

# Automatic Texture Synthesis for Face Recognition from Single Views

Xiaozheng Zhang, Yongsheng Gao  
Computer Vision & Image Processing Lab,  
School of Engineering, Griffith University,  
Australia  
{x.zhang, yongsheng.gao}@griffith.edu.au

Maylor K.H. Leung  
School of Computer Engineering,  
Nanyang Technological University,  
Singapore  
asmkleung@ntu.edu.sg

## Abstract

*One possible solution for pose- and illumination-invariant face recognition is to employ appearance-based approaches, which rely greatly on correct facial textures. However, existing facial texture analysis algorithms are suboptimal, because they usually neglect specular reflections and require numerous training images for virtual view synthesis. This paper presents a novel texture synthesis approach from a single frontal view for face recognition. Using a generic 3D face shape, facial textures are analyzed with consideration of all of the ambient, diffuse, and specular reflections. Virtual views are synthesized under different poses and illuminations. The proposed approach was evaluated using the CMU-PIE face database. Encouraging results show that the proposed approach improves face recognition performances across pose and illumination variations.*

## 1. Introduction

Recently, pose and illumination variations in face recognition captured growing interests of researchers in computer vision and pattern recognition. One strategy to pose- and illumination-invariant face recognition is to employ appearance-based methods [1, 2, 3, 4] in which pixel intensities of face images are directly used for classifications. These methods are faithful to original images and easy to implement, though they usually require numerous training images for virtual view synthesis.

Textural information plays an important role in the appearance-based face recognition approach. Most existing texture analysis methods [1, 2, 4] assume face surfaces to be Lambertian surface, which reflects incident lights only diffusely. However, human faces have both specular and diffuse reflections [5] and Lambertian assumption results in suboptimal facial textures. Few researchers considered non-Lambertian

texture analysis in virtual view synthesis. Cootes *et al.* [3] modeled facial textures using three constraints across pose changes. Ip and Yin [6] synthesized novel textures by interpolating frontal and profile textures. Blanz and Vetter [7] applied a stochastic optimization on facial textures, which is highly non-convex and unstable. A stable face texture modeling approach from single views is still elusive.

This paper presents a novel stable and automatic texture analysis approach for pose- and illumination-invariant face recognition from single frontal views. It takes into consideration the specular reflection as well as the diffuse and ambient reflections. The proposed approach uses single frontal views and generates virtual textures and virtual views under different illumination conditions and poses. Those virtual views are then used in an appearance-based recognition system to improve recognition performance.

## 2. Texture analysis

As reviewed in Sec. 1, the most popular model to analyze facial textured is Lambertian model, which approximately considers that face surfaces only have diffuse reflection. However, this approximation is not accurate, because specular reflection contributes a significant part in the total reflection of face surfaces. Phong model [8] is expressed as

$$I = k_a I_a + k_d I_{po} (\vec{L} \cdot \vec{N}) + k_s I_{po} (\vec{R} \cdot \vec{V})^n, \quad (1)$$

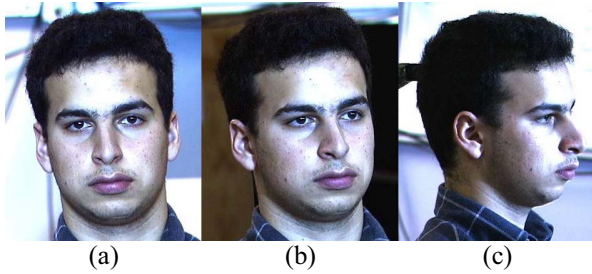
where  $I$ ,  $I_a$ ,  $I_{po}$  are image, ambient, incident light intensities,  $k_a$ ,  $k_d$ ,  $k_s$  are ambient, diffuse, specular albedos,  $\vec{L}$ ,  $\vec{N}$ ,  $\vec{V}$ ,  $\vec{R}$  are light, normal, viewing, reflecting directions, and  $n$  is specular exponent. As in Eq. 1, Phong model takes into considerations all of the ambient, diffuse, and specular reflections, while Lambertian model, on the other hand, has only diffuse reflection which is independent to viewpoints. In real

face photographs, reflections especially highlights differ in different viewpoints as evident in Fig. 1.

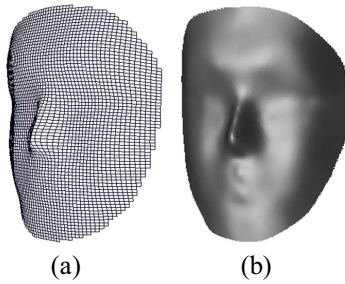
Surface norms in Eq. 1 are calculated from a generic 3D face model, since a personalized 3D model can hardly be generated accurately with only a single view. The 3D shape was obtained by averaging 40 face shapes [9] from the CMU-PIE database [10] as shown in Fig. 2. Surface norm  $\vec{N}$  is calculated using B-spline surface approximation on 16 neighbor points in a facial area as

$$\vec{N} = \frac{\partial}{\partial x} Q(x, y) \times \frac{\partial}{\partial y} Q(x, y), \quad (2)$$

where  $Q(x, y)$  is the B-spline surface matrix [12].



**Fig. 1. Face images under a common illumination but from different viewpoints in CMU-PIE database. Note the highlight changes because of specular reflections.**



**Fig. 2. The generic 3D face shape averaged from 40 shapes in [9]. (a) The wire frame, and (b) a shaded shape under Phong reflectance assumption with a uniform albedo.**

Let  $H_a^i = I_a^i k_a$ ,  $H_d^i = I_{po}^i k_d$ ,  $H_s^i = I_{po}^i k_s$ , and assume  $n \equiv 2$  [6]. The image intensity on point  $i$  of a frontal  $m$ -pixel image is then calculated as

$$I_F^i = H_a^i + H_d^i (\vec{V}_F \cdot \vec{N}^i) + H_s^i (\vec{R}^i \cdot \vec{V}_F)^2. \quad (3)$$

The facial reflection components are estimated in two steps. First all the surface points are assumed to have a uniform reflection property. The reflection components of Phong model are approximated as uniform  $H_a$ ,  $H_d$ ,

and  $H_s$ . For all  $m$  pixels of the frontal training image, we have

$$\begin{cases} I_F^1 = H_a + H_d (\vec{V}_F \cdot \vec{N}^1) + H_s (\vec{R}^1 \cdot \vec{V}_F)^2 \\ \vdots \\ I_F^m = H_a + H_d (\vec{V}_F \cdot \vec{N}^m) + H_s (\vec{R}^m \cdot \vec{V}_F)^2 \end{cases}. \quad (4)$$

Eq. 4 is solved in a least square sense using the gradient projection method [11]. In the second step, a refinement is performed to estimate pixel-wise reflection components for individual surface points.

Frontal intensity estimation  $I_F^i$  is generated provided  $H_a$ ,  $H_d$ , and  $H_s$  as

$$I_F^i = H_a + H_d (\vec{V}_F \cdot \vec{N}^i) + H_s (\vec{R}^i \cdot \vec{V}_F)^2. \quad (5)$$

Since  $H_a$ ,  $H_d$ , and  $H_s$  are approximate solutions of Eq. 4, the virtual texture  $I_F^i$  is almost always different from the gallery view  $I_F^i$ . To obtain exact individual reflection components, a refinement ratio  $c_i$  is introduced as

$$I_F^i = c_i \{ H_a + H_d (\vec{V}_F \cdot \vec{N}^i) + H_s (\vec{R}^i \cdot \vec{V}_F)^2 \}, \quad (6)$$

where  $c_i = I_F^i / I_F^i$ . The refined and pixel-wise reflection components are computed as  $H_a^i = c_i H_a$ ,  $H_d^i = c_i H_d$ , and  $H_s^i = c_i H_s$ .

Novel virtual textures under different illuminations are synthesized based on the refined reflection components. Suppose the new illumination has ambient light  $I_a^{new}$  and  $l$  point lights  $I_{po,1}^{new} \vec{L}_1^{new}$ ,  $I_{po,2}^{new} \vec{L}_2^{new}$ , ...,  $I_{po,l}^{new} \vec{L}_l^{new}$ . The virtual texture under the new illumination is calculated as

$$I^i = H_a^i I_a^{new} + H_d^i \sum_{j \leq l} I_{po,j}^{new} (\vec{L}_j^{new} \cdot \vec{N}^i) + H_s^i \sum_{j \leq l} I_{po,j}^{new} (\vec{R}^i \cdot \vec{V})^2. \quad (7)$$

If the point is in cast shadows where  $j$ th point light doesn't reach, z-buffering [12] is performed to generate those shadows. Though the illumination conditions in gallery views is limited to be frontal, the algorithm can generate virtual views in all possible illuminations and hence is a generative method. Similarly, by modifying viewing direction  $\vec{V}$ , virtual texture from an arbitrary viewpoint can also be synthesized.

### 3. Experiments

In experiments, the proposed approach is evaluated and applied in face recognition across illumination and poses. The CMU-PIE face database [10] consisting of

68 individuals was selected. 28 persons wearing glasses were excluded from our experiment, since the proposed algorithm doesn't deal with transparent facial occlusions. The experiment includes two parts: virtual view synthesis and face recognition.

Virtual textures of virtual views of the 40 subjects were generated using our proposed texture analysis approach. The frontal view under a frontal illumination in Fig. 1(a) is used for training in the proposed texture analysis and synthesis. Virtual views under new illuminations were automatically generated and illustrated in Fig. 3. The point light directions change from approximately  $-70^\circ$  to  $70^\circ$  in yaw and  $0^\circ$  to  $20^\circ$  in tilt. Each row contains synthesized views under horizontally different illuminations and different columns reflect vertical illumination changes. The ambient light and the point light intensity are the same with those of the training frontal view. Virtual views from novel viewpoints (non-frontal) were generated from rotating the generic 3D shape and apply new viewing directions to calculate facial textures pixel-wisely. Fig. 4 shows 15 synthesized virtual views in both yaw and tilt under a frontal illumination.



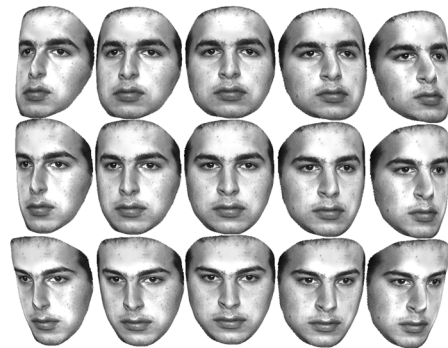
**Fig. 3. The synthesized views under varying illuminations. The 21 illumination conditions were used in the CMU-PIE database.**

In Fig. 5, the synthesized textures using the proposed approach are compared with the virtual textures generated based on Lambertian model. Fig. 5(a) is the training frontal view, and (b, d, f) are the virtual textures generated based on Lambertian reflectance model, (c, e, g) are the virtual views generated using the proposed approach and based on Phong model. Because Lambertian assumption is view-independent, all the textures based on Lambertian model in different poses are identical. On the contrary, the textures generated using the proposed approach have correct highlights and hence more realistic. There is incorrectness of textures around the eyes, especially in virtual views near profile view. That's because only single frontal view is used for training and textural information changes greatly from frontal direction to profile direction. The synthesized views with Lambertian reflectance model also have the same

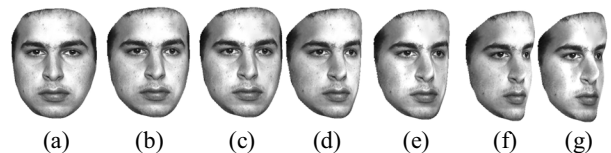
incorrectness. The introduction of more gallery views might be helpful for the virtual view generation under large pose variations.

In recognition, the face recognition system is augmented with the virtual views generated using the proposed approach to cope with pose and illumination variations. The following experiments are conducted.

1. Face recognition across pose variations.
  - a. Frontal views as gallery;
  - b. Use the proposed approach to augment gallery.
2. Frontal face recognition across 21 different illumination conditions.
  - a. Frontal views as gallery;
  - b. Use the proposed approach to augment gallery.



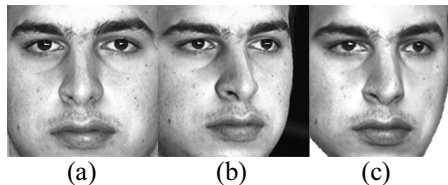
**Fig. 4. The synthesized views from different viewpoints.**



**Fig. 5. Comparison of synthesized views using the proposed approach (c, e, g) and Lambertian-based method (b, d, f) from a real training view (a). Note that the synthesized views using the proposed method have better highlights.**

First, the proposed approach was tested to perform pose-invariant face recognition. 40 face views under a rotated pose were used as probe and 40 frontal views were taken as gallery. The illumination includes ambient light and a flash from the frontal direction. Virtual views were first synthesized for every person with the same illumination and viewing conditions as to the gallery images as shown in Fig. 6. Principle component analysis [13] is then applied for recognition across pose. The rotated probe views (Fig. 6b) are matched directly with the real frontal views (Fig. 6a). The system achieved 37.5% accuracy on 40 persons.

Then the system was augmented with synthesized views and 40 virtual views (Fig. 6c) were used as gallery. After the augmentation, the recognition accuracy was 97.5% (see Table 1). The accuracy has been greatly improved. In the second face recognition experiment, the proposed approach was applied to augment face recognition across illumination. 21 frontal views under different illuminations per person were used as probe images and 840 images were tested. The augmentation improved the recognition accuracy from 79.64% to 87.62% as shown in Table 1.



**Fig. 6. Real and synthesized face views for recognition.**

**Table 1. Face recognition accuracy using direct match and the proposed approach**

Image variations	Pose	Illumination
Number of probe views	40	840
Direct match	37.5%	79.64%
<b>The proposed approach</b>	<b>97.5%</b>	<b>87.62%</b>

The experimental results show that the proposed virtual view synthesis approach improves the face recognition performance across pose and illumination. The experiment of face recognition across pose indicates that it is possible to recognize views in different poses, when only a single frontal view is available. The recognition accuracy of faces under different illuminations is also improved using the proposed non-Lambertian approach.

#### 4. Conclusion

In this paper, an automatic and stable texture analysis and synthesis approach is proposed to generate virtual face views under varying illuminations and poses for face recognition. The proposed approach utilizes single frontal face views for training and synthesizes virtual views to augment view-based face recognition to cope with pose and illumination variations. For optimal texture analysis, Phong model is introduced and specular reflection is taken into consideration. Provided a generic 3D face shape, the albedos of face surface points are computed from a single frontal face view using a gradient projection method and a reflection component refinement process. With the estimated textural coefficients, new illumination and viewing conditions are imposed onto

face shapes to generate virtual views to augment gallery views in face recognition. In experiments, virtual views under 21 different illuminations and from 15 different viewpoints were generated. To evaluate the effectiveness of the proposed texture synthesis approach for face recognition, 40 persons and 880 face images from the CMU-PIE database were tested in a PCA face recognition system. Encouraging recognition performance shows that the proposed approach significantly improves the recognition accuracies across illumination and pose variations. One of the future works is to extend the proposed approach to handle shadows.

#### References

- [1] D. Beymer and T. Poggio, "Face Recognition from One Example View," *ICCV*, 1995, pp. 500–507.
- [2] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose," *PAMI*, 23(6), 2001, pp. 643–660.
- [3] T.F. Cootes, G.V. Wheeler, K.N. Walker, and C.J. Taylor, "View-Based Active Appearance Models," *Image and Vision Computing*, 20(9-10), 2002, pp. 657–664.
- [4] R. Gross, I. Matthews, and S. Baker, "Appearance-Based Face Recognition and Light-Fields," *PAMI*, 26(4), 2004, pp. 449–465.
- [5] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the Reflectance Field of a Human Face," *SIGGRAPH*, 2000, pp. 145–156.
- [6] H.H.S. Ip and L. Yin, "Constructing a 3D Individualized Head Model from Two Orthogonal Views," *The Visual Computer*, 12(5), 1996, pp. 254–266.
- [7] V. Blanz and T. Vetter, "Face Recognition Based on Fitting a 3D Morphable Model," *PAMI*, 25(9), 2003, pp. 1063–1074.
- [8] B.T. Phong, "Illumination for Computer-Generated Pictures," *Comm. ACM*, 18(6), 1975, pp. 311–317.
- [9] X. Zhang, Y. Gao, and M.K.H. Leung, "Multilevel Quadratic Variation Minimization for 3D Face Modeling and Virtual View Synthesis," *Proc. IEEE MMM Conf.*, 2005, pp. 132–138.
- [10] T. Sim, S. Baker, and M. Bsat, "The CMU Pose, Illumination, and Expression Database," *PAMI*, 25(12), 2003, pp. 1615–1618.
- [11] Luenberger, D.G., *Introduction to Linear and Non-linear Programming*, Addison-Wesley Publishing Company, 1973.
- [12] Foley, J.D., A. van Dam, S.K. Feiner, and J.F. Hughes, *Computer Graphics: Principles and Practice*, Addison-Wesley Publishing Company, 2nd ed., 1995.
- [13] M. Kirby and L. Sirovich, "Application of the Karhunen-Loève Procedure for the Characterization of Human Face," *PAMI*, 12(1), 1990, pp. 103–108.