10-2005

Automatic 3D Face Modeling Using 2D Active Appearance Models

Jianke ZHU
Chinese University of Hong Kong

Steven HOI
Singapore Management University, chhoi@smu.edu.sg

Michael R. LYU
Chinese University of Hong Kong

Automatic 3D Face Modeling Using 2D Active Appearance Models

Jianke Zhu, Steven C.H. Hoi, Edward Yau and Michael R. Lyu
Department of CS&E, The Chinese University of Hong Kong, Hong Kong
E-mail: {jkzhu,chhoi,edyau,lyu}@cse.cuhk.edu.hk

Abstract

Although a lot of promising research findings have been studied on 3D face modeling in the past years, it is still a challenge to generate realistic 3D human face models and facial animations. This paper presents a novel approach to model 3D faces automatically from still images or video sequences without manual interactions. Our proposed scheme comprises three steps. First, we offline construct 3D shape models using Active Appearance Models (AAMs), which saves large computation costs for online modeling. Second, based on the computed 3D shape models, we propose an efficient algorithm to estimate the parameters of 3D pose and non-rigid shape via local bundle adjustment. Third, we employ the recovered 3D face shape to deform the high resolution face mesh through scattered data interpolation, and extract the face texture maps for rendering the reconstructed face model from various viewpoints. Since the correspondence is built by AAMs fitting, the 3D face model can be constructed effectively even from a single image.

Keywords: Face Modeling and Animation, Active Appearance Models, Non-Rigid Shape Recovery, Bundle Adjustment

1. Introduction

Face models are essential to computer games, film making, online chat, virtual presence, and video conferencing, etc. The goal of face modeling is to build an automated modeling system that can systematically find features and correspondences with minimal user intervention. Face modeling has been widely studied in the past years [1][2][5][7]. A comprehensive overview of face modeling and animation can be found in [4].

The key challenge of face modeling from still images or video frames is the difficulty in establishing accurate and reliable correspondences between facial images since faces are non-rigid and their images have a high degree of variability in shape, texture, pose, and imaging conditions. Moreover, obtaining accurate estimation results often requires huge computational cost, which can be critical for many online applications.

In this paper we propose a novel cost-effective scheme to model 3D faces automatically from images and video sequences. The rest of this paper is organized as follows. Section 2 briefly reviews the Active Appearance Models and discusses our improved method for better performance. Section 3 provides the overview of our proposed scheme. Section 4 describes our method for non-rigid shape and pose recovery. Section 5 presents the issues of experimental implementations and our experimental results. Section 6 concludes our work.

2. Active Appearance Models

The Active Appearance Models (AAMs) [3] are a successful method for matching statistical models of appearance to new images. This approach has been applied in numerous different applications. AAMs establish compact parameterizations of object variability, as learned from a training set by estimating a set of latent variables. The modeled object properties are usually shape and pixel intensities.

Instead of using a traditional approach for AAM matching in [3], we implement a modified AAM fitting algorithm for quicker convergence and better matching accuracy similar to the approach in [3]. The proposed iterative AAM matching algorithm predicts shape directly from texture [3]. In our experiments, the AAMs are built up with 140 still face images belonging to 20 individuals, seven images for each. Each image is manually labeled with 100 points.

3. Overview of Our Proposed Scheme

Toward an automatic scheme for 3D face modeling from still images or video sequences, we propose and develop a novel cost-effective system, illustrated in Figure 1. To enable an efficient online solution, the 2D AAMs are trained offline before launching the online face modeling.

The first stage is to obtain the facial images of the human to be modeled. The facial images can be captured either from a camera or a video clip. In general, only one camera is required to capture the images for building face models, although our scheme can also be applied to multiple cameras for better results.

The second stage is to fit the 2D face mesh using

AAMs for the given facial images. The 2D face mesh is found by the extended AAM matching algorithm provided in Section 2, which is more accurate and robust than regular approaches.

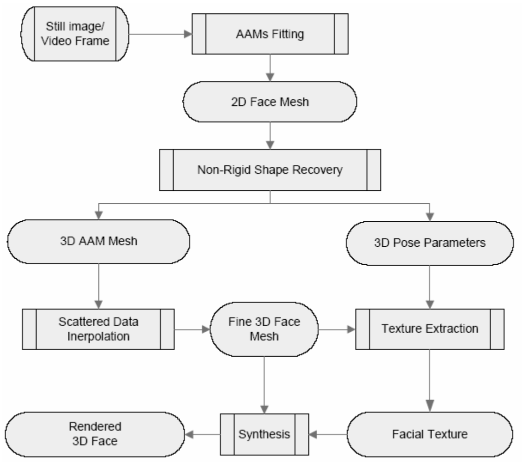

Figure 1: Overview of our proposed system

The third stage is the key component in our scheme which estimates the 3D pose and shape parameters from 2D face mesh results. To solve the challenging problem, we propose a novel two-stage solution comprised of offline construction of 3D shape basis and online estimation of 3D shape and pose parameters. To make it feasible for real-time applications, we propose an efficient online algorithm to simultaneously compute the 3D pose and 3D shape parameters.

The fourth stage is to construct a fine 3D face mesh and extract facial texture on the 3D model. To obtain a more detailed 3D model, we suggest applying the scattered data interpolation [5] technique to build the fine 3D model. Based on the computed fine 3D model and the 3D pose parameters, we can easily extract an effective facial texture of the face model. Note that a blending procedure on multiple texture images from different views is typically needed to enable the rendering in the appearance of large rotations [5].

After the fine 3D face mesh and facial texture map are available, we can directly apply standard rendering techniques to synthesize them for generating realistic 3D face animation or other applications. In the following sections, we describe the details of the main techniques and algorithms in our proposed scheme.

4. Non-Rigid Shape and Pose Recovery

In this section, we describe how to estimate the 3D pose parameters and the non-rigid shape parameters simultaneously, and how to develop an efficient algorithm for real-time applications. Traditional techniques may be neither flexible and powerful enough for model representations nor efficient enough for real-time purposes. For tackling the challenges, we solve the problem by a two-stage scheme via AAM techniques. The overall idea of our proposed scheme can be summarized as follows:

We acquire the 2D shape of objects using the AAM fitting algorithm described in Section 2 firstly, and then construct the 3D shape basis offline based on the AAM fitting results. We estimate the 3D pose and 3D shape parameters online simultaneously via local bundle adjustment by building up the point correspondences of 2D and 3D. The optimization procedure can converge quickly within a couple of iterations when it begins with a good initial estimation.

The above proposed solution differs from the regular approach in [6], which estimates the pose of an object through point matching. To exploit the representational power of AAMs, instead of matching points between two frames, we propose a novel approach to set up the point correspondences between 2D and 3D shapes via AAM fitting to a single image. This procedure needs no manual initialization.

5. The 3D Face Modeling

5.1. Experimental Setup

The first step is to construct the 3D face model for a specific person. Based on the built 3D face model, the second step is to extract the texture maps for rendering photorealistic images in different views.

To build a fine 3D face model, we propose to apply the scattered data interpolation technique [5] to obtain a more detailed 3D face model based on the 3D AAM mesh from the non-rigid shape and pose recovery stage. Given the fine 3D face model, view-dependent texture maps can be extracted with the computed 3D pose information. Figure 2 shows two examples of blended face texture images extracted by our approach. All the rendering operations are done using graphics hardware, and the approach is very fast. Based on the fine 3D face model and the texture maps, we can further synthesize new views from given images or video sequences.

5.2. Experimental Results

To evaluate the performance of the 2D AAM fitting algorithms and the 3D face modeling scheme, Figure 3 shows the experimental results of selected frame in the input video sequence. The figure contains four blocks: the recovered 3D non-rigid shape (top left), the rendered novel view rotated from the current pose by -15° on both X-axis and Y-axis (top right), the rendered novel view rotated from the current pose by 20° on Y-axis (bottom right), and the fine 3D mesh of the bottom right view (bottom left). From the experimental results, we can observe that our AAM matching algorithm can fit perfectly on the given frame and the recovered 3D non-rigid mesh matches very well in the sequences.

Figure 2: Blended face texture images

In order to demonstrate our proposed 3D face model is effective and promising for 3D facial animation purposes, we employ the constructed 3D face model to render the novel views by mapping different textures on the 3D face model. Figure 4 shows the experimental result by replacing face texture of a person with another one. In the figure, the top left one is the modeled person and the bottom left is the constructed 3D mesh in which 3D pose information is available. The top right one is the front face of the replaced person and the bottom right shows the generated results by replacing the texture using the built 3D model and pose parameters. The generated view fits well on the 3D model.

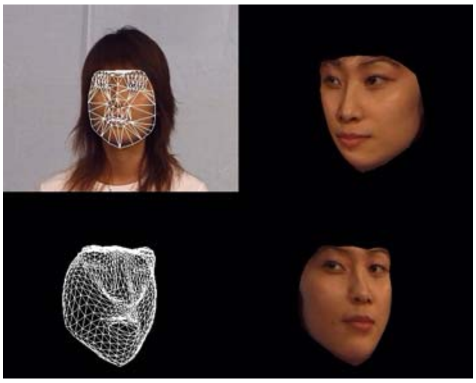

Figure 3: The modeling results of selected frame in the female video sequence

6. Conclusion

This paper proposed a novel solution for modeling 3D faces automatically from still images or video clips. Our main contributions can be summarized as follows. First of all, we formulated a unified scheme for automatic 3D face modeling and animations. In our scheme, a two-stage solution was suggested for the recovery of non-rigid 3D shape and pose parameters simultaneously via local bundle adjustment, along with the extended AAMs to build up point correspondences. Moreover, we implemented the high-resolution mesh for modeling the face through interpolating the recovered 3D shape and extracted the facial texture map on a virtual cylinder enclosing the face model. Finally, we conducted extensive experiments to evaluate the performance of our scheme and algorithms, in which promising results have shown our scheme is effective and promising for automatic 3D face modeling and animation applications.

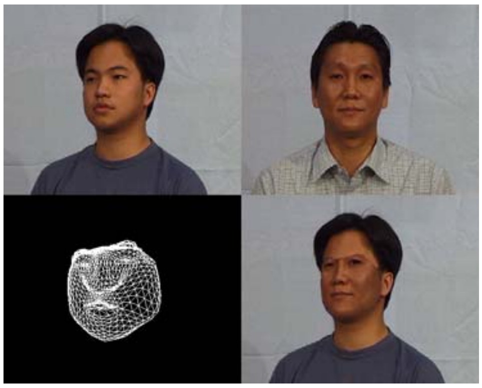

Figure 4: Synthesis result using different texture

Acknowledgments

The work described in this paper was fully supported by two grants: Innovation and Technology Fund ITS/105/03, and the Research Grants Council Earmarked Grant CUHK4205/04E.

References

[1] V. Blanz and T. Vetter, “Face recognition based on fitting a 3d morphable model”, *IEEE Trans. PAMI*, vol. 25, 2003.

[2] C.Bregler, A.Hertzmann, and H.Biermann., “Recovering non-rigid 3d shape from image streams”, *IEEE CVPR’2000*, vol. 2, pp. 690–696, 2000.

[3] T.F. Cootes and P. Kittipanya-ngam, “Comparing variations on the active appearance model algorithm”, *BMVC’2002.*

[4] F. I. Parke and K.Waters, *Computer Facial Animation*. AKPeters, Wellesley, Massachusetts, 1996.

[5] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. Salesin, “Synthesizing realistic facial expressions from photographs”, *SIGGRAPH’98,* pp. 75–84, 1998.

[6] L. Vacchetti, V. Lepetit, and P. Fua, “Stable realtime 3d tracking using online and offline information”, *IEEE Trans. PAMI*, vol. 26(6), October 2004.

[7] Jing Xiao, S. Baker, I. Matthews, and T. Kanade, “Realtime combined 2d+3d active appearance models”, *CVPR’2004,* vol. 2, pp. 535–542, 2004.