Multi-Lighting 3D Face Morphable Model Based on Mesh Resampling

Yongli Hu, Baocai Yin, Yanfeng Sun
Multimedia and Intelligent Software Technology Laboratory
Beijing University of Technology, Beijing 100022, China
huyongli@emails.bjut.edu.cn

Abstract

This work describes a multi-lighting 3D face morphable model based on mesh resampling. Despite having realistic results and modeling automatically in 3D face synthesis, the morphable model depends on the unstable optical flow algorithm for model construction, and the model matching is not fitting for complex illumination. To improve modeling results, the mesh resampling method is used to overcome the key problem of model construction, the pixel-to-pixel alignments of prototypic faces. A multi-lighting model is also proposed to evaluate the illumination of the input facial image. The experimental results show these measurements have good performance.

1. Introduction

Modeling human faces has challenged researchers in computer graphics since its beginning. Since the pioneering work of Parke [1], various methods have been reported for modeling the shape and texture of faces and for animating them [2,3]. In the past years, a novel face modeling method, the morphable model is proposed by T.Vetter [4,5]. Matching the model to given facial image, even single image, the realistic 3D face could be reconstructed automatically. Comparing to the traditional face modeling methods, the morphable model has many advantages, such as automation and good results. The most important is that the model gives light on the difficult problems, the illumination and poses evaluation. On the other hand, the construction of the model depends on the unstable algorithm, the optical flow algorithm, which does the pixel-wise correspondences among 3D face. And the model is not fit for the complex illuminations. We propose the mesh resampling method to do the correspondences and construct a multi-light lighting model to overcome the model matching problems under complex illumination. The experimental results show the proposed methods are effective.

We will introduce the morphable model in the following section. Then the alignments of the prototypic faces based on the mesh resampling will give in section 3. A multi-lighting morphable model is constructed in section 4. Finally, we give the facial synthesis results using the multi-lighting morphable model.

2. 3D face morphable model

The 3D face morpable model is composed of two components, model construction and model matching.

To construct the morphable model, a set of prototypic 3D faces are acquired by Cyberware Scanner. Since he morphable model is a type of linear model, the prototypes must be aligned to have the linear operation. So the point-to-point mapping of the prototypic faces must be constructed according the facial feature. But the registration of the dense 3D faces is a difficult problem. T.Vetter projects 3D face onto a cylinder coordinate getting the texture and shape 2D images and computes the correspondence of the 2D images by optical flow algorithm [6]. From the alignments of 2D images to get the correspondence of the 3D faces.

Once aligned pixel-to-pixel, the prototypic faces can be represented as shape and texture vectors:

$$\begin{array}{l} S_i = (X_{i1}, Y_{i1}, Z_{i1}, \cdots, X_{in}, Y_{in}, Z_{in})^T \in R^{3n} \\ T_i = (R_{i1}, G_{i1}, B_{i1}, \cdots, R_{in}, G_{in}, B_{in})^T \in R^{3n} \end{array} \quad (1)$$

where $1 \le i \le N$, N is the number of the faces,n is the number of points of the 3D faces, and $(R_{ij}, G_{ij}, B_{ij})$is the color values of the point (X_{ij}, Y_{ij}, Z_{ij}).

The linear combination operation of these shape and texture vectors will produce new 3D faces:

$$S_{new} = \sum_i^N a_i S_i \quad T_{new} = \sum_i^N b_i T_i \quad (2)$$

where $\sum\limits_i^N a_i = \sum\limits_i^N b_i = 1$.As a number of high dense points in the 3D faces lead too much computation and there are correlation among the faces, the technique of principal component analysis is used to give the final morphable model[4]:

$$S_{model} = \overline{S} + \sum_{i}^{m-1} \alpha_i s_i \quad T_{model} = \overline{T} + \sum_{i}^{m-1} \beta_i t_i \qquad (3)$$

where $\overline{S}, \overline{T}$ are average shape and texture, s_i,t_i are the main components in descending order according to their eigenvalues. $\vec{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_{m-1})$,$\vec{\beta} = (\beta_1, \beta_2, \cdots, \beta_{m-1})$ are the shape and texture combination coefficients. To produce new faces the coefficients must be determined.

The matching of the morpable model is to find the suitable combination coefficients which produce the closest 3D face to the given facial image. Since the 3D model and the 2D facial image can not measure directly, the camera model and illumination model are used to project the 3D model into image plane and the error between the projective image I_{model} and the input image I_{input} is defined:

$$E = \sum_{x,y} [I_{input}(x, y) - I_{model}(x, y)]^2 \qquad (4)$$

The Phone Illumination model and perspective projection are used to produce the model projecting image. Supposing the point (X, Y, Z) with texture value (R, G, B) on the model has the projective point(x, y) in image plane, I_{model} will have the following R value at (x, y) :

$$I_{R,model}(x, y) = R(I_{aR}+I_{dirR}(L\cdot N))+K_s I_{dirR}(F\cdot V)^n \qquad (5)$$

where I_{aR}, I_{dirR} are amble light and directly light. K_s is the reflectance.L, N, F, V are light direction, normal, reflective direction and the direction of the view at (X, Y, Z). $n$is the surface shinningness. The computations of G, B values are same to R.

If the camera and the illumination parameters are denoted as $\vec{\rho}$, the error in (4) can be looked as a function $E(\vec{\alpha}, \vec{\beta}, \vec{\rho})$. Then the matching problem changes to a minimal optimal problem and can be overcome by optimization method.

3. Alighnment based on mesh resampling

The alignment of the prototypic faces is the key step to construct the morphable model, but the correspondence or registration of 3D dense objects is a challengeable problem. The optical flow algorithm is used to do the alignment presently [6]. Generally optical flow algorithm is used in perception of object movement in video sequence [7]. As the likeness of human faces we can look two facial images as neighboring frames in video. So the alignment of faces can be computed by optical flow algorithm. While the difference between two facial images is too large to satisfy the hypothesis, the change between two sequent frames is continuous and trivial, the optical algorithm will fail with obvious error. We propose a mesh resampling method to do the face alignments. Generally mesh resampling is used to construct the surface from discrete points cloud [8]. Here the resampling method is used to re-organize the prototypic faces data and uniform its format. By mesh resampling, all the faces will have the same number of points and triangles with same topology. The point-to-point alignments are constructed naturally.

The Krishnamurthy mesh resampling method [8] is adapted to process the data of the prototypic faces. Firstly, the faces are manually divided into patches according its feature. To get uniform mesh lastly all the patch are divided into quadrangle and almost have same area. Then the isolines are initialized. Finally, the faces are resampled reduplicatively until the points density nears the original. As all faces are divided into patches with same number and all patches are resampled with same times, the alignments of the faces can be constructed by the alignments of the patches and its resampled points. The details of resampling refer to [9]. The mesh resampling procedure is shown in Figure 1.

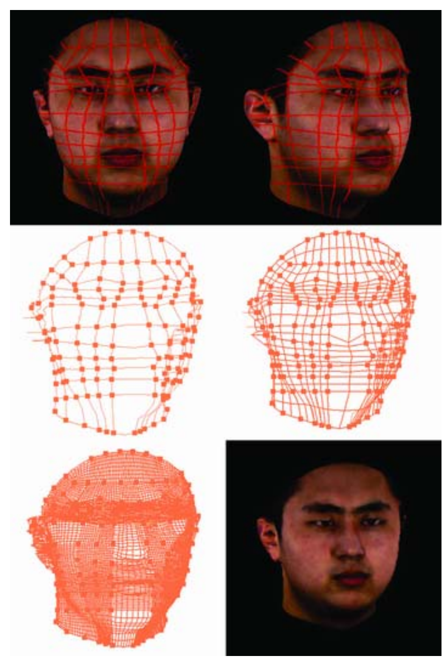

Figure 1. Mesh resampling. Top: the divided prototypic faces; Middle: the initialized isolines and one times resampling; Bottom: three times resampling and the final mesh resampling result.

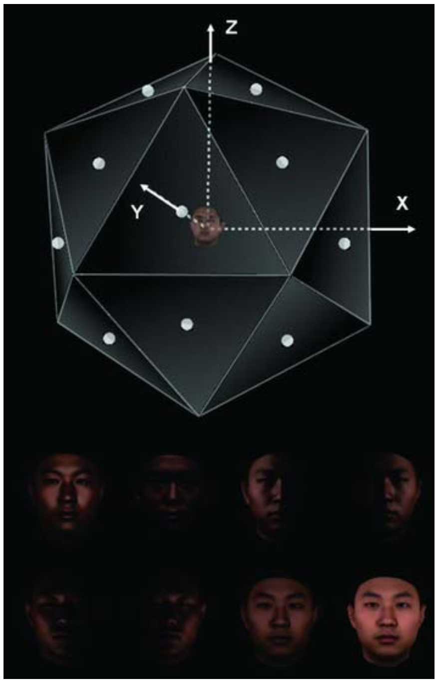

Figure 2. Multi-lighting model and facial images with different illumination when changing the brightness of lights in the model.

4. Multi-lights illumination

To synthesize 3D face from facial images in complex illuminations, we design a multi-lighting model in the model matching procedure stimulated by the illumination evaluation and lighting reproduce work by Debevec [10]. Intuitively, if one set enough number of lights around an object and the brightness of each light can vary independently, the arbitrary illumination of the object can be simulated. But the more lights will produce more computation in the model matching. Trading off between the computation and the last evaluation result, we use the icosahedron to set the lights. Let the icosahedron center at the center of the model face and select the front 10 triangles of the icosahedron to set lights. The lights are located in the center of the triangles. The multi-lighting model is shown in Figure 3. When changing the brightness of the lights, there will be different facial images with different illumination, shown in the bottom in Figure 3.

Using the multi-lights model in face model matching, the computation of the model projective image in (5) will replaced by the following formula.

$$I_{R,model} = RI_{aR} + \sum_{h}^{H} [RI^{h}_{dirR}(L^{h} \cdot N) + K_s I^{h}_{dirR}(F^{h} \cdot V)^{n}] \quad (6)$$

where H is the number of the lights.

By the matching optimization the brightness of each lights are determined, so the illumination of the given facial image is evaluated. The illumination evaluation is very important to eliminate the illumination disturbance to 3D face synthesis. From the evaluation the reconstructed 3D face even can be rendered in normal illumination.

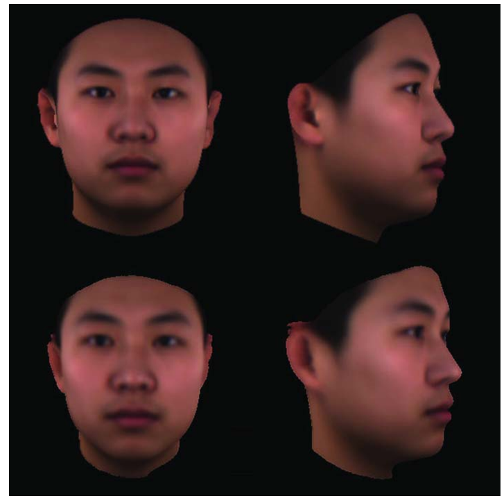

Figure 3. Average face of the aligned prototypic 3D faces. Top: the average face of mesh resampling method aligned faces; Bottom: the average face of optical flow algorithm aligned faces.

5. Results

To construct the morphable model, we scan 200 3D faces by Cyberware scanner. To align the prototypic faces, every faces are divided into 122 patches and the isolines are initialized. Then the faces are aligned pixel-to-pixel by 4 times mesh resampling. Based on the aligned prototypes a 3D face morphable model are constructed. Matching the models to the given facial images the 3D face will reconstructed automatically.

Comparing with the optical flow algorithm, the mesh resampling method obtains better aligning results. As showing in Figure 4, the former average face image is clearer than the latter especially in eye and ear areas. Since those areas are too different to satisfy the continuous assumption of the optical flow, the computed errors lead to the blur on the average face. Although mesh resampling method needs much manual work while the optical flow is automatic, to get more precision of alignments the offline interaction work is acceptable.

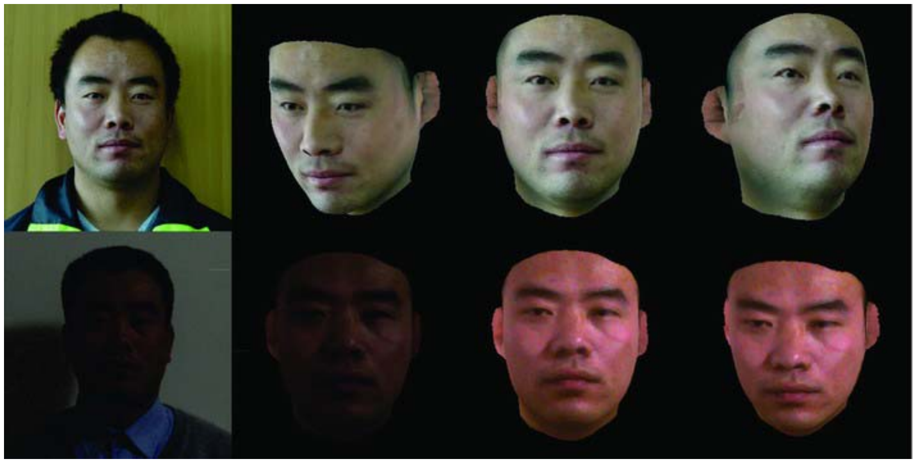

Figure 4. 3D face synthesis. The top row is the normal illumination results. The first is the input image; the others are the synthesis face in different views. The bottom row is the abnormal illumination results. The first is the input image, the second is the synthesis face, the other two is the face images rendering in normal illumination.

The multi-lighting morpahble model is used to match facial images in different illumination. The results are shown in Figure 5.By the multi-lighting model the illumination of the given facial image is evaluated and the reconstructed 3D face can be rendered with different illumination. The experimental results show the multi-lights model has much contribution in light evaluation and expands the morphable model into complex illumination situations.

Acknowledgement: This work is supported by the National Natural Science Project of China (No.60375007) and the Beijing Natural Science Foundation of China (No. D070601-01).

References

[1] F.I.Parke. Computer generated animation of faces. ACM National Conference, November 1972.

[2] K.Waters. A muscle model for animating three dimensional facial expression. Computer Graphics,1987, 21 (4): 17- 24

[3] B.Guenter, C.Grimm, D.Wood, et al. Making faces. Proceedings of SIGGRAPH'98, Orlando, 1998, 55-66

[4] V.Blanz, T.Vetter. A morphable model for the synthesis of 3D faces. Proceeding of SIGGRAPH'99, Los Angeles, 1999, 187-194

[5] V.Blanz, T.Vetter. Face recognition based on fitting a 3D morphable model. IEEE Trans. on Pattern Analysis and Machine Intell. 2003, 25(9): 1063-1074

[6] T.Vetter, M.J.Jones, T.Poggio. A bootstrapping algorithm for learning linear models of object classes. Proceeding of Computer Vision and Pattern Recognition, Puerto Rico, USA: 1997. 40-47

[7] J.L.Barron, D.J.Fleet, S.S.Bauchemin. Performance of Optical Flow Techniques. Internation Journal of Computer Vision, 1994, 12(1): 43-77

[8] V.Krishnamurthy, M.Levoy. Fitting smooth surfaces to dense polygon meshes. Proceedings of SIGGRAPH'96, New Orleans, 1996, 313-324

[9] Chunliang Gu, Baocai Yin, Yongli Hu, Shiquan Cheng. Resampling Based Method for Pixel-wise Correspondence between 3D Faces. International Conference on Information Technology (ITCC'04). April 5 -7, 2004 Las Vegas, NV, USA

[10] Paul Debevec, Andreas Wenger, Chris Tchou, et al. Lighting Reproduction Approach to Live-Action Compositing, Proceedings of SIGGRAPH'02, San Antonio, 2002, 547-556