A 2D Morphable Model of Craniofacial Profile and Its Application to Craniosynostosis

Hang Dai[1(✉)], Nick Pears[1], and Christian Duncan[2]

[1] Department of Computer Science, University of York, York, UK
{hd816,nick.pears}@york.ac.uk
[2] Alder Hey Craniofacial Unit, Liverpool, UK
Christian.Duncan@alderhey.nhs.uk
https://www-users.cs.york.ac.uk/~nep/research/LYHM/

Abstract. We present a fully automatic image processing pipeline to build a 2D morphable model of craniofacial saggital profile from a set of 3D head surface images. Subjects in this dataset wear a close fitting latex cap to reveal the overall skull shape. Texture based 3D pose normalization and facial landmarking are applied to extract the sagittal profile from 3D raw scan. Fully automatic profile annotation, subdivision and registration methods are used to establish dense correspondence among sagittal profiles. The collection of sagittal profiles in dense correspondence are scaled and aligned using Generalised Procrustes Analysis (GPA), before applying Principal Component Analysis to generate a morphable model. Additionally, we propose a new alternative alignment called the Ellipse Centre Nasion (ECN) method. Our model is used in a case study of craniosynostosis intervention outcome evaluation and the evaluation reveals that the proposed model achieves state-of-the-art results. We make publicly available both the morphable model with matlab code and the profile dataset used to construct it.

Keywords: Morphable · Craniofacial · Fully automatic · Craniosynostosis

1 Introduction

In the analysis of head shape, the sagittal profile is often the most revealing and informative dimension to look for deviations from population norms and it is often useful, in terms of visual clarity and attention focus, for the clinician to examine shape from such a canonical viewpoint. Therefore, we have developed a novel image processing pipeline to generate a 2D morphable model of craniofacial saggital profile from a set of 3D head surface images. Other profiles are of course useful, as is a full 3D morphable model of the entire craniofacial region, but these are beyond the scope of this paper.

A morphable model is constructed by performing some form of dimensionality reduction, typically Principal Component Analysis (PCA), on a training dataset

M. Valdés Hernández and V. González-Castro (Eds.): MIUA 2017, CCIS 723, pp. 731–742, 2017.
DOI: 10.1007/978-3-319-60964-5_64

of shape examples. This is feasible only if each shape is first re-parametrised into a consistent form where the number of points and their anatomical meaning are made consistent. Shapes satisfying these properties are said to be in dense correspondence with one another. Once built, the morphable model provides two functions. Firstly, the it is powerful prior on 2D profile shapes that can be leveraged in fitting algorithms to reconstruct accurate and complete 2D representations of profiles. Secondly, the proposed model provides a mechanism to encode any 2D profile in a low dimensional feature space; a compact representation that makes tractable many 2D profile analysis problems in the medical domain.

Contributions: We propose a new pipeline to build a 2D morphable model of craniofacial sagittal profile. A new pose normalisation scheme is presented called *Ellipse Centre- Nasion* (ECN) normalisation. Extensive qualitative and quantitative evaluations reveal that the proposed normalisation achieves state-of-the-art results We use our morphable model to perform craniosynostosis intervention outcome evaluation on a set of 25 craniosynostosis patients. For the benefit of the research community, we will make publicly available the sagittal profile dataset, and our 2D morphable model with matlab code.

Paper Structure: In the following section, we discuss related literature. Section 3 discusses our new pipeline used to extract sagittal profiles and construct 2D morphable models. The next section evaluates several variants of the constructed models both qualitatively and quantitatively, while Sect. 5 illustrates the use of the morphable model in intervention outcome assessment for a population of 25 craniosynostosis patients. A final section concludes the work.

2 Related Work

In the late 1990s, Blanz and Vetter built a '3D morphable model' (3DMM) from 3D face scans [1] and employed it in a 2D face recognition application [2]. Two hundred scans were employed (young adults, 100 males and 100 females). Dense correspondences were computed using a gradient-based optic flow algorithm - both shape and colour-texture is used. The model is constructed by applying PCA to shape and colour-texture (separately).

There are very few publicly available morphable models of the human face and, to our knowledge, none that include the full cranium. The Basel Face Model (BFM) is the most well-known and widely used face model and was developed by Paysan et al. [3]. Again 200 scans were used, but the method of determining corresponding points was improved. Instead of optic flow, a set of hand-labelled feature points is marked on each of the 200 training scans. The corresponding points are known on a template mesh, which is then morphed onto the training scan using underconstrained per-vertex affine transformations, which are constrained by regularisation across neighbouring points [4].

Other deformable template methods could be used to build morphable models, such as the well-known method of Thin Plate Splines (TPS) [17] or the work of Li et al. (2008). Their global correspondence optimization method solves simultaneously for both the deformation parameters as well as the correspondence

positions [5]. Myronenko and Song (2009) consider the alignment of two point sets as a probability density estimation [6] and they call the method Coherent Point Drift(CPD), and this is remains a highly competitive template morphing algorithm.

Template morphing methods need an automatic initialisation to bring them within the convergence basin of the global minimum of alignment and morphing. To this end, Active Appearance Models (AAMs) [7] and elastic graph matching [8] are the classic approaches of facial landmark and pose estimation. Many improvements over AAM have been proposed [9,10]. Recent work has focused on global spatial models built on top of local part detectors, sometimes known as Constrained Local Models (CLMs) [11,12]. Zhu and Ramanan [13] use a tree structured part model of the face, which both detects faces and locates facial landmarks. One of the major advantages of their approach is that it can handle extreme head pose and we exploit this directly in our model building pipeline.

Another relevant model-building technique is the minimum description length method (MDL) [18], which selects the set of parameterizations that build the 'best' model, where 'best' is defined as that which minimizes the description length of the training set.

3 Model Construction Pipeline

Our pipeline to build a 2D morphable model is illustrated in Fig. 1. It employs a range of techniques in both 3D surface image analysis and 2D image analysis and has three main stages: (i) Profile extraction: The raw 3D scan from the *Headspace dataset* undergoes pose normalization, preprocessing to remove redundant data, and profile detection to find the sagittal profile; (ii) Dense correspondence establishment: A collection of sagittal profiles are reparametrised into a form where each sagittal profile has the same number of points joined into a connectivity that is shared across all sagittal profiles. Furthermore, the semantic or anatomical meaning of each point is shared across the collection; and (iii) Similarity alignment and statistical modelling: The collection of sagittal profiles in dense correspondence are subjected to Generalised Procrustes Analysis (GPA) to remove similarity effects, leaving only shape information. The processed meshes are statistically analysed, typically with PCA, generating a 2D morphable model expressed using a linear basis of eigenshapes. This allows for the generation of novel shape instances.

Each profiles is represented by m 2D points (y_i, z_i) and is reshaped to a $2m$ row vector. Each of these vectors is then stacked in a $n \times 2m$ data matrix, and each column is made zero mean. Singular Value Decomposition (SVD) is applied from which eigenvectors are given directly and eigenvalues can be computed from singular values. This yields a linear model as:

$$\mathbf{x_i} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b_i} = \bar{\mathbf{x}} + \sum_{i=1}^{k} \mathbf{p^k} b_i^k \tag{1}$$

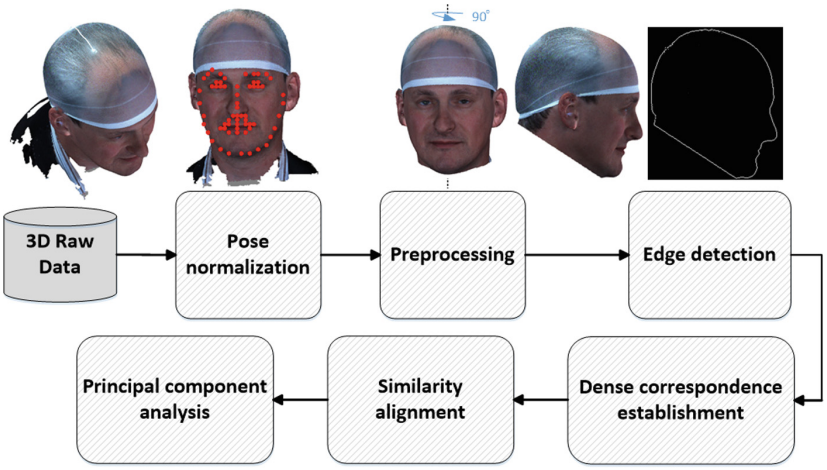

Fig. 1. The pipeline for 2D morphable model construction.

where $\bar{x}$ is the mean head profile shape vector and $\mathbf{P}$ is a matrix whose columns $\mathbf{p}^k$ are the eigenvectors of the covariance matrix (after pose alignment), describing orthogonal modes of head profile variation. The vector $\mathbf{b}$ holds the shape parameters $\{b^k\}$, that weight the shape variation modes which, when added to the mean shape, model a shape instance $\mathbf{x}_i$. The three main stages of the pipeline are described in the following subsections.

3.1 Profile Extraction

Profile extraction requires three stages, namely (i) pose normalisation, (ii) cropping and (iii) edge detection. Each of these stages is described in the following subsection.

Pose Normalisation. Using the colour-texture information associated with the 3D mesh, we can generate a realistic 2D synthetic image from any view angle. We rotate the scan over 360° in pitch and yaw (10 steps of each) to generate 100 images. Then the Viola-Jones face detection algorithm [14] is used to find the frontal face image among this image sequence. A score is computed that indicates how frontal the pose is. The 2D image with the highest score is chosen to undergo 2D facial landmarking. We employ the method of Constrained Local Models (CLMs) using robust Discriminative Response Map Fitting [15] to do the 2D facial image landmarking. The CLMs are trained using data from the Biwi Kinect Head Pose Database, which is equipped with ground truth rotation angles (pitch, yaw and roll). Then the trained system is used to estimate the three angles for the image with facial landmarks. Finally, 3D facial landmarks are captured by projecting the 2D facial landmarks to 3D scan. By estimating the rigid transformation matrix T from the landmarks of a 3D scan to that of

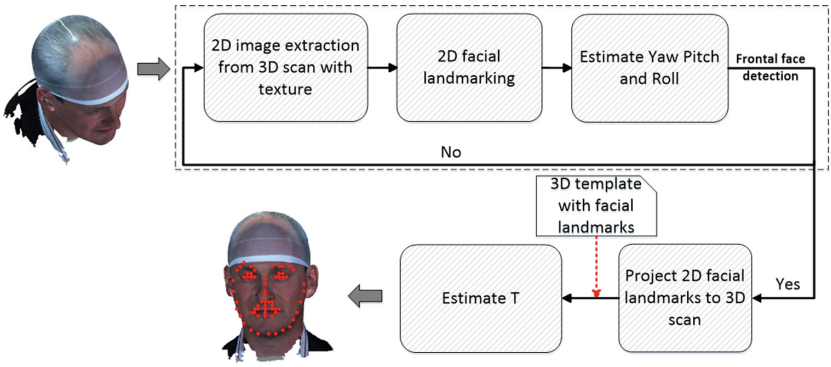

Fig. 2. 3D pose normalization using the texture information

a template, a small adjustment of pose normalization is implemented by transforming 3D scan using T^{-1}.

Cropping. 3D facial landmarks can be used to crop out redundant points, such as the shoulder area and long hair. The face landmarks delineate the face size and its lower bounds on the pose normalised scan, allowing any of several cropping heuristics to be used. We calculate the face size by computing the average distance from facial landmarks to their mean. Subsequently a plane for cropping the 3D scan is generated by moving the cropping plane downward an empirical percentage of the face size. We use a sloping cropping plane so that the chin area is included, but that still allows us to crop close to the base of the latex skull cap at the back of the neck to remove the (typically noisy) scan region, where the subject's hair emerges from under the cap (see Fig. 2).

Edge Detection. The scan is rotated 90° to reveal the head profile and we can generate a 2D image of this profile by orthogonal projection. Then the Canny edge detector [16] is employed to find the edges of this 2D image. The threshold is chosen to be the average of the set of pixels that excludes those that are white space (i.e. off the profile).

3.2 Dense Correspondence Establishment

To extract profile points using subdivision, we have an interpolation procedure that ensures that there is a fixed number of evenly-spaced points between any pair of facial profile landmarks. However, it is not possible to landmark the cranial region and extract profile model points in the same way. This area is smooth and approximately elliptical in structure and so we project vectors from the ellipse centre and intersect a set of fitted cubic spline curves, starting at the nasion, and incrementing the angle anticlockwise in small steps (we use one degree) over a fixed angle.

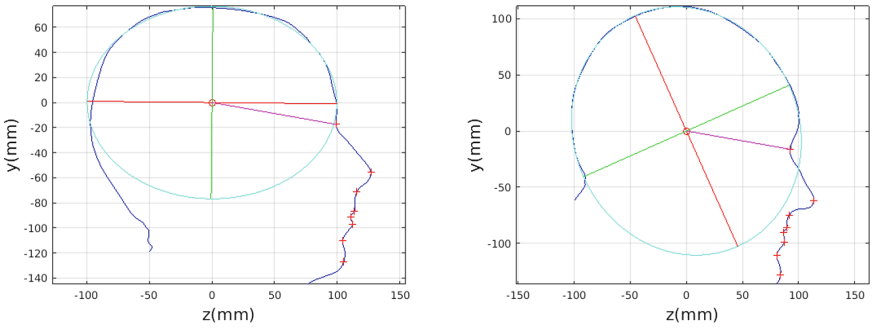

Fig. 3. Head tilt pose normalisation based on ellipse centre and nasion position. The extracted head profile is shown in blue, red crosses show facial landmarks and the ellipse fitted to the cranial profile is shown in cyan. Its major axis is red and its minor axis green. (Color figure online)

As well as using subdivision points directly in model construction, we form a model template as the mean of the population of subdivided and aligned profiles and we use template deformation on the dataset. The resulting deformed templates are re-parameterised versions of each subject that are in correspondence with one another. In this paper, we apply Subdivison, Thin Plate Splines (TPS) [17] nonrigid ICP (NICP) [4], Li's method [5], Coherent Point Drift (CPD) [6] and Minimum Description Length (MDL) [18] to the proposed pipeline for comparative performance evaluation.

3.3 Profile Alignment

A profile alignment method is needed before PCA can be applied to build the 2DMM. We use both the standard GPA approach and a new *Ellipse Centre - Nasion* (ECN) method. Ellipse fitting was motivated by the fact that large sections of the cranium appeared to be elliptical in form, thus suggesting a natural centre and frame origin with which to model cranial shape. One might ask, why not just use GPA over the whole head for alignment. One reason is because variable facial feature sizes (e.g., the nose's Pinocchio effect) induce displacements in the cranial alignment, which is a disadvantage if we are primarily interested in cranial rather than facial shape. We use the nasion's position to segment out the cranium region from the face and use a robust iterative ellipse fitting procedure that rejects outliers.

Figure 3 shows examples of the robust ellipse fit for two head profiles. The centre of the ellipse is used in a pose normalisation procedure where the ellipse centre is used as the origin of the profile and the angle from the ellipse centre to the nasion is fixed at $-10°$. We call this Ellipse Centre - Nasion (ECN) pose normalisation and later compare this to GPA. The major and minor axes of the extracted ellipses are plotted as red and green lines respectively in Fig. 3.

Figure 4 shows all the profiles overlaid with the same alignment scheme. We noted regularity in the orientation of the fitted ellipse as is indicated by the clustering of the major (red) and minor (green) axes in Fig. 4 and the histogram of ellipse orientations in Fig. 4. A minority of heads (9%) in the training sample have their major ellipse axes closer to the vertical (brachycephalic).

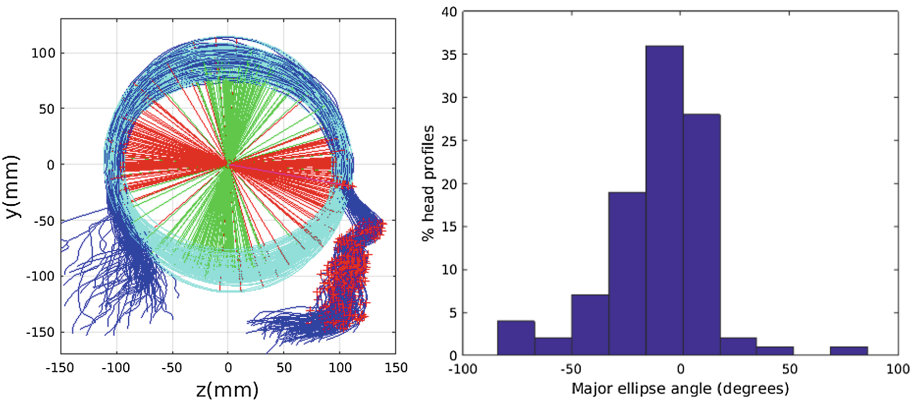

Fig. 4. (1)All training profiles after ECN normalisation; (2)Major axis ellipse angles with respect to an ECN baseline of -10°: median angle is -6.4° (2sf). (Color figure online)

4 Morphable Model Evaluation

We built four main 2DMM variants using 100 adult males from the Headspace dataset and animated shape variation along the principal components. The four model variations correspond to full head, scale normalised and unscaled, and cranium only, scale normalised and unscaled.

As an example, when ECN is used (Fig. 5, 1st row), the following three dominant (unscaled) modes are observed: (i) *Cranial height* with *facial angle* are the main shape variations, with small cranial heights being correlated with a depression in the region of the coronal suture; (ii) The overall size of the head varies: surprisingly this appears to be almost uncorrelated with craniofacial profile shape. This was only found in the ECN method of pose normalisation; (iii) The length of the face varies - i.e. there is variation in the ratio of face and cranium size. The second row of Fig. 5, shows the model variation using GPA alignment for comparison.

For quantitative evaluation of morphable models, Styner et al. [19] gives detailed descriptions of three metrics: compactness, generalisation and specificity, now used on our *scale-normalised* models.

Compactness: This describes the number of parameters (fewer is better) required to express some fraction of the variance in the training set. As illustrated in Fig. 6, the compactness using ECN is superior to that of GPA, for

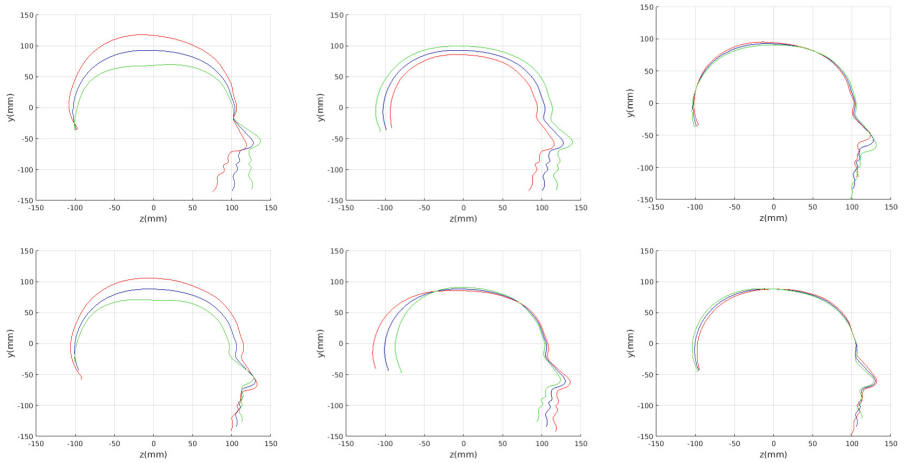

Fig. 5. 1st row: The dominant four modes (left: mode 1; right: mode (3) of head shape variation using automatic profile landmark refinement and ECN similarity alignment. Mean is blue, mean + 3SD is red and mean − 3SD is green. 2nd row: GPA similarity alignment. (Color figure online)

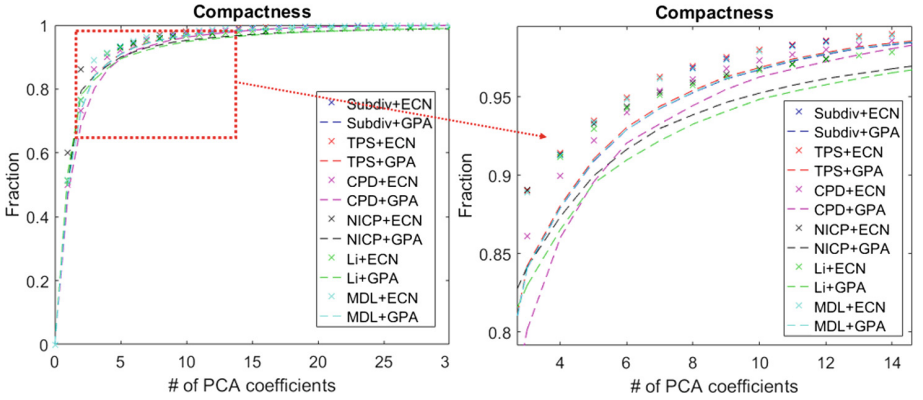

Fig. 6. Compactness

the same correspondence method. Among these methods, subdivision, TPS and MDL, all aligned with ECN are able to generate the most compact models.

Specificity: Specificity measures the model's ability to generate shape instances of the class that are similar to those in the training set. We generate 1000 random samples and take the average Euclidean distance error to the closest training shape for evaluation, lower is better. We show the specificity error as a function of the number of parameters in Fig. 7. Across all correspondence methods with GPA, it gives better specificity against all correspondence methods

with ECN. This suggests that GPA helps improve the performance of modelling the underlying shape space. NICP with GPA captures the best specificity.

Generalisation: Generalisation measures the capability of the model to represent unseen examples of the class of objects. It can be measured using the *leave-one-out* strategy, where one example is omitted from the training set and used for reconstruction testing. The accuracy of describing the unseen example is calculated by the mean point-to-point Euclidean distance error, the lower the better. Generalization results are shown in Fig. 7 and for more parameters, the error decreases, as expected. NICP with GPA performs better in terms of Euclidean distance once less than 7 model dimensions are used. Between 7 and 20 model dimensions, TPS with ECN outperforms other methods. When more than 20 model dimensions are used, CPD with GPA has the best generalization ability. Overall, GPA is able to help more successfully model the underlying shape against ECN for the same correspondence method, thereby generating better reconstructions of unseen examples.

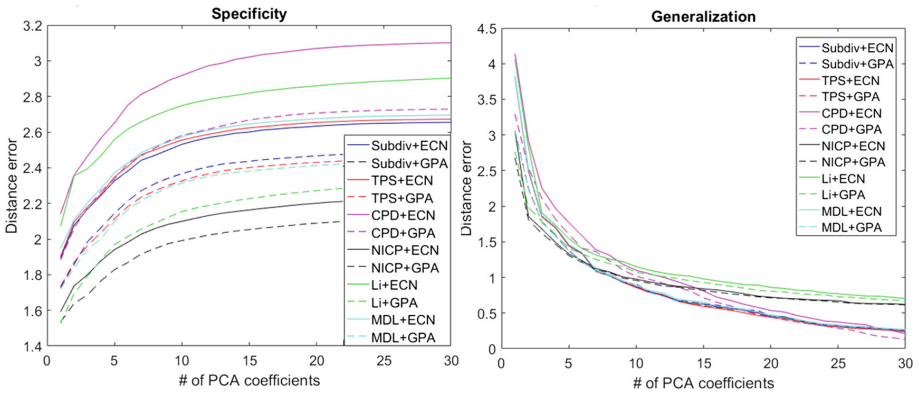

Fig. 7. Left: specificity; right: generalization.

5 Craniosynostosis Intervention Outcome Evaluation

Craniosynostosis is a skull condition whereby, during skull growth and development, the sutures prematurely fuse, leading to both an abnormally shaped head and increased intracranial pressure. We present a case study of 25 craniosynostosis patients (all boys), 14 of which have undergone one type of corrective procedure called *Barrel Staving* (BS) and the other 11, another corrective procedure called *Total Calvarial Remodelling* (TCR). The intervention aim is to remodel the patient's skull shape towards that of an adult and we can employ our model in assessing this.

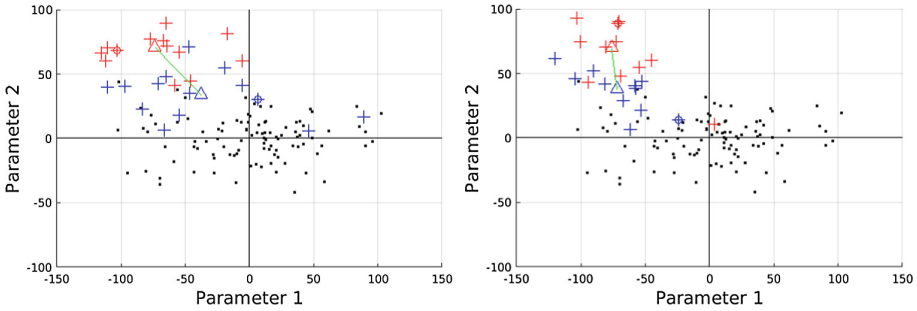

Fig. 8. Patient cranial profile parameterisations, BS (left) and TCR (right) intervention: pre-operative (red crosses) and post-operative (blue crosses) in comparison to the training set (black dots). The circled values represent an example patient (Color figure online)

We build a *scale normalised, cranium only* (to the nasion) 2D morphable model (2DMM) using 100 male subjects, without cranial conditions. Note that both facial structure and overall scale are now irrelevant and that major cranial shape changes are not thought to occur after 2 years old. The patients' scale normalised profiles are then parameterised using the model, indicating the distance from the mean cranial shape along in terms of the model's eigenstructure. The comparisons of pre-operative and post-operative parametrisations show the shapes moving nearer to the mean of the training examples, see Fig. 8.

For the BS patient set, the Mahalanobis distance of the mean pre-op parameters (red triangle in Fig. 8) is 4.670, and for the mean post-op parameters (blue triangle) is 2.302. For shape parameter 2 only (the dominant effect), these figures are 4.400 and 2.156. For the TCR patient set, the Mahalanobis distance of the mean pre-op parameters (red triangle in Fig. 8) is 4.647, and for the mean post-op parameters (blue triangle) is 2.439. For shape parameter 2 only these figures are 4.354 and 2.439. We note that most of this change occurs in parameter 2, which corresponds to moving height in the cranium from the frontal part of the profile to the rear. In these figures, we excluded one patient, who preoperatively already had a near-mean head shape (see red cross near to the origin in Fig. 8, so any operation is unlikely to improve on this (but intervention is required in order to relieve potentially damaging intracranial pressure).

It is not possible to make definitive statements relating to one method of intervention compared to another with these relatively small numbers of patients. However, the cranial profile model does show that both procedures on average, lead to a movement of head shape towards the mean of the training population. An example of analysis of intervention outcome for a BS patient and a TCR patient are given in Fig. 9. The particular example used is highlighted with circles on Fig. 8 to indicate pre-op and post-op parametrisations.

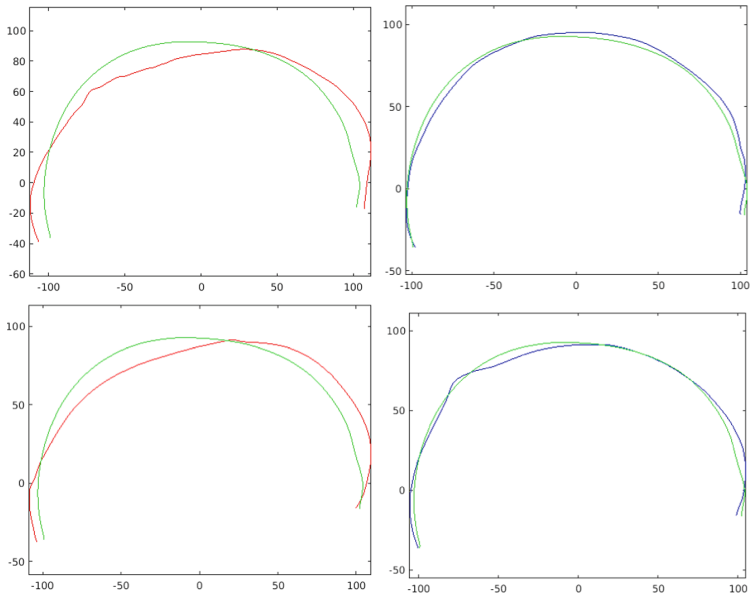

Fig. 9. 1st row: Pre-op and post-op profiles for a BS patient; 2nd row: Pre-op and post-op profiles for a TCR patient. The red and blue traces show the extracted sagittal profiles of the patient pre-operatively and post-operatively respectively, whilst the green shows the mean profile of the training set. (Color figure online)

6 Conclusions

We have presented a fully automatic general and powerful sagittal profile modelling pipeline. Alignment using the Ellipse-Centre Nasion method was introduced. The proposed model has been demonstrated to capture profile shape and assess the intervention outcomes. ECN builds more compact sagittal profile models when compared to GPA. Subdivision, TPS and MDL with ECN are recommended for a more compact sagittal profile model, while NICP with GPA is recommended to capture more specificity. NICP with GPA is able to generate better reconstructions of unseen profiles when fewer than 7 model dimensions are used. If using between 7 and 20 model dimensions, TPS with ECN is recommended for a better generalisation ability. When more than 20 model dimensions are used, CPD with GPA builds a model with better ability to reconstruct unseen examples.

Acknowledgments. The authors wish to thank the Royal Academy of Engineering and Leverhulme Trust for funding the University of York Senior Research Fellowship that primed this work. The Headspace project was funded through the Quality Improvement, Development and Innovation Scheme (QIDIS) from the National Commissioning Group (NCG) between 2011 and 2014. The authors also wish to thank Rachel Armstrong, the Headspace project coordinator.

References

1. Volker, B., Vetter, T.: A morphable model for the synthesis of 3D faces. In: Proceedings of the 26th Annual Conference on Computer Graphics and Interactive Techniques, pp. 187–194 (1999)
2. Volker, B., Vetter, T.: Face recognition based on fitting a 3D morphable model. IEEE Trans. Pattern Anal. Mach. Intell. **25**(9), 1063–1074 (2003)
3. Paysan, P., Knothe, R., Amberg, B., Romdhani, S., Vetter, T.: A 3D face model for pose and illumination invariant face recognition. In: Sixth IEEE International Conference on AVSS 2009, pp. 296–301 (2009)
4. Amberg, B., Romdhani, S., Vetter, T.: Optimal step Nonrigid ICP algorithms for surface registration. In: IEEE Conference on CVPR 2007. IEEE (2007)
5. Li, H., Sumner, R.W., Pauly, M.: Global correspondence optimization for nonrigid registration of depth scans. Comput. Graph Forum **27**(5), 1421–1430 (2008)
6. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. **32**(12), 2262–2275 (2010)
7. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**(6), 681–685 (2001)
8. Wiskott, L., Krger, N., Kuiger, N., Von Der Malsburg, C.: Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 775–779 (1997)
9. Sauer, P., Cootes, T.F., Taylor, C.J.: Accurate regression procedures for active appearance models. In: British Machine Vision Conference (2011)
10. Tresadern, P.A., Sauer, P., Cootes, T.F.: Additive update predictors in active appearance models. In: British Machine Vision Conference (2010)
11. Smith, B.M., Zhang, L.: Joint face alignment with nonparametric shape models. In: 12th European Conference on Computer Vision (2012)
12. Zhou, F., Brandt, J., Lin, Z.: Exemplar-based graph matching for robust facial landmark localization. In: 14th ICCV (2013)
13. Zhu, X., Ramanan, D.: Face detection, pose estimation, and landmark localization in the wild. In: Computer Vision and Pattern Recognition (2012)
14. Viola, P., Jones, M.J.: Robust real-time face detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
15. Asthana, A., et al.: Robust discriminative response map fitting with constrained local models. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2013)
16. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **6**, 679–698 (1986)
17. Bookstein, F.L.: Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Mach. Intell. **11**(6), 567–585 (1989)
18. Davies, R.H., Twining, C.J., Cootes, T.F., Waterton, J.C., Taylor, C.J.: A minimum description length approach to statistical shape modeling. IEEE Trans. Med. Imaging **21**(5), 525–537 (2002)
19. Styner, M.A., et al.: Evaluation of 3D correspondence methods for model building. In: Taylor, C., Noble, J.A. (eds.) IPMI 2003. LNCS, vol. 2732, pp. 63–75. Springer, Heidelberg (2003)