5-2006

Real-Time Non-Rigid Shape Recovery via Active Appearance Models for Augmented Reality

Jianke ZHU
Chinese University of Hong Kong

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

Michael R. LYU
Chinese University of Hong Kong

Real-Time Non-Rigid Shape Recovery via Active Appearance Models for Augmented Reality

Jianke Zhu, Steven C.H. Hoi, and Michael R. Lyu

Department of Computer Science & Engineering,
Chinese University of Hong Kong, Shatin, Hong Kong
{jkzhu,chhoi,lyu}@cse.cuhk.edu.hk

Abstract. One main challenge in Augmented Reality (AR) applications is to keep track of video objects with their movement, orientation, size, and position accurately. This poses a challenging task to recover non-rigid shape and global pose in real-time AR applications. This paper proposes a novel two-stage scheme for online non-rigid shape recovery toward AR applications using Active Appearance Models (AAMs). First, we construct 3D shape models from AAMs offline, which do not involve processing of the 3D scan data. Based on the computed 3D shape models, we propose an efficient online algorithm to estimate both 3D pose and non-rigid shape parameters via local bundle adjustment for building up point correspondences. Our approach, without manual intervention, can recover the 3D non-rigid shape effectively from either real-time video sequences or single image. The recovered 3D pose parameters can be used for AR registrations. Furthermore, the facial feature can be tracked simultaneously, which is critical for many face related applications. We evaluate our algorithms on several video sequences. Promising experimental results demonstrate our proposed scheme is effective and significant for real-time AR applications.

1 Introduction

1.1 Augmented Reality

The objective of Augmented Reality (AR) is to integrate virtual objects into real-world video sequences, enabling computer generated objects to be overlaid on the video in such a manner as to appear part of the viewed 3D scene. Recently, some well-known AR toolkits have been developed for AR applications [1]. Although these tools have facilitated the AR applications to obtain good registration data automatically and robustly, it is still a challenging and open issue to keep track of objects with their movement, orientation, size, and position accurately in AR applications. This critical requirement also results in an important problem, i.e., determining the position and orientation of an object, which plays an important role in many research areas such as robotics, computer vision, computer graphics.

In the subsequent part we describe some recent advances of technologies for object tracking and shape recovery in the computer vision community. Along with the introduction of previous work, we provide the motivation and brief introduction of our work in this paper particularly for AR applications.

1.2 Previous Work and Motivation

L. Vacchetti et al. [2] proposed an efficient real-time solution for tracking rigid objects in 3D scene using a single camera. They demonstrated that the virtual glasses and masks can be added on to the head. Since they employed a rigid 3D model, the local facial feature was not able to be located and tracked. In addition, a few keyframes were required to make the tracker more robust. L. Vacchetti et al. pointed that it was very convenient to estimate the camera position from a single image in order to initialize the tracker and to recover the failure automatically. Active Appearance Models based approaches [3–5] provide a good solution to recover the 2D affine pose parameters along with the feature points from single image. Recently, researchers [6–8] have attempted to build the AAM with three dimensions.

P. Mittrapiyanumic [6] proposed two AAMs algorithms for rigid object tracking and pose estimation. The first method is to utilize two instances of AAM to track landmark points in a stereo pair of images and perform 3D reconstruction of the landmarks followed by 3D pose estimation. The second method, i.e., AAM matching algorithm, is an extension of the original AAM that incorporates the full six degrees of freedom pose parameters as part of the parameters for the minimization. The results showed that the accuracy in pose estimation of appearance based methods is better than the methods using the geometric approach. J. Ahlberg [7] proposed an approach using the 3D AAM for face and facial feature tracking, in which the depth information of 3D shape was acquired by fitting a generic model. In addition, the pose parameters were estimated from a motion tracker, then the shape model parameters were recovered by AAM fitting.

Jing Xiao et al. [8] proposed a non-rigid structure-from-motion algorithm that could be used to convert a 2D AAM into a 3D face model. They then described how a non-rigid structure-from-motion algorithm was able to be employed to compute the corresponding 3D shape models from a 2D AAM. Their method did not require 3D range data in [9] and also shared fast fitting speeds. They then showed how the 3D modes could be used to constrain the AAM so that it could only generate model instances, but also could be generated with the 3D modes. Their fast fitting algorithm mainly benefited from the projection-out method and Inverse Compositional update strategy, thus the Jacobi matrix was constant. However, the approximation that the shape Jacobi matrix was made orthogonal to the texture Jacobi matrix, was only valid for few texture modes. Only shape parameters were recovered iteratively, and the texture parameters were recovered linearly in one step. In addition, the recovered pose parameters were not accurate enough, mainly because the pose parameters were compensated by the shape variations. A weak perspective camera model was employed in order to decrease the computational cost, and the full perspective camera model was necessary for the common AR applications. These may limit their applications particularly for AR applications.

This paper presents a novel scheme of real-time non-rigid shape recovery via active appearance models for augmented reality applications. The rest of this paper is organized as follows. Section 2 reviews the AAM algorithm and

describes an extended AAM matching algorithm which predicts shape directly from texture for improving the accuracy of AAM searching. Section 3 presents our proposed scheme. We first provide an overview of our scheme in the context of augmented reality applications in Section 3.1. Then Section 3.2 describes how to construct the 3D shape models based on the 2D AAM tracking results. Section 3.3 presents a novel and efficient algorithm for online estimation of 3D pose and non-rigid shape parameters simultaneously via local bundle adjustment. Section 3.4 gives our experimental results and the details of our experimental implementation. Section 4 discusses the critical requirements of real-time AR applications, several major differences of our proposed scheme compared with previous work, and the advantages of our scheme particularly for AR applications as well as the disadvantages and our future work. Section 5 sets out our conclusion.

2 An Extended AAM Matching Algorithm

The Active Appearance Models (AAMs) [3–5, 7] have been proven as a successful method for matching statistical models of appearance to new images. AAMs are taking the analysis-through-synthesis approach to the extreme. This approach has been successfully applied in numerous different applications. AAMs establish a compact parameterizations of object variability, as learned from a training set by estimating a set of latent variables. The modelled object properties are usually shape and pixel intensities. There are several modifications for the basic AAM algorithm [4]. One approach was the Direct Appearance Model (DAM) for improving the convergence speed and searching accuracy by predicting the shape directly from the texture [10].

The algorithm of AAM Matching
1. Generate texture vector $\mathbf{g}_m$ from model
2. Sample image below the model shape $\mathbf{g}_i$
3. Evaluate error vector $\mathbf{r} = \mathbf{g}_i - \mathbf{g}_m$ and error $\mathbf{E} = \lvert\mathbf{r}\rvert$
4. Compute displacements in pose $\delta\mathbf{t} = \mathbf{R_t r}$
5. Compute displacements in texture $\delta\mathbf{b_t} = \mathbf{R_g r}$
6. Update pose and texuture parameters with initial $k = 1$
7. Transform the shape by the estimated parameters
8. Repeat step 1-3 to form a new error $\mathbf{E}^{'}$
9. If $\mathbf{E}^{'} < \mathbf{E}$ accept the new estimate, otherwise goto step 6 to try other k=0.5, 0.25, ...

Fig. 1. An extended AAM matching algorithm

The AAM matching algorithm tries to minimize the residual between the model and image $\mathbf{r} = \mathbf{g}_i - \mathbf{g}_m$, where $\mathbf{g}_i$ is the sampled image below model shape, and $\mathbf{g}_m$ is the model texture. During the DAM training phase, one learns the relationships

$$\delta\mathbf{t} = \mathbf{R_t r} \ ,$$

$$\delta \mathbf{b_t} = \mathbf{R_g r} \ .$$

Instead of using a traditional approach for AAM matching in [3], we implement a modified AAM fitting algorithm for quicker convergency and better matching accuracy similar to the approach in [5]. The proposed iterative AAM matching algorithm which predicts shape directly from texture is given in Fig. 1.

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 labelled with 100 points. As shown in Fig. 2, the matching experiment is performed on an AAM with 14 shape parameters, 68 texture parameters, and 36335 color pixels. Fig 2 respectively show (a) the original single image, (b) the initialization of our AAM fitting, (c) the result after 10 iterations and the final converged result after 21 iterations. In each case the rendered model images and estimation errors are displayed in the figures.

(a) Original (b) Initialized (c) 10 iterations (d) Converged

Fig. 2. An example of our AAM fitting to a single image. The estimated errors are displayed in each case.

3 Real-Time Non-Rigid Shape Recovery for AR

3.1 Overview of Our Solution

Our scheme tries to attack the critical problems of pose and non-rigid shape recovery. 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:

- We acquire the 2D shape of objects using the AAM fitting algorithm described in Section 2 firstly, 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 between 2D and 3D.

The above proposed solution differs from the regular approach in [2] which estimated 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 setup the point correspondences between the 2D and 3D shape via AAM fitting to a single image. This procedure needs no manual initialization. The details of our approach are described as follows.

3.2 Offline Construction of 3D Shape Basis

Bregler et al. [11] proposed a solution for recovering 3D non-rigid shape models from image sequences. Their technique is based on a non-rigid model, where the 3D shape in each frame is a linear combination of a set of basis shapes. By analyzing the low rank of the image measurements, they proposed a factorization-based method that enforces the orthonormality constraints on camera rotations for reconstructing the non-rigid shape and motion. Torresani et al. [12] extended the method in [11] to initialize the optimization process. By using the extended AAM matching algorithm in Section 2, we first acquire the 2D shapes of objects. With the trained 2D shapes, we are able to construct the 3D shape basis due to the powerful representational capability of AAMs [8].

The 3D shape can be described as a set of key-frame basis $S_1, S_2, \cdots, S_m$. Each key-frame S_i is a $3 \times n$ matrix. The 3D shape of a specific configuration is a linear combination of the following basis set:

$$\mathbf{S} = \mathbf{S_0} + \sum_{i=1}^{m} p_i \mathbf{S_i} \qquad \mathbf{S}, \mathbf{S_i} \in R^{3\times n}, p_i \in R \tag{1}$$

where the coefficients p_i are the 3D shape parameters, and $\mathbf{S_i}$ are the 3D coordinates: $\mathbf{S} = \{\mathbf{M_1}, \mathbf{M_2}, \cdots, \mathbf{M_n}\}, \mathbf{M_i} \in R^{3\times 1}$. Under a weak perspective projection, the n points of $\mathbf{S}$ are projected into 2D image points (u_i, v_i):

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \\ v_1 & v_2 & \cdots & v_n \end{bmatrix} = \mathbf{R} \cdot (\sum_{i=0}^{m} p_i \mathbf{S_i}) + \mathbf{T} \tag{2}$$

$\mathbf{R}$ contains the first 2 rows of the full 3D camera rotation matrix, and $\mathbf{T}$ is the camera translation. The scale of the projection is coded in $p_1, p_2, \cdots, p_m$. The camera translation $\mathbf{T}$ is eliminated by subtracting the mean of all 2D points, and henceforth one can assume that $\mathbf{S}$ is centered at the origin.

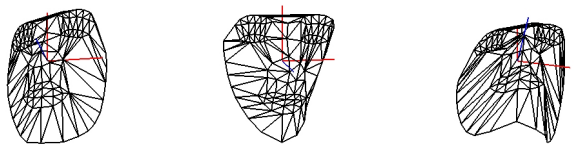
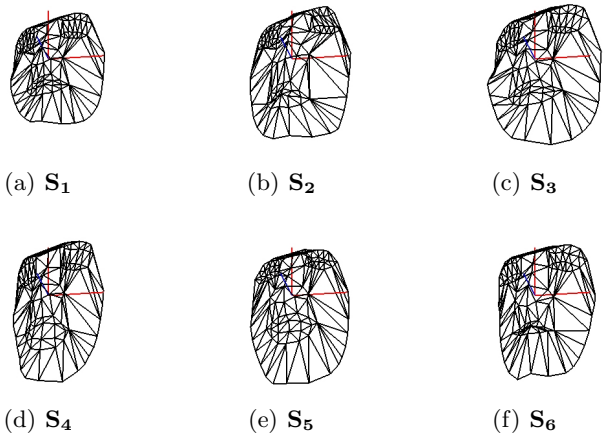


Fig. 3. An example of 3D mean shape of three views $\mathbf{S_0}$



(a) $\mathbf{S_1}$ (b) $\mathbf{S_2}$ (c) $\mathbf{S_3}$

(d) $\mathbf{S_4}$ (e) $\mathbf{S_5}$ (f) $\mathbf{S_6}$

Fig. 4. An example of the first six 3D shape modes (a-f) from an AAM.

If the AAMs are tracked through a sequence of N images, 2D points of the AAM shape in each frame can be obtained. Let us add a temporal index to each 2D point, and denote the tracked points in frame t as (u_i^t, v_i^t). All points of AAM shape in all N images are stacked into one large measure $2N \times n$ matrix W. The number of 3D shape verities equals to the number of 2D AAM vertices n, it can be rewritten as follows:

$$W = \begin{bmatrix} u_1^1 & u_2^1 & \cdots & u_n^1 \\ v_1^1 & v_2^1 & \cdots & v_n^1 \\ \vdots & \vdots & \vdots & \vdots \\ u_1^N & u_2^N & \cdots & u_N^n \\ v_1^N & v_2^N & \cdots & v_n^N \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{R}_1 & p_1^1\mathbf{R}_1 & \cdots & p_m^1\mathbf{R}_1 \\ \mathbf{R}_2 & p_1^2\mathbf{R}_2 & \cdots & p_m^2\mathbf{R}_2 \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{R}_N & p_1^N\mathbf{R}_N & \cdots & p_m^N\mathbf{R}_N \end{bmatrix}}_{\mathbf{M}} \cdot \underbrace{\begin{bmatrix} \mathbf{S_0} \\ \mathbf{S_1} \\ \vdots \\ \mathbf{S_m} \end{bmatrix}}_{\mathbf{B}} \tag{3}$$

where $\mathbf{M}$ is a $2N \times 3(m+1)$ scaled projection matrix and $\mathbf{B}$ is a $3(m+1) \times n$ shape matrix. In the noise-free case, $\mathbf{W}$ has a rank $r \leq 3(m+1)$, which can be factorized into the product of a $2N \times 3(m+1)$ matrix $\tilde{\mathbf{M}}$ and a $3(m+1) \times n$ matrix $\tilde{\mathbf{B}}$. This decomposition is not unique, which can be determined by a linear transformation. Any non-singular $3(m+1) \times 3(m+1)$ matrix $\mathbf{G}$ and its inverse could be inserted between $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{B}}$. In addition, their product still remains equal to $\mathbf{W}$. Namely, we have the following equations

$$\mathbf{M} = \tilde{\mathbf{M}} \cdot \mathbf{G} \tag{4}$$

$$\mathbf{B} = \mathbf{G}^{-1} \cdot \tilde{\mathbf{B}} \tag{5}$$

where the corrective matrix $\mathbf{G}$ can be found by solving a least square optimization problem [11]. Thus, given 2D tracking data $\mathbf{W}$, a non-rigid 3D shape matrix

with r degrees of freedom can be estimated, along with the corresponding camera rotations and configuration weights for each time frame.

In our experiments, we implement the AAM matching algorithm given in Section 2 and run it to fit the short video sequences of of 20 individuals (2678 frames in total). The training results are employed to construct the 3D shape basis in our experiments. Fig. 3 shows an example of the computed 3D mean shape modes of three views from AAM. Fig. 4 shows the first six 3D shape modes from an AAM.

3.3 Real-Time Non-Rigid Shape and Pose Recovery for AR

To make it flexible and general for wide applications, we employ the perspective camera model, in which a 3D point $\mathbf{Q}$ is re-projected based on the 2D point $\mathbf{q}$:

$$\mathbf{q} = \mathbf{A}[\mathbf{R}|\mathbf{T}] \cdot \mathbf{Q}$$

where the camera rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{T}$ estimated from the current frame are expressed in the object coordinate system, and $\mathbf{A}$ is the intrinsic camera matrix. The intrinsic parameters of the camera can be calculated offline. This does not require to be done precisely, and typically an approximate configuration is sufficient. Hence, we can assume the intrinsic parameters are fixed. Moreover, in order to allow some deformation, the rigid shape model is replaced by the 3D linear shape model. We now describe how to in real-time estimate the 3D pose parameters and non-rigid shape parameters simultaneously.

Given the constructed 3D shape basis via AAM training algorithm, we can build up the 2D-3D correspondences. Based on the established correspondences, an efficient way for estimating the parameters of camera position and the 3D shape coefficients can be turned into minimizing the re-projection error:

$$\min_{\mathbf{R},\mathbf{T},p} \rho\left(\mathbf{s}, \phi\left(\mathbf{A}[\mathbf{R}|\mathbf{T}], \mathbf{S}\right)\right) \tag{6}$$

Let $\mathbf{S} = \mathbf{S_0} + \sum_{i=1}^{m} p_i \mathbf{S_i}$, the optimization problem can be written as

$$\min_{\mathbf{R},\mathbf{T},p} \rho\left(\mathbf{s}, \phi\left(\mathbf{A}[\mathbf{R}|\mathbf{T}], \mathbf{S_0} + \sum_{i=1}^{m} p_i \mathbf{S_i}\right)\right) \tag{7}$$

with respect to the orientation and translation parameters $\mathbf{R}$ and $\mathbf{T}$, where

- ρ is the robust M-estimator [13] in consideration of outliers which can be given as follows:

$$\rho(u) = \begin{cases} \frac{\alpha^2}{6}[1 - (1 - (\frac{u}{\alpha})^2)^3], & |u| \leq \alpha \\ \frac{\alpha^2}{6}, & |u| > \alpha \end{cases} \tag{8}$$

- $\phi\left(\mathbf{A}[\mathbf{R}|\mathbf{T}], \mathbf{S_0} + \sum_{i=1}^{m} p_i \mathbf{S_i}\right)$ denotes the projection of 3D shape given the parameters $\mathbf{A}$, $\mathbf{R}$ and $\mathbf{T}$.

The above optimization procedure can converge quickly within a couple of iterations when it begins with a good initial estimation.

3.4 Experimental Results

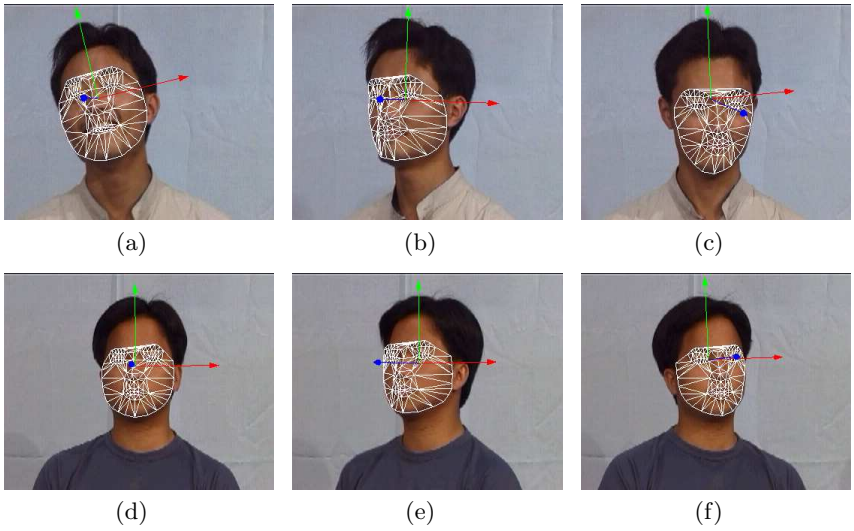


Fig. 5. Tracking faces using proposed method in the augmented video sequences, the axis in the displayed frames indicates the current 3D pose of tracked subject.

The results of estimated 3D shapes of two individuals are depicted in Fig. 5, which are extracted from two video clips with total 300 frames. We can see that the 3D shapes are successfully fitted to the face image. The face deformation can be well described by 6 3D shape parameters, for example, fitting to different individuals with the same AAM model in Fig. 5(a-f). The algorithm can handle large pose variations and displacements, as shown in Fig. 5(a,b,e,f). Fig. 5(a,c) revealed that the proposed approach can handle tilt pose, and Fig. 5(d-f) displayed the results which deal with out-of-plane rotation. In each result image, the axis indicates the current orientation and translation. Since the intrinsic and extrinsic camera matrices are computed, the virtual rigid and deformable objects can be inserted into the scene. Fig. 6 shows that a rigid virtual glasses and a deformable beard are added into the video sequences. From the results, we can observe that the beard can be deformed along with the expression changes. The added virtual objects are tightly overlaid on the subject. We use the results of previous frames as the initial values for the optimization, thus, only 3-4 iterations per frame required for AAM convergence. Since no relation with image information, the 3D pose and 3D shape parameters are computed efficiently.

Fig. 7 plots the re-projection error in the online non-rigid shape recovery step when varying number of 3D shape basis $\mathbf{m}$. The experiment is performed on a video clip with 65 frames. As shown in Fig. 7, large error occurred only rigid shape is used for pose estimation, and the error reaches $700/100 = 7$ per

(a) (b) (c)

Fig. 6. Adding glasses and beard to the subject in the augmented video sequence, the beard is deformed along with the expression changes.

point. The re-projection error decreases significantly when introducing the 3D linear shape model, additionally, it becomes smaller when $\mathbf{m}$ grows up. When six 3D shape basis are used, the average re-projection is below $100/100 = 1$ each point. However, large number of nonlinear parameters would affect the convergence speed of the object function, there is a trade-off between the accuracy and efficiency. Furthermore, large number of 3D shape basis may decreases the number of optimization iterations.

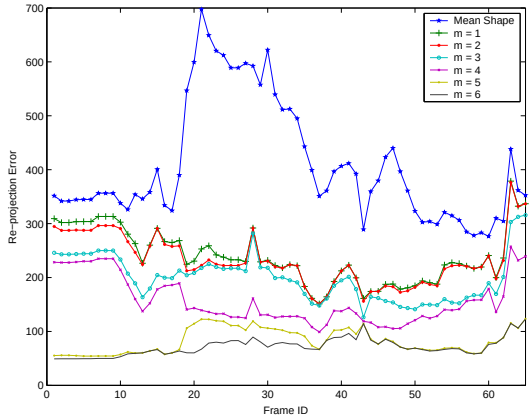
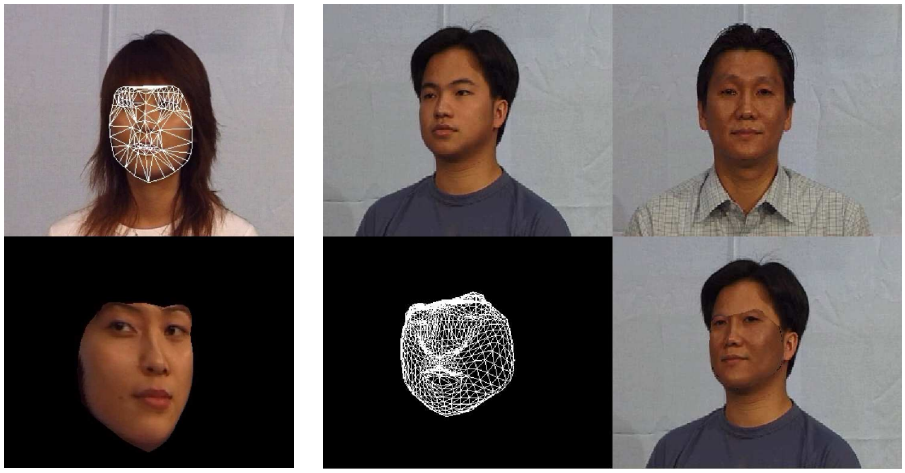


Fig. 7. The re-projection error with various number of 3D shape basis

In order to demonstrate our proposed nonrigid shape and pose recovery approach is effective and promising for generating novel view and 3D facial animation purposes, we first map the recovered 3D nonrigid shape into high resolution mesh via interpolation [14], then render the novel views by mapping different texture and with different poses. Fig. 8(a) shows rendered enlarged novel view rotated from the current pose by 20° on Y-axis. Fig. 8(b) shows the experimental result by replacing face texture of a person with anther person. In the

Fig. 8(b) , the top left one is the modelled 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 3D pose parameters. The generated view fits well on the 3D model. But one may find the skin is not smooth since we do not consider the lighting condition; this can be easily improved by adding smooth operations and lighting adjustment. But the experimental results can answer our question that our constructed 3D model are effective and promising for 3D facial animations.



(a) Rendering novel view (b) Replacing face texture

Fig. 8. Applications of non-rigid shape and pose recovery

We evaluate the computational cost of the proposed method on a Pentium III 1GHz CPU. It runs at 200ms per image of size 352×288. AAM fitting takes 40ms and 3D recovery step takes 74ms. The AAM with 10 shape parameters, 52 texture parameters. The non-rigid shape recovery step with 6 camera parameters and 6 3D shape parameters.

4 Discussions

In this section, we discuss several major differences and advantages of our proposed scheme compared with previous work from several aspects in which we show that our proposed scheme is particularly flexible and powerful for augmented reality applications. Finally, we also mention the disadvantages and some improvements in our future work.

Rigid vs. Non-Rigid. The prior model employed by L. Vacchetti et al. [2] is only for rigid objects or deformable objects with small variations. P. Mit-

trapiyanumic et al. [6] do not take full advantage of AAM's deformation power, the AAM is just used to estimated the 3D pose of rigid objects. The proposed method can deal with 3D deformation through introducing 3D linear shape models. In addition, large variation can be obtained by increasing the number of 3D shape basis. The facial feature can be located accurately by the power of AAM fitting, thus, the added virtual beard can be deformed with the facial expressions in Fig. 6. The novel view can be generated from the current view, even the facial texture of different individuals can be exchanged, as shown in Fig. **??** and Fig. 8. Additionally, the proposed approach provides a solution for building the 2D-3D correspondence from single image. Thus, the tracker can be initialized without manual intervention. In addition, the failure can be recovered automatically.

Offline vs. Online. Many methods [11, 12] have been presented for offline non-rigid shape recovery from image sequences through factorizing analysis on the 2D tracked points. Different from these approaches, our proposed method is able to work online by exploiting the 3D shape models that can be constructed offline effectively by using AAM tracking. This enables us to online acquire 3D non-rigid shape and pose which can be applicable for many AR applications.

Advantages for AR applications. In [8], the model and the fitting algorithm are person specific. The generic AAM is slower than the person specific AAM, but provides good accuracy in the case of large texture variations[15]. In addition, the inverse compositional update strategy is good for smooth shape, and not for the non smooth ones. The proposed extended AAM is a generic model with additive update method rather than person specific model with inverse compositional approach. Thus, it can handle large texture variations, fitting to different individuals. The weak-perspective model used in "Combined 2D+3D AAM" is not suitable for augmented reality applications, moreover, the optimization procedure of the algorithm is complicated. We optimize AAM and 3D pose parameters respectively. Virtual objects can be added to the scene by the estimated camera, orientations and translations information. In addition, the proposed approach is more flexible. The AAM fitting step can be replaced with other algorithms, such as Active Shape Models based approaches [13].

Disadvantages and Future Work. The proposed approach does not take full advantage of 3D information for speeding up AAM convergence. The accuracy of AAM fitting is critical to the 3D pose output. Large rotation may be compensated by the 3D linear mode, therefore, the estimated pose is not so accurate. In the future, problem mentioned above will be solved by training the 3D AAM with the aligned 3D shapes instead of 2D shapes.

5 Conclusions

In this paper we presented a novel scheme for non-rigid shape recovery in real-time augmented reality applications. Our scheme first builds the 3D shape models offline using an effective AAM algorithm. Given the constructed 3D shape models, an efficient online algorithm is suggested to estimate both the 3D pose and non-rigid shape parameters simultaneously. One of our main contributions

is the introduction of 3D linear shape model to estimate the 3D pose parameters and non-rigid shape simultaneously via local bundle adjustment. Moreover, we suggested an updating scheme to predict the shape directly from texture that can improve the accuracy of AAM searching. The promising experimental results validate our proposed scheme is effective for real-time AR applications.

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

References

1. Kato, H., Billinghurst, M.: Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In: Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality. (2004)
2. Vacchetti, L., Lepetit, V., Fua, P.: Stable real-time 3d tracking using online and offline information. IEEE Trans. PAMI **26** (2004)
3. Cootes, T., Edwards, G., Taylo, C.: Active appearance models. IEEE Trans. PAMI **23** (2001)
4. Cootes, T., Kittipanya-ngam, P.: Comparing variations on the active appearance model algorithm. (In: British Machine Vision Conference)
5. Stegmann, M., Ersboll, B., Larsen, R.: Fame-a flexible appearance modeling environment. IEEE Trans. Medical Imaging **22** (2003)
6. Mittrapiyanumic, P., DeSouza, G., Kak, A.: Calculating the 3d-pose of rigid-objects using active appearance models. In: Proceedings of the International Conference in Robotics and Automation. Volume 5. (2004) 5147–5152
7. Ahlberg, J.: Using the active appearance algorithm for face and facial feature tracking. In: Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems, 2001. Proceedings. IEEE ICCV Workshop on. (2001) 68–72
8. Xiao, J., Baker, S., Matthews, I., Kanade, T.: Real-time combined 2d+3d active appearance models. In: IEEE CVPR'2004. Volume 2. (2004) 535–542
9. Blanz, V., Vetter, T.: Face recognition based on fitting a 3d morphable model. IEEE Trans. PAMI **25** (2003)
10. Hou, X., S.Z., L., Zhang, H., Cheng, Q.: Direct appearance models. In: IEEE CVPR'2001. Volume 1. (2001) 828–833
11. C.Bregler, A.Hertzmann, H.Biermann: Recovering non-rigid 3d shape from image streams. In: IEEE CVPR'2000. Volume 2. (2000) 690–696
12. Torresani, L., Yang, D., Alexander, E., Bregler, C.: Tracking and modeling non-rigid objects with rank constraints. In: IEEE CVPR'01. Volume 1. (2001) 493–500
13. Medioni, G., Kang, S.B.: Emerging topics in computer vision. Prentice Hall (2004)
14. Nielson, G.: Scattered data modeling. IEEE Computer Graphics and Applications **13** (1993)
15. Gross, R., Matthews, I., Baker, S.: Generic vs. person specific active appearance models. In: British Machine Vision Conference. (2004)