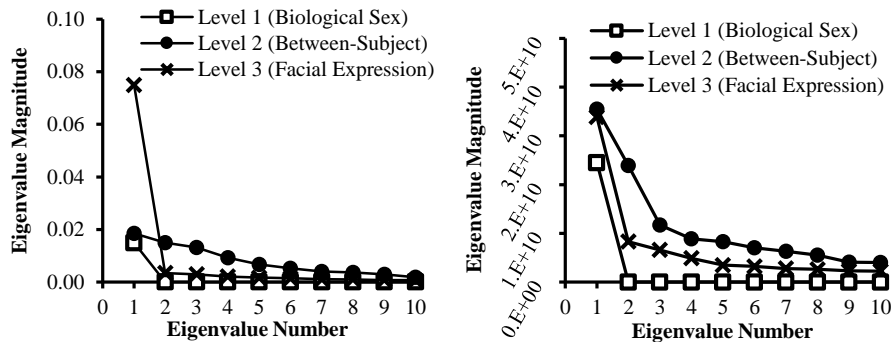


Fig. 3. Eigenvalues from mPCA level 1 (biological sex), level 2 (between-subject variation), and level 3 (within-subject variation: facial expression). (Left) shape data; (Right) image texture data. (All shapes have been scaled so that the average point-to-centroid distance is of order 1.)

Eigenvalues for both shape and also image texture via mPCA are shown in Fig. 3. These results for mPCA demonstrated a single non-zero eigenvalue for the level 1 (biological sex), which indicates that all of the variation in shape and image texture due to sex can be reduced to a single component (mode). A single large eigenvalue for the level 3 (facial expression) is evident for shape and also for image texture, although many non-zero (albeit of much smaller magnitude) eigenvalues are seen for image texture. These results suggest also that variation

in shape and image texture due to facial expression (smiling or neutral) is dominated by a single component / mode. Finally, the level 2 (between subjects) variation tends to have the largest number of non-zero eigenvalues for both shape and image texture. However, we see the first two eigenvalues for image texture are (relatively) large. mPCA results suggest that biological sex seems to contribute about 7.4% of variation relating to shape and image texture, whereas facial expression accounts for about 47.5% of variation of shape but only 37.1% variation of image texture. (Percentages were found by dividing the sum of eigenvalues at a given level by the sum over all levels; a similar approach was used in Ref. [17] to explain within-cluster variability.)

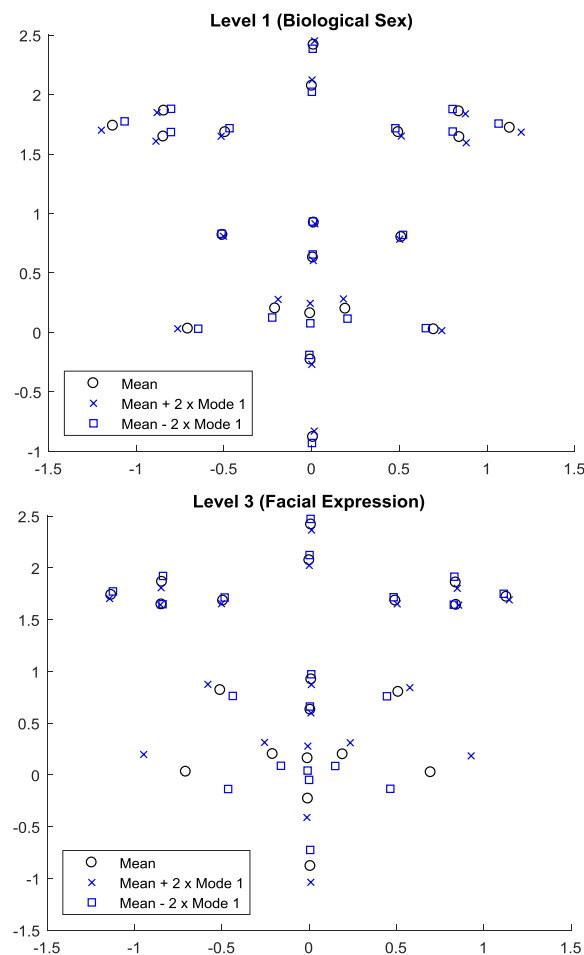


Fig. 4. Shape Variation: first mode of variation / principal components for mPCA at levels 1 and 3. (All shapes have been scaled so that the average point-to-centroid distance is of order 1.)

Modes of variation for shape via mPCA are presented in Fig. 4. We see that the first mode at level 3 (facial expression) for mPCA does indeed capture changes in facial expression (i.e., neutral to smiling and vice versa). Strong changes in mouth shape can be seen that relate clearly to the act of smiling. However, even subtle effects such as a slight increase in the width of the nose and narrowing of the eyes are also seen to occur during smiling for this mode of variation. We see that the eyes become further apart (relatively) and that the mouth becomes larger for the first mode at level 1 (biological sex) for mPCA. As all shapes have been scaled so that the average point-to-centroid distance is of order 1, this result makes sense because it has been seen previously (see, e.g., Refs. [5,18,19]) that women have a generally shorter and rounder faces than men. Finally, the first mode at level 2 (between-subject variation) for mPCA (not shown here) seems to correspond also to changes in the relative thinness / width of the face, which presumably can occur irrespective of sex.



Fig. 5. Image texture: first mode of variation / principal components for mPCA. (Top row) level 1 (biological sex); (Middle row) level 2 (between-subjects variation); (Bottom row) level 3 (facial expression). (Left column) mean - SD; (Middle column) mean; (Right column) mean + SD.

Modes of variation for image texture via mPCA are presented in Fig. 5. We see that mode 1 for level 1 (biological sex) mPCA does indeed correspond to

changes in appearance due to biological sex (e.g., as seen also in Ref. [19]: “females tend to have more prominent eyes and cheeks”), as required. Mode 1 for level 3 (facial expression) mPCA corresponds to changes due to the act of smiling (i.e., mean $-$ SD = not smiling, mean = half smile, and mean $+$ SD = full smile). We see clear evidence of a smile that exposes the teeth in this mode. Furthermore, subtle effects are again also seen for mPCA at this level such as: increased prominence of the cheeks; increased nose width; and, narrowing of the eyes. We note that there is no obvious evidence of changes due to change of facial expression or biological sex in modes at level 2 of the mPCA model, as expected. Mode 1 at level 2 of the mPCA model appears to account for residual changes in illumination, whereas modes 2 and 3 at level 2 (not show here) appear to account for small residual changes in facial position and appearance that is probably due to some subjects not facing directly into the camera.

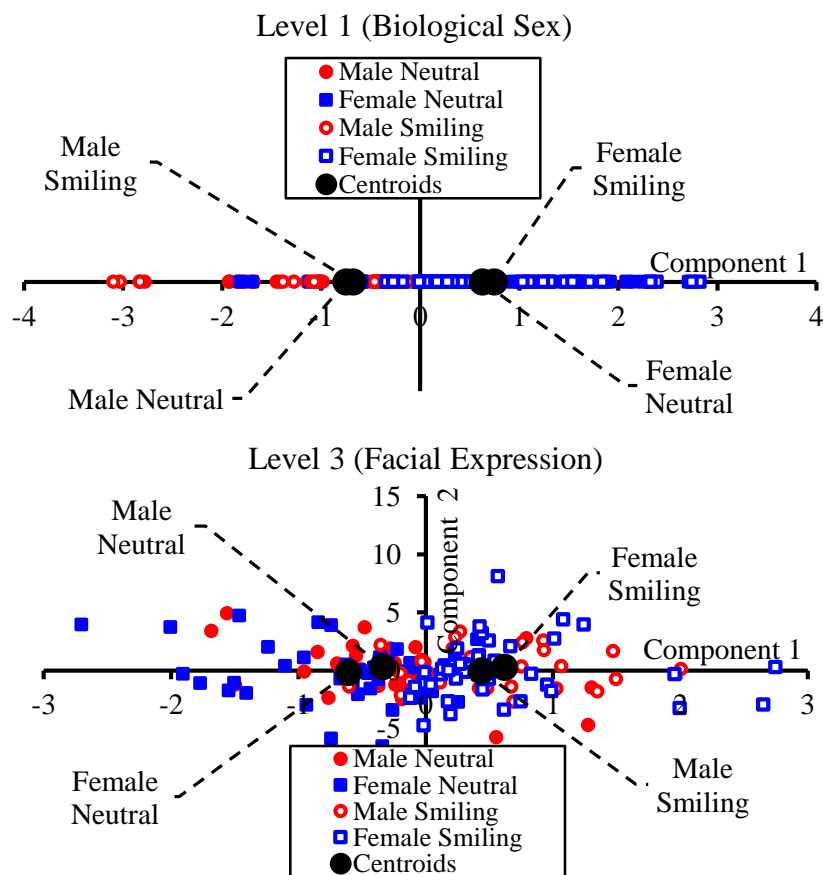


Fig. 6. Standardized component scores with respect to shape for levels 1 and 3 for mPCA ($l_1 = 1$, $l_2 = 26$; $l_3 = 3$).

Results for the standardized component “scores” for mPCA for shape are shown in Fig. 6. (Note that no constraints are placed on these coefficients in this case.) We see that the centroids in Fig. 6 at level 1 (biological sex) mPCA are strongly separated by biological sex, although not by facial expression. Hence, strong clustering by biological sex (alone) is observed for shape at level 1 (biological sex) mPCA, as required. By contrast, we see that the centroids in Fig. 5 at level 3 (facial expression) mPCA are strongly separated by facial expression (neutral, smiling), although not by biological sex. Hence, strong clustering by facial expression (alone) is observed at level 3 (facial expression) mPCA, also as required. Note that no strong clustering by facial expression or biological sex is seen at level 2 (between-subject variation) mPCA (not shown here), also as required. Very similar results were observed also for image texture, i.e., strong clustering by biological sex or facial expression (neutral or smiling) at the correct levels of the model.

4 Conclusions

Here we considered a set of (2D frontal) photographic facial images from a group of 80 Finnish subjects (34 male; 46 female) with two different facial expressions (smiling and neutral) per subject. A three-level mPCA model was constructed that included variation due to biological sex, between-subject variation (i.e., all other variations that are not dependent on sex or facial expression), and within-subject variations (i.e., facial expression: neutral and smiling) at different levels of the model. The major modes of variation for the facial expression level of the mPCA model clearly reflected changes in mouth, chin, nose, and eye shape and texture during smiling. Strong changes in mouth, exposure of teeth, and increased prominence of the cheeks in terms of shape and image texture were seen clearly. However, even very subtle effects as a result of the act of smiling were observed such as a slight flattening or widening of the nose (shape) and narrowing of the eyes. Facial expression explained 47.5% of variation of shape in our dataset and it explained 37.1% of variation of image texture. mPCA has thereby allowed us to answer the question: “what’s in a smile?” Similarly, the major mode for the biological sex level via mPCA gave results that made sense, e.g., it showed clearly that males have longer / thinner faces on average than women [5]. Model fits yielded “standardized scores” for each principal component / mode of variation that show strong clustering for both shape and texture by biological sex and facial expression at appropriate levels of the model. We conclude that mPCA correctly decomposes sources of variation due to biological sex and facial expression (etc.). This study has therefore been an excellent initial test of the usefulness of mPCA in terms of modelling either shape or image texture.

The complexity of the modelling scenario can be extended in two ways: e.g., by adding in more groups at a given level; or, by adding in further levels to the model. Both of these are straightforward to incorporate into the mPCA model. An example of adding more groups is given in Ref. [3] where mouth expressions (smiling, sad, and neutral) were explored for Monte Carlo simulated data that was analysed using mPCA at the “expression” level of this model. A similar extension to other facial expressions for the dataset considered in this paper would be achieved by allowing more than two groups at level 3 of the models for shape and image texture. An example of requiring more levels is presented in Ref. [5], where ethnicity for four groups (English, Welsh, Finnish, and Croatian) is included at a separate level of the model because this is completely distinct to variations due to, say, biological sex. Further studies will concentrate on correlating and modelling facial shape and texture to other variations such as body-mass index, age, and genetic profiles, although this lies beyond the scope of this initial study.

Clearly, the extension of the mPCA method to AAMs (e.g., of the type discussed in Refs. [11-16]) is readily achievable now that image texture has been shown to be tractable via this approach in these initial studies. Indeed, it is worth noting that linear discriminant analysis (LDA) has been applied to modelling facial appearance previously (see, e.g., Refs. [11,13,20,21]), and this approach is similar in spirit to mPCA in that it creates a distinction between- and within-group effects. In particular, LDA was used in Ref. [11] to “isolate the changes in appearance due to different sources of variability (person, expression, pose, lighting).” A strength of the mPCA approach, in particular, is that it allows us to design bespoke multilevel models that isolate specific factors that might affect the shapes or images. We can then judge if these effects are important or not quantitatively by inspecting eigenvalues. A model can thereby be created that can adjust or account for (confounding) factors such as subject age, gender, ethnicity, or (in principle) even time for video data. This approach might yield much insight in future into applications of facial morphometrics where (often subtle) changes that can occur in facial and cranial shape due to orthognathic surgery, sex, ethnicity, age, fetal alcohol syndrome, Treacher Collins syndrome, (etc.).

References

1. Gelman A, Hill J (2007) Data analysis using regression and multilevel/hierarchical models (Cambridge University Press, Cambridge, UK).
2. Lecron F, Boisvert J, Benjelloun M, Labelle H, Mahmoudi S (2012) Multilevel statistical shape models: A new framework for modeling hierarchical structures, 9th IEEE International Symposium on Biomedical Imaging (ISBI) 1284-1287.

3. Farnell DJJ, Popat H, Richmond S (2016) Multilevel principal component analysis (mPCA) in shape analysis: A feasibility study in medical and dental imaging, *Computer Methods and Programs in Biomedicine* **129**: 149-159.
4. Farnell DJJ, Galloway J, Zhurov A, Richmond S, Perttiniemi P, Katic V (2017) Initial Results of Multilevel Principal Components Analysis of Facial Shape, *Annual Conference on Medical Image Understanding and Analysis*. Springer, Cham: 674-685. ISBN 978-3-319-60964-5.
5. Farnell DJJ, Galloway J, Zhurov A, Richmond S, Perttiniemi P, Katic V (2018) An Initial Exploration of Ethnicity, Sex, and Subject Variation on Facial Shape – in preparation.
6. Cootes TF, Hill A, Taylor CJ, Haslam J (1994) Use of Active Shape Models for Locating Structure in Medical Images. *Image and Vision Computing* **12**: 355-365.
7. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active Shape Models - Their Training and Application, *Computer Vision and Image Understanding* **61**: 38-59.
8. Hill A, Cootes TF, Taylor CJ (1996) Active shape models and the shape approximation problem, *Image and Vision Computing* **14**: 601-607.
9. Cootes TF, Taylor CJ (1999) A mixture model for representing shape variation, *Image and Vision Computing* **17**: 567-573.
10. Allen PD, Graham J, Farnell DJJ, Harrison EJ, Jacobs R, Nicopolou-Karayianni K, Lindh C, van der Stelt PF, Horner K, Devlin H (2007) Detecting reduced bone mineral density from dental radiographs using statistical shape models, *IEEE Trans.Inf.Technol.Biomed* **11**: 601-610.
11. Edwards GJ, Lanitis A, Taylor CJ, and Cootes T. Statistical Models of Face Images: Improving Specificity. In *British Machine Vision Conference 1996*, Edinburgh, UK, 1996.
12. Taylor CJ, Cootes TF, Lanitis A, Edwards G, Smyth P, Kotcheff ACW (1997) Model-based interpretation of complex and variable images, *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **352**: 1267-1274.
13. Edwards GJ, Cootes TF, Taylor CJ (1998) Face recognition using active appearance models. In: Burkhardt H., Neumann B. (eds) *Computer Vision — ECCV'98*. ECCV 1998. *Lecture Notes in Computer Science* **1407**: 581-595.
14. Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**: 681-685.
15. Cootes TF, Taylor, CJ (2004) Anatomical statistical models and their role in feature extraction, *British Journal of Radiology* **77**: S133-S139.
16. Matthews I, Baker S (2004) Active Appearance Models Revisited *International Journal of Computer Vision* **60**: 135-164.
17. Di C-Z, Crainiceanu CM, Caffo BS, Punjabi NM (2009) Multilevel functional principal component analysis, *The Annals of Applied Statistics* **3**, 458-488.
18. Velemínská J, Bigoni L, Krajiček V, Borský J, Šmahelová D, Cagáňová V, Peterka M (2012) Surface facial modelling and allometry in relation to sexual dimorphism, *HOMO - Journal of Comparative Human Biology* **63**: 81-93.
19. Toma AM, Zhurov A, Playle R, Richmond S (2008) A three-dimensional look for facial differences between males and females in a British-Caucasian sample aged 15 ½ years old. *Orthodontics and Craniofacial Research* **11**: 180-185.

20. Etemad K, Chellappa R (1997) Discriminant analysis for recognition of human face images. *Journal of the Optical Society of America A* **14**: 1724-1733.
21. Kim T-K, Kittler J (2005) Locally Linear Discriminant Analysis for Multi-modally Distributed Classes For Face Recognition with a Single Model Image. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27**: 318–327.