Initial Results of Multilevel Principal Components Analysis of Facial Shape

DJJ Farnell[1]† (ORCID: 0000-0003-0662-1927), J Galloway[1] (ORCID: 0000-0003-4800-5970), A Zhurov[1] (ORCID: 0000-0002-5594-0740), S Richmond[1] (ORCID: 0000-0001-5449-5318), P Perttiniemi[2] (ORCID: 0000-0003-4514-836X), and V Katic[3] (ORCID: 0000-0003-3299-8306)

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

[2]Faculty Of Medicine, P.O.Box 5000, 90014, University Of Oulu, Finland

[3]Department of Orthodontics, School of Medicine, University of Rijeka, Kresimirova 40, HR-51000, Rijeka, Croatia

FarnellD@cardiff.ac.uk († corresponding author)
GallowayJL@cardiff.ac.uk
ZhurovAI@cardiff.ac.uk
RichmondS@cardiff.ac.uk
visnja.katic@gmail.com
pertti.pirttiniemi@oulu.fi

Abstract. Traditionally, active shape models (ASMs) do not make a distinction between groups in the subject population and they rely on methods such as (single-level) principal components analysis (PCA). Multilevel principal components analysis (PCA) allows one to model between-group effects and within-group effects explicitly. Three dimensional (3D) laser scans were taken from 240 subjects (38 Croatian female, 35 Croatian male, 40 English female, 40 English male, 23 Welsh female, 27 Welsh male, 23 Finnish female, and 24 Finnish male) and 21 landmark points were created subsequently for each scan. After Procrustes transformation, eigenvalues from mPCA and from single-level PCA based on these points were examined. mPCA indicated that the first two eigenvalues of largest magnitude related to within-groups components, but that the next largest eigenvalue related to between-groups components. Eigenvalues from single-level PCA always had a larger magnitude than either within-group or between-group eigenvectors at equivalent eigenvalue number. An examination of the first mode of variation indicated possible mixing of between-group and within-group effects in single-level PCA. Component scores for mPCA indicated clustering with country and gender for the between-groups components (as expected), but not for the within-group terms (also as expected). Clustering of component scores for single-level PCA was harder to resolve. In conclusion, mPCA is viable method of forming shape models that offers distinct advantages over single-level PCA when groups occur naturally in the subject population.

Keywords: multilevel principal components analysis; active shape models; facial shape

1 Introduction

Active shape models (ASMs) and active appearance models (AAMs) [1-8] are common techniques in image processing that are used to search for specific features or shapes in images. However, if clustering or multilevel data structures exist naturally in the data set, e.g., as illustrated by the flowchart in Fig. 1, the eigenvectors and eigenvalues from principal components analysis (PCA) will only be partially reflective of the true variation in the set of images / shapes. Multilevel principal components analysis (mPCA) provides a convenient method of modelling both the underlying structures within the images and also any groupings between images. mPCA carries out PCA at both within-group and between-group levels independently. Note that the within-group level might be thought of as being "nested" within the broader between-group level, e.g., as shown in Fig. 1 for human facial expression. This approach also retains the desirable feature that any segmentation can still be constrained so that a fit of the model never "strays too far" from the training set used in forming the model (described in the methods section below).

Fig. 1. Flowchart illustrating the "nested" nature of multilevel data

A previous application of mPCA to form ASMs related to the segmentation of the human spine [9]. The results of this study showed that mPCA offers more flexibility and allows deformations that classical statistical models cannot generate. Another recent application of using mPCA to form ASMs related to the field of dental imaging [10]. Proof-of-principle was tested by applying mPCA to model basic peri-oral expressions that were approximated to the junction between the mouth/lips. Monte Carlo simulation was used to create the data set, where a simple quadratic function $y = cx^2$ was used to represent the centreline of the lips and the value of c controlled "expression." Different expressions (i.e., –ve c = sad; $c \approx 0$ = neutral; +ve c = happy) were modelled correctly at the between-group level of the model and changes in lip width were modelled correctly at the within-group level. Some evidence was seen that those cases that were extreme (yet still possible) in the training set in terms of both the within-group variation (width of lips) and also the between-group variation (expression) were modelled adequately by mPCA but not by standard (single-level) PCA. mPCA was also used to analyse a dataset that had landmark points placed on panoramic mandibular radiographs

by two different clinicians (see also Ref. [8]), thus leading to two sets of such landmark points for the set of images. Variations in the shape of the cortical bone were modelled by one level of mPCA (within-group) and variations between the experts at another (between-group). Not surprisingly, eigenvalues indicated that variation due to changes the shape of the cortical bone were much more important than variation due to any disagreements in placement between the clinicians. Indeed, these clinicians had reported anecdotally [8] that placement of the point along the boundaries was difficult and it was observed the first mode of variation for the between-group level correctly reflected this type of variation. The authors concluded [10] that mPCA was found to provide more control and flexibility than standard "single-level" PCA when multiple levels occurred naturally in the dataset.

Here we apply mPCA to study landmark points of three-dimensional (3D) laser scans of the heads of English, Croatian, Finnish, and Welsh subjects who were of both genders. Details of the mathematics that underpins mPCA for ASMs and also of the 3D laser scanning procedure are presented in the methods section. Results are then presented for the eigenvectors and eigenvalues, and component scores are found by fitting the mPCA model to each set of points for each subject in the dataset. Results of mPCA are compared to those results of standard (single-level) PCA. The major modes of variation are explored. The implications of our research are presented in the discussion.

2 Methods

2.1 Mathematical Formalism

The ASM method has been extensively documented in the literature (see, e.g., Refs. [1-8]), and therefore this topic is not discussed here. One carries out PCA for the covariance matrix as discussed in Ref. [8], and the eigenvalues and eigenshapes (i.e., eigenvectors) are found readily enough for this matrix using standard software. Landmark points (i.e., mark-up points) are represented by a vector, z_i, and the k^{th} element of this vector is given by z_{ik}. The total number of such points is n, and the mean shape vector (averaged over all N subjects) is given by $\bar{z}$. The covariance matrix is found by evaluating

$$C_{k_1,k_2} = \frac{1}{N-1}\sum_{i=1}^{N}(z_{ik_1} - \bar{z}_{ik_1})(z_{ik_2} - \bar{z}_{ik_2}) \tag{1}$$

where k_1 and k_2 indicate elements of the covariance matrix. We find the eigenvalues λ_l and eigenvectors u_l of this matrix. Note that all of the eigenvalues are non-negative, real numbers because covariance matrices are symmetric and (indeed) positive semi-definite. We rank all of the eigenvalues λ_l into descending order and we choose the largest m eigenvalues to be retained in the model. Any new shape is given by

$$z = \bar{z} + \sum_{l=1}^{m} a_l u_l \quad . \tag{2}$$

The eigenvectors u_l are orthonormal and so we can determine the coefficients, a_l, for a fit of the model to a new shape vector, z, readily by using

$$a_l = u_l \cdot (z - \bar{z}) \qquad . \tag{3}$$

Constraints may be placed on these a-coefficients, such as $|a_l| \le 3\sqrt{\lambda_l}$, which ensures that subsequent model fits to a new shape vector never "strays too far" from the cases in the training set.

The formalism is slightly more complicated for mPCA and details are presented in Ref. [10]. However, we remark here that we form two covariance matrices for a two-level model, namely: a *within-group* covariance matrix which is the covariance matrix evaluated over all subjects with a group and with respect to their local group means or centroids, and this matrix is then averaged all groups; and, a *between-group* covariance matrix that is covariance matrix of the centroids of the groups with respect to an "grand" mean shape $\bar{\bar{z}}$ of the average of these centroids. The rank of this matrix is limited by the number of groups.

We carry out PCA for the (positive semi-definite) *within-group* covariance matrix of the above equation and the eigenvalues are non-negative, real numbers. The l^{th} eigenvalue is denoted λ_l^w and its eigenvector is denoted by u_l^w. Independently, we carry out PCA also for the (positive semi-definite) *between-group* covariance matrix given above and the eigenvalues are non-negative, real numbers. The l^{th} eigenvalue is denoted λ_l^b and its eigenvector is denoted by u_l^b. We rank all of the eigenvalues λ^b and λ^w into descending order for the between- and within-group levels separately, and then we retain the m_b and m_w largest such eigenvectors, respectively. Any new shape is now given by

$$z = \bar{\bar{z}} + \sum_{l_1=1}^{m_w} a_{l_1}^w u_{l_1}^w + \sum_{l_2=1}^{m_b} a_{l_2}^b u_{l_2}^b \qquad . \tag{4}$$

Constraints may again be placed on these a-coefficients, such as $\left|a_l^b\right| \le 3\sqrt{\lambda_l^b}$ and $\left|a_l^w\right| \le 3\sqrt{\lambda_l^w}$, which ensures that subsequent model fits to a new shape vector again never "stray too far" form the cases in the training set with respect to both within-group variation *and* between-group variation.

The covariance matrices are symmetrical and so all "within" eigenvectors u_l^w are orthogonal to all other "within" eigenvectors (and similarly for the "between" eigenvectors). However, the eigenvectors u_l^w and u_l^b do not necessarily have to orthogonal with respect to each other, and so an equivalent projection to Eq. (3) for mPCA becomes problematic. A fit of the model given by Eq. (4) to a set of candidate points is achieved

by minimising the overall (squared) error with respect to the coefficients a_l^w and a_l^b. A gradient descent method (see Ref. [10] for details) may be implemented straightforwardly to solve this problem iteratively. Importantly, note that the within and between components of variation are fitted to the set of candidate points *simultaneously*. All analyses were carried out using MATLAB R2014a.

2.2 3D Laser Scanning

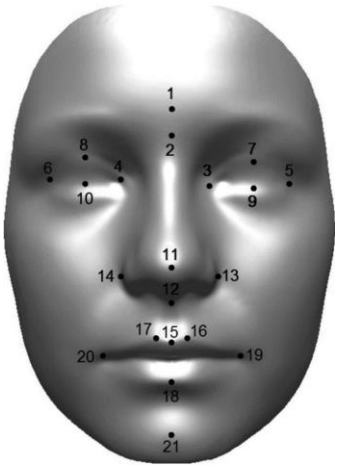

Fig. 2. Twenty-one anthropometric landmarks which were identified on facial laser scans of participants (shown in the coronal plane). (1) Glabella (g); (2) Nasion (n); (3) Endocanthion left (enl); (4) Endocanthion right (enr); (5) Exocanthion left (exl); (6) Exocanthion right (exr); (7) Palpebrale superius 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). Definitions by Farkas [14] were used. Reprinted from the author's previous publication with permission from 'John Wiley and Sons'.

Two Konica Minolta Vivid laser cameras were used to capture the images of the subjects and this has been reported extensively in the literature [11-13]. Twenty-one reliable facial landmarks were manually identified for each subject. Each landmark point vector z was of size 63 (= 21 × 3). These facial landmarks are shown in Fig. 2. The numbers of subjects in 8 groups were: Croatian female (n = 38), Croatian male (n = 35), English female (n = 40), English male (n = 40), Welsh female (n = 23), Welsh male (n = 27), Finnish female (n = 23), and Finnish male (n = 24). Landmark points were scaled by Procrustes transformation so that all sets of points were on broadly the same scale.

3 Results

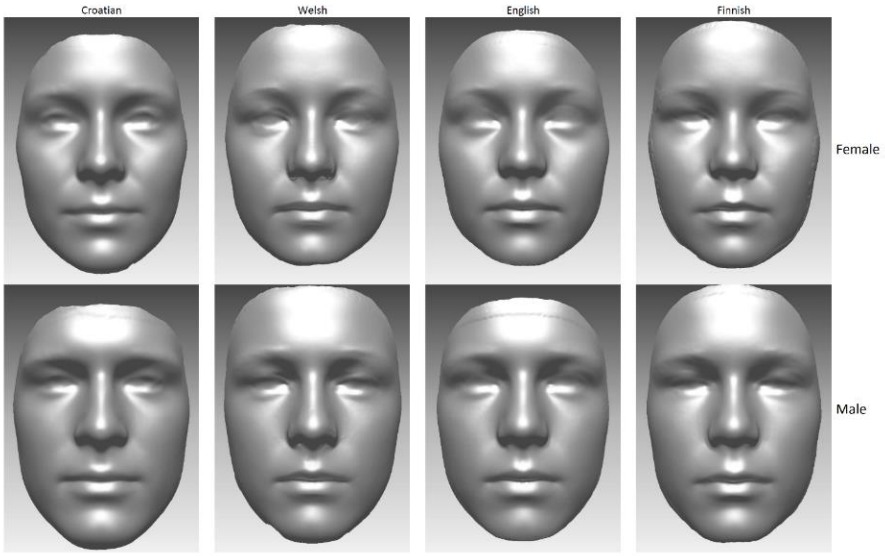

Fig. 3. 3D facial scans averaged over all subjects in each group (country and gender)

Results for the 3D facial scans averaged over all subjects in each group (country and gender) are shown in Fig. 3. As one might expect [15], strong differences in facial shapes can be seen qualitatively by both country and gender. Hence, we might reasonably expect to see commensurate differences in landmark points between groups.

Eigenvalues from standard (single-level) PCA and also within-group and between-group eigenvalues from mPCA are presented in Fig. 4. Results of between-groups mPCA demonstrate that between-group eigenvalues are non-zero for the first seven eigenvalues with $\lambda_l \leq 10^{-15}$ for $l \geq 8$. Note that within-group eigenvalues are clearly non-zero to much higher eigenvalue numbers, although we find that $\lambda_l \leq 2\times10^{-8}$ for $l \geq 58$. The first two largest eigenvalues from mPCA are due to within-groups effects and then the first and second between-groups eigenvalues are broadly of the same magnitude is the third and fourth within-group eigenvalues. These results indicates that both within-groups and between-groups effects are important. Eigenvalues from standard (single-level) PCA lie above those of eigenvalues from both within-group and between-group mPCA at equivalent eigenvalue number.

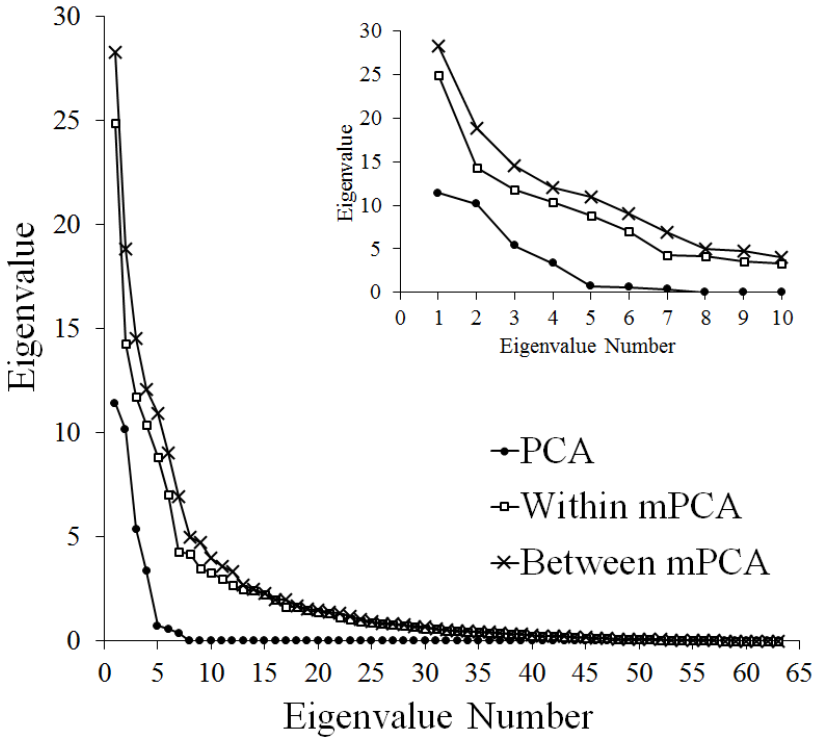

Fig. 4. Eigenvalues from standard (single-level) PCA and also within-group and between-group eigenvalues from mPCA. (Inset: the first ten eigenvalues in more detail.)

The effects of the principal components of within-group and between-groups mPCA can be investigated by considering the mean shapes ± 2 × standard deviations (i.e., $\lambda^{0.5}$) multiplied by its corresponding eigenvector. The first mode for single-level PCA and the first mode for the within-group mPCA were found to be similar in the coronal plane, e.g., deviations from the mean shape are large at the exocanthion (right and left) positions, as shown in Fig. 5. Broadly, one might equate this mode to the aspect ratio of the face. Variations due to the first mode of variation for the between-group mPCA appeared to vary little from the mean shape in this plane. For the transverse and sagittal planes (note shown here), deviations from the mean shape were large at the exocanthion (right and left) positions; again, the first mode for single-level PCA and the first mode for the within-group mPCA are similar here. By contrast, the first mode for between-groups mPCA had strong deviations from the mean at many points for the transverse and sagittal planes, and again it is quite different to the first mode for within-groups mPCA. Particularly, strong deviations from the mean were seen for the first mode for between-groups mPCA for the pronasale, and this was also seen at this point for the first mode for PCA. This hints that this mode might govern the length and shape of the nose and / or face in this plane. In any case, it is clear though that the first modes for within-group mPCA and between-group mPCA are quite different. We might also speculate that the first mode of single-level PCA might mix the effects of modes from

within-group and between-group mPCA, although this is difficult to judge in 2D. Indeed, many such subtle effects occur even in the first major modes and visualising such subtle changes in 2D plots is difficult. However, the 3D visualisation of these modes, and their subsequent interpretation, lies beyond the scope of this initial analysis.

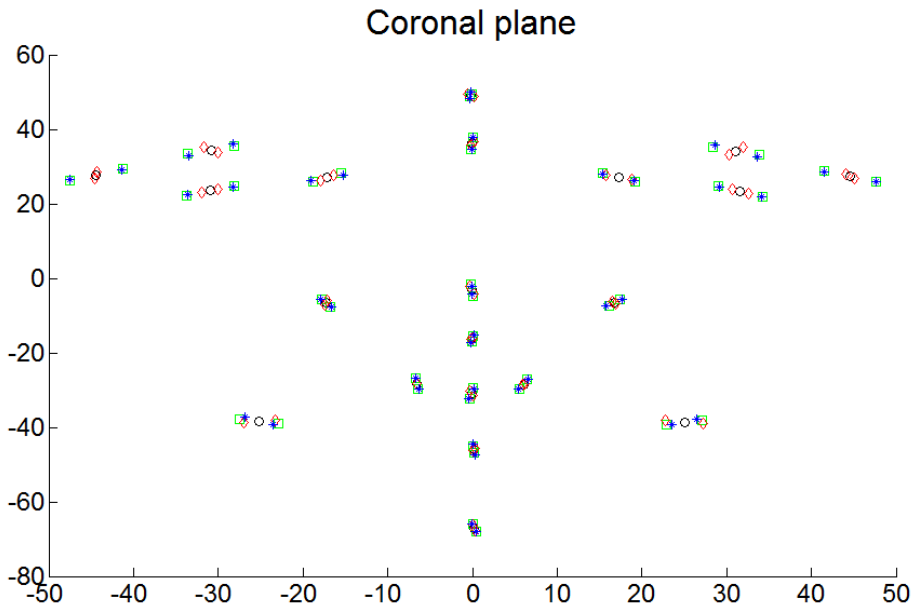

Fig. 5. First modes of variation in the coronal plane. (○ = mean; □ = mode 1 for PCA; ◊ = mode 1 for between-group mPCA; * = mode 1 for within-group mPCA.)

Component "scores" (i.e., coefficients a in Eqs. (2) and (4)) may be found by fitting the single-level PCA and mPCA models to each set of landmark points for each subject in the data set. (Note that no constraints are placed on these coefficients in this case.) Results for single-level PCA, shown in Fig. 6, indicate some evidence of clustering for the different groups. Centroids for males are on the right-had side of the figure for a_1 versus a_2 for PCA and those for females are on the left-hand side. Furthermore, centroids by country tend to be quite close to each other for a_1 versus a_2 and this is shown by the solid lines connecting centroids of the different genders for the same country. However, there is considerable overlap between the groups in the scores for individual subjects for single-level PCA for these component scores (i.e., a_1 versus a_2).

Results for between-groups components of mPCA are shown in Fig. 7 for $m_b = 7$ and $m_w = 40$ and distinct indications of clustering for the different groups is seen. (These results were typical of mPCA results generally for $m_b \geq 3$ and $m_w \geq 3$.) It is remarkable that males and females are connected by a vector that is of similar direction and magnitude for all countries. Note that this result is not imposed by assumptions of the model (as far as we are aware) and that it seems to emerge naturally from the data. This result is shown by the solid lines connecting centroids of the different genders for the same

country. Furthermore, we see that centroids are being separated quite strongly by country more clearly for the between-group mPCA. The centroids of each group are certainly much easier to resolve for single-level PCA. Although there is overlap in the individual scores between the groups, this overlap appears to be less than for single-level PCA.

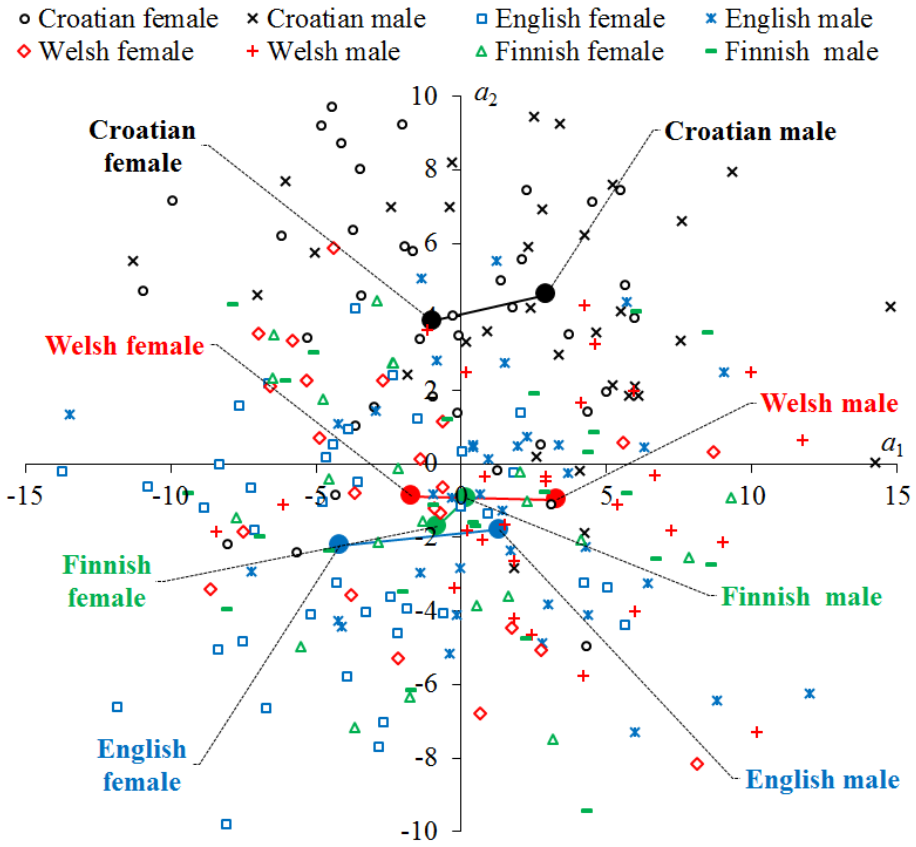

Fig. 6. Component scores from standard (single-level) PCA. Large circles indicate the centroids and results for females and males from the same country are linked by a line.

Results for within-groups mPCA (not shown here) indicate no evidence of clustering for the different groups and all of the centroids were found to lie very close to each other (i.e., at the origin); this is exactly what one would expect as differences between groups ought to be accounted for by between-group components alone. The scatter of points for individual scores about their centroids was found be broadly uniform and of similar magnitude for all of the groups. Indeed, one would also expect the centroids to lie near to the origins because subjective variations ought to be equally spaced both "above" and "below" the group averages.

Results for component scores for the first versus the third components or the second versus third components also showed strong clustering for the between mPCA scores. The centroids for the within mPCA component scores for the first versus the third components or the second versus third components were near to the origin. Some evidence of clustering was seen for the first versus the third components and the second versus third components for single-level PCA.

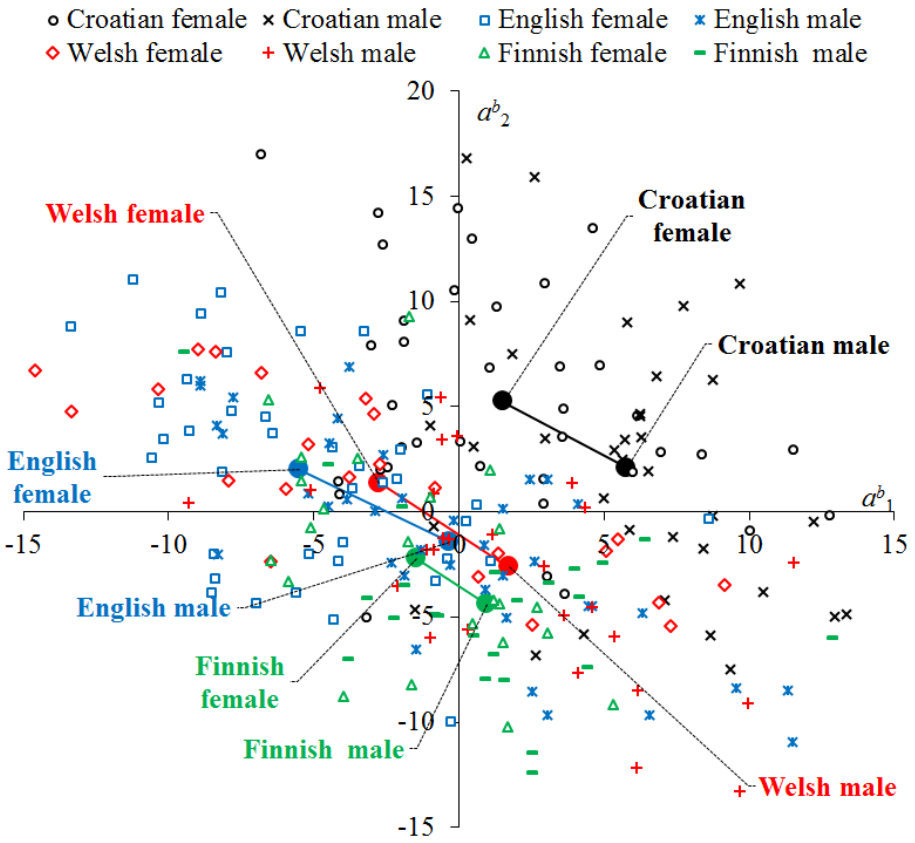

Fig. 7. Component scores from between-group mPCA. Large circles indicate the centroids and results for females and males from the same country are linked by a line.

Results for the point-to-point errors are shown in Table 1. Note that this was not an image search as such, rather a model fit using Eq. (3) for PCA and the iterative procedure for mPCA. Indeed, "test" cases were taken to be the same as those used the training set. As expected, point-to-point errors are found (trivially) to reduce for both PCA and mPCA as we increase the number of eigenvectors retained. However, this is still a reasonable check of mPCA because it demonstrates that the correct iterative solution is probably being identified. Point-to-point errors for mPCA are of the same magnitude

(or slightly lower) than single-level PCA at broadly "equivalent" numbers of eigenvectors retained, which is also reasonable. If we include all possible eigenvectors then point-to-point errors appear to go to zero for both PCA and mPCA.

Table 1. Point-to-point errors with respect to all 21 points and all subjects

		Mean	Standard Deviation	Maximum
PCA	$m = 3$	1.851	0.333	2.993
	$m = 5$	1.617	0.271	2.452
	$m = 7$	1.421	0.242	2.299
	$m = 20$	0.781	0.127	1.199
	$m = 40$	0.262	0.059	0.446
mPCA	$m_b = 3$ and $m_w = 3$	1.580	0.290	2.637
	$m_b = 5$ and $m_w = 5$	1.289	0.227	2.084
	$m_b = 7$ and $m_w = 7$	1.106	0.189	1.876
	$m_b = 7$ and $m_w = 20$	0.636	0.121	1.049
	$m_b = 7$ and $m_w = 40$	0.167	0.048	0.340

4 Conclusions

The effect of naturally occurring groups in the subject population for facial shape data has been explored in this article. The formalism for mPCA has been described briefly and we have shown that mPCA allows us to model variations at different levels of structure in the data (i.e., between-group and within-group levels). Examination of eigenvalues showed that both between and within-group sources of variation were important. Modes of variation appeared to make sense: the first mode of the within-group mPCA seemed to govern width of the face in the coronal plane and the first mode of the between-group mPCA seemed to govern length and shape of the nose and / or face in the transverse and sagittal planes. We have also demonstrated that initial results of mPCA for facial shape data appear to show evidence clustering in the between-group component scores. Indeed, a consistent relationship between genders was observed for each country. Evidence of clustering was also observed for single-level PCA, although the nature of this clustering was less clear. Point-to-point errors of model fits for both mPCA and PCA reduced with the number of eigenvectors retained, as expected. This research is an excellent first step in evaluating the usefulness and feasibility of mPCA in analysing facial shape.

References

1. 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.
2. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active Shape Models - Their Training and Application, Computer Vision and Image Understanding **61**: 38-59.

3. Hill A, Cootes TF, Taylor CJ (1996) Active shape models and the shape approximation problem, Image and Vision Computing **14**: 601-607.
4. 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.
5. Cootes TF, Taylor CJ (1999) A mixture model for representing shape variation, Image and Vision Computing **17**: 567-573.
6. Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models, Ieee Transactions on Pattern Analysis and Machine Intelligence **23**: 681-685.
7. Cootes TF, Taylor, CJ (2004) Anatomical statistical models and their role in feature extraction, British Journal of Radiology **77**: S133-S139.
8. 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.
9. 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.
10. 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.
11. Kau CH, Cronin A, Durning P, Zhurov AI, Sandham A, Richmond S (2006) A new method for the 3D measurement of postoperative swelling following orthognathic surgery. Orthodontics and Craniofacial Research, **9**: 31-37.
12. Kau CH, Hartles FR, Knox J, Zhurov A I, Richmond S (2005) Natural head posture for measuring three-dimensional soft tissue morphology. In: Middleton J , Shrive M G , Jones M L (eds). Computer methods in biomechanics and biomedical engineering – 5 First Numerics Ltd, Cardiff University.
13. Kau CH, Richmond S (2008) Three-dimensional analysis of facial morphology surface changes in untreated children from 12 to 14 years of age. American Journal of Orthodontics and Dentofacial Orthopedics **134**: 751-760.
14. Farkas LG (1994) Anthropometry of the head and face. New York: Raven Press, pp 21-5.
15. Hopman SM, Merks JH, Suttie M, Hennekam RC, Hammond P (2014) Face shape differs in phylogenetically related populations. Eur J Hum Genet. **22**: 1268-71.