Eigen-Harmonics Faces: Face Recognition under Generic Lighting

Laiyun Qing[1,2], Shiguang Shan[2], Wen Gao[1,2]
[1]*Graduate School, CAS, Beijing, China, 100080*
[2]*ICT-ISVISION Joint R&D Laboratory for Face Recognition, CAS, Beijing, China, 100080*
Emails: {lyqing, sgshan, wgao}jdl.ac.cn

Abstract

The performances of face recognition systems are heavily subject to the variations in lighting. We propose a novel approach for face recognition under generic illumination conditions, named as Eigen-harmonics faces in this paper. First, using bootstrap set consisting of 3D face models with texture, we render the spherical harmonic images for every face and train the PCA harmonics faces model. During registration, given a novel face image under arbitrary illumination, we estimate the lighting of the image and recover the PCA coefficients of the spherical harmonics images for this face. During testing, we recognize the face using the PCA coefficients. The experimental results on the images under a wide range of illumination conditions in the public CMU-PIE database are promising.

1. Introduction

Much progress in face recognition has been made in the past few years [23]. However, face recognition remains a difficult, unsolved problem in general [11, 12]. The images of face depend not only on the identity of the person, but also on parameters such as head pose and illumination. Variations in pose and illumination, which may produce changes larger than the differences between different people's images, are the main challenge for face recognition. FRVT test 2002 shows that even for the best face recognition systems, the recognition rate for faces captured outdoors is still very low [12]. One of the characters of the outdoor images is the changing directional sunlight illumination.

The goal of recognition algorithms is to separate the characteristics of a face, which are determined by the intrinsic shape and texture of the facial surface, from the extrinsic imaging conditions of image generation, such as lighting and pose. In this paper we propose an eigen-harmonics faces method, which recovers the PCA coefficients of the spherical harmonic images of a novel a face from just one image taken under arbitrary illumination conditions. The eigen based methods can represent the common and differences of the subclasses in a larger class, such as human face. The eigen based methods have been applied successfully in many class based vision system, such as Eigenfaces [4, 9, 19], 3D Morphable Model [6] and SSFS (statistical shape from shading) [1]. The spherical harmonic images are very convenient for processing images under generic lighting [2, 3, 14, 15, 21, 22] and they capture the intrinsic shape and texture of the facial surface.

According how to deal with the extrinsic imaging parameters, the methods in face recognition can be classified into two fundamental strategies: model-based approach and statistics-based approach. The model-based approaches treat the extrinsic parameters as separate variables and model their functional role explicitly. The statistics based approaches analyze images directly using statistical methods and does not formally distinguish between intrinsic and extrinsic parameters. The former includes SSFS (statistical shape from shading) [17], Symmetric Shape-from-Shading [24], Illumination Cone [8], Quotient Image [16] and 3D Morphable Model [6]. The latter includes Eigenfaces [19] and FisherFaces [4, 9]. Our method is a model-based approach.

Many of the earlier model-based methods assume simple light models. Using spherical harmonics and signal-processing techniques, Basri et al [2] and Ramamoorthi [14] have shown that the set of images of a convex Lambertian object obtained under generic lighting can be approximated accurately by a nine dimensional linear subspace. Furthermore, a simple scheme for face recognition with excellent results was described in [2]. However, to use this recognition scheme, the basis images spanning the illumination space for each face are required. These images can be rendered from a 3D scan of the face or can be estimated by applying PCA to a number of images of the same subject under different illuminations [3]. An effective approximation of this basis by 9 single light source

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE

images of a face was reported in [10]. The above-mentioned methods need a number of images and/or 3D scans of each subject in the database, thus requiring specialized equipment and procedures for the capture of the gallery set, thus limiting their applicability.

In this paper we propose a method that recovers the nine spherical harmonic images of the illumination space from just one image taken under arbitrary illumination conditions. Our method computes a statistical model of the harmonic images during a bootstrap phase, which encapsulates texture and shape information, similar with 3D Morphable Model [6]. The two methods need 3D scans of the subjects in training set. But for test, only a face image is needed. Compared with the 3D Morphable Model [6], Our method has following differences: First, we treat the shape and texture together as harmonic images to consider the correlation between shape and texture data rather than to recover its texture and shape parameters separately; Second, Our method can process arbitrary illumination environment rather than illumination environments with a ambient light and a directional light; Third, we speed the system by simplifying the model (no specular reflection, no optimization). Of course, the performance of the system will be a little lower. But we see from the experimental results that the performance is still applicable.

This paper is organized as follows. In the next section, we will show how to construct the eigen-harmonic faces. In section 3, we explain how to recover the parameters of the spherical harmonic images from a novel face image. In section 4, we describe our experiments and the results. The final Section presents the conclusions and the future work.

2. Eigen-harmonic faces

Assumed the surface of human faces is convex Lambertian surface, the set of images of face under varying lighting can be approximated by a 9D linear subspace spanned by harmonic images [2, 14]. The harmonic images of face, here called as *harmonic faces*, are images of the face seen under harmonic lights. Harmonic light is a virtual light in which only one harmonic component is included. Let λ_i denote the albedo of a point p_i on the face surface and (θ,ϕ) denote its normal, the harmonic faces are constructed as:

$$b_{lm}(p_i) = \lambda_i \rho_l Y_{lm}(\theta,\phi)\,, \tag{1}$$

where $\rho_l\ (\rho_0 = \pi, \rho_1 = \frac{2\pi}{3}, \rho_1 = \frac{\pi}{4})$ is the spherical harmonic coefficients of Lambertian reflectance [2, 14], Y_{lm} is the spherical harmonic function. The real forms of the first nine harmonic faces in are:

$$\begin{gathered}
(x, y, z,) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\
b_{oo}(p_i) = \lambda_i \sqrt{\frac{\pi}{4}}\,, \\
(b_{1-1}, b_{10}, b_{11})(p_i) = \lambda_i \sqrt{\frac{\pi}{3}}(y, z, x)\,, \\
(b_{2-2}, b_{2-1}, b_{21}) = \lambda_i \sqrt{\frac{15\pi}{64}}(xy, yz, xz)\,, \\
b_{20}(p_i) = \lambda_i \sqrt{\frac{5\pi}{256}}(3z^2 - 1)\,, \\
b_{22}(p_i) = \lambda_i \sqrt{\frac{15\pi}{256}}(x^2 - y^2)\,.
\end{gathered} \tag{2}$$

Note that $\mathbf{b}_{00}$ is an image obtained under constant, ambient light, and so it contains simply the surface albedo (up to a scaling factor). The other harmonic images $\mathbf{b}_{lm}$ contain both information of surface normal and the albedo.

We select 130 aligned 3D face scans with texture information from USF Human ID 3D database [5] as our bootstrap set. We render nine harmonic images per face model using Equation set (2).

Let $\mathbf{H}_j = (\mathbf{b}_{00}, \mathbf{b}_{1-1}, \mathbf{b}_{10}, \mathbf{b}_{11}, \mathbf{b}_{2-2}, \mathbf{b}_{2-1}, \mathbf{b}_{20}, \mathbf{b}_{21}, \mathbf{b}_{22})$ denote the vector composed by the nine harmonics images of the j^{th} person. We perform Principal Component Analysis [7] on the bootstrap set of $\mathbf{H}_j$, $j = 1,2,\ldots,m, m = 130$. We subtract the average $\overline{\mathbf{h}} = \frac{1}{m}\sum_{j=1}^{m}\mathbf{H}_j$ from each harmonic vector, $\mathbf{a}_j = \mathbf{H}_j - \overline{\mathbf{h}}$, and define a data matrix $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m)$. Then we compute the eigenvector $\mathbf{h}_1, \mathbf{h}_2, \ldots$ of the covariance matrix $\mathbf{C} = \frac{1}{m}\mathbf{A}\mathbf{A}^T = \frac{1}{m}\sum_{j=1}^{m}\mathbf{a}_j\mathbf{a}_j^T$, which can be achieved by a Singular Value Decomposition [13] of $\mathbf{A}$. The eigenvalues of $\mathbf{C}$, $\sigma_{s1}^2 \geq \sigma_{s2}^2, \ldots$, are the variances of the data along each eigenvetor. We select the first $n (n \leq m)$ eigenvectors form an orthogonal basis,

$$\mathbf{H} = \overline{\mathbf{h}} + \sum_{i=1}^{n}\alpha_i\mathbf{h}_i\,. \tag{3}$$

$n = 60$ is used in our experiments, in which covers 90% of the variances. The average vector, the first eigenvector and the last eigenvector are visualized as four harmonic images in Figure 1. The first eigenvectors have the shape of face, while the values of the last eigenvectors are random.

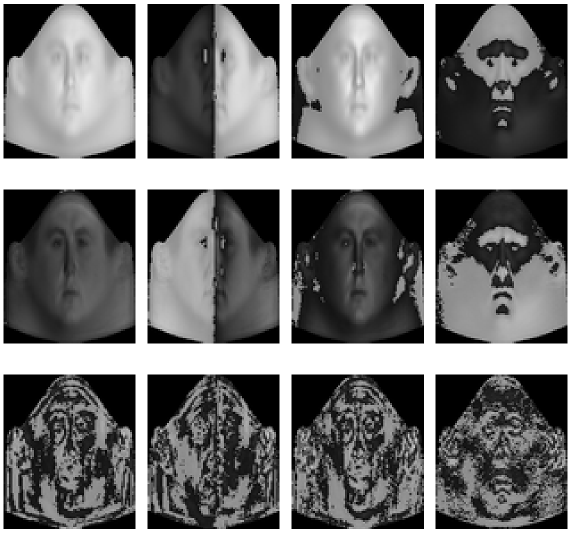


Figure 1. The visualized harmonic images of the eigenvectors. The first row is the average vector. The second row is the first eigenvector and the third row is the last eigenvector. Positive values and red indicates negative values. Lighter gray indicates positive values and darker gray indicates negative values (for positive values, the range of pixel intensities is from 140 to 240 and for negative values, the range is from 20 to 120). For page limit, only the first four harmonic images are given.

3. Eigen-harmonics faces based image analysis

The goal of the eigen-harmonic faces based image analysis is to represent a novel face in an image by model the PCA coefficients α_i in Equation (3).

The experiments in psychology have shown that human know the lighting direction of a face image before he recognizes the face. This indicates that we can use the average face model rather than the specific face model in estimating lighting. This assumption has also been used in many vision systems [16, 21, 24]. As long as the lighting is known, we can recover the parameters of the specified spherical harmonic images of the given face.

3.1. Face Alignment of the frontal face images

Before the lighting estimation, we need to align the face in the input image to the model. The alignment is achieved by morphing based on feature points.

Given a 2D image, to create the correspondence between the vertices of the average face model and the 2D image, we first create correspondence between the feature points on the average face and the 2D image. Then the rest of the vertices are aligned with image warping technique.

The feature points on the 2D image are marked with an enhanced Active Shape Model [20]. An example is illustrated in Figure 2.

(a) (b)

Figure 2. The input face image (a) and its aligned image (b) (the white pixels is invisible). The face image is from CMU-PIE database.

3.2. Harmonic images recovering

Given an aligned face image $\mathbf{I}$, let $\bar{\mathbf{h}}$ and $\mathbf{h}_j$ denote the matrix form of the average spherical harmonic images and the PCA base harmonic images in this subsection (every column is spherical harmonic images), then by solving the least squares problem

$$\min\left\|\bar{\mathbf{h}}\mathbf{L}-\mathbf{I}\right\|, \qquad (4)$$

we get the vector of the nine illumination coefficients $\mathbf{L}$, which approximates the Lambertian part of the illumination.

After $\mathbf{L}$ is known, we recover the PCA coefficients α_i of the given face. These PCA coefficients α_i can be achieved by solving another least squares problem

$$\min\left\|\bar{\mathbf{h}}\mathbf{L}+\sum_{i=1}^{n}\alpha_i(\mathbf{h}_i\mathbf{L})-\mathbf{I}\right\|. \qquad (5)$$

Once the PCA coefficients of the harmonic images are recovering, the harmonic images can be reconstructed with Equation (3). Figure 3 shows the spherical harmonic images recovered from face images under two different illuminations of two persons (not the face in bootstrap set). We can see that the harmonic images recovered for the same person from different images are almost the same while the differences between different persons are still preserved.

4. Experimental Results

In the real world, illumination usually consists of an ambient light with one or more possible point lighting sources. To obtain representative images of such cases, CMU-PIE database [18] 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

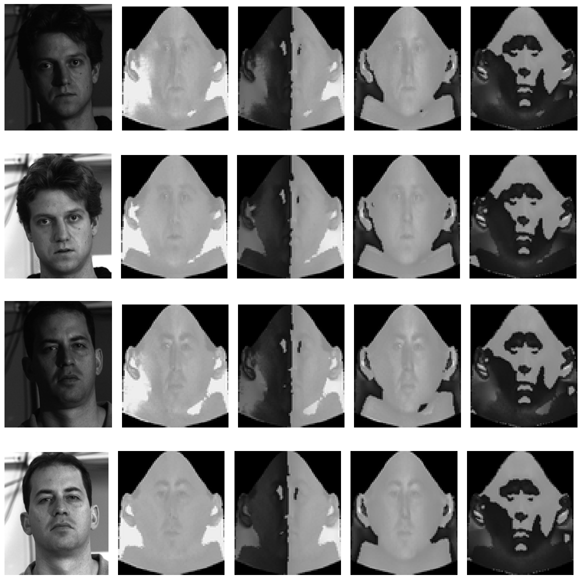


Figure 3. The results of reconstructed spherical harmonic images recovered from images of the different persons under various lighting. The first column is the images we used for the recovery followed by the set of spherical harmonic images.

included in PIE face database. For more details about the CMU-PIE database, please refer to [18].

The images are divided into two sets, the images in set *a* are images with ambient lights off and the images in set *b* are images with ambient lights on. The images in each set are divided into 4 subsets according to the greater of the longitudinal and latitudinal angles of the flash direction from the frontal face 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). Figure 4 shows some of the images we used in our experiments.

Because the current ASM model can only handle frontal face image, we select only the frontal images for experiments.

Figure 4. The real face image under various lighting used in the experiments.

4.1. Results of lighting estimation

We compute the irradiance environment map with light spherical harmonic coefficients $\mathbf{L}$ as:

$$E(\theta,\phi)=\sum_{l=0}^{2}\sum_{m=-l}^{l}\rho_l L_{lm} Y_{lm}(\theta,\phi)\,, \qquad (6)$$

where ρ_l and $Y_{lm}(\theta,\phi)$ are the same as in Equation (1). Some examples of the estimated irradiance environment maps are illustrated in Figure 5.

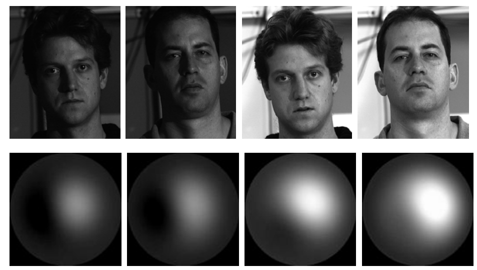


Figure 5. The irradiance environment maps of estimated lighting. The first row is the input images and the second row is the irradiance environment maps.

By selecting the irradiance maps of one person under every light as gallery and all the other irradiance spheres as probe, we classify the estimated lighting to evaluate the results of lighting estimation. The normalized correlation is exploited as the image similarity between two irradiance maps. Classification is achieved by finding a nearest neighbor based on the image similarity.

The classification rate is given in Figure 6. The results of classification with feature points automatically labeled and manually labeled are little different, which indicate the light estimation is stable even if the feature points are not labeled exactly. The results of classification are not very nice because the difference between neighboring light are little, especially after blurred as irradiance. Some other metrics for evaluating the lighting results are welcome.

4.2. Results of face recognition

To compare two faces, we use cosine of the angle between two PCA coefficients vectors of the spherical harmonic images c_1, c_2 as face similarity,

$$d_A=\frac{\langle c_1,c_2\rangle}{\|c_1\|\cdot\|c_2\|}\,. \qquad (7)$$

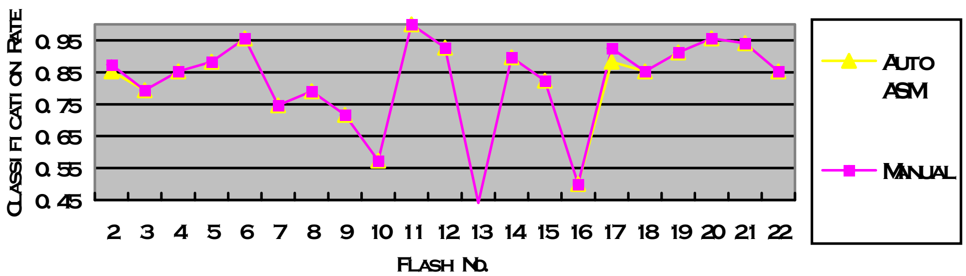
Face recognition is achieved by finding a nearest neighbor based on the face similarity. The experimental results are given in Table 1. To test the effect of lighting variation only, the gallery and the probe are the same images set (there are little differences in glasses in set *a* and set *b*). We have not tried the LDA metric of the PCA coefficients yet, which is declared better than angle metric in [6]. We can expect its performance will be better than the angle metric in our experiments.

Since we have already estimated the PCA coefficients of the spherical harmonic images, we can reconstruct the spherical harmonic with Equation (3). Then faces can be recognized using the 9D linear subspace in [2].

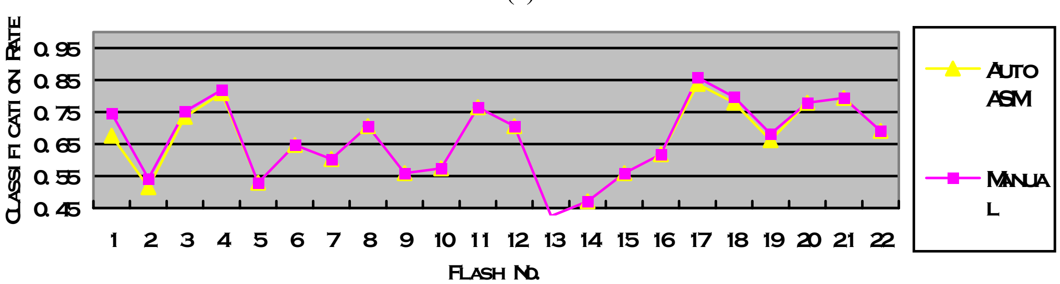
5. Conclusion

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 light source model. In this paper, we propose an eigen-harmonic faces technique for face recognition under generic illumination based on this discovery. The spherical harmonic images are very convenient for processing images under generic lighting and they capture the intrinsic shape and texture of the facial surface. The eigen based methods can represent the common and differences of human faces. Tested on CMU-PIE database of images covering large variations in illumination (the illuminations range point light to ambient light, from extreme direction to frontal direction), our algorithm achieved promising results.

So far, we have not addressed non-frontal view images. If the pose of a non-frontal face image is known, the average 3D face mesh can be rotate to that pose. Then we can detect every pixel whether it is visible. The visible pixels are used in recovering the PCA coefficients of the harmonic images. For the extreme pose, i.e., profile, only half of the face is visible. As the number of PCA coefficients is much less than the visible pixels, the half face is still enough for reconstructing the harmonic images of the face. Face recognition with various lights and poses are one of our next works.



(a)



(b)

Figure 6. The results of lighting estimation. (a) Classification rate of on image set a; (b) Classification rate of on image set b;

Table 1. Recognition rate comparisons using various metric on CMU-PIE face database

Gallery	Probe	Metric	Performance of Subset No. (%)			
			1	2	3	4
a(11)	a	Co-rrelation	96	58	24	9
		Angle distance	98	93	87	65
b(11)	b	Co-rrelation	100	97	85	45
		Angle distance	99	97	93	85

6. 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. Portions of the research in this paper use the USF Human ID 3-D database and the CMU-PIE database. The authors wish to thanks everyone involved in collecting these data.

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. 1321-340, 1996.

[2] R. Basri and D. Jacobs, "Lamberian reflectance and linear subspaces", *Proc.ICCV'2001*, vol. II: pp.384-390, 2001.

[3] R. Basri and D. Jacobs, "Photometric Stereo with General, Unknown Lighting", *Proc. CVPR'2001*, vol. II: pp. 374-381, 2001.

[4] P.N. Belhumeur and J.P. Hespanha and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection", *IEEE Trans. on PAMI*, vol.19, no.7, pp.711-720, July, 1997.

[5] V. Blanz, V. and T. A. Vetter, "orphable Model for the Synthesis of 3D Faces", Proc. *SIGGRAPH'99*, pp. 187-194, 1999.

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

[7] R.O. Duda, P.E. Hart, and D.G. Stork, *Pattern Classification*, second edition, John Wiley & Sons, 2001.

[8] 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," *IEEE Trans. on PAMI*, vol.23, no.6, pp.643-660, June 2001.

[9] R. Gross, I. Matthews and S. Baker, "Eigen Light-Fields and Face Recognition Across Pose", *Proc. FG'2002*, 2002.

[10] K. C. Lee, J. Ho and D. Kreigman, "Nine Points of Light: Acquiring Subspaces for Face Recognition under Variable Lighting," *Proc. CVPR'2001,* vol. I, pp.519-526, 2001.

[11] P. J. Phillips, H. Moon, D. M. Blackburn, E. Tabassi, and J. M. Bone, "The FERET Evaluation Methodology for Face-Recognition Algorithms," *IEEE Trans. on PAMI*, vol.22, no.10, pp.1090-1104, 2000.

[12] P. J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and J. M. Bone, "FRVT 2002: Evaluation Report," March 2003, http://www.frvt.org/DLs/FRVT_2002_Evaluation_Report.pdf

[13] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C*, Cambridge Univ. Press, 1992.

[14] R. Ramamoorthi and P. Hanrahan, "An efficient representation for irradiance environment maps", *Proc. SIGGRAPH'2001*, pp.497-500, 2001.

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

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

[17] T. Sim, T. Kanade, "Combining Models and Exemplars for Face Recognition: An Illuminating Example," *Proc. CVPR Workshop on Models versus Exemplars in Comp. Vision*, December, 2001.

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

[19] M. Turk and A. Pentland, "Eigenfaces for Recognition," *Cognitive Neuroscience*, vol.3, no.1, pp.71-96, 1991

[20] Left Blank for Blind Review

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

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

[23] W. Y. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face recognition: A literature survey", *UMD CfAR Technical Report CAR-TR-948*, 2000.

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