Some Issues of Biological Shape Modelling with Applications

Rasmus Larsen, Klaus Baggesen Hilger, Karl Skoglund, Sune Darkner, Rasmus R. Paulsen, Mikkel B. Stegmann, Brian Lading, Hans Henrik Thodberg, and Hrafnkell Eiriksson

Informatics and Mathematical Modelling, Technical University of Denmark
Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, Denmark
rl@imm.dtu.dk, http://www.imm.dtu.dk/image

Abstract. This paper illustrates current research at Informatics and Mathematical Modelling at the Technical University of Denmark within biological shape modelling. We illustrate a series of generalizations to, modifications to, and applications of the elements of constructing models of shape or appearance. These elements are correspondence analysis, analysis and decomposition of variability, alignment, and visualisation.

1   Introduction

Statistical models of shape and appearance are estimated from a series of training examples. These shape examples sometimes come in the form of manually identified point sets, more often the training examples are provided as delineated object boundaries. In the latter case we first need to infer the correspondences of these curves and surfaces. The minimum description length principle has successfully been applied to this task [1]. We show how local shape characteristics of 2D curves can aid in the process. The same principle of regularising the correspondence analysis using local shape characteristics is applied in a Markow random field formulation of iterative point matching [2].

The next step is to filter out nuissance parameters, e.g. Euclidean similarity transformations. This is done using generalised Procrustes analysis [3, 4]. This is a least squares methods and thus not necessarily resistant to outliers. We describe an alternative based on minimising the sum of absolute deviations. Next we need tools to analysis and help understand the shape variability. It is common to decompose shape tangent space variability using principal components analysis (PCA) [5] and inspect the resulting modes of variations. We show how canonical correlation analysis [6] and cluster analysis can be used to understand the interrelations between landmarks.

Principal components is the favored method of decomposing shape variability. However, in many situations regression type techniques would be more obvious. We show an example of using partial least squares regression [7] to find modes of variation that describe bone growth.

In appearance models [8] shape and intensity information is modelled simultaneously. We show that it can be advantageous to consider intensity variation

not in the original pixel domain but rather in a wavelet domain as proposed in [9]. Furthermore, we introduce the use of wedgelets [10, 11] to represent image date in triangulated domains. In a final issue regarding shape and appearance models we show how a mixture of Gaussians can be used to represent multi-modal distributions effectively in a case of modelling the human heart across the heart cycle in a MR perfusion imaging experiment. Finally, we present a case of making 3d shape models of human faces from a series of 3D scannings as proposed by [12, 13].

2 Optimising landmark correspondence using minimum description length

Often training data for 2D shape analysis come in the form of delineated object boundaries. Prior to shape analysis it is therefore necessary to establish correspondences between such curves. A successful solution to this problem is the minimum description length (MDL) approach proposed in [1]. However, this MDL approach does not always capture important shape characteristics related to the curvature of the curves, and occasionally it places marks in obvious conflict with the human notion of point correspondence. The standard MDL method seeks compact description of the positions of the shape points. Curvature is proposed as another salient piece of information. By requiring the model to describe both position and curvature we get a different optimum with point correspondences that matches both positions and curvatures. The correspondences across the data set are then optimised with respect to the following costs

$$\mathrm{DL} = \sum_{\lambda > \lambda_{\mathrm{cut}}} 1 + \log \frac{\lambda_m}{\lambda_{\mathrm{cut}}} + \sum_{\lambda \leq \lambda_{\mathrm{cut}}} \frac{\lambda_m}{\lambda_{\mathrm{cut}}} \qquad [\text{MDL positions cost}]$$

$$\mathrm{NC} = \sum_i (a_i^{\mathrm{average}} - a_i^{\mathrm{target}})^2 \qquad [\text{Node positions cost}]$$

$$\mathrm{CC} = \sum_{i,r} (k_{ir} - k_i^{\mathrm{mean}})^2 \qquad [\text{Curvature cost}]$$

λ_m are the eigenvalues of a principal components decomposition of landmark positions, λ_{cut} is a preset threshold; a_i^{average} and a_i^{target} are average and target parameters for the ith landmark. The node cost secures that landmark does not pile up in some regions and dilute in others; k_{ir} is the curvature at ith landmark on rth shape, k_i^{mean} is the mean curvature at the ith landmark across all shapes. This curvature cost ensures the same curvature signature across shapes. In Fig. 1 we see how the inclusion of the curvature cost let landmarks fall to rest at more biologically intuitive points.
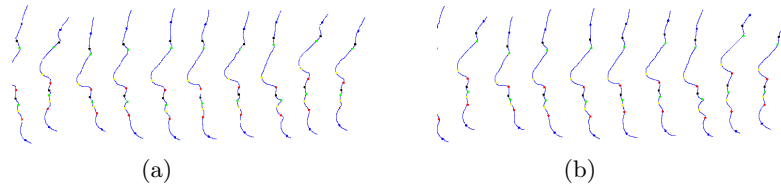
Fig. 1. Correspondences estimated using MDL without curvature cost; (b) correspondences using curvature cost.

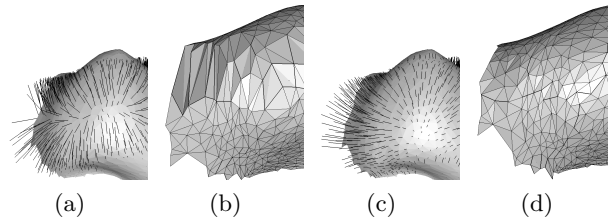(a)   (b)   (c)   (d)

Fig. 2. (a) The correspondence vector field using point to surface projection; (b) resulting mesh representation; (c) the correspondence vector field using using the Markov random field restoration; (d) improved mesh representation.

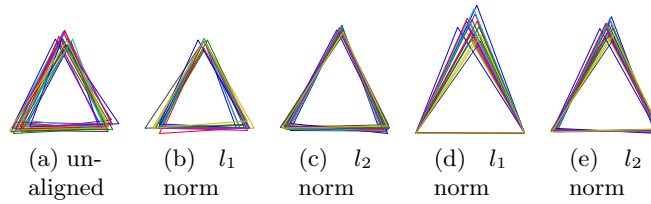(a) un-aligned   (b) l_1 norm   (c) l_2 norm   (d) l_1 norm   (e) l_2 norm

Fig. 3. (a) 10 triangles. (b-c) alignment based on the 3 corner points using the l_1 and l_2, respectively. (d-e) aligned as (b-c) but with an additional 19 landmarks distributed equidistantly on the lower side included.
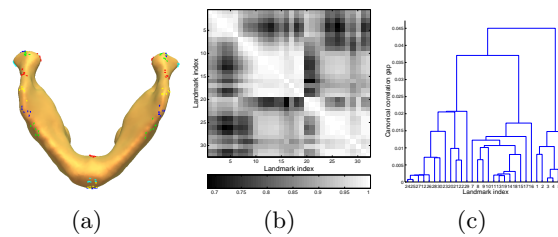

(a)   (b)   (c)

Fig. 4. (a) landmarks from 6 individuals projected onto a reference mandibular; (b) landmark canonical correlations; (c) hierarchical clustering of anatomical landmarks using single linkage, the canonical correlation are used as similarity measure.

3   Markov Random Field Correspondences

A method is developed for building statistical shape models based on a training set with an initial sparse annotation of corresponding landmarks of vary-

ing confidence [14]. A model mesh is aligned to all shapes in the training data using the thin plate spline (TPS) transformation based on a few anatomical landmarks. From the deformed model mesh and a target shape we derive a dense registration field of point correspondences. Applying a Markov random field (MRF) restoration we obtain a dense, continuous, invertible registration field (i.e. a homeomorphism) [15]. The stochastic restoration acts as a relaxation on the TPS constrained model mesh with respect to the biological landmarks. The landmarks are identified with varying confidence and the MRF relaxation allows for a data driven enhancement of the object correspondences. Using a site-prior, the algorithm converges to the most simple deformation field which has a tendency to match points of similar geometry since the field otherwise must be more complex. In comparison to applying point to surface projection the MRF regularization provides i) improved homogeneity of the target shape polygonization free of surface folds, ii) better reconstruction capabilities, and iii) a more compact Active Shape Model description of all the training data. The MRF regularization produces a reduction of the total variance contained in shape tangent space. The reduction is explained by increased collinearity between semi-landmarks distributed over the entire shape.

4 L_1 Generalized Procrustes 2D Shape Alignment

Using linear penalty instead of the quadratic penalty of least squares is a standard way of achieving resistance to outliers. Let there be given L training examples represented by (x_{ij}, y_{ij}), $i = 1, \ldots, L$, $j = 1, \ldots, k$. The alignment problem consists of estimating a modal shape, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_k, \nu_1, \ldots, \nu_k)^T$, and a set of Euclidean similarity parameters for each shape. Let these parameters be scale: $\beta_i \in \mathbb{R}_+$, rotation: $\psi \in [0; 2\pi[$, and translation: $\boldsymbol{\gamma}_i = (\gamma_{xi}, \gamma_{yi})^T \in \mathbb{R}^2$. This is conveniently formulated as the minimisation of a vector function wrt. all $\boldsymbol{\Theta}_i = [\beta_i \cos \psi_i, \beta_i \sin \psi_i, \gamma_{xi}, \gamma_{yi}]^T$ and $\boldsymbol{\mu}$.

$$
\mathbf{F}(\phi) = \begin{bmatrix} \mathbf{Z}_1 & \mathbf{0} & \ldots & \mathbf{0} & -\mathbf{I} \\ \mathbf{0} & \mathbf{Z}_2 & \ldots & \mathbf{0} & -\mathbf{I} \\ \vdots & & \ddots & & \vdots \\ \mathbf{0} & \ldots & \ldots & \mathbf{Z}_L & -\mathbf{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Theta}_1 \\ \boldsymbol{\Theta}_2 \\ \vdots \\ \boldsymbol{\Theta}_L \\ \mu \end{bmatrix} \text{ where } \boldsymbol{Z}_i = \begin{bmatrix} x_{i1} & -y_{i1} & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ x_{ik} & -y_{ik} & 1 & 0 \\ y_{i1} & x_{i1} & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ y_{ik} & x_{ik} & 0 & 1 \end{bmatrix} \tag{1}
$$

Necessary linear restrictions to avoid degenerate solutions are $\sum_{i=1}^L \gamma_{xi} = 0$, $\sum_{i=1}^L \gamma_{yi} = 0$, $\sum_{i=1}^L \beta_i \cos \psi_i = L$ and $\sum_{i=1}^L \beta_i \sin \psi_i = 0$. Minimising the L_2-norm of F results in the usual least squares 2D GPA. Using the L_1-norm results in the minimization of the city block distance between corresponding landmarks of all examples and a reference shape. This minimisation can be formulated as finding bounds of the values of F, such that the sum of these bounds is minimized. Expanding the indices can be stated as the following additional linear

constraints

$$-o_{xij} \le x_{ij}(\beta_i \cos \psi_i) - y_{ij}(\beta_i \sin \psi_i) + \gamma_{xi} - \mu_j \le o_{xij}$$

$$-o_{yij} \le y_{ij}(\beta_i \cos \psi_i) + x_{ij}(\beta_i \sin \psi_i) + \gamma_{yi} + \nu_j \le o_{yij}.$$

This is an LP (linear programming) problem. However, the use of the city block distance obviously introduces dependence of the orientation of the chosen coordinate system. Rotational independence is achieved by using a linear approximation to Euclidean distance based on averages of city block distances in equiangularly rotated coordinate systems.

In Fig. 3 the alignment of triangles in the L_1 and L_2 norms is shown based on the 3 corners points only, and using an additional 18 points added by linear interpolation between the to lower corner points. We see that the L_1 as well as the L_2 norm alignment are sensitive to the chosen representation, In the case of densely sampling the lower side the L_1 alignment regards the top corner point as an outlier and disregards it thereby achieving perfect alignment of the the lower sides. Initial work on L_1 Procrustes is presented in [16].

5   Analysis and visulization of landmark distributions

The intercorrelations of landmark coordinates may be visualized by showing the correlation matrix of the coordinates of all landmarks across the dataset. However, this representation is not invariant with respect to rotation of the frame of reference. Such a rotation will shift the correlations between the x, y, and z coordinates. Also, it is convenient to study the correlations in terms of landmarks. In order to express the correlation between a univariate quantity and a set of covariates we use the multiple correlation coefficient, and to express the correlation between two sets of multivariate quantities (i.e. point coordinates) we may use the canonical correlation. The canonical correlation is determined by use of a canonical correlation analysis [6]. This procedure rotates the scattering of to sets of variables (in casu landmarks i and j across the dataset) individually, such that maximum correlation is obtained along a projection axis. This correlation is the canonical correlation.

This visualization is performed for a set of 32 anatomical landmarks identified on the mandibular bone of 6 patients CT scanned 3-4 times from age 1 months to 14 years. In Fig. 4(a) the landmarks identified on the oldest mandible of each patient are projected onto a reference surface model using correspondences found by geometry constrained diffusion [17]. The scatter of each landmark is small and serves as a clinical validation of geometry constrained diffusion. In Fig. 4(b) the inter-(canonical)-correlations of the 32 landmarks are shown. We see that the landmarks are organised in 3 clusters. This is further revealed by a formal cluster analysis based on using the canonical correlations as similarity measure in a single linkage hierarchical clustering of the 32 landmarks, see Fig. 4(c). The 3 clusters are related to 1) the chin; 2 and 3) the ramus, the condylar and the coronoid processes of the left and right side of the mandible.

6   Shape modelling using partial least squares regression

A set of biological landmarks annotated of mandibular bones is applied in a statistical shape analysis in pursuit of a growth model. An ASM is typically constructed based on a principal components (PC) analysis of the Procrustes aligned shapes. Figure 5 shows a pairs plot of the subject, centroid size, and PC1–3 component scores. PC1 correlates to centroid size with a coefficient of 76%. By PC we solely focus on obtaining a basis that has the maximum likelihood of reconstructing the training data. In the present case, PC correlates to growth. However, this is not by design. Instead, we propose to decompose the dynamics in shape space by turning to regression based techniques. In particular, we wish to search for dominating shape variations that show high correlation to growth measured by centroid size. The method of partial least squares (PLS) thus becomes a natural choice for shape variation decomposition. In PLS $R = \mathrm{Cov}\{\boldsymbol{a}^T\mathbf{x}, \boldsymbol{b}^T\boldsymbol{y}\}$ (here $\mathbf{x}$ is an observation in the tangent space and $\boldsymbol{y}$ often a scalar response variable) is maximized with the following constraints: $\boldsymbol{a}^T\boldsymbol{a} = \boldsymbol{b}^T\boldsymbol{b} = 1$ leading to

$$R^2 = \frac{\boldsymbol{a}^T\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{21}\boldsymbol{a}}{\boldsymbol{a}^T\boldsymbol{a}} = \frac{\boldsymbol{b}^T\boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_{12}\boldsymbol{b}}{\boldsymbol{b}^T\boldsymbol{b}}, \tag{2}$$

[18]. The first pair of canonical variates (or latent variables) are calculated and the response CV is regressed on the predictor CV. If more information is present in the residuals these are subtracted from the original response variables by means of multiple regression, and the predictor variables are projected into a subspace orthogonal to the solution found, and more iterations are performed, see also [7, 19]. In Figure 5 the scores for the PLS based decomposition is shown using centroid size as the dependent variable in the regression. Notice that PLS1 has almost a one-to-one correspondence to PCA1, but possesses slightly higher correlation to centroid size of 77%. The PLS study thus confirms that growth of the mandible is primarily contained in a one-dimensional linear component in Procrustes tangent space. This is in agreement with the findings in [20] and does not conflict with the non-linear growth observed in biological coordinate references systems, [21]. A novel Active Shape Model is thus proposed targeting growth modelling by applying Partial Least Squares regression in decomposing the Procrustes tangent space. Here shape centroid size is applied as dependent variable but the method generalizes to handle other, both uni- and multivariate, effects probing for high covariation wrt. shape variation.
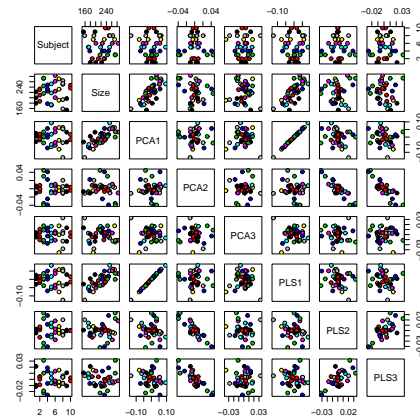
Fig. 5. Pairs plot of the patient index, the centroid size, the PCA1–3 and the PLS1–3 components.
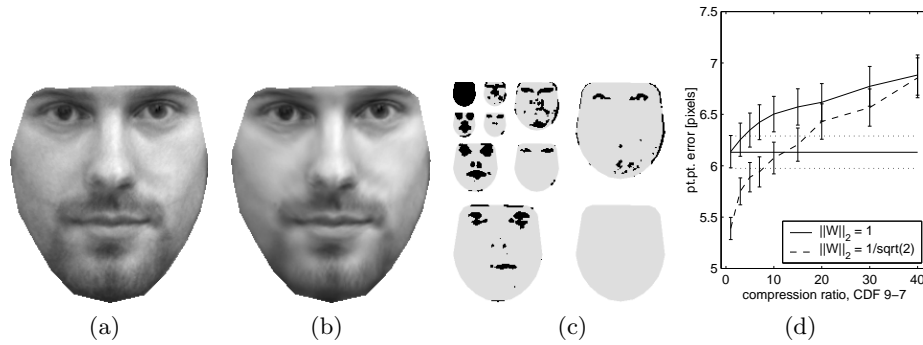

Fig. 6. (a) AAM synthesis; (b) Wavelet AAM synthesis, compression ratio 1:10; (c) Selected wavelet coefficients, ratio 1:10; (d) Segmentation accuracy using orthonormal wavelets (full), weighted wavelets (dashed) and a normal AAM (straight lines).
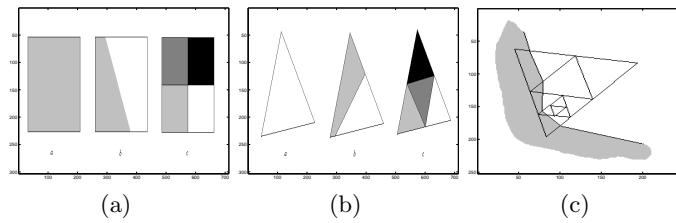

Fig. 7. (a) dyadic representaions; (b) triangle representations; (c) triangulated base representation of a simple neighbourhood.
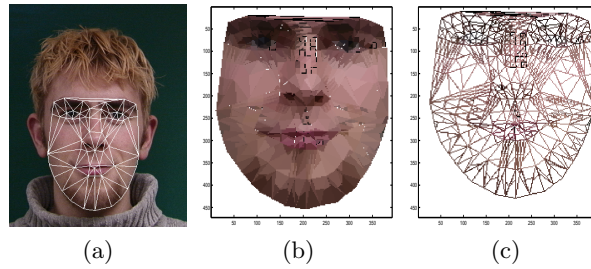
Fig. 8. Facial triangulated wedgelet quadtree representation derived using penalized residual sum-of-squares. (a) original triangulation; (b) wedgelet representation; (c) estimated edges.
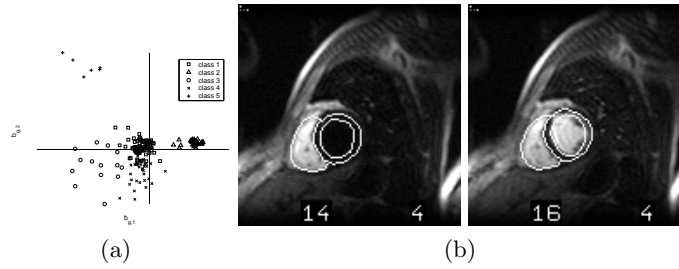

Fig. 9. (a) Unsupervised classification of perfusion texture vectors obtained from five subjects; (b) Registration of heart ventricles in two frames during the bolus passage.
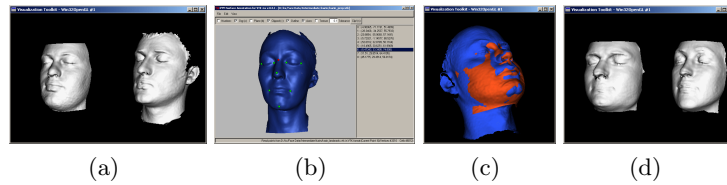

Fig. 10. (a) The template shape next to an unregistered face; (b) Landmarked face; (c) Thin-plate spline warped template (red) over face to be registered (blue); (d) Two registered faces
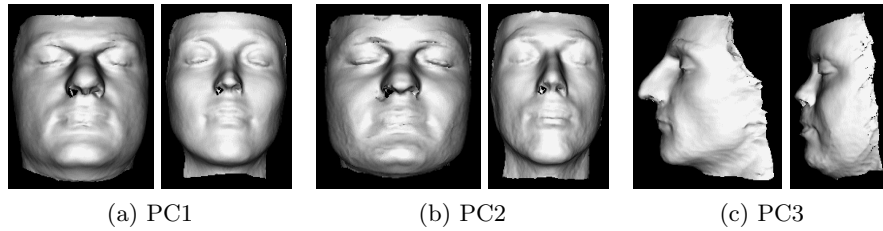

(a) PC1   (b) PC2   (c) PC3

Fig. 11. Extremal 3D faces corresponding to +/- 2 standard deviations of the three principal modes of variation.

7   Wavelet Enhanced Appearance Modelling

Generative methods such as the Active Appearance Models [8] establish dense correspondences by modelling variation of shape and pixel intensities. Explicitly modelling the value of every pixel covering an object is feasible for low-resolution 2D images. Alas, for high-resolution 2D images, 3D images and 3D time-series, this approach is rendered infeasible due to excessive storage and computational requirements. This is very unfortunate due to the many information-rich medical imaging modalities emerging and maturing these days. We address the problem by modelling wavelet coefficient subsets rather than pixel intensities. Figures 6(a-c) show that even moderate compression ratios leaves the wavelet AAM synthesis almost unaffected, since emphasis is put on perceptual important structures, in this case of face images. When localising unknown faces, Figure 6 (d) shows that the segmentation accuracy degrades gracefully with increasing compression ratio. Further, when utilising the frequency seperation of the wavelet decomposition by weighting the wavelet filters, we see that the model size can be decreased *while* increasing segmentation accuracy. For further details see [22].

8   Wedgelet compresion of triangulated domains

As an alternative to wavelets we propose to use wedgelets [10] to represent textural components in the training set of an active appearance model. Due to the nature of the AAM the challenge is to move from a dyadic domain to a triangulated domain. The basic idea of wedgelets is illustrated in Fig. 7(a), i.e. a dyadic square is represented by its mean value, an edge and two mean values, or 4 squares at higher resolution, leading to a quadtree representation. We define an equivalent triangular representation as shown in Fig. 7(b). For each triangle we run through all possible divisions of that triangle by a line connecting two points on the sides and compare the resulting representations to a simple mean over the entire triangle, and a subdivision into four triangles. The comparison is carried using a penalized residual sum-of-squares

$$PRSS = \| \, \boldsymbol{y} - \boldsymbol{\mu} \, \|^2 + \lambda \cdot \#P,$$

where y are original pixel values in the triangle, $\boldsymbol{\mu}$ the model, and $\#P$ a complexity term. An example of this triangulated wedgelet representation is shown in Fig. 7(c). In Fig. 8 an example of representing a face from an initial Delauney triangulation of a series of landmarks is shown.

9   Multimodal Appearance Models

Within the last decade magnetic resonance imaging has been proven able to assess myocardial perfusion. By injecting a bolus of contrast the myocardial perfusion mechanism can be quantified, which is essential in ischemic heart diseases. As the contrast agent tags the blood stream and amplifies the MR signal, areas

of the myocardium served by diseased arteries show a delayed and attenuated response. To quantify such areas, corresponding points need to be determined for all frames of a perfusion time sequence. The Active Appearance Model holds great promise for carrying out such a task for many reasons. However, the radical changes in intensity during the contrast uptake violate the assumption that image intensities (texture) are well modelled by a multivariate Gaussian. This yields a less specific texture model, potentially leading to false positives during tracking of perfusion sequences. We propose to replace the AAM texture model by an ensemble of local models estimated from a training set using unsupervised clustering. The result of such a clustering carried out on five subjects is shown in Figure 9 (a). A corresponding registration of an unseen perfusion sequence is shown in Figure 9 (b). For further details see [23, 24]

10   Statistical Shape Model from 3D Face Scans

In the face analysis examples above models have been built based on 2D projections. A more natural modelling is attained using 3D scans. Using a Minolta Vivid 900 laser scanner 15 individuals are scanned. The face shapes are represented by meshes consisting of approximate 20000 points. We register each face (study) to a template face following [12], i.e. by thin plate spline warping a template using a set of 9 manually identified landmarks, projecting the template landmarks to each of the studies, discarding the old study landmarks, and inverse thin plate spline warping these new landmarks to the original space. This is shown in Fig. 10. As a by-effect the pruning of the template is transferred to the studies. After a generalized Procrustes alignment a principal component analysis of tangent space coordinates reveals the principal modes of variation shown in Fig. 11. The Minolta scanner also provides a texture scan. Following [13] this will be included to generate a 3D facial appearance model.

References

1. R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," *Medical Imaging, IEEE Transactions on*, vol. 21, no. 5, pp. 525–537, 2002.
2. P. J. Besl and N. D. McKay, "A method for registration of 3d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, 1992.
3. J. C. Gower, "Generalized Procrustes analysis," *Psychometrika*, vol. 40, pp. 33–50, 1975.
4. 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.
5. 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, pp. 38–59, Jan. 1995.
6. H. Hotelling, "Relations between two sets of variables," *Biometrika*, vol. 28, pp. 321–377, 1936.

7. A. Höskuldsson, "PLS regression methods," *Jour. of Chemometrics*, vol. 2, pp. 211–228, 1986.

8. T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. on Pattern Recognition and Machine Intelligence*, vol. 23, no. 6, pp. 681–685, 2001.

9. C. B. H. Wolstenholme and C. J. Taylor, "Wavelet compression of active appearance models," in *Medical Image Computing and Computer-Assisted Intervention, MICCAI*, pp. 544–554, 1999.

10. D. Donoho, "Wedgelets: Nearly minimax estimation of edges," *Annals of Statistics*, vol. 27, pp. 859–897, 1999.

11. X. Huo and D. Donoho, "Beamlets and multiscale image analysis," in *Multiscale and Multiresolution Methods* (J. Barth, T. Chan, and R. Haimes, eds.), Lecture Notes in Computational Science and Engineering, pp. 149–196, Springer, 2001.

12. T. J. Hutton, B. R. Buxton, and P. Hammond, "Dense surface point distribution models of the human face," *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA 2001)*, pp. 153–60, 2001.

13. V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in *Proceedings of SIGGRAPH 99: 26th International Conference on Computer Graphics and Interactive Techniques*, pp. 187–194, ACM, 1999.

14. R. R. Paulsen, R. Larsen, S. Laugesen, C. Nielsen, and B. K. Ersbøll, "Building and testing a statistical shape model of the human ear canal," in *Medical Image Computing and Computer-Assisted Intervention - MICCAI*, pp. 373–380, Springer, 2002.

15. R. R. Paulsen and K. B. Hilger, "Shape modelling using markov random field restoration of point correspondences (to appear)," in *Information Processing in Medical Imaging, IPMI*, 2003.

16. F. Er, *Robust methods in Statistical Shape Analysis*. PhD thesis, University of Leeds, Department of Statistics, Leeds, UK, 1998. x + 86 pp.

17. P. R. Andresen and M. Nielsen, "Non-rigid registration by geometry-constrained diffusion," *Medical Image Analysis*, vol. 5, no. 2, pp. 81–88, 2001.

18. K. B. Hilger, *Exploratory Analysis of Multivariate Data*. PhD thesis, Department of Informatics and Mathematical Modelling, Technical University of Denmark, 2001.

19. I. E. Frank and J. H. Friedman, "A statistical view of some chemometrics regression tools," *Technometrics*, vol. 35, no. 2, pp. 109–135, 1993.

20. P. R. Andresen, F. L. Bookstein, K. Conradsen, B. K. Ersbøll, J. Marsh, and S. Kreiborg, "Surface-bounded growth modeling applied to human mandibles," *IEEE Transactions on Medical Imaging,*, vol. 19, pp. 1053–1063, dec 2000.

21. A. Björk and V. Skieller, "Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric implant studies over a period of 25 years," *Eur. J. Orthodontics*, vol. 5, pp. 1–46, 1983.

22. M. B. Stegmann, S. Forchhammer, and T. F. Cootes, "Wavelet enhanced apperance modelling," (in preparation).

23. M. B. Stegmann and H. B. Larsson, "Motion-compensation of cardiac perfusion MRI using a statistical texture ensemble," in *Functional Imaging and Modeling of the Heart, FIMH 2003*, LNCS, (Lyon, France), Springer, 2003 (to appear).

24. M. B. Stegmann and H. B. W. Larsson, "Fast registration of cardiac perfusion MRI," in *Proc. International Society of Magnetic Resonance In Medicine – ISMRM 2003, Toronto, Ontario, Canada*, (Berkeley, CA, USA), ISMRM, 2003 (to appear).