EXAMPLE BASED 3D FACE RECONSTRUCTION

George Stylianou *
Cyprus College
P.O. Box 22006, Nicosia, Cyprus.
gstylianou@cycollege.ac.cy

Andreas Lanitis
Department of Multimedia and Graphic Arts, Cyprus University of Technology
P.O. Box 50329, Limassol, Cyprus
andreas.lanitis@gmail.com

ABSTRACT

The use of 3D data in face image processing applications can be of utmost importance in developing truly robust face image processing systems. However, a limiting factor towards the widespread use of 3D face image processing systems is the failure of the 3D scanning technology to become small, robust and cheap. As an alternative, considerable effort has been invested in developing 3D face reconstruction methods that allow the generation of a 3D face model based on a single face image. In this paper we review example-based 3D face reconstruction methods reported in the literature, discuss the limitations of existing methods and outline future trends for further work in this area.

KEYWORDS

3D Face Models, Deformable Models, Face Reconstruction.

1. INTRODUCTION

Face image processing figures in a wide range of real life applications including security related applications, access control, human machine interaction, animation and low-bit communication. Most of the applications stated above require the use of robust algorithms capable of dealing with all possible destructing factors that may be encountered in face images such as variation due to different imaging conditions and facial expressions, the introduction of occlusions and variation in individual facial appearance.

The increasing requirements for improved and more robust systems initiated significant interest towards the development of 3D (instead of 2D) face image processing applications. 3D face data provide an accurate representation of the geometry of a face image that is potentially useful for face interpretation tasks. Usually facial geometry is coupled with the face texture that provides supplementary information so that a comprehensive overall face representation is achieved.

A limiting factor towards the wide-spread use of 3D face image processing systems is the failure of the 3D scanning technology to become small, robust and cheap. For this reason along with advances in 3D face image processing, considerable effort has been invested in the development of 3D face reconstruction systems capable of generating a face 3D model based on a single 2D image obtained using cheap and widely available technology such as digital cameras and camcorders. The availability of robust 3D face reconstruction methods will enable virtually everybody to implement and use 3D face processing applications.

In the present study we describe the state of the art in the example based 3D face reconstruction that uses PCA[1]-based 3D deformable models. The review is organized as follows: In section 2 we present the

[1] PCA = Principal Component Analysis

PCA 3D deformable models reported in the literature. Subsequently we describe the main methods developed for example-based 3D face reconstruction. A discussion and future trends of research in this area are presented in section 4.

2. PCA-BASED 3D MODELS

PCA-based models are generated by applying PCA on the covariance matrix of the deviations of training samples from the mean sample, in order to extract the eigenvectors that define the main modes of variability within a training set. As a result of the analysis, training examples can be coded into a low dimensional space representing the weights applied to each eigenvector. Eigenvector weights control the deviation of a sample from the mean example among the training set in terms of the main modes of variability within the training set. New model instances can be generated based on the equation $\mathbf{X} = \overline{\mathbf{X}} + \sum_{\iota=1}^{m} \alpha_\iota \mathbf{x}_\iota$ (1) where **X** is a new model instance, $\overline{\mathbf{X}}$ is the mean example among the training set, $\mathbf{x_i}$ is the ith eigenvector, α_i are the weights applied to each eigenvector (or model parameters) and m is the number of significant eigenvectors within a training set.

The basic method for building PCA-based models has been used for generating shape, texture and combined shape and texture deformable 2D or 3D models (Edwards et al. 1998, Blanz and Vetter 1999). In the case of a shape model each training sample is represented by the 3D coordinates of all vertices in the triangle mesh and in the case of texture models training samples are represented by the RGB intensities of each vertex. Figure 1 demonstrates the method for building 3D deformable face models.

Variations of PCA-based models have been used in various applications including the recovery of 3D facial structure based on a 2D face image (Blanz and Vetter 1999, Blanz and Vetter 2002), face recognition (Blanz and Vetter 2003), face tracking (Baker et al. 2004, Blanz, Basso et al. 2003}, reconstruction based on incomplete data (Basso et al. 2005, Basso et al. 2006, Blanz and Vetter 2002), synthesis of 3D faces (Blanz and Vetter 1999), caricaturing (O' Toole et al. 1997), age progression (Hutton et al. 2001) and face animation (Blanz and Vetter 2003).

3. 3D FACE RECONSTRUCTION

Example-based 3D face reconstruction methods are statistical methods that use a database of 3D example faces to produce a generic morphable model that is deformed to match a novel image of a face (Atick et al. 1996, Vetter and Blanz 1998, Blanz and Vetter 1999). These methods use usually a single face image as an input. Figure 2 demonstrates the general PCA-based 3D face reconstruction method.

Atick et al. (1996) proposed the first example-based method for reconstructing a 3D face from a single image. First, they acquire a database of 200 scans of 3D faces, parameterize all faces in cylindrical coordinates and align them using 3D rigid transformations. Principal Component Analysis (PCA) is used to compute the eigenvalues and eigenvectors of the training set so that a novel 3D face can be represented using the equation $r(\theta,l) = r_0(\theta,l) + \sum_i \alpha_i \Psi_i(\theta,l)$ where $r_0(\theta,l)$ is the average face, Ψ_ι are the eigenvectors and α_ι are the coefficients that specify the novel 3D face. The authors suggest that for better results training samples are grouped in clusters, i.e. males or females so that a gender-specific model is used during the reconstruction process. In order to reconstruct the 3D face they use the conjugate gradient optimization method for minimizing an objective function that measures the difference between the input image and the image produced by the rendered 3D face model instance.

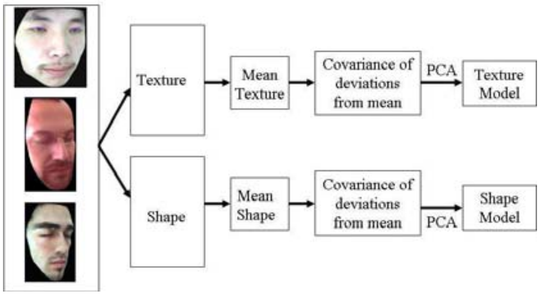

Figure 1. Block diagram of the method used for building 3D PCA-based models.

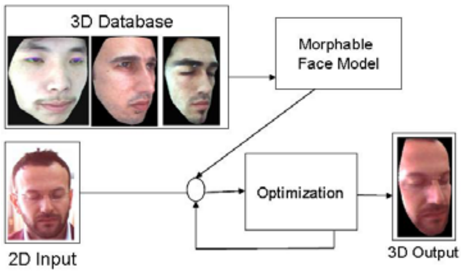

Figure 2. The framework for example based 3D face reconstruction.

Vetter and Blanz (1999) also use a PCA-based model for 3D face reconstruction. During the reconstruction process they compute the weights of the shape and texture eigenvectors required for synthesizing a face similar to the one shown in a previously unseen face image. The method requires manual initialization so that the 3D morphable model and the novel face image are roughly aligned. During the reconstruction process 22 rendering parameters such as pose, focal length of the camera, light intensity, color and direction are estimated. The reconstruction is done using an analysis by synthesis method that minimizes the distance function $E = \sum_{x,y} \| I_{input}(x,y) - I_{model}(x,y) \|^2$ where $I_{input}(x,y)$ is the input face image and $I_{model}(x,y)$ is the 2D face image generated by projecting a 3D deformable model instance in the two dimensional space and E is the reconstruction error. Even though the method focuses on 3D face reconstruction from a single face image, the authors show that it is trivial to extend it such that reconstruction is done using multiple face images in an attempt to obtain more accurate results.

Hu et al. (2004) propose a method quite similar to Vetter and Blanz (1999). PCA was used, as in Blanz and Vetter (1999) on all faces in the database in order to generate a PCA-based morphable model. The vital difference is the fact that this method is completely automatic. Frontal face detection and 2D alignment are used to locate the face and 83 facial feature points such as the contour points of the face, left and right eyes, mouth and nose. Then the 3D face is reconstructed based on the feature points and the morphable model. A major limitation of this method is that it does not attempt to estimate rendering parameters; hence the proposed method will not perform well under arbitrary lighting conditions.

In order to deal with the illumination problem exhibited by all previous work, Lee et al. (2005) use two face databases. The first database stores 205 3D face models and the second database stores 2336 faces images captured under different illuminations and 3D orientations. All 3D faces are registered to each other using 40 feature points and the illumination 2D images are registered to the 3D geometry of the subject faces. The first database is used to generate the PCA based model. The second database is used to recover as accurately as possible the illumination conditions of the input images.

4. DISCUSSION AND CONCLUSION

Reconstructing 3D faces using PCA-based models has proved to be a promising 3D face reconstruction method. The main advantage of this method is the ability to reconstruct the 3D structure of a subject's face using a single face image of the subject. Although in principle the problem of 3D reconstruction from a single 2D image is an ill-posed problem, the use of a PCA model enables the estimation of the 3D structure of a face by augmenting information from a single face image with information pertaining to the 3D geometry of a face as learned from the training samples. As a result it is possible to obtain reasonable reconstructions using a single face image. However, PCA-based 3D face reconstructions methods exhibit the following disadvantages:

Generation of smoothed faces: In PCA-based models high frequency modes of variations are eliminated in favor of more subtle and systematic sources of variation within the training set. As a result PCA-based deformable face models generate smoothed faces. In order to overcome this problem and increase the level of

detail expressed in PCA-based reconstructions Blanz and Vetter (1999) generate local models for different areas of the face such as the mouth, nose and eyes. Further work is required in order to deal with this issue.
Generalization ability: PCA-based models are generated based on a training set; hence they are able to reconstruct faces similar to the ones encountered in the training set. For example a PCA-based model will not be able to accurately reconstruct the face a subject belonging to a certain ethnic origin, unless subjects from the same ethic origin are included in the training set. The issue of generating face PCA-models that can efficiently generalize and cope with any faces still needs to be investigated. The non-existence of truly diverse 3D face databases is a limiting factor in this effort.
Mathematical framework: PCA-models rely on PCA analysis, which is a linear transformation. However the 3D facial appearance often undergoes non-linear transformations hence for better results it is required to use non-linear statistical analysis techniques (instead of PCA) in order to produce more specific face models.
Computational load: Due to the high dimensional search space and due to the considerable time required for assessing a candidate solution, the computational load involved in analysis by synthesis 3D face reconstructions is significant. As a result PCA-based 3D face reconstruction is not performed in real time. Further work is required in order to develop more efficient optimization algorithms in an attempt to minimize the computation load involved.
Performance evaluation: In most cases researchers who investigated 3D face reconstruction evaluated their systems in terms of the looks of reconstructed faces. It is desirable to use quantitative evaluation measures for assessing the ability of different methods to reconstruct accurately the 3D geometry of a face, rather than the ability to produce aesthetically nice results. Work in this area will also enable the accurate comparison of different techniques reported in the literature.
Despite the problems encountered, PCA-based 3D face reconstruction is a promising approach to the problem. Due to the increased need for real-time accurate 3D data acquisition using low cost hardware, it is expected that work in this area will be reinforced in order to deal with the limitations stated above.

ACKNOWLEDGEMENTS

This work was supported by a Cyprus Research Promotion Foundation research grant (EPYAN/0205/08) and Cyprus College.

REFERENCES

Atick, J.J., Griffin, P.A., Redlich, N.A., 1996. Statistical Approach to Shape from Shading: Reconstruction of 3D Face Surfaces from Single 2D Images. *In Neural Computation*, Vol 8, pp 321-1340.

Baker, S., Matthews, I., Xiao, J., Gross, R., Kanade, T., Ishikawa, T., 2004. Real-Time Non-Rigid Driver Head Tracking for Driver Mental State Estimation. *Proceedings of 11th World Congress on Intelligent Transportation Systems*.

Basso, C. and Vetter, T., 2005. Statistically Motivated 3D Faces Reconstruction. *Proceedings of the 2nd International Conference on Reconstruction of Soft Facial Parts*, Vol 31, No 2. BKA Research Series, Luchterhand Publishers.

Basso, C., Paysan, P., Vetter, T., 2006. Registration of Expressions Data using a 3D Morphable Model. *Proceedings of the 7th IEEE International Conference Automatic Face and Gesture Recognition*.

Blanz, V. and Vetter, T., 1999. A Morphable Model For The Synthesis Of 3D Faces. *Proceedings of Siggraph*, pp 187-194, ACM.

Blanz, V. and Vetter, T., 2002. Reconstructing the Complete 3D Shape of Faces from Partial Information. Computer Graphics Technical Report, University of Freiburg.

Blanz, V. and Vetter, T., 2003. Face Recognition Based on Fitting a 3D Morphable Model. *In IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol 25, pp 1063-1074.

Blanz, V., Basso, C., Poggio, T., Vetter, T., 2003. Reanimating Faces in Images and Video. *Computer Graphics Forum*, Vol 22, No 3, pp 641-650.

Edwards, G.J., Lanitis, A., Taylor, C.J., Cootes, T.F., 1998. Statistical Face Models: Improving Specificity. In *Image and Vision Computing*, Vol 16, no 3, pp 203-211.

Hutton, T.J., Buxton, B.F., Hammond, P., 2001. Dense Surface Point Distribution Model of the Human face. Proceedings of the CVPR Workshop on Mathematical Methods in Biomedical Image Analysis.

Hu, Y., Jiang, D., Yan, S., Zhang, L., Zhang, H., 2004. Automatic 3D Reconstruction for Face Recognition *Proceedings of the IEEE Int. Conf. on Automatic Face and Gesture Recognition*.

Lee, J., Machiraju, R., Pfister, H., Moghaddam, B., 2005. Estimation of 3D Faces and Illumination from Single Photographs Using a Bilinear Illumination Model. *Proceedings of the Eurographics Symposium on Rendering Techniques*, pp 73-82.

O'Toole, A.J., Vetter, T., Volz, H., Salter, E., 1997. Three-Dimensional Caricatures of Human Heads: Distinctiveness and the Perception of Age, *In Perception*, Vol 26, pp 719-732.

Vetter, T. and Blanz, V., 1998. Estimating coloured 3D face models from single images: An example based approach. *Proceedings of the European Conference on Computer Vision*, Vol II, pp 499-513.