FACE RELIGHTING FOR FACE RECOGNITION UNDER GENERIC ILLUMINATION

Laiyun, Qing[1,2], *Shiguang Shan*[2], *Xilin Chen*[2,3]

[1]Graduate School, CAS, Beijing, 100039, China,
[2]ICT-ISVISION Joint R&D Laboratory for Face Recognition, CAS, Beijing, China, 100080
[3]Department of Computer Science, Harbin Institute of Technology, Harbin, China, 150001

ABSTRACT

The performance of current face recognition systems suffers heavily from the variations in lighting. To deal with this problem, this paper presents a novel illumination normalization approach by relighting faces to a canonical illumination based on harmonic images. Benefit from the observations that the human faces share similar shape and the albedo of the face surface is quasi-constant, we first estimate the nine low-frequency components of the lighting from the input face image. Then the face image is normalized by relighting it to canonical illumination based on illumination ratio image. For face recognition purpose, two kinds of canonical illumination, uniform and frontal point lighting, are considered, among which the former encodes merely texture information, while the latter encodes both texture and shading information. Our experimental results show that the proposed relighting normalization can significantly improve the performance of a face recognition system.

1. INTRODUCTION

Much progress in face recognition has been made in the past few years. However, face recognition remains a difficult, unsolved problem in general. The performance of the face recognition systems is subject to the variations in lighting conditions, especially under the outdoor environments [7, 8].

There has been much work dealing with illumination variation in face recognition. The low dimensionality of the face image subspace under lighting variations is the main cue. The subspace based methods such as Eigenface [15] and Fisherface[3] learn the low dimensional images subspace statistically from training images under different illuminations. These methods have demonstrated their easy implementation and accuracy. However, only under the similar imaging conditions to those of the training images, can they perform well enough.

The image space of a Lambertian surface is determined by a basis of three images, if shadows are ignored [11]. Therefore it is practical to synthesize novel images given a small number of model images. Illumination Cone [8] and Quotient image [12] are based on the 3D linear subspace, whose results are acceptable when the illumination environment is one or several point light sources.

However, the natural illumination in the real world (for example, the outdoor environment) is highly complex, consisting of reflected light from every direction as well as distributed and localized primary light sources. In this case, the rendered images are unrealistic if shadows are ignored. Basri et al [2] and Ramamoorthi et al [9] have got the analytic nine dimensional lighting space of a convex Lambertian surface, expressed in terms of spherical harmonics, considering attach-shadows. With the discovery that the effect of illumination on diffuse object is low dimension with analytic analysis, it will not be more difficult to deal with generic illumination than to deal with simple light model. Recent face recognition algorithms such as 9D linear subspace [2] and Harmonic Exemplars [18] can recognize face images under generic lighting conditions based on this model.

In this paper, we propose an algorithm for recognizing face under generic illumination with face relighting based on harmonic images [2, 9]. The problem of face relighting is stated as follows: given a face image under unknown lighting, re-render the face to the predefined canonical lighting to get its canonical form. Then face recognition is achieved by matching the canonical forms of the gallery images and that of the probe. Because usually only single image is available in face recognition systems, the classed-based method is exploited. The class-based method uses the prior knowledge of the “human face class”. The prior class knowledge used here is the similar face shape and quasi -constant face albedo, which has been used in many vision systems [1, 4, 12, 13, 17, 18, 19]. Differing from REM Relighting [17], we apply the method to face recognition rather than to computer graphics.

The remainder of this paper is organized as follows. We introduce the proposed face relighting method based on harmonic images in section 2; Section 3 gives the experimental results of face recognition; Section 4 concludes our paper and discusses the future work.

2. FACE RELIGHTING WITH SPHERICAL HARMONIC IMAGES

The reflection equation is a rotational convolution and it is natural to analysis it in space-frequency domain. It has been proved in [2, 9] in space-frequency domain that Lambertian surface is a low-pass filter and the set of images of a Lambertian object under varying lighting can be approximated by a 9D linear subspace spanned by harmonic images b_{lm} ($0 \le l \le 2$). The harmonic images are defines as

$$b_{lm}(x,y) = \rho(x,y) A_l Y_{lm}(\theta(x,y),\phi(x,y)), \quad (1)$$

where $\rho(x,y)$ is the albedo at pixel (x,y) and Y_{lm} is spherical harmonic at the surface normal (θ,ϕ) corresponding to pixels is (x,y) . The image under arbitrary lighting can be written as

$$I(x,y) = \sum_{l=0}^{2} \sum_{m=-l}^{l} L_{lm} b_{lm}, \quad (2)$$

where L_{lm} is the spherical harmonic coefficients of the lighting.

2.1. Lighting estimation

Lighting estimation is to estimate the lower nine spherical harmonic coefficients L_{lm} ($0 \le l \le 2$).

Given an input image $\mathbf{I}$ (a column vector of n elements, n is the number of pixels in the image) of an object, if the albedo map and normal map of the object are known, then the codfficents $\mathbf{L}$ can be gotten by solving the least squares problem [2]

$$\min \|\mathbf{BL} - \mathbf{I}\|, \quad (3)$$

where $\mathbf{B}$ denotes the basis images, arranged as a $n \times 9$ matrix. Every column of $\mathbf{B}$ contains one harmonic image b_{lm} as in equation (1).

To compute the harmonic images b_{lm}, the albedo map $\rho(x,y)$ and normal map $\theta(x,y),\phi(x,y)$ are needed to be known. In face recognition system, usually only single image is available. Therefore the albedo map and normal map are gotten with the "human face class" constrains.

If the pose of the image is identical, the normal map of a given face can be obtained by aligning the face image with an average face normal map of that pose, because human faces share similar shape. First the feature points in the face image are labeled automatically with an improved ASM method [16]. Then the face image and the normal map are aligned with image warping technique based on feature points. An example of the feature points labeled is given in Fig. 1.

Fig.1. An example of feature points labeling. The face image is from CMU-PIE database [14].

For albedo map, we assume quasi -constant face albedo, i.e., the albedo map contains no low frequency components ($1 \le l \le 4$), except for DC component [1]. Though the albedo of faces does not strictly satisfy this constrain, we find that we can still obtain good results in practice.

Ramamoorthi [10] shows that only 5 or 6 coefficients can be stably recovered given only one image. As we need only one new image under canonical illumination, this does not affect us.

2.2. Face relighting

Once we have estimated the lighting of the original image, it is straightforward to relight it to the canonical illumination with illumination ratio image. Illumination ratio image between the canonical image and original image the is defined as

$$IRI(x,y) = \frac{I_{can}(x,y)}{I_{org}(x,y)} = \frac{E_{can}(\theta(x,y),\phi(x,y))}{E_{org}(\theta(x,y),\phi(x,y))}, \quad (4)$$

where the subscripts are index of illumination, E is the incident irradiance. Image relighting with illumination ratio image can be written as

$$I_{can}(x,y) = IRI(x,y) \times I_{org}(x,y) . \quad (5)$$

We discuss two kinds of canonical illuminations in this paper: uniform and frontal point lighting. The intensity of light is equal in every direction in uniform illumination environment and uniform illumination can be represented with DC harmonics component. Therefore only the texture information of the face is persevered in the canonical form. In the frontal point lighting, there is as much as possible shading information (corresponding to shape information) and texture information are encoded. The irradiance environment maps of the two canonical illuminations and the original illumination are as following separately,

---

[1] The spherical harmonic coefficients of a product of irradiance and albedo, as in the basis functions Y_{lm}, are determined by a Clebsch-Gordan expansion of the product of spherical harmonics. To ensure that orders 0, 1 and 2 of the image correspond to irradiance coefficients scaled by the DC term of the albedo, assuming the only relevant irradiance coefficients are orders 0,1 and 2, we require orders 1-4 of the albedo vanish.

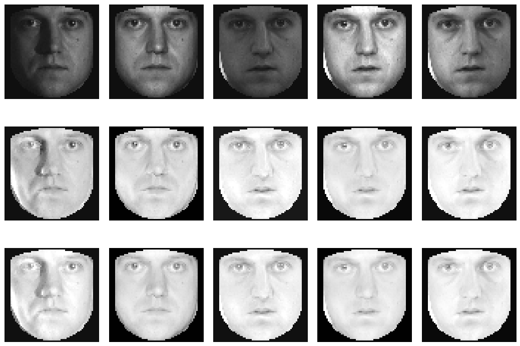


Fig.2. Relighting results: the images in the first row are the original face images, the images in the second and the last row are relit images under uniform light and frontal point light. The face images are from CMU-PIE face database [14].

$$E_{can\_uniform} = A_0 L_{00}\,, \tag{6}$$

$$E_{can\_frontal}(\theta,\phi) = \sum_{l=0}^{2} A_l L_{l0} Y_{lm}(\theta,\phi)\,, \tag{7}$$

$$E_{org}(\theta,\phi) = \sum_{l=0}^{2}\sum_{m=-l}^{l} A_l L_{lm} Y_{lm}(\theta,\phi)\,. \tag{8}$$

Given the original image $I_{org}(x,y)$, the two canonical images can be gotten with the equation (5). An example of the relighting results is given in Fig.2.

The actual albedo map of the face can be gotten by

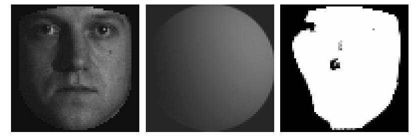
$$\rho(x,y) = I_{org}(x,y) / E_{org}(\theta,\phi)\,. \tag{9}$$

With the albedo map and normal map, we can compute the harmonic images b_{lm} of the face, which can be used for face recognition using 9D linear subspace [2].

2.3. Limited dynamic range of image

Because digitized image has limited dynamic range, ratio based relighting would have artifacts where skin pixel values are too low or nearly saturated. These pixels are called outliers. If we do no process for these pixels, they may introduce confusion in recognition.

Wen et al [17] used constrained texture synthesis [6] to alleviate the problem, which can also be applied for our purpose. As the synthesized texture for outliers adds no more valuable information for recognition, we just discard these pixels in face recognition. This is like portion face recognition, except that the portion is automatically decided. We use the frontal half irradiance environment map E_{org} to detect the outliers. If the intensity of its corresponding pixel in irradiance map is too low, the pixel in the image must be in dark shadows and it is declared as an outlier. If the energy of the DC component of the lighting is larger enough compared with that of the first order, the detection of outliers are not necessary because the pixels in the irradiance map will not be very low. An example of the detected outliers is illustrated in Fig.3.



(a) (b) (c)

Fig.3. Outliers detection. (a) input image, (b) estimated frontal. irradiance environment map, (c) detected outliers (black pixels).

3. EXPERIMENTAL RESULTS

We select CMU-PIE database [14] for the test dataset because it includes face images both with the ambient lights on and with them off. There are totally 43 different lighting conditions, 21 flashes with ambient light on or off. The images of 68 persons are included in CMU-PIE database. Only the frontal images are used in the experiments for we do not consider pose variation here. For more details about the CMU-PIE database, please refer to [14].

The simplest normalized correlation is exploited as the similarity between two images. Face recognition is achieved by finding a nearest neighbor based on the image similarity. The images are divided into 4 subsets according to the greater of the longitudinal and latitudinal angles of the flash direction from the camera axis—Subset 1(f06~f09, f11, f12, f20), Subset 2(f05, f10, f13, f14, f19, f21), Subset 3(f04, f15, f18, f22), and Subset 4(f02, f03, f16, f17).

The experimental results are illustrated in Table 1. The images in set *a* are images with ambient lights off and the images in set *b* are images with ambient lights on. The numeral in the blank is the number of the flash No. To test the effect of lighting variation only, the gallery and the probe are the same images set (there are little differences in glasses in set *a* and set *b*).

Table 1. Recognition rate comparisons between different relighting techniques on CMU-PIE database

gallery	probe	relight	Performance of Subset No. (%)			
			1	2	3	4
a(*11*)	*a*	No	96	58	24	9
		uniform	100	99	97	73
		uniform-outliers	100	99	97	78
		frontal	100	100	89	45
		frontal-outliers	100	100	91	51
b(11)	*b*	No	100	97	85	45
		uniform	100	100	100	100
		frontal	100	100	100	99

We can see that the recognition rates are higher after relighting in all the cases. The recognition rates of image set *b* are perfect for all the four subsets. The results of face recognition without outliers for image set *b* are not listed in the Table 1 because the illuminations include ambient light. The performance is a little better if the outliers are discarded for image set *a*. We had hoped that the recognition rate would higher when the canonical illumination is frontal point lighting. But from experiments we found that the performance was poorer in some cases. One possible reason is the inaccurate automatic feature points labeling, because inaccurate landmarks will lead to incorrect 3D shape and result in wrong face matching.

4. CONCLUSION

With the discovery that the effect of illumination on diffuse object is low dimension with analytic analysis, it will not be more difficult to deal with generic illumination than to deal with simple point light source model. Based on these observations, we propose a technique for face relighting under generic illumination based on harmonic images, i.e., calibrating the input face image to the canonical illumination, to reduce the negative effect of poor illumination in face recognition. A comparison study is performed between two kinds of relit images and original face. The two kinds of canonical illuminations are uniform and frontal point lighting, in which texture information, and both texture and shape information are encode respectively. Our experimental results show that the proposed lighting normalization method based on face relighting can significantly improve the performance of a face recognition system.

The performance of relit images under uniform lighting is a little better than that of frontal point lighting, contrary to what we had expected. This indicates that there are risks in using shape information if it is not very accurate. The results of feature points labeling under good is practical while those of extreme lighting are not good enough with current technology. Therefore maybe the recovered harmonic base images for gallery images under good lighting with the 9D linear subspace [2] is more applicable for face recognition systems under extreme lighting.

ACKNOWLEDGEMENTS

This research is partly sponsored by Natural Science Foundation of China (under contract No. 60332010), National Hi-Tech Program of China (No.2001AA114190 and No. 2002AA118010), and ISVISION Technologies Co., Lt.. The authors would also give thanks to those who provided the public face databases.

REFERENCES

[1] J. Atick, P. Griffin, and A. Redlich, "Statistical Approach to Shape from Shading: Reconstruction of 3D Face Surfaces from Single 2D Images", *Journal of Neural Computation*, vol. 8, no. 6, pp. 1321—1340, 1996.

[2] R. Basri, D. Jacobs, Lambertian Reflectance and Linear Subspaces, Proc. ICCV'2001, 2001.

[3] P. N. Belhumeur, J. P. Hespanha and D. J. Kriegman, "Eigenfaces vs Fisherfaces: recognition using class specific linear projection". *IEEE TPAMI*, vol.20, No.7, 1997.

[4] V. Blanz and T. Vetter, "Face Recognition Based on Fitting a 3D Morphable Model", *IEEE TPAMI*, vol. 25, no. 9, pp.1-12, 2003.

[5] A.S. Georghiades, P.N. Belhumeur and D.J. Kriegman, "From Few to Many: Illumination Cone Models for Face Recognition under Differing Pose and Lighting", *IEEE TPAMI*, Vol.23, No.6, pp643-660, 2001.

[6] A. Hertzmann et al., Image analogies, Proc. SIGGRAPH'2001, pp. 327–340, 2001.

[7] P.J. Phillips, H. Moon, et. al., "The FERET Evaluation Methodology for Face-Recognition Algorithms", *IEEE TPAMI*, Vol.22, No.10, pp.1090-1104, 2000.

[8] P.J. Phillips, P. Grother, R.J. Micheals, et. al., FRVT 2002: Evaluation Report, http://www.frvt.org/DLs/FRVT_2002_ Evaluation_Report.pdf, March 2003.

[9] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object", *Journal of the Optical Society of America A*, vol.18, no.10, pp.2448-2459, 2001.

[10] R. Ramamoorthi, Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object. *IEEE TPAMI*, vol. 24, no. 10, pp. 1-12, 2002.

[11] A. Shashua. "On photometric issues in 3D visual recognition from a single 2D image". *Iernational Journal of Computer Vision*, vol.21, no.1-2, pp.99–122, 1997.

[12] A. Shashua and T. Riklin-Raviv, "The Quotient Image: Class-Based Re-Rendering and Recognition With Varying Illuminations", *IEEE TPAMI*, vol.23, no.2, pp.129-139, 2001.

[13] T. Sim, T. Kanade, Combining Models and Exemplars for Face Recognition: An Illuminating Example, Proc. of Workshop on Models versus Exemplars in Computer Vision, Proc. CVPR'2001, 2001.

[14] T. Sim, S. Baker, and M. Bsat, The CMU Pose, Illumination, and Expression (PIE) Database, Proc. FG'2002, 2002.

[15] M. Turk and A. Pentland, "Eigenfaces for Recognition", *Journal of cognitive neuroscience*, vol. 3, no. 1, pp71-86, 1991.

[16] W. Wang, S. Shan, W. Gao, et. al., An Improved Active Shape Model For Face Alignment, Proc. ICMI'2002, pp523-528, 2002.

[17] Z. Wen, Z. Liu and T.S. Huang, Face Relighting with Radiance Environment Maps, Proc. CVPR'2003, vol II, pp.158-165, 2003.

[18] L. Zhang, D. Samaras, Face Recognition Under Variable Lighting using Harmonic Image Exemplars, Proc. CVPR'2003, vol. I, pp.19-25, 2003.

[19] W. Zhao and R. Chellappa, Illumination-insensitive Face Recognition using Symmetric Shape-from-Shading, Proc. CVPR'2000, 2000.