World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:6, No:9, 2012

2D Spherical Spaces for Face Relighting under Harsh Illumination

Amr Almaddah, Sadi Vural, Yasushi Mae, Kenichi Ohara, Tatsuo Arai

Abstract—In this paper, we propose a robust face relighting technique by using spherical space properties. The proposed method is done for reducing the illumination effects on face recognition. Given a single 2D face image, we relight the face object by extracting the nine spherical harmonic bases and the face spherical illumination coefficients. First, an internal training illumination database is generated by computing face albedo and face normal from 2D images under different lighting conditions. Based on the generated database, we analyze the target face pixels and compare them with the training bootstrap by using pre-generated tiles. In this work, practical real time processing speed and small image size were considered when designing the framework. In contrast to other works, our technique requires no 3D face models for the training process and takes a single 2D image as an input. Experimental results on publicly available databases show that the proposed technique works well under severe lighting conditions with significant improvements on the face recognition rates.

Keywords—Face synthesis and recognition, Face illumination recovery, 2D spherical spaces, Vision for graphics.

I. INTRODUCTION

FACE recognition has gained significant amount of attention in recent years as one of the most successful applications of image analysis and understanding. Due to its passive nature, face recognition has clear advantages over other biometric recognition techniques requiring subjects cooperation such as fingerprint [1] and iris recognition [2].

We can generalize most of the literature face recognition approaches into feature based and appearance based methods. In geometric feature based methods [3], [4], [5], facial features such as eyes, mouth, and nose are detected. Characteristics and relations between the detected features are used as descriptors in the face recognition frameworks. In contrast, appearance based methods [6], [7] perform directly on an image based representation (i.e., pixel intensity array).

One of the main challenges encountered by current face recognition techniques lies in the difficulties of handling varying and harsh illumination conditions. In their work, Adini et al. [8] showed that in a face recognition system, image variation due to lighting changes is more signicant than that due to different personal identities. Furthermore, the problem of face relighting and recognition becomes even more difficult when we are dealing with only a single harshly illuminated image of a target face. Many attempts have been made to overcome this issue yet it still remains an active area of research in the vision community.

Amr Almaddah, Sadi Vural, Yasushi Mae, Kenichi Ohara and Tatsuo Arai are with the Graduate School of Engineering Science, Department of Systems Innovation, Osaka University, Japan. Tel./Fax: +81-6-6850-6366 email: (amr, sadi, k-oohara, mae, arai)@arai-lab.sys.es.osaka-u.ac.jp

This work was supported in part by Grant-in-Aid for Scientific Research (C) 23500242.

Given the wide range of face recognition techniques available, rather than proposing another face classifier, we propose a face relighting technique that can handle harsh illumination conditions and can be used to increase the success rates of current face recognition approaches.

In the work performed by Shashua et al. [9], face images were successfully recognized under varying illumination by using the Quotient class based re-rendering. Unfortunately, in addition to the necessity of 3 to 9 images per each subject, their approach is valid only when we have a single or finite number of known light sources which is not the case in most of uncontrolled or outdoor environments. The same issue goes for the illumination cone approach proposed by Belhumeur et al. [10]. In the photometric stereo approach, Georghiades et al. [11] successfully reconstructed 3D face geometry and albedo from seven frontal images under arbitrary illuminations.

Assuming faces to be convex Lambertian objects, Basri and Jacobs [12] proposed that the set of images of an object under all possible lighting conditions forms a polyhedral illumination cone and can be approximated by a 9-dimensional linear subspace. Lee et al. [13] successfully acquired subspaces for facial recognition under variable lighting as they built the face spherical harmonic bases for each test subject by requiring 9 images of each target subject in the training process in addition to the 3D face scans for the internal training. In another spherical harmonic approach, Zhang et al [14] introduced a 3D spherical harmonic bases morphable model (SHBMM). By enrolling the spherical harmonics illumination coefficients into a morphable model framework, they achieved a successful and robust 2D face relighting from a single image. Regrettably, SHBMM approach requires 3D face scans for the morphable model training in addition to the manual interventions required for marking the face features. Furthermore, due to the estimation of illumination coefficients out of images synthesized from 3D scans, an unresolved approximation error is found.

With current state of the art techniques and hardware development, the cost of acquiring 3D geometry decreased dramatically. Still, the majority of current face databases are made out of 2D images. Moreover, for a dynamic and adaptive training database that can be processed automatically in real time, 2D faces training database is essential.

In this research, we propose a robust face relighting technique by using spherical space properties. First, an illumination database is generated by computing face albedo and face normal out of 2D images. Based on the generated database, we analyze the target face pixels and compare them with the database by using pre-generated tiles. Our approach takes a single 2D image as an input and returns the relighted

Fig. 1. Face relighting by using 2D spherical spaces on a single input image of size 50×50 pixels.

face automatically and in real time processing speed, figure.1.

Compared with previous spherical harmonics works [12–14], our 2D spherical spaces relighting technique has the following distinctions:

- Complete processing in the 2D spherical spaces in real time, as all of our training and testing data are 2D images.
- No prior information requirement. Our approach does not require any prior knowledge regarding the lighting conditions or the 3D geometries of the faces and takes only a single image as an input.
- Realistic spherical harmonics parameters, due to the fact that our training data is not artificially synthesized and it is made out of naturally illuminated 2D faces.

The remainder of this paper explains our approach and experimental procedures in details.

II. Spherical harmonics representation

Spherical harmonics are the angular portion of an orthogonal set of solutions to Laplaces equation defined on the unit sphere. It has been proven by Basri et al. [12] that, under the assumption of varying distant isotropic lights, a Lambertian surface will act as a low pass filter and the reflectance function of a convex Lambertian object can be successfully presented by the second order spherical harmonics approximation as shown in eq.1.

$$I = \sum_{l=0}^{2} \sum_{m=-l}^{l} \alpha_m^l B_m^l \tag{1}$$

where I is a face image, B_m^l respresents the face spherical harmonic bases, and α_m^l are the illumination coefficients. Using the results of [12] the first 9 spherical harmonics bases up to the second harmonic order, (B_{00}, $\ldots$,B_{22}^o), can be described as follow:

$$\begin{aligned}
B_{00} &= \rho\frac{1}{\sqrt{4\pi}} & B_{10} &= \rho\sqrt{\frac{3}{4\pi}}n_z \\
B_{11}^e &= \rho\sqrt{\frac{3}{4\pi}}n_x & B_{11}^o &= \rho\sqrt{\frac{3}{4\pi}}n_y, \\
B_{20} &= \frac{\rho}{2}\sqrt{\frac{3}{4\pi}}(3n_z^2-1) & B_{21}^e &= 3\rho\sqrt{\frac{5}{12\pi}}n_x n_z \\
B_{21}^o &= 3\rho\sqrt{\frac{5}{12\pi}}n_y n_z & B_{22}^e &= 3\rho\sqrt{\frac{5}{12\pi}}n_x^2 - n_y^2, \\
B_{22}^o &= 3\rho\sqrt{\frac{5}{12\pi}}n_x n_y
\end{aligned} \tag{2}$$

where ρ is the face albedo, and (n_x, n_y, n_z) represent the unit surface normal values in the x, y, and z directions, respectively. The superscript *e* and *o* are used to identify the even and odd spherical harmonic bases.

Any image *I* of a face under random lighting conditions can be approximated by a weighted combination of the previously mentioned spherical harmonic bases as described in eq.3.

$$I \approx B\alpha \tag{3}$$

B is a spherical harmonic bases matrix of the size ($p \times 9$) where *p* is the number of *I* surface points. α is a 9 elements vector representing the spherical illumination coefficients.

III. *2D* Statistical training models

To successfully relight a novel 2D face image, a training dataset with accurate 9 spherical harmonic bases is required. To build our statistical training models from a set of 2D images, we will have to go through the following 3 processes:

1) Surface points albedo and normal estimation.
2) Spherical harmonic bases recovery.
3) Spherical illumination coefficients extraction.

In the next sections we will discuss the previously mentioned steps in details.

A. Surface points albedo and normal estimation

Successful and fast 3D transformation of 2D images still remains as one of the unresolved challenges in the field of image processing. In this research, we will concentrate on extracting the required information from the 2D images without trying to completely transform them into 3D faces. As previously stated in eq.2, in addition to the surface albedo a 3D understanding of the face is essential to calculate the spherical harmonic bases. In the internal training process, the input is set to be a group of 2D images. In other words, we only have the *x* and *y* positions of the face pixels. We define the face surface function as:

$$z = f(x, y) \tag{4}$$

Then a surface normal can be represented by the vector:

$$[\frac{\delta f(x,y)}{\delta x}, \frac{\delta f(x,y)}{\delta y}, -1] \tag{5}$$

Let Ĩ $= [I_1, I_2, I_3]$ be the column vector of intensities recorded at a point (*x*, *y*) in each of three views of a spherical spaces internal training face. Further, let $\tilde{s}_1$, $\tilde{s}_2$ and $\tilde{s}_3$ be unit vectors defining the directions of the incident illumination on the faces 3 images, eq.6.

$$\begin{aligned}
\tilde{S}_1 &= [S_1 1, S_1 2, S_1 3]' \\
\tilde{S}_2 &= [S_2 1, S_2 2, S_2 3]' \\
\tilde{S}_3 &= [S_3 1, S_3 2, S_3 3]'
\end{aligned} \tag{6}$$

$$\tilde{S} = \begin{bmatrix} S_1 1 & S_1 2 & S_1 3 \\ S_2 1 & S_2 2 & S_2 3 \\ S_3 1 & S_3 2 & S_3 3 \end{bmatrix} \tag{7}$$

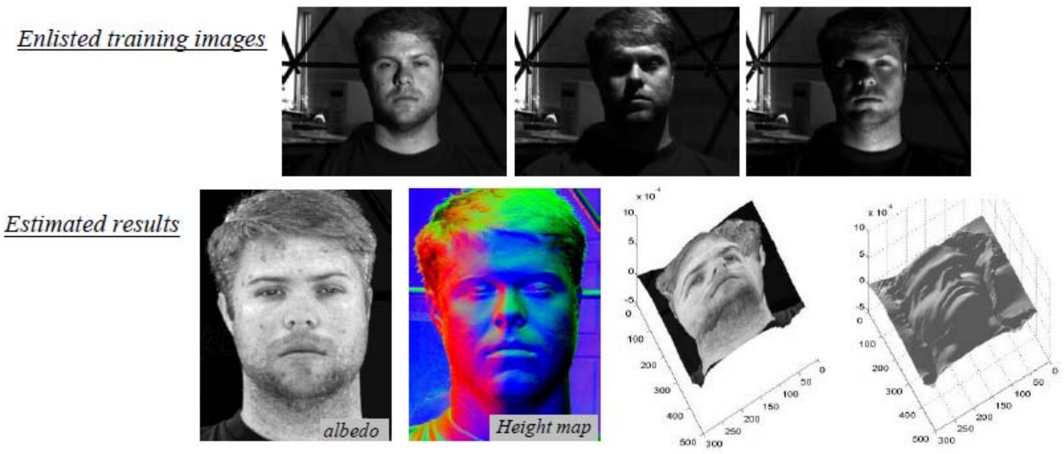

Fig. 2. Extracted surface albedo and height maps.

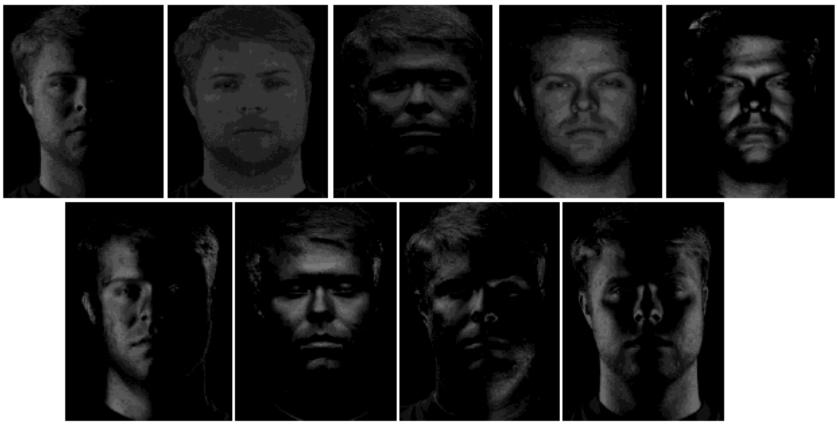

Fig. 3. Recovered 9 Spherical Harmonics Bases for a test subject.

Where $\tilde{S}$ can be rewritten as,

$$\tilde{S} = [\tilde{S_1}\tilde{S_2}\tilde{S_3}]' \tag{8}$$

The face albedo ρ can be expressed in term of $\tilde{S}$ and $\tilde{I}$ as stated in eq.9.

$$\rho = |\tilde{S}^{-1} * \tilde{I}| \tag{9}$$

The value of $\tilde{n}$, the surface normal at a point (x,y) of the training face, can be determined according to eq.10.

$$\tilde{n} = \frac{1}{\rho}\tilde{S}^{-1}\tilde{I} \tag{10}$$

As $\tilde{S}$ remains unknown in eq.9 and eq.10, we will make a rough estimation of the light direction by placing a chrome sphere on the image where the highlight will indicate the light source position. Next, a stiff deformable model is fitted to the image, assuming the previously estimated lighting position. Finally, a least–squares estimate of the light source position is derived from the model using the Levenberg Marquardt method as described by Samaras et al. [15]. In the case of only 1 or 2 images enlisted in a face training folder, we use the average face data to substitute the missing I_2 and I_3 and then we calculate the approximation error. Figure 2 shows successful estimations of surface albedo and normal values from 3 images of a face under arbitrary lightening conditions.

B. *Spherical harmonic bases recovery*

In section 3.1, we successfully estimated the surface albedo and height maps for all the enrolled internal training faces. Recovering the first 9 spherical harmonic for the training faces will be a direct application in eq.2. As a result, we will get the $(p \times 9)$ spherical harmonic bases matrix B_k , where p is the number of pixels in the enrolled face and k is superscript identifying the face between all the other faces enrolled in the internal training database. Every column in B_k contains 1 spherical harmonics base vector, eq.11

$$\tilde{B}_k = \begin{bmatrix} B_{00_1} & B_{10_1} & \dots & B^o_{22_1} \\ \vdots & \vdots & \vdots & B^o_{22_2} \\ B_{00_{(p-1)}} & \vdots & \vdots & \vdots \\ B_{00_p} & \dots & \vdots & B^o_{22_p} \end{bmatrix} \tag{11}$$

Finally, we build a $(p \times 9 \times k)$ B array containing all the training faces 9 spherical harmonic bases. Figure 3 shows the recovered spherical bases for an enrolled test subject.

Fig. 4. Recovered 9 Spherical harmonic bases for a single input test image.

Fig. 5. Recovered images for variant faces from a single 50×50 input image.

C. Spherical illumination coefficients extraction

As stated in eq.1, α represents the illumination coefficients for an input image. The value of α can be found by solving the least square problem shown in eq.12.

$$min\|B_k a_j - I_j\| \tag{12}$$

The spherical harmonic coefficient α_j is a 9 elements vector correlated to the face lighting condition. The superscript j identifies the image between all enlisted images for a training face. For each enrolled face we build a $(1 \times 9 \times j)$ α_k array containing all the illumination coefficient vectors for all enlisted images. k is the face identifier between all the internal training faces.

IV. Face relighting for a single 2D image

To relight a single image of an input face in the testing process, we perform a pixel by pixel comparison with all the enrolled images in the internal training dataset.

We assume that the spherical harmonics bases for a novel input test image at a point (x_1, y_1) can be approximated by the spherical harmonics bases at a point (x_2, y_2) in a training image with the closest pixel intensity and feature values.

$$B_{input}(x_1, y_1) = B_{training}(x_2, y_2) + er \tag{13}$$

er is an approximation correction term equal to intensity difference between the test and training images.

For the lighting coefficient α estimation, we build a bootstrap composed of ten training images with the closest intensities to the input image. α coefficients for the test image will be equal to

$$\alpha_{input} = \gamma * [\alpha_{training} + \alpha_{av}] \tag{14}$$

where

$$\gamma = \frac{\sum_{i=1}^{i=p} I_{av}(i)}{\sum_{i=1}^{i=p} I_{input}(i)} \tag{15}$$

$\alpha_{training}$ is the lighting coefficients of the training image with the closest intensity and light direction to the input test image. α_{av} is the average values of all lighting coefficients stored in the training bootstrap. I_{av} is the bootstrap average intensity image. We use the previously constructed B_{input} and α_{input} to generate the new relighted test image I_{result}, eq.16.

$$I_{result} = \gamma * B_{input} * \alpha_{input} + (1 - \gamma) * I_{input} \tag{16}$$

Figure 4 presents the recovered 9spherical harmonic bases for a single test image. Figure 5 shows a set of recovered intensity images from single input face images with size of 50×50 pixels and arbitrary lighting conditions.

V. Face relighting and recognition framework

Our proposed 2D spherical spaces relighting technique is a preprocessing component that simply takes a single image of size $(m \times n)$ and returns a relighted face image with the same dimensions.

To evaluate our approach we need to combine the relighting technique with state of the art recognition algorithm. In the next sections, we present our framework when combined successfully with several well–known face recognition algorithms.

A. Recognition by using Gabor descriptors

Gabor descriptors can robustly capture salient visual characteristics such as spatial localization, orientation selectivity, and spatial frequency characteristic. The Gabor feature vectors at facial feature points of face images are obtained by convolving Gabor wavelet kernels with the intensity of the pixel at the facial feature points. In this work we used the following Gabor wavelet kernels:

$$W(x, y, \theta, \lambda, \sigma) = e^{-\frac{1}{2\sigma^2}(\bar{x}.\bar{x})} e^{i\bar{k}.\bar{x}} \tag{17}$$

TABLE I
ERROR PERCENTAGES OF OUR RELIGHTING TECHNIQUE WHEN COMBINED WITH VARIOUS RECOGNITION ALGORITHMS

Recognition Algorithm	Error before SHB	Error after SHB
Gabor	4.9 %	3.8 %
PCA	13.0 %	8.7 %
KPCA	11.2 %	6.2 %
Sparse	13.4 %	9.7 %

TABLE II
ERROR PERCENTAGES OF FACES RELIGHTING AND RECOGNITION TASK PERFORMED ON SEVERAL DATABASES USING OUR GABOR COMBINED FRAMEWORK

Database	Error before SHB	Error after SHB
YALE B	4.90 %	3.80 %
Multi-PIE	5.25 %	4.14 %
PIE	6.72 %	4.61 %

λ represents the wave length, and , σ is the size of the Gaussian filter. The value of $theta$ is the wavelet direction, while $\bar{k}$ value can be calculated by:

$$\bar{k} = (\frac{2\pi cos\theta}{\lambda}, \frac{2\pi sin\theta}{\lambda}) \tag{18}$$

In our experiments, we consider 32 Gabor wavelet kernels obtained by the following parameters:

$$\begin{aligned} \theta &\in (0, 30, 45, 60, 75, 90, 120, 150) \\ \lambda &\in (2, 3, 4, 5) \\ \sigma &= \lambda \end{aligned} \tag{19}$$

We classify our 32 Gabor wavelet kernels by using the enhanced fisher linear discriminant model for face recognition as described by Liu et al. [16]. Our combined frame work performs 2D spherical spaces relighting and then apply Gabor wavelet kernels on the resulted images.

B. Recognition by principle component analysis

PCA is mathematically defined as an orthogonal linear transformation which projects the data along the directions where the data varies the most. The projection directions are determined by the eigenvectors of the covariance matrix corresponding to the largest eigenvalues. The magnitude of the eigenvalues corresponds to the variance of the data along the eigenvector directions. The covariance matrix for the data $I_k, k = 1, \ldots, l, \sum_{k=1}^{l} I_k = 0$ defines as:

$$C = \frac{1}{l} \sum_{j=1}^{l} I_j I_j^T \tag{20}$$

Yang et al. [17] described their framework for achieving fast and robust face recognition by using two-dimensional appearance based PCA. After we successfully relight the input test images, we calculate the PCA projections of the enrolled faces as described in [17].

C. Recognition by KPCA

KPCA, kernel principal component analysis, is a nonlinear extension of PCA. In KPCA, we map the input space into a feature space via nonlinear mapping and then calculate the principal components in that feature space. To build our KPCA face recognition system, we followed the work presented by Kim et al. [18].

D. Recognition via sparse representation

In sparse representation based classification, the testing image is represented as a sparse linear combination of the training samples. The representation fidelity is measured by the $l_2 - norm$ or $l_1 - norm$ of spares residual. Following the work of Wright et al. [19], we built a sparse representation framework to recognize relighted images.

VI. EXPERIMENTAL RESULTS

In this section, we demonstrate that 2D spherical spaces relighting technique improves the performance of current face recognition algorithms. We also show that our combined relighting and recognition framework is an appropriate method for face recognition if compared with other state of the art techniques.

For the experiments spherical harmonics internal training, we enroll 150 faces by an average of 20 images per face from the FRGC1.0 Database. For the face relighting and recognition experiments, we use images from Yale Face Database B that contains 10 subjects each under 64 different lighting conditions and images from CMU Multi-PIE database with 337 subjects under 19 illumination conditions. We also perform experiments on 68 subjects from PIE database, each under 43 different illumination conditions. In all experiments, face images are automatically aligned and cropped by using eye coordinates. The intensity values of each vectorized test image were normalized to have zero mean and unit variance.

In the experiment process, we enroll one image per test subject in the gallery folder while the rest of the subject images are listed in the probe folder. The target of this experiment is relighting all enrolled images and identifying each test subject.

Table 1 shows the improvement of the error percentages for state of the art face recognition algorithms when successfully combined with our 2D spherical spaces relighting technique. From table 1 we can see that the best performance was found

TABLE III
ERROR PERCENTAGES FOR A FACIAL RELIGHTING AND RECOGNITION TASK PERFORMED BY VARIOUS METHODS

Method	No. of required images	Error percentage
Correlation	1	33.5 %
Histogram fitting	1	41.8 %
Eigen faces	1	28.7 %
Liner subspace	7	4.9 %
Quotient Image	3	24.6 %
9 points of light	9	1.2 %
2D spherical spaces	1	3.8 %

in the Gabor relighting and recognition framework and tested on YaleB database.

Table 2 displays the reduction of error percentages for images relighting and recognition task performed on several databases using our Gabor combined framework. The best performance was acheived in YALE.B database as it provides the most challanging illumination conditions.

Table 3 shows the error comparison for a facial relighting and recognition task when performed by our approach and other face relighting and recognition techniques.

VII. CONCLUSION

In this work, we proposed our novel 2D spherical spaces face relighting technique. Our proposed approach was able to successfully recover the lighting, shape, and albedo from a single face image under varying illumination conditions. We validated our technique by presenting promising experimental results for single 2D faces relighting and recognition processes. In our proposed method, no 3D data for the training or testing process were used. The computational process is performed in real time, and with low input size requirements. In the future, we plan to further improve the results by increasing α sensitivity to the lighting condition and expand the current model to be able to recover and recognize faces under occlusion and big pose variation.

REFERENCES

[1] P. S, "On the individuality of fingerprints," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1010–1025, 2002.

[2] L. Ma, T. Tan, and D. Zhang, "Personal identification based on iris texture analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1519–1533, 2003.

[3] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," *In Procceedings of IEEE Conference on Computer Vision and Pattern Recogition*, pp. 995–999, 1994.

[4] A. Samil and P. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Journal of Foo*, pp. 65–75, 1992.

[5] L. Wiskott, J. Fellous, N. Kruger, and C. Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775–779, 1997.

[6] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–96, 1991.

[7] H. Murase and S. Nayar, "Visual recognition of 3-d objects from appearance," *International Journal of Computer Vision*, vol. 14, no. 1, pp. 5–24, 1995.

[8] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for hanges in illumination directions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 3, pp. 721–732, 1997.

[9] A. Shashua and T. Riklin, "The quotient image: Class-based re-rendering and recognition with varying illuminations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 129–139, 2001.

[10] P. Belhumeur, "What is the set of images of an object under all possible lighting conditions?" *IEEE conference on Computer Vision and Pattern Recognition*, 1996.

[11] A. Georghiades and D. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[12] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE International Conference on Computer Vision*, pp. 383–390, 2001.

[13] K. Lee, J. Ho, and D. Kriegman, "Nine points of light: Acquiring subspaces for face recognition under variable lighting," *IEEE Conference on Vision and Pattern Recognition*, pp. 519–525, 2001.

[14] L. Zhang and D. Samaras, "Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 3, pp. 351–363, 2006.

[15] D. Samaras and D. Metaxas, "Coupled lighting direction and shape estimation from single images," *IEEE International Conference on Computer Vision*, vol. 2, pp. 868–874, 1999.

[16] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.

[17] J. Yang, D. Zhang, and A. Frangi, "Two-dimensional pca: a new approach to appearance-based face representation and recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.

[18] K. Kim, K. Jung, and H. J. Kim, "Face recognition using kernel principal component analysis," *IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 40–42, 2002.

[19] J. Wright, A. Yang, A. Ganesh, S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210–227, 2009.