Advanced Science and Technology Letters
Vol.120 (CGAG 2015), pp.778-783
http://dx.doi.org/10.14257/astl.2015.120.153

Physical Based Inverse Rendering of 3D Faces from a Single Image

Hongyi Zhang, Shuangjiu Xiao

Digital Art Lab, School of Software, Shanghai JiaoTong University
Shanghai, China
zhywdtc@163.com, xsjiu99@cs.sjtu.edu.cn

Abstract. In this paper, we present an automatic framework for generating 3D faces in real-time VR systems or games from a single image. The framework are divided into reconstruction stage and rendering stage. In the first stage, we propose a linear machine learning model which is faster than previous works because features of illumination are learned from images in advance. In the second stage, we propose a physical based method to generate skin details and an approximation rendering algorithm. Finally, our work is verified in a real-time graphic application where we can view the face model with strong realism and incredible resemblance to the origin face photo.

Keywords: Image Feature Learning, Realistic Face Reconstruction, Physical Based Rendering

1 Introduction

Human face modeling and rendering is one of the most difficult and challenging problems in the fields of computer vision and graphics. Existing research work is broadly divided into two categories: one is using 3D digital scanning device to obtain facial geometry and texture data [1, 2, 3], and the other is based on large historical data [4, 5]. Scanning methods pursues more accuracy and details. However, it is very expensive to acquire those devices. In contrast, data-based methods provide more convenience because only several images or video sequences are needed to recover the original shape and texture of the face. Eigenfaces [6] and 3D morphable model [7] are well-known methods based on PCA model. Moreover, few researchers around the world have ever considered the problem of realistic rendering after reconstruction. Recently, the only work is from Oswald [8] who presents a complete framework to inverse render faces. Despite the excellence of this research, global shading is very simple and many skin details such as pores and wrinkles are lost.

In this paper, a complete automatic framework for 3D face reconstruction from a single image and realistic rendering is introduced and implemented. There are two key points in our research. First, machine learning algorithms are widely applied in the process of image-processing and face modeling. In brief, we can ensure that everything comes from learning. Second, physical based methods are used to the implement real-time face rendering. Our work are verified in a real-time graphic

ISSN: 2287-1233 ASTL

application, where we show a 3D face model with incredible resemblance to the origin photo, along with dynamic environment lighting from arbitrary point of view.

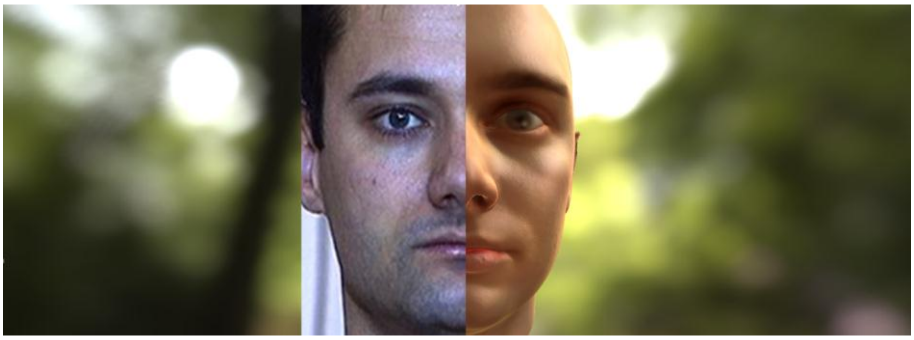

Fig. 1. Result of our method. Left side is the original photo. Right side is inverse rendering of the face with dynamic environment lighting in our real-time graphic application.

2 Framework

Our inverse rendering process contains two stages. In the first stage, we reconstruct a low-resolution model containing vertex position and color. In the second stage, the realistic 3D Face are rendered online. Everything needed for real-time rendering is pre-calculated offline, including all the graphic data and the rendering equation after approximation.

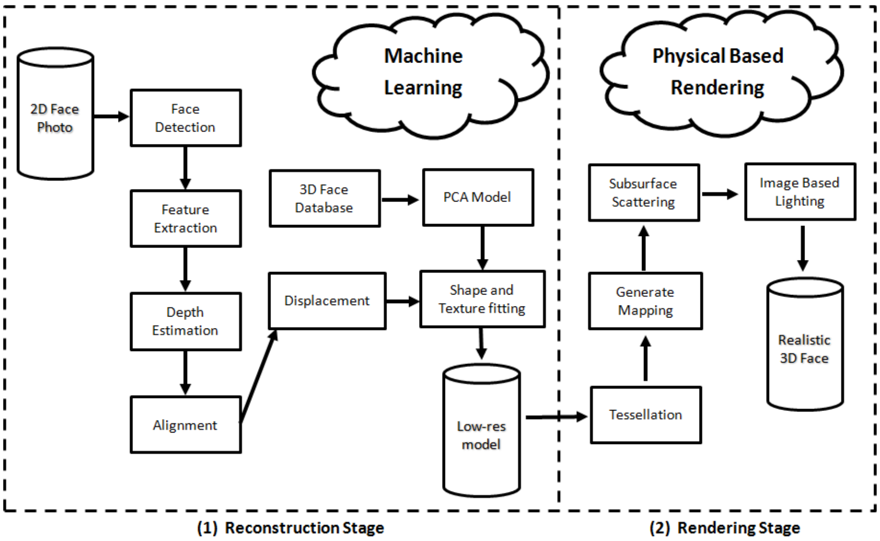

Fig. 2. Complete framework of our inverse rendering process.

3 Linear Feature Learning for Illumination Parameters

A linear machine learning model which performs more efficiency is presented in our work. We start from defining landmarks on human faces and find them through active shape model. The face model is parameterized in shape and texture and we treat the distance between a temp frame and original input image as an energy function. During the time of minimizing the energy function, suitable coefficient vectors are found so that the mean shape and texture will be morphed to the target.

3.1 Illumination Parameters

Our reconstructed face model needs to be the same for different photos of the same person when taken in various illumination conditions. Therefore we needs to estimate those illumination parameters besides shape and texture coefficients. The illumination parameters include light intensity I, light direction d, view direction v, and material shininess s. They are represented by a vector L = (I, d, s, v).

In previous works, nonlinear least squares fitting method are used to calculate illumination parameters. This is because the formula to simulate image formation process is based on Phong lighting model. It becomes a non-linear energy function when morphing.

3.2 Linear Fitting Model

To improve the accuracy of this non-linear function mappings, we design an artificial neural network with double hidden layers. Lavenberg-Marquardt algorithm is used for training the neural network.

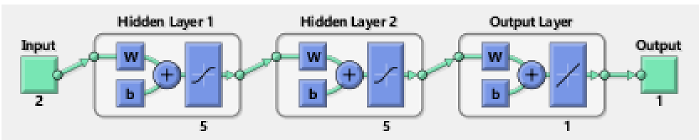

Fig. 3. Artificial Neural Network Designed for Estimating Illumination Parameters.

The illumination vector L can be estimated when we input photos with different lightings. Therefore, the energy function become linear and we can solve this least squares problem with linear regression algorithms. Finally, we manage to reduce the time of model fitting greatly.

4 Physical Based Face Shading Model

Phong lighting model is too simple for skin shading since the geometry of structure of human face is so complex. To improve the realism, a physical based shading model considering micro-facet reflecting and subsurface scattering is presented in our work. What's more, we approximate the shading equation to meet the demand of real-time rendering.

4.1 Generating Skin Details

First of all, tessellation is done to the low-resolution model generated in face reconstruction stage. Then, we do convolution to the high-resolution model using custom skin displacement maps. Figure 4 shows examples of our pre-defined displacement maps such as pores and scars. With face features detected in the image, we can sculpt the face with displacement maps. If there is no pre-defined displacement map available, Gaussian noise is still applicable for simulating pores on the face.

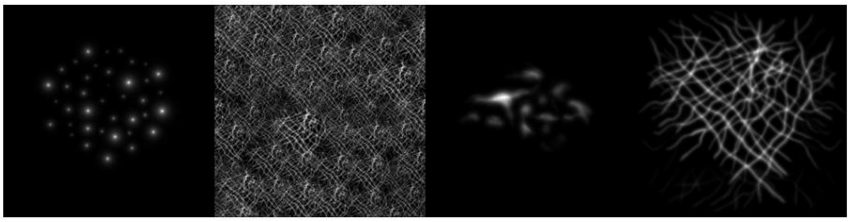

Fig. 4. Examples of Pre-defined face skin displacement maps.

4.2 Shading Approximation

We split the integration to calculate specular lighting offline. Diffuse lighting is represented by the first term which is pre-filtered by spherical harmonics basis with the irradiance map. Irradiance of the environment is also pre-filtered offline which is represented by a cube-map. The second term is approximated using Monte Carlo sampling.

5 Results

We show effects of reconstruction and rendering separately.

5.1 Results of Reconstruction

Figure 5 shows our reconstruction results. For each test case from (a) to (h), left side is the 2D input face photos from CMU-PIE Face Image Database, and right side is the output 3D face model using our feature learning method.

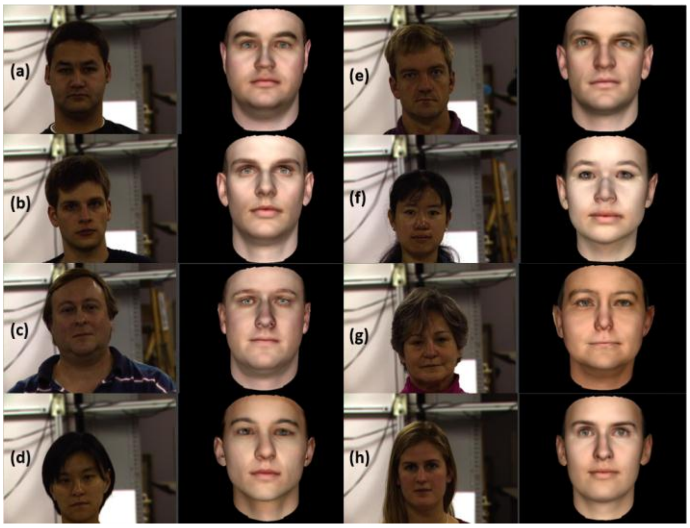

Fig. 5. 8 examples for 3D face reconstruction from a single image.

5.2 Results of Rendering

Figure 6 shows our rendering results. We can see a number of details on the cheek and mouth lips. Some areas of the face blush because of subsurface scattering property of translucent objects. Diffuse and specular reflection are simulated based on environment lighting and roughness of the face. As we can see from the snapshots, physical based methods improve the face realism to a great extent.

Fig. 6. Snapshots for physical based face rendering.
(1): In the evening. (2): On a sunny day. (3): In the forest. (4): With shadows

5.3 Details

Our work is faster because of the linear feature learning model. Moreover, generated skin details add more reality. From figure 7, we can even see wrinkles on the virtual face.

Fig. 7. Simulating face skin details

Acknowledgement. This work is supported in part by the Natural Science Foundation of China (No. 61173105, 61373085), and the National High Technology Research and Development Program of China (No. 2015AA016404).

References

1. Xu C, Quan L, Wang Y. Adaptive multi-resolution fitting and its application to realistic head modeling[C]. Geometric Modeling and Proeessing, Beijing, China (2004)
2. Debevec P, Hawkins T, Tchou C. Acquiring the Reflectance Field of a Human Face[C]. SIGGRAPH 2000 Conference, New Orleans (2000)
3. Tarini M, Yamauchi H, Haber J. Texturing Faces[C]. Graphics Interface 2002 Conference, Calgary, Canada (2002)
4. SenguptaK,Ko C C.Scanning face models with desktopcameras[J]. IEEETransactionson Industrial Electronics (2001)
5. Dimitrijevic M, Ilic S F P. Accurate face models from uncalibrated and i11 一 lit video sequence[C]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Jun 27-Jul 2, 2004, Washington, DC, USA (2004)
6. Iwasa T, Shima T, Sai M. 3D Eigenfaces for face modeling[C]. Proceedings 5th Asian Conference on Computer Vision (2002)
7. Blanz V, Vetter T. A morphable model for the synthesis of 3D face[C]. Proeeedings of the 26th annual conference on Computer graphics and interactive techniques (1999)
8. Aldrian, O. Smith, W.A.P., Inverse Rendering of Faces with a 3D Morphable Model[J], Pattern Analysis and Machine Intelligence, IEEE Transactions on Biometrics Compendium, 35(5): 1080-1093 (2013)