An Automatic Approach for 3-D Facial Shape Change Analysis by Combination of ASM and Morphometric Tools

Z. Mao and A. J. Naftel

Department of Physics, Astronomy and Mathematics
University of Central Lancashire, Preston PR1 2HE, UK

Abstract. This paper presents a new approach for automatic facial shape change analysis by combining a stereo-assisted active shape model (ASM) with morphometric tools. We firstly describe how to extend the 2-D ASM to 3-D space by stereo correspondence search and disparity map interpolation for automatic facial landmark extraction. Morphometric techniques such as generalized Procrustes analysis, principal component analysis and thin plate spline decomposition are then reviewed and applied to the quantitative analysis of shape changes in 2-D facial midline profiles and 3-D facial landmarks. The proposed method is validated both statistically and visually by characterizing bite-block induced shape changes in a heterogeneous sample of young orthodontic patients.

1 Introduction

Morphological changes in facial surface shape may occur as a result of growth processes, pathology or from soft tissue movements following corrective treatment. Techniques for three-dimensional analysis of morphological changes resulting from surgical correction have recently been proposed [1,2]. Among the existing non-landmark based shape analysis methods, the most successful approach applied to facial imaging is the use of surface curvature methods [3]. However, the calculation of surface curvature is unavoidably noise-sensitive and pre-processing techniques such as smoothing and local surface fitting must be carried out beforehand, which may remove surface detail. Recent advances in the rapidly maturing field of geometric morphometrics [4,5] offer an effective means for landmark-based shape analysis. Yet to date, morphometric tools have largely relied on expert knowledge for the manual detection of landmarks or curving outlines to segment anatomical regions of interest. In this paper, we present a new approach for automatic 3-D facial shape change analysis by combining 2-D active shape models (ASM) [6] with stereo disparity analysis for the automatic extraction of 3-D facial landmarks. These landmarks can then be subjected to morphometric analysis. The proposed method is applied to the shape analysis of group differences in a heterogeneous sample of young patients attending an orthodontic clinic. Anterior and lateral facial landmark shapes are analyzed using morphometric techniques with a view to characterising the differences in a geometrically and statistically meaningful way.

2 Materials

We collected frontal view facial stereo image pairs of 61 patients (aged 10 to 26 years; mean 16 years) attending a weekly orthodontic clinic at the Royal Preston Hospital. To test the efficacy of the proposed method, a deliberate gross change in the shape of the jaw of the patient was produced by means of a bite block. This induces 3-D morphological changes in the soft tissue having a similar order of magnitude to that which could be achieved clinically by functional orthodontic appliances or by surgical correction. Twenty-five facial images of patients were used to train a 2-D ASM for feature extraction. Metric 3-D facial surface reconstructions were obtained using an area-based stereo matching technique [7]. The 3-D data sets were then divided into two groups; group I to denote shape data prior to insertion of the bite block and group II after insertion. A morphometric analysis of facial shape change was then carried out using randomly selected patients to evaluate differences between groups I and II.

3 Method

3.1 Automatic 3-D Facial Landmark Detection by ASM and Stereo Correspondence

ASM [6] is a statistical-based technique for building compact models of the shape and appearance of a flexible object to automatically locate new instances of the object in 2-D images.

Fig. 1. Shape model points for 2-D feature detection

Initially, 107 facial landmark points were defined to represent the 2-D facial features as shown in Fig. 1. They were manually located on 25 grey-level training images from our facial database and Procrustes aligned to compute a mean shape and build the point distribution model (PDM). Principal component analysis on the covariance matrix of deviations from the mean shape was carried out to yield a set of basis vectors describing the main modes of shape variation in the training data. At the same time, the grey-level appearance models about each shape point were constructed in the same way as for the PDM. After the training procedure

was completed, the average grey level profiles were used to drive and deform the shape model iteratively to locate 2-D facial features for stereo intensity images outside the training set.

To ensure correct convergence, an initial position of the PDM, defined by the scale and translation parameters, was calculated by locating the horizontal and vertical edges of the facial borders (Fig. 2) using the integral projection of the intensity image. The vertical integral projection $V(x)$ and horizontal projection $H(y)$ of an image $I(x, y)$ in a rectangular region $[x_1, x_2] \times [y_1, y_2]$ are defined by:

$$V(x) = \sum_{y=y_1}^{y_2} I(x, y), \quad H(x) = \sum_{x=x_1}^{x_2} I(x, y) \quad (1)$$

A Sobel horizontal edge filter was applied to the intensity image followed by the vertical projection of the filtered image. Fig. 3 depicts the vertical projection of the horizontal edge, where L and R indicate the left and right facial borders. By finding the minimum and maximum values along the horizontal axis, the left and right facial borders were located. A similar operation was carried out to locate the upper facial border and the vertical position of the eyes and eyebrows (Fig. 2) and this information was then used to determine the initial scale and pose. Based on this initialization, 2-D facial landmark detection in the left stereo image was carried out. A maximum limit of 50 iterations was imposed when searching for the optimum position.

Fig. 2. Intensity image used for Figure feature detection and stereo matching with detected borders

Fig. 3. The vertical integral projection of the filtered intensity image

After accurately locating the 2-D features in a left stereo image, area-based stereo matching [7] was performed to obtain the correspondence between stereo image pairs. If the corresponding point for a left feature point could not be found in the right image by the matching algorithm, the missing match was interpolated from the known neighbouring disparity values using an adjustable matching window and a thin-plate spline (TPS) mapping function. A smaller window size can be used in regions where the disparity map is dense, whilst a larger window is needed for regions containing sparse data. In our experiments, if the number of known disparity values was less than 40% of the total in a selected

window, then the window size was automatically increased until the number of the known value reached a preset minimum of 20 values. Once all the disparity values had been obtained, these point correspondences were converted into 3-D space using a set of camera parameters obtained from a metric calibration rig [8]. Here, only 90 3-D facial feature landmarks (Fig. 4) are used for further morphometric analysis, whilst the detected landmarks on the facial border are used purely for the purpose of segmenting the face from the background.

Fig. 4. Automatically detected facial landmarks on a 3-D facial model

3.2 Morphometric Techniques for the Shape Analysis

Three main approaches for analysing facial shape change were employed; 1) Generalized Procrustes analysis (GPA), 2) principal components analysis (PCA) in the tangent space and 3) thin plate spline (TPS) interpolation for visualizing shape deformation.

Generalized Procrustes analysis is primarily used for estimating an average shape [5]. It registers sets of landmark configurations by removing translational and rotational differences and scaling them until they most closely match. The registered landmark configurations constitute Kendall's shape manifold [5] with Procrustes distance as metric. The dimension of Kendall's shape space is lower than that of the original figure space because the translation, rotation and scale differences have been removed. For k landmarks in 2-D, the shape space has dimension of $2k - 4$ and in 3-D, $3k - 7$. After Procrustes registration, standard multivariate analyses such as tests of group differences and PCA can be carried out directly on the tangent coordinates.

Principal components analysis in the tangent plane provides an effective means of investigating the structure of shape variability [5] and reducing the dimensionality of the shape data prior to multivariate analysis. Since the PCs are orthogonal, they each represent statistically independent modes of variation in the sample. When these PCs are sorted in descending order of eigenvalues for the covariance matrix, main modes of shape variation can be described by the first few significant PCs.

Shape deformation can be further interpreted using the TPS together with a Cartesian transformation grid to exactly map the biological shape change

between two homologous 2-D data sets with minimum bending energy. The associated principal warps [4] decompose shape change into local and global components of successively increasing scale analogous to a Fourier analysis.

4 Result

ASM-assisted landmark detection was tested on 25 facial images outside the training set and a success rate of more than 80 percent was obtained. These landmarks were then converted into 3-D coordinate space by stereo correspondence information and disparity map interpolation. Metric reconstruction accuracy have previously been established using a calibration object with precisely determined ground truth. The average root mean square (RMS) error for stereo matching is 0.26 pixel yielding a RMS error of 0.28mm in 3-D spatial coordinates for target check points. Ninety three-dimensional landmarks were automatically located for both group I and group II shapes. Midline profiles were then obtained by a facial symmetry plane registration method [9] and resampled to produce 39 pseudo-landmark positions for each 3-D image model.

The morphometric analysis of facial shape change was first carried out on the 2-D midline profiles of 16 randomly selected patients. Fig. 5 illustrates the scatter plots of these profiles which were partially Procrustes aligned using the unchanged (upper 25) landmarks. The solid lines indicate the Procrustes averaged profiles for groups I and II. As expected, the mouth and lower jaw region for the group II sample has been displaced forward and slightly downward by the bite block. This shape change can be depicted more clearly in Fig. 6 by a Cartesian transformation grid calculated using TPS deformation. The left and right pictures are the average profiles for group I and group II respectively.

GPA was then used to align the set of midline profiles followed by PCA on the Procrustes tangent coordinates to reveal the main modes of shape variation. The shape variation represented by the first three PCs is illustrated in Fig. 7. The largest shape differences were found in the mouth region captured in PC 1. This is largely due to the effect of bite block insertion and hence the simulated surgical correction. PCs 2 and 3 reveal combined effects of variation in nose shape which is expected given the age range in our sample (10-26 years). The first two PCs resulting from this analysis are plotted in Fig. 8. It can be seen that the two groups have been partially separated by the plot with group I ('o') largely occupying the bottom right half and group II ('*') the top left half of the figure. The percentages of variability accounted for by the first 3 PCs are 35.1%, 16.9% and 16.1%. Performing a two-sample Hotelling's T^2 test on Procrustes tangent coordinates on the first two PCs (sample size 20) yields a test statistic of 8.673 ($p = 0.025$). Hence, there is good evidence that the mean shapes are different and that this has not been masked by the heterogeneous nature of the patient sample. This indicates that the apparatus is sufficiently sensitive to detect relatively small changes in facial shape.

However, the midline profiles only utilise 2-D shape information. For a more revealing picture, the full 3-D landmark data set derived from the stereo cor-

Fig. 5. Scattered plot of mic profiles. Solid lines indicate the averaged profiles.

Fig. 6. The deformation of midline profiles viewed using a thin-plate spline deformation.

Fig. 7. Shape variation explained by first 3 PCs. Each row denotes the i th PC ($i = 1,2,3$) and each column displays a multiple c of the standardised PC scores ($c = -3, -2, -1, 0, 1, 2, 3$). $c = 0$ corresponds to the Procrustes averaged form. Row i corresponds to i th PC.

Fig. 8. Plot of PC 1(horizontal axis) vs PC 2 for group differences in midline profile, 'o' denotes group I (before bite block insertion) and '*' denotes group II (after insertion).

Fig. 9. Plot of PC 1(horizontal axis) vs PC 2 for group differences in 3-D landmarks, 'o' denotes group I (before bite block insertion) and '*' denotes group II (after insertion).

respondence search are chosen. Assigning landmark data to groups I and II as before, the extracted data were submitted to three-dimensional GPA and PCA. In this case the first 3 PCs account for 33.92%, 12.19% and 9.28% of the total variation in the combined sample. As shown in Fig. 9, the groups I and II are well separated signifying a clear difference in shape. Hotelling's T^2 test on the first 3 PCs (sample size 20) gave a statistically significant group difference at $p \sim 0.00029$. The evidence of difference in mean shape for 3-D landmark data is not in doubt and is clearly stronger than for 2-D data.

Investigating 3-D landmarks in the mouth region alone and projecting into the frontal facial plane, we can attempt to visualise the shape differences as a deformation from group I to group II, using TPS and principal warps decomposition. The results are illustrated in Fig. 10 and 11. Fig. 11 depicts the decomposition of deformation (i.e. shape difference) at successively larger scales. It provides an alternative picture of the basis space to PCA in tangent space coordinates. It was found that principal warp 3 accounts for 63.1% of the total bending energy. The weights for each principal warp (partial warp scores) and the affine contribution was calculated. Hotelling's T^2 test on these partial warp scores gave a statistically significant group difference at $p < 0.00001$.

Fig. 10. Deformation from group I to group II as a thin plate spline. Averaged landmark configuration of group I(left), averaged landmark configuration of group II(right).

Fig. 11. Principal warps for the landmark configuration shown in Fig.10 (left).

5 Conclusion

We have developed an effective means for automatic landmark-based 3-D facial shape change analysis which combines stereo-assisted ASM with morphometric tools. The experiment results signify a relative displacement of the average shape of the mouth and lower jaw forwards and slightly downwards after insertion of the bite block. This is the desired effect of surgical treatment on patients presenting class II and class III abnormalities. We have shown that such changes can be validated statistically. It suggests that the proposed method may be useful for auditing clinical treatment of class II and class III abnormalities with respect to its effects on facial soft tissue morphology.

Acknowledgements

We would like to thank Dr Michael Trenouth of Royal Preston Hospital for providing access to orthodontic patients and the University of Central Lancashire for funding this work through an overseas research studentship award.

References

1. A. M. McCance, J. P. Moss, W. R. Fright, A. D. Linney and D. R. James. Three-dimensional analysis techniques. II Laser scanning: A quantitative 3-D soft-tissue analysis using a colour-coding system. *Cleft Palate-Craniofacial Journal*, 34(1):46–51, 1997.
2. J. P. Moss, A. M. McCance, W. Fright, A. D. Linney and D. R. James et al. A 3-Dimensional soft-tissue analysis of 15 patients with class-II, division-1 malocclusions after bimaxillary surgery. *American Journal of Orthodontics and Dentofacial Orthopedics*, 105(5):430–437, 1994.
3. A. M. Coombes, R. Richards, A. D. Linney, V. Bruce and R. Fright. Description and recognition of faces from 3D data. *Proceedings of SPIE - The International Society for Optical Engineering*, 1766:307–319, 1992.
4. F. L. Bookstein. Shape and the information in medical images: A decade of the morphometric synthesis. *Computer Vision and Image Understanding*, 66(2):99–118, 1997.
5. I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis*, John Wiley & Sons, 1998.
6. T. F. Cootes, A. Hill, C. J. Taylor and J. Haslam. The use of active shape models for locating structures in medical images. *Image and Vision Computing*, 12(6):355–366, 1994.
7. Z. Mao and A. J. Naftel. Improved area-based stereo matching using an image segmentation approach for 3-D analysis of facial soft tissue changes. *Medical Image Understanding and Analysis '97*, University of Oxford, pp.209–212, 1997.
8. R. Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323–44, 1987.
9. Z. Mao and A. J. Naftel. 3D image analysis of facial shape changes using depth from stereo. *Nobless Workshop on Non-linear Model Based Image Analysis*, Glasgow, pp.283–288, 1998.