A Minimalist Approach to Facial Reconstruction

M. Gotla and Z. Huang †

National University of Singapore‡

Abstract

We propose a minimalist approach to 3D facial reconstruction. By applying cognitive and anatomical heuristics, we show that a realistic face model can be reconstructed from a generic face model using only one or two real facial images and minimal user interaction.

1. Introduction

Much work has already been done in the area of facial reconstruction [14, 17, 22, 10, 8, 11, 3, 4, 6, 16]. An excellent summary of the tools and techniques used in facial reconstruction and animation can be found in [15]. However, most work emphasizes photo realistic quality that requires extensive input, a large number of sample points, and intensive user interaction.

We propose a minimalist approach to 3D facial reconstruction. By applying cognitive and anatomical heuristics, we show a realistic facial model can be reconstructed from a generic facial model using only two real facial images and minimal user interaction. There is no need for camera calibration. The method is very suitable to rapidly customize facial models used as avatars in VR walk-through applications.

The main idea is to apply cognitive and anatomical heuristics to the algorithm. We focus only on those aspects of the face, which are cognitively significant. Cognitive psychologists have demonstrated, through phenomena such as the differential inversion and the caricature advantage [21], that humans appear to pay special attention to contours of facial features such as the eyes, lips, etc. These are believed to form the dimensions of face space [7]. In addition, our solution's feature extraction component takes advantage of anatomical expertise in the form of redundancies in the human face to help trace the features of interest. This expertise is primarily in the form of rules that all human faces conform to and have been compiled in [15].

From our knowledge, all reconstruction frameworks using images need camera calibration. It is a very tedious process though significant progress has been made on techniques of self-calibration in computer vision society [18]. In our reconstruction solution, only two facial images from two special camera poses are used as input - a front and a profile image. Thus we are able to avoid the camera calibration process. From our experiments the reconstruction results are still acceptable, even when these two images are not taken exactly from the front and side views.

The most relevant work to ours is [12]. We follow the same trend towards automatic reconstruction. The major difference is that we introduce cognitive and anatomical heuristics into processes of feature extraction and deformation. By doing this, we can segment the extraction and deformation to multi-processes with each sub-process working on its own important feature of face model. The accuracy and efficiency of algorithm have been improved.

2. The Solution

Our solution consists of three steps - feature extraction, 3D deformation, and texture mapping. We apply cognitive and anatomical heuristics in feature extraction and deformation. Texture mapping is used to enhance the final results(Figure 1).

2.1. Feature Extraction

The function of this process is to obtain cognitively significant features represented by contours of the face space di-

† Email: meherzad | huangzy@comp.nus.edu.sg

‡ School of Computing, NUS, Singapore 119260. This work was supported partly by the NUS under grant RP3982704.

Published by the Eurographics Association, ISSN 1017-4656.

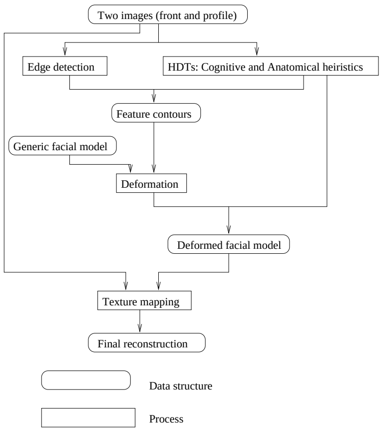

Figure 1: *Function model of reconstruction process.*

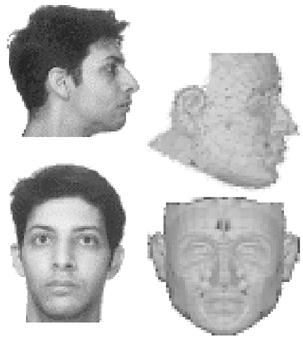

Figure 2: *Two images of the first author's face (left) used as input and a generic facial model (right)*

mensions - e.g., the eyes, lips, etc. We do this by the technique of Hierarchical Deformable Templates (HDTs) which is an extension of the snake method introduced by [9]. An HDT represents hierarchically encoded anatomical expertise to help finding the cognitively significant features. Essentially, this involves taking a pre-defined face template containing contours of a generic face model and allowing these contours to be deformed by facial images in a constrained fashion. These constraints are represented in a hierarchical manner and codify the anatomical expertise required to constrain the search space for individual contours.

A top-down approach is used to extract the facial contours. First a face outline template finds the approximate location of the face. It then enforces approximate feature locations for the eyes, for instance, which are the child nodes of the face HDT. Finally the third level HDTs enforce their own set of constraints on its own children, e.g., the eye HDT enforces its set of constraints on the eye-ball (pupil, cornea, and iris) and eyelid.

For each image, we use a separate HDT to extract its contours in 2D. The results of this process are then used to deform the generic 3D model in next step.

2.2. 3D Deformation

A common technique in reconstructing the 3D structure of the face is to deform a generic 3D face instead of recreating the model from scratch (e.g. using Cyberware scanners [2]). Most prior work deforms the generic face using structural information derived from photogrammetric techniques. Our approach to deformation is different. We perform a two-stage deformation on the generic face, once for each image. We pre-define a feature-ring of vertices for each facial feature on the generic facial model corresponding to those extracted by HDTs in the extraction phase. These features represent the most cognitively significant parts of the face. Each of these features is then deformed using the same snake-based technique to match the derived contours. The rest of the face is deformed by interpolation based on their locations relative to the feature vertices of the feature-ring.

2.3. Texture Mapping

To provide additional realism, we texture-map the two original images onto the deformed model. We use a very simple squared-inverse-distance blending function to create the texture map and then set up the correspondence by using a cylinder to project the map onto the face.

3. Sample Results

A sample result is shown on the next page. Input is shown in Figure 2: two real facial images and a generic 3D face model. In Figure 3, the reconstruction results of the top row are obtained using only one image - the front view of the face. The results in bottom row are obtained using two images - front and profile. The 3D features are better reconstructed using two images. But if only the front view image is available for a face, e.g., from a web page, ID or passport photo, the 3D-reconstruction result is still very distinguishable.

4. Ongoing Work

To achieve full automation of reconstruction is our final goal. At present, user interaction is required under some circumstances when we do not obtain optimal results from the automatic feature extraction process, e.g., noisy images. Therefore, we are trying to make this process more robust. At present, we only make use of edge-gradient information and anatomical heuristics. We are exploring methods of introducing a probabilistic framework over multiple criteria. It will encode more useful target information into HDTs.

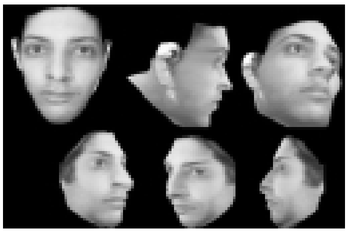

Figure 3: *The reconstruction results using one image (top) and two images (bottom) after texture mapping*

Figure 4: *The reconstruction result using two images, Copyright 1995 University of Bern*[5].

Figure 5: *The reconstruction result using only one portrait of the second author.*

References

1. T. Akimoto, Y. Seunaga, and R. S. Wallace, Automatic Creation of 3D Facial Models. *IEEE Computer Graphics and Applications*, Vol. 13, No. 5, pp. 16-22.
2. Cyberware Laboratory, 4020/RGB 3D Scanner with Color Digitizer. *Cyberware Laboratory, Inc*, Monterey, California, 1990.
3. I. Essa, S. Basu, T. Darrell, and A. Pentland, Tracking and Interactive Animation of Faces and Heads Using Input from Video. *Computer Animation'96*, pp. 68-79.
4. P. Fua and C. Miccio, From Regular Images to Animated Heads: A Least Squares Approach. *ECCV'98*, Freiburg, Germany.
5. *ftp://iamftp.unibe.ch/pub/Images/FaceImages/README*.
6. B. Guenter, C. Grimm, D. Wood, H. Malvar, and F. Pighin, Making Faces. *ACM SIGGRAPH'98*, pp. 55-66.
7. R. A. Johnston and H. D. Ellis, The development of face recognition. *In Cognitive and Computational Aspects of Face Recognition: Explorations in Face Space*. Routledge, 1995.
8. P. Kalra and N. Magnenat Thalmann, Modeling of Vascular Expressions in Facial Animation. *Computer Animation'94*, pp. 50-58.
9. M. Kass, A. Witkin, and D. Terzopoulos, Snakes: Active Contour Models. *ICCV'87*, pp. 259-268.
10. T. Kurihara and K. A. Arai, Transformation Method for Modeling and Animation of Human Face from Photographs. *Computer Animation'91*, pp. 45-58.
11. Y. Lee, D. Terzopoulos, and K. Waters, Realistic Modeling for Facial Animation. *ACM SIGGRAPH'95*, pp. 55-62.
12. W. S. Lee, P. Kalra, and N. Magnenat Thalmann, Model Based Face Reconstruction For Animation. *MMM'97*, pp. 323-338.
13. L. Moccozet and N. Magnenat Thalmann, Dirichlet Free-Form Deformation and their Application to Hand Simulation. *Computer Animation'97*, pp. 60-70.
14. F. I. Parke, A Model for Human Faces that allows Speech Synchronized Animation. *Computer and Graphics*, Pergamon Press, Vol. 1, No. 1, 1975. pp. 1-4.
15. F. I. Parke and K. Waters, Computer Facial Animation. *AK Peters, Wellesley*, Massachussetts, 1996.
16. F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. Salesin, Synthesizing Facial Expressions from Photographs. *ACM SIGGRAPH'98*, pp. 75-84.
17. S. Platt and N. Badler, Animating Facial Expressions. *ACM SIGGRAPH'81*, pp. 245-252.
18. M. Pollefeys, R. Koch, and L. Van Gool, Self-Calibration and Metric Reconstruction in spite of Varying and Unknown Internal Camera Parameters. *ICCV'98*, Bombay, 1998, pp. 90-95.
19. M. Prosmans and L. Van Gool, Getting Facial Features and Gestures in 3D. *In the Proceedings of the NATO Advanced Study Institute on Face Recognition: From Theory to Applications*, Stirling, Scotland, UK, June 23-July 4, 1997.
20. T. W. Sederburg, Free-Form Deformation of Solid Geometric Models. *ACM SIGGRAPH'86*, pp. 151-160.
21. S. V. Stevenage, Expertise and the Caricature Advantage. *In Cognitive and Computational Aspects of Face Recognition: Explorations in Face Space*. Routledge, 1995, pp. 24-46.
22. K. Waters, A Muscle Model for Animating Three-Dimensional Facial Expression. *ACM SIGGRAPH'87*, pp. 17-24.