2019 International Conference on Energy, Power, Environment and Computer Application (ICEPECA 2019)
ISBN: 978-1-60595-612-1

A Multi-View Texture Fusion Approach for High Quality 3D Face Modelling

Bing-chuan LI[1], Yu-ping YE[1,2], Zhan SONG[1,3,*], Ling-sheng KONG[2] and Su-ming TANG[1,*]

[1]Shenzhen Key Laboratory of Virtual Reality and Human Interaction Technology, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

[2]Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, China

[3]The Chinese University of Hong Kong, Hong Kong, China

*Corresponding author

Keywords: Texture mapping, Multi-view reconstruction, Visibility analysis.

Abstract. This paper presents an accurate means for the 3D face texture mapping based on a multiple-view 3D scanning system. Given a reconstructed 3D mesh model and a set of calibrated images, a high-quality texture mosaic of the surface can be created by the proposed method. The proposed method aims to avoid noticeable seams, color discontinuity and ghosting problems. The method performs in two steps, i.e. occlusion areas filtering and the fusion of different texture images. The filtering is performed to identity invisible areas from the camera viewpoint and leaves them not textured. To perform the occluded areas filtering, we combine the geometrical features of the objects and the visual range of the camera. Experimental results showed that, texture mapping quality and precision can be improved in comparison with conventional means.

Introduction

3D reconstruction is an important research topic in both computer vision and computer graphics domains. It is a technology to study how to obtain the three-dimensional information of objects in real world via the passive or active optical means. 3D shape reconstruction and texture reconstruction from photographs are the two major aspects in the modeling of real objects. There have been a lot of techniques which can generate image-based 3D models with high accuracy [1]. However, existing texture mapping methods are usually lack of precision, especially for the texture images captured with various viewpoints or the change of lighting conditions [2].

In the past decades, many sophisticated texture generation approaches have been proposed. Early works have focused on different weighting heuristics to average overlapping textures. According to the weighting function, these approaches can be generally categorized as Weighted Blending [3], Multi-Band Blending [4] and Super-Resolution Maps [5] methods. Underlying principle of these methods is to attach a color attribute to each vertex. Then blend the color intensities from multiple views by calculating the weights of all visible views for each vertex. Some other methods generate a texture map by collecting all texture patches together. The combination of texture patches has much in common with the stitching and texture synthesis of planar images. Graph cut optimization and gradient-domain techniques [6] are often used to improve the performance over intensity blending. Instead of blending the color intensities from multiple views, some methods introduced Markov Random Field (MRF) [7] into texture patch registration to optimize the texture alignment, which considers both image visibility and color continuity, and computes the texture of each facet from only one viewpoint. After the texture patch registration, the Poisson fusion [8] approach to deal with the chromatic aberration between the sides of the patch joint. These methods have achieved good results in texture fusion objects described in their respective papers. But there are still some problems like mosaic error in the applications of 3D human face reconstruction which have high precision requirements for texture fusion.

In this paper, an accurate texture mapping method is investigated based on a structured light 3D scanning system. In order to map images on 3D models, a two-step method is adopted including a) occluded areas filtering and b) fusion of different images in a unique texture. According to the geometrical features of the model face, we can synthesize a person's head to a cylinder. Then, we can decide each triangular patch on the surface of 3D face model belong to the visual range of which view by calculating the angle between the viewing direction of multiple views and each triangle mesh. Also, we use the cosine of this angle in the weighting function to fuse different view images in the overlapping regions. Experimental results showed that, the proposed method can realize accurate model and texture fusion with improved visual effect compared with traditional means.

Methodology

Since our concern is only put on the front face, images from three perspectives can cover completely the front of the face. Each point corresponds to one pixel in the texture image, and each point is assigned three weights, each weight corresponding to one perspective. Therefore, the texture mapping problem is transformed into the design of the weight function. For each point p_i on the surface, we use a weighting function to calculate the weights of the three input images. We use l_i, m_i and r_i to represent the input weights of the left, middle, and right perspectives, respectively i=1, 2,..., k (k is the total number of points). Our goal is to make higher quality input image weights larger in overlapping areas and avoid seams. As shown in Fig. 1, images of three perspectives to be used as texture, a detailed face model and their registration in the same reference frame. So, our work contains two almost independent steps: the visibility analysis and the texture assignment.

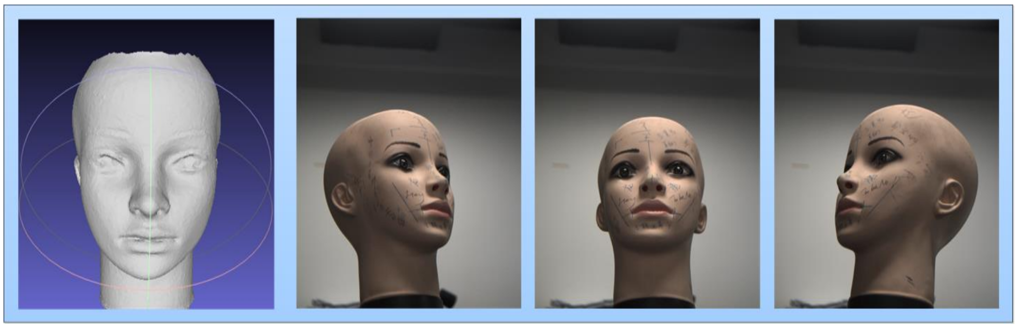

Figure 1. Face model and images captured from three viewpoints.

Visibility Analysis

Considering the geometrical features of the face, we can synthesize a person's head to a cylinder. As shown in Fig. 2, the x-axis of the coordinate system points to the front of the face, and the z-axis is the vertical axis of the face. From the cross section of face, each perspective texture covers a fixed range of faces. The area covered by the left perspective texture map is not visible in the texture map of the right perspective. In summary, for each point p_i, if φ represents half of the angle of middle perspective texture covering the face, we have conclusions as follows: a) when $\alpha\in(-\varphi, \varphi)$, $0°<\varphi<90°$ we should insure $l_i = 0$, $m_i > 0$ and $r_i = 0$ that is takes only the input image of the middle perspective when α varies in this range; b) when $\alpha\in(-90°, \varphi)$, we should insure $l_i > m_i$, $m_i > 0$ and r_i =0. In other word, we take only the input images of the left perspective and the middle perspective; c) when $\alpha\in(\varphi,90°)$, we should insure $l_i = 0$, $m_i > 0$ and $r_i > m_i$. With above conclusions, we can introduce the weighting function as equations (1-3).

$$m_i=max(0, cos\alpha). \tag{1}$$

$$l_i=max(0, cos(\alpha+\theta)). \tag{2}$$

$$m_i = max(0, cos(\alpha\text{-}\theta)). \tag{3}$$

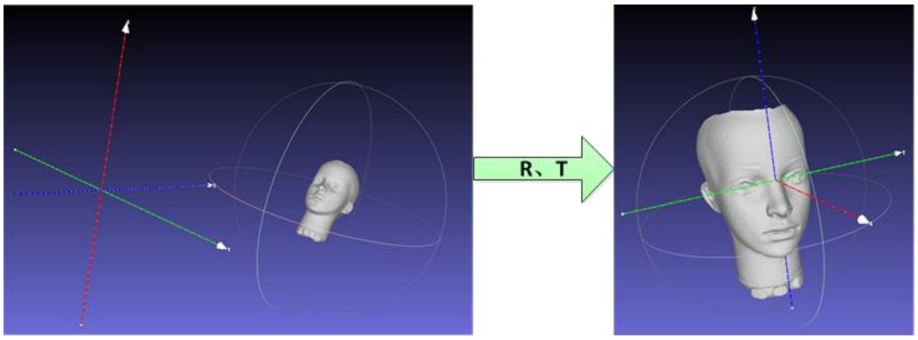

Figure 2. Original model and the transformed model to cylindrical coordinate system.

The definition of α and θ of the formula is shown in Fig. 3. If p represents one point on the face, p_{xy} is the projection of p on the xoy plane and α represents the angle between vector op_{xy} and vector i(the x-axis unit vector). The middle perspective contains almost all the 68 key points of face obtained from the face detection of the input image using the dlib c++ library [9]. So we can make the y-axis go through the projection of the 3D points corresponding to the key points 1 and 17 of the face on the plane oxy, then let φ= 90°.

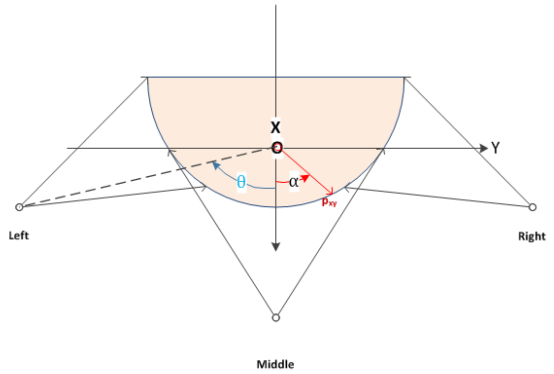

Figure 3. Cross section of the face.

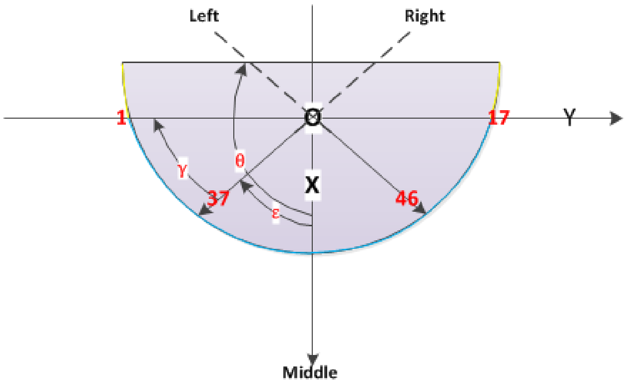

Figure 4 Definition of parameters based on face cross section.

Texture Assignment

Let θ=φ+90°- γ, to avoid the blurring of eyes, we can initialize γ=90°-ε, so θ=φ+ε=90°+ε, where ε is the absolute value of the angle between the left (right) eye and the x-axis, the value will be adjusted upward according to the experimental effect. However, the human face is not an ideal cylindrical surface [10]. When $\alpha \in (-\varphi, \varphi)$, if make $l_i = 0$, $m_i > 0$ and $r_i = 0$, due to some parts of the nose area in the middle view image are not photographed by the camera will cause holes if we create texture mapping according to Eqn. (1-3).

Using β to represent the angle between x-axis and the vector n_{xy}, which is the projection of vector n on the oxy plane and n refers to the normal vector of i-th point p_i. For example, on the left side of the nose, there is $\beta \approx$-90° and $cos(\beta+\theta) \approx cos(-90°+90°+\varepsilon) \approx cos\varepsilon$. It is obviously that $cos\beta < cos(\beta+\theta)$ and $cos\alpha > cos(\alpha+\theta)$ are both satisfied. In order to make $l_i >> m_i$, we can let m_i=min(cosα, cosβ) and l_i =max($cos(\alpha+\theta)$, $cos(\beta+\theta)$), so that the holes on the left side of the nose can be filled. And so on, we need to let r_i =max(cos (α-θ), cos (β-θ)). So Eqn. (1-3) can be rewritten as:

$$m_i = max(0, min(cos\alpha, cos\beta)). \tag{4}$$

$$l_i = max(0, max(cos(\alpha+\theta), cos(\beta+\theta))). \tag{5}$$

$$r_i = max(0, max(cos(\alpha\text{-}\theta), cos(\beta\text{-}\theta))). \tag{6}$$

Experimental Results

The proposed method is implemented by three steps: 1) fitting a cylindrical coordinate system according to the key points of human face; 2) calculating l_i, m_i and r_i according to Eqn. (4-6); 3) adjusting γ according to the experimental results. In order to fit cylindrical coordinate system, we calculated the texture coordinates for the triangles by back-projecting the triangle coordinates in the object space to the viewing images. According to the nearest neighbor method, we can find the nearest three-dimensional point corresponding to each face key point.

In order to verify the effect of our weight function, we calculate the m_i according Eqn. (4) and let l_i=0, r_i=0. Then, we map the image of middle perspective onto our model to obtain the middle fragment. In the same way, the left and right segments are obtained. In many weighted fusion methods [11, 12], pixel with small angle of normal to the mesh is selected as the input for each point. In other words, the most orthogonal image is selected as the best input. But they consider the normal as the main factor of the weight function. Our weight function takes the projection of the coordinates on the *oxy* plane as the main factor and combines the projection of the normal on the *oxy* plane.

For comparison, we calculated the texture map respectively using the cylindrical coordinate method and the classical method [11, 12]. The results are shown in Fig. 5. By comparison, we can see that, the cause of ghosting and blurring artifacts by classical method can be avoided in the proposed method. In addition, the traditional methods have seamsand holes, and our method solves this problem. Fig. 6 shows the experimental result with real human face. The human face is scanned three times from various scanning angle. In each scan, a digital camera is use to capture one color image and used as the texture image. A high quality face model can be obtained by the proposed texture mapping and alignment method.

Figure 5. The result obtained by our method (1st row) and classical method (2nd row).

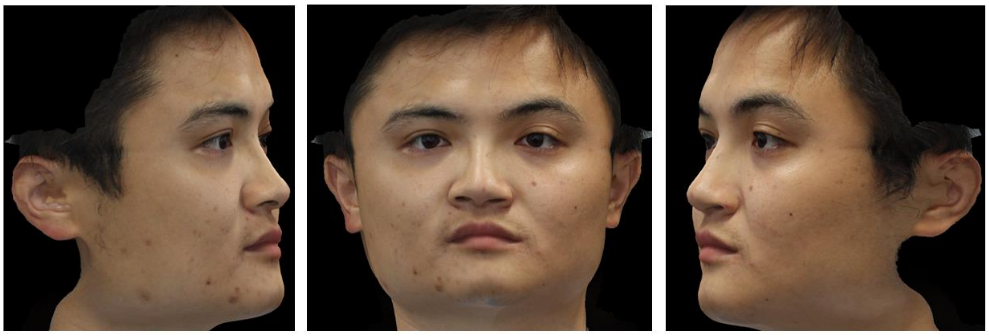

Figure 6. Experimental results with real human face.

Conclusion and Future Work

This paper presents an accurate texture mapping method based on a structured light 3D scanning system. A two-step method is adopted including a) occluded areas filtering and b) fusion of different images in a unique texture. According to the geometrical features of the model face, a cylindrical coordinate system is established for the face model. Then, we can decide each triangular patch on the 3D model by calculating the angle between the viewing direction and each triangle mesh. And cosine of this angle in the weighting function is also used to fuse different view images in the overlapping regions. Experimental results on the model and real human faces showed that, the proposed method can realize accurate model and texture fusion with improved visual effect compared with classical means. The future work can consider illumination equalization of the input images and to further improve the visual quality of textures.

Acknowledgment

This work was supported in part by the National Key R&D Program of China (No. 2017YFB1103602), Shenzhen Key Laboratory (No. ZDSYS201605101739178), and the National Natural Science Foundation of China (Nos. 61775172, 61773363, 51705513).

References

[1] M.D. Zhang et al. "Imaged-based 3D face modeling." IEEE International Conference on. Computer Graphics, Imaging and Visualization, 2004.

[2] J.Y. Zhang and S.Q. Luo. "Image-Based texture mapping method in 3D face modeling." IEEE/ICME International Conference on Complex Medical Engineering, 2007.

[3] C. Marco, et al. "Masked photo blending: Mapping dense photographic data set on high-resolution sampled 3D models." Computers & Graphics 32.4 (2008): 464-473.

[4] A. C. J-P. Pons and R. Keriven. "Seamless image-based texture atlases using multi-band blending." International Conference on Pattern Recognition, 2008.

[5] Z.L. Chen et al. "3D texture mapping in multi-view reconstruction." International Symposium on Visual Computing. Springer, Berlin, Heidelberg, 2012.

[6] K. Johannes, et al. "Image-based rendering in the gradient domain." ACM Transactions on Graphics (TOG) 32.6 (2013): 199.

[7] S.Y. Dou and H. Nagahashi. "High Quality Texture Mapping for Multi-view Reconstruction." International Conference on Multimedia and Image Processing, 2017.

[8] D. Arnaud, et al. "Seamless texture stitching on a 3D mesh by Poisson blending in patches." IEEE International Conference on Image Processing, 2014.

[9] K. Davis. "Dlib c++ library." Access on: http://dlib.net (2015).

[10] P. Mattia, L. Barazzetti, and M. Scaioni. "An automated and accurate procedure for texture mapping from images." International Conference on Virtual Systems and Multimedia, 2012.

[11] C. Marco, P. Cignoni, and R. Scopigno. "Reconstructing Textured Meshes from Multiple Range RGB Maps." International Fall Workshop on Vision, Modeling, and Visualization, 2002.

[12] L. Hendrik PA, W. Heidrich, and H-P. Seidel. "Automated texture registration and stitching for real world models." The Eighth Pacific Conference on Computer Graphics and Applications, 2000.