Face Synthesis and Recognition from a Single Image under Arbitrary Unknown Lighting using a Spherical Harmonic Basis Morphable Model

Lei Zhang, Sen Wang, Dimitris Samaras
Computer Science Department,
State University of New York at Stony Brook, NY 11794-4400, USA
{lzhang, swang, samaras}@cs.sunysb.edu

Abstract

*Understanding and modifying the effects of arbitrary illumination on human faces in a realistic manner is a challenging problem both for face synthesis and recognition. Recent research demonstrates that the set of images of a convex Lambertian object obtained under a wide variety of lighting conditions can be approximated accurately by a low-dimensional linear subspace using spherical harmonics representation. Morphable models are statistical ensembles of facial properties such as shape and texture. In this paper, we integrate spherical harmonics into the morphable model framework, by proposing a 3D Spherical Harmonic Basis Morphable Model(**SHBMM**) and demonstrate that any face under arbitrary unknown lighting can be simply represented by three low-dimensional vectors: shape parameters, spherical harmonic basis parameters and illumination coefficients. We show that, with our SHBMM, given one single image under arbitrary unknown lighting, we can remove the illumination effects from the image (face "de-lighting") and synthesize new images under different illumination conditions (face "re-lighting"). Furthermore, we demonstrate that cast shadows can be detected and subsequently removed by using the image error between the input image and the corresponding rendered image. We also propose two illumination invariant face recognition methods based on the recovered SHBMM parameters and the de-lit images respectively. Experimental results show that using only a single image of a face under unknown lighting, we can achieve high recognition rates and generate photo-realistic images of the face under a wide range of illumination conditions, including multiple sources of illumination.*

1. Introduction

Understanding arbitrary illumination effects on human faces has been a fascinating yet challenging problem both for face synthesis and for recognition. Great progress has been made in generating photo-realistic images of objects including human faces [7][8][10][22] and face recognition under different lighting conditions [1][17][25][11][13][20]. However, when only one single image under unknown lighting is available, both face synthesis and recognition become particularly challenging. In this paper, we propose a new 3D spherical harmonic morphable model (SHBMM). We demonstrate that with our model, we can generate images under new lighting conditions with remarkable quality even if only one single image under unknown lighting is available. We also provide two illumination invariant face recognition methods which achieve high recognition rates for images under a wide range of illumination conditions. [1]

Previous research suggested that illumination variability in face images is low-dimensional e.g. [12][3][1][18]. The illumination cone method [11] requires at least three images per subject to build the illumination cone. Recently, using spherical harmonics, it has been shown [2][16] that the set of images of a convex Lambertian object obtained under a wide variety of lighting conditions can be approximated accurately by a 9 dimensional linear subspace. This led to face recognition with excellent results [2] using basis images that span the illumination space. These images can be rendered from a 3D scan of the face or can be approximated by acquiring a number of images of the same subject under different illuminations [15][13]. This requirement for a number of training images and/or 3D scans of the subjects in the database necessitates specialized equipment and procedures for the capture of the training set. In a single image based approach [23] for face recognition under arbitrary illumination conditions, a statistical model of spherical harmonics is based on a bootstrap collection of 2D basis images. To recover a new set of basis images, the input image should be accurately aligned with the bootstrap images. Furthermore, this method[23] cannot perform

[1] We would like to thank Sudeep Sarkar and Simon Baker for providing databases and Thomas Vetter and Sami Romdhani for helpful discussions. This research was supported by grants from U.S. Department of Justice(2004-DD-BX-1224) and National Science Foundation(ACI-0313184)

recognition for images obtained under different viewpoints efficiently. Zhou et al. [26] extend photometric stereo algorithms to handle the appearance of the class of human faces and propose a method of recovering albedos and surface normals from one image under unknown illumination conditions. This method assumes that the human face is lit by a distant illumination, thus, it is bound to images taken under a single directional illuminant.

Inverse rendering methods [14][7][8] suggest that one can generate photo-realistic renderings of objects under new lighting conditions by capturing the lighting environment and recovering surface reflectance properties, requiring a number of images to model the environment map and face reflectance. In the Quotient image method[21], the set of images generated by varying lighting conditions on a collection of Lambertian objects (same shape, different texture) can be characterized using images of a prototype object and a illumination invariant "signature" image per object of the class. In our work, instead of assuming that all faces have the same shape, we use a set of shape parameters to represent shape information for each face and we use the spherical harmonic basis which has been proven to be illumination invariant[2][15] as our "signature" basis. Similar to our work in [22], spherical harmonics approximate the radiance environment map for any given image. The lighting conditions of one person's face can be modified so that it matches the lighting conditions of a different person's face image by assuming two faces have similar skin albedos and using a generic face model. In this work, given an image of a face, we explicitly recover the shape information and estimate the spherical harmonic basis of the face (containing the albedo information of that face). Thus the input image and target image can have different skin albedos and poses. Another benefit of our model is that by detecting the shadow errors from the image difference between the input image and the rendered image, we can remove the cast shadows from the input image and add cast shadows to the re-lit image to generate more photo-realistic images.

In 3D face Morphable Models [5], each face can be represented by linear combinations of a set of 3D face exemplars. Fitting the Morphable model to the input image was used succefully in face recognition [6] and generated impressive face synthesis results [5][4]. In this work, we integrate a more general illumination representation into the Morphable Model approach. The method in [6] is bound to images taken under directional illuminants and requires the knowledge of light direction which is difficult to know in most cases. In our method, the illumination variations are captured by the spherical harmonic basis, thus, there is no illumination limitation on the input images. Another important difference lies in the process of face synthesis. In the face synthesis applications [5][4] of the Morphable Model, new faces were synthesized by setting different shape and texture parameters, i.e. a new face was represented by a linear combination of a set of 3D face exemplars resulting to the possible loss of detail for the specific face.

The major contribution of this paper is a 3D Spherical Harmonic Basis Morphable Model, an integration of spherical harmonics into the morphable model framework. We demonstrate that any face under arbitrary illumination conditions can be represented simply by three low dimensional vectors: shape parameters, spherical harmonic basis parameters and illumination coefficients. We call these three vectors the SHBMM parameters. We show that, given one single image under unknown lighting, we can recover the set of SHBMM parameters, thus removing the illumination from the image (face "de-lighting") and generate images under new illumination conditions (face "re-lighting"). In our method, we combine our Spherical Harmonic Basis Morphable Model and a concept similar to Ratio images [21] to generate photo-realistic face images. Given an input image for one subject and a target image of another subject, we relight the input image according to the illumination condition in the target images. Since our model is constructed in 3D, the input and target images can be taken under different poses. Furthermore, by detecting the shadow errors from the image difference between the input image and the rendered image, we can remove or add cast shadows. We also provide two illumination invariant recognition methods based on the recovered SHBMM parameters and the de-lit images respectively. Experimental results show that we achieve high recognition rates for images under a wide range of illumination conditions; including multiple sources of illumination which are not easily handled by previous methods.

2. Spherical Harmonic Basis Morphable Model

Here, we will explain the spherical harmonic illumination representation [2][16] and introduce the Spherical Harmonic Basis Morphable Model.

2.1. Spherical Harmonic Illumination Representation

Spherical harmonics are a set of functions that form an orthonormal basis for the set of all square-integrable functions defined on the unit sphere. They are the sphere analog of the Fourier basis on the line or circles. Let p_i denote the ith object point. Let λ denote the vector of the object's albedos, that is, λ_i is the albedo of p_i. Similarly, let n_x, n_y, n_z denote three vectors of the same length that contain the x, y and z components of the surface normals. Further, let n_{x^2} denote a vector such that $n_{x^2,i} = n_{x,i}n_{x,i}$. We define $n_{y^2}, n_{z^2}, n_{xz}, n_{yz}, n_{xy}$ similarly. We use $\lambda_. * \mathbf{v}$ to denote the component-wise product of λ with any vector

$\mathbf{v}$. Using this notation, the first nine harmonic images of the objects are:

$$\begin{array}{ll} b_{00} = \frac{1}{\sqrt{4\pi}}\lambda, & b_{10}^{e} = \sqrt{\frac{3}{4\pi}}\lambda.*n_z, \\ b_{11}^{o} = \sqrt{\frac{3}{4\pi}}\lambda.*n_y, & b_{11}^{e} = \sqrt{\frac{3}{4\pi}}\lambda.*n_x, \\ b_{20} = \frac{1}{2}\sqrt{\frac{3}{4\pi}}\lambda.*(2n_{z^2} - n_{x^2} - n_{y^2}), & \\ b_{21}^{o} = 3\sqrt{\frac{5}{12\pi}}\lambda.*n_{yz}, & b_{21}^{e} = 3\sqrt{\frac{5}{12\pi}}\lambda.*n_{xz}, \\ b_{22}^{o} = 3\sqrt{\frac{5}{12\pi}}\lambda.*n_{xy}, & b_{22}^{e} = \frac{3}{2}\sqrt{\frac{5}{12\pi}}\lambda.*(n_{x^2} - n_{y^2}) \end{array} \tag{1}$$

where the superscripts e and o denote the even and the odd components of the harmonics respectively. With this set of basis images, any image under arbitrary illumination conditions can be approximately represented by the linear combination of the basis as Eq. 2.

$$I \approx b\ell \tag{2}$$

Eq. 2 states that any image I under arbitrary illumination conditions can be represented by the weighted combination of the basis images.

2.2. Morphing the Spherical Harmonic Basis

Face morphable models [5] were successfully applied in both face recognition and face synthesis applications [6][24] where a face was represented by a shape vector and texture vector. Inspired by the idea of morphing, we propose a new 3D Spherical Harmonic Basis Morphable Model (SHBMM) which integrates Morphable Models and the Spherical Harmonic illumination representation. Thus any face under arbitrary illumination conditions can be represented simply by three low dimensional vectors. This low dimensional representation greatly facilitates both face recognition and synthesis especially when only one input image under unknown lighting is provided.

A Spherical Harmonic Basis Morphable Model is a 3D model of faces with separate shape and spherical harmonic basis models that are learnt from a set of exemplar faces. Morphing between faces requires complete sets of correspondences between the faces. Similarly to [5], when building such a model, we transform the shape and spherical harmonic basis spaces into vector spaces. We used a collection of 3D faces supplied by USF [5] to construct our model. For each 3D face, we computed a set of 9 spherical harmonic bases. We present a face of with a shape-vector $S = (X_1, Y_1, Z_1, X_2, ..., Y_n, Z_n)^T \in \Re^{3n}$ and a Spherical Harmonic basis vector $B = (B_1^1, ..., B_1^9, ..., B_n^1, ..., B_n^9)^B \in \Re^{9n}$, where n is the voxel numbers of the face. The Spherical Harmonic basis morphable model can be constructed using a data set of m exemplar faces; exemplar i is represented by the shape-vector S_i and SHB-vector B_i. New shapes s and spherical harmonic bases b can be generated by convex combinations of the shapes and textures of the m exemplar faces as Eq. 3 shows:

$$s = \bar{s} + S\alpha; b = \bar{b} + B\beta \tag{3}$$

Combining Eq. 2 and 3, we see that, any face under arbitrary illumination conditions can be represented by three low dimensional vectors (SHBMM parameters): $\{\alpha, \beta, \ell\}$ with α representing the geometry parameters, β the spherical harmonic basis parameters and ℓ representing the illumination coefficients.

3. Fitting SHBMM to Images under Arbitrary Unknown Lighting

Here, we will explain the recovery of those parameters from a single image under arbitrary unknown lighting.

3.1. Shape Parameters Recovery

As shown in [6], a realistic face shape can be generated by:

$$s_{2d} = fPR(\bar{s} + S\alpha + t_{3d}) + t_{2d} \tag{4}$$

where f is a scale parameter, P an orthographic projection matrix and R a rotation matrix with ϕ, γ and θ the three rotation angles for the three axes. t_{3d} and t_{2d} are translation vectors in 3D and 2D respectively. Eq. 4 relates the vector of 2D image coordinates s_{2d} and the shape parameters α. The shape parameters of the Morphable Model can be estimated from single or multiple input images automatically as described in [6][9]. In our implementation, we do not focus on automatic recovery of shape parameters and so we choose to use a simple semi-automatical feature correspondence selection method [23]. This simple feature based method performed sufficiently well to demonstrate the strength of our approach for face synthesis and recognition.

Given an input image of a face, we first initialize all the parameters in Eq. 4: the shape parameter α is set to 0 and pose parameters f,ϕ,γ,θ and t_{2d} are initialized manually. For the set of hand-picked feature points in the SHBMM, we find the correspondence s_f^{img} in the training image semi-automatically. The set of feature points contains major and secondary features, see Figure 1. After the correspondences of major features are manually set, the secondary feature correspondences are updated automatically. Thus, the pose parameters f, ϕ, γ and θ can be recovered by minimizing the error between s_f^{img} and the model feature points: $argmin_{f,\phi,\gamma,\theta,t_{2d}} \|s_f^{img} - (fPR(\bar{s}_f + S_f\alpha + t_{3d}) + t_{2d})\|^2$ where $\bar{s}_f$ and S_f is the corresponding shape information of the feature points in the SHBMM in Equation 3.

The shape error of the feature points, δs_f^{2d}, is defined as the difference between s_f^{img} and the new shape information

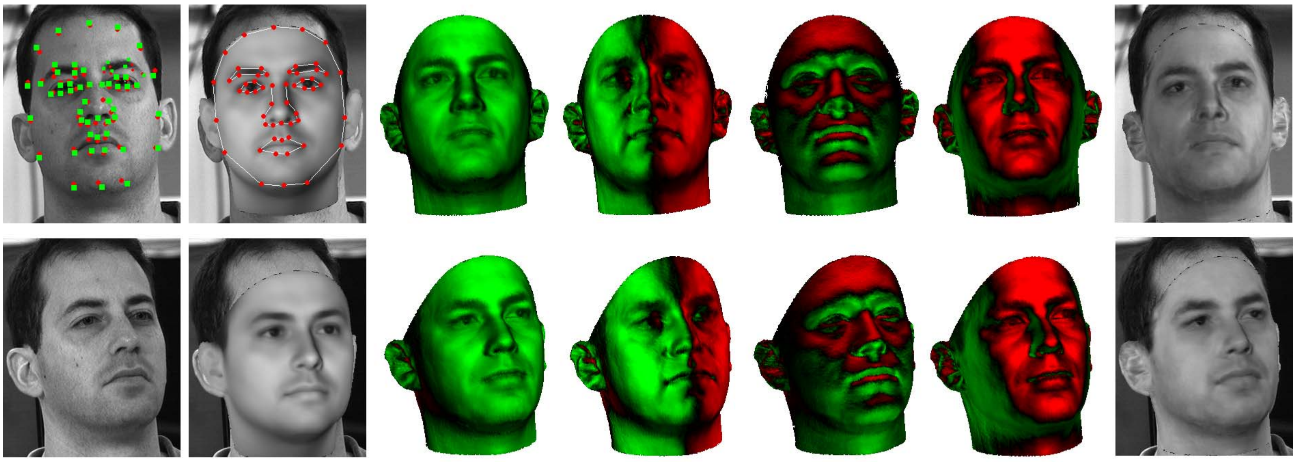

Figure 1. Fitting SHBMM to images: In each row, the first image is the input image followed by initial fitting and recovered spherical harmonic basis. The last image is the rendered image using the recovered parameters. In the first row, red points are hand-picked major features, green points are the corresponding features and the points lying in the white line are secondary features.

of feature points in the model that was rendered by recovered parameters f, ϕ, γ and θ. Thus, the vector of shape parameters α can be updated by solving a linear system of equations:

$$\delta s_f^{2d} = fPRS_f\delta\alpha \tag{5}$$

Please refer to [6][23] for the details of forward and inverse face shape rendering and shape parameters recovery.

3.2. Estimating Spherical Harmonic Basis Parameters and Illumination Coefficients

According to Eq. 2 and 3, a realistic face image can be generated by:

$$I = (\bar{b} + B\beta)\ell \tag{6}$$

where $\bar{b} + B\beta$ is the spherical harmonic basis component of the SHBMM and ℓ is the vector of illumination coefficients. Given an input image I_{input} of a face, the spherical harmonic basis parameters β and the illumination coefficients ℓ can be estimated by minimizing the difference between the input image and the rendered image from SHBMM:

$$min_{\beta,\ell}\|(\bar{b} + B\beta)\ell - I_{input}\|^2 \tag{7}$$

We solve Eq. 7 iteratively as follows:

1) Step 0: initialize the spherical harmonic basis parameter β_0 as 0 and define $b_0 = \bar{b} + B\beta_0 = \bar{b}$. Set step index $i = 1$.

2) For each step i, estimate the illumination coefficients ℓ_i by solving a linear equation: $b_{i-1}\ell_i = I_{input}$.

3) The image error δI_i of step i is the difference between input image I_{input} and the rendered image I_i in step i where $I_i = b_{i-1}\ell_i$. The spherical harmonic basis parameters can be updated: $\beta_i = \beta_{i-1} + \delta\beta$ where $\delta\beta$ is computed by solving the linear equation: $\delta I_i = B\delta\beta\ell_i$

4) Update the new spherical harmonic basis $b_i = \bar{b} + B\beta_i$ and increase step index i by 1.

5) Perform steps 2) to 4) iteratively until $\|\delta I\| < \xi_I$ or $\|\delta\beta\| < \xi_\beta$ where ξ_I and ξ_β are pre-selected constants.

The core of the above described process is the minimization of the image error as shown in Eq. 7 where two variables β, ℓ need to be recovered iteratively. Thus, there are two methods to start the iteration: initialize β as 0 and compute ℓ afterwards as we described above, or start with a random ℓ to achieve the global minimum. We chose the first method since our experiments on synthetic data showed that, the illumination coefficients ℓ computed by using the mean spherical harmonic basis $\bar{b}$ were close to the actual values, which made the whole recovery fast and accurate. Another important point is that, the spherical harmonic basis cannot capture specularities and cast shadows. Thus, for better recovery results, we employed two thresholds to avoid using the image errors caused by specularities and cast shadows. Figure 1 shows the fitting process and results.

4. SHBMM for Synthesis and Recognition

In this section, we will demonstrate how to apply our Spherical Harmonic Basis Morphable Model to face synthesis and recognition. Section 4.1 will explain how to combine our SHBMM and Ratio image technique for photo-realistic face synthesis. In Section 4.2, we will propose two face recognition methods based on the recovered SHBMM parameters and de-lit images respectively.

4.1. Face Synthesis

The face synthesis problem we will discuss can be stated as following: given one single image under unknown lighting, can we remove the effects of illumination from the image (face "de-lighting") and generate images of the object consistent with the illumination conditions of the target images(face "re-lighting")? The input image and target images can be acquired under different unknown lighting conditions and poses. In the following, I_s represents the input image, I_t represents a target image, I_d represents the de-lit image from I_{input} and I_r represents the re-lit image.

In the previous section, we demonstrated how to recover the set of SHBMM parameters $\{\alpha_s, \beta_s, \ell_s\}$ from an input face I_s. Inversely, we can render a face $I_s^{'}$ using the recovered parameters to approximate I_s: $I_s^{'} = (\bar{b}+B\beta_s)\ell_s$. Thus, the face texture (de-lit face) can be directly computed from the estimated spherical harmonic basis according to Eq. 1. Hence, face re-lighting can be performed by setting different values to the illumination parameters ℓ similar to [2]. However, in this method, a face was represented by a linear combination of a set of 3D face exemplars which results to possible loss of detail for the specific face. Alternatively, ignoring cast shadows and specularities, we notice that:

$$\frac{I_s}{I_d} = \frac{H(n_t)\lambda_t \ell}{\lambda_t} \approx \frac{H(n_e)\lambda_e \ell}{\lambda_e} = \frac{I_s^{'}}{\lambda_e} \tag{8}$$

where $H(n)\ell$ is the spherical harmonic basis, n_t and n_e are the actual and estimated surface normals and λ_t and λ_e are the actual and estimated face textures.

Eq. 8 states that the intensity ratio of the input image to the de-lit image should be approximately equal to that of the rendered face and the corresponding face texture (albedo). The face texture (albedo) of the rendered face can be easily computed from Eq. 1 as: $\lambda = \sqrt{4\pi} b_{00}$.

Rewriting Eq. 8, an input image can be de-lit as:

$$I_d = \frac{I_s \times \sqrt{4\pi} b_{00}}{(\bar{b} + B\beta_s)\ell_s} \tag{9}$$

Given the two images I_s, I_t with recovered parameters $\alpha_s, \beta_s, \ell_s$ and $\alpha_t, \beta_t, \ell_t$ respectively, we have:

$$\frac{I_s}{I_d} = \frac{(\bar{b} + B\beta_s)\ell_s}{\sqrt{4\pi} b_{00}^s} \tag{10}$$

$$\frac{I_r}{I_d} = \frac{(\bar{b} + B\beta_s)\ell_t}{\sqrt{4\pi} b_{00}^s} \tag{11}$$

Combining Equation 10 and 11, the re-lit image can be computed:

$$I_r = \frac{(\bar{b} + B\beta_s)\ell_t \times I_s}{(\bar{b} + B\beta_s)\ell_s} \tag{12}$$

4.1.1 Cast Shadows

In the above described face de-lighting and re-lighting steps, we initially ignored cast shadows and specularities. Previous work [2] has shown that a human face can be approximately treated as a Lambertian surface. However, if the input image has any cast shadows, the de-lit and re-lit images will be unrealistic, see Figure 4. The image difference (image error E) between the input image I_{input} and the re-rendered image by the SHBMM provides a simple solution to the cast shadow problem.

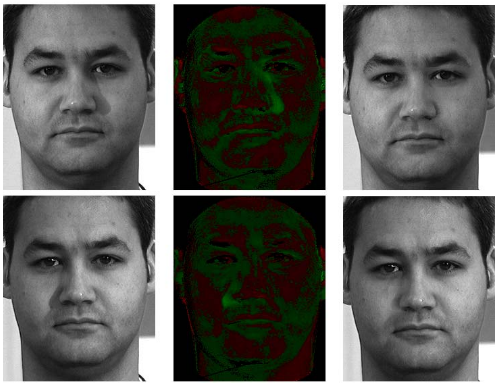

Figure 2. The images in the left column are input images and those in the middle are the error images where the cast shadow errors are significant thus easy to be removed. The right column shows the images after shadow removal. In error images, green and red represent negative and positive errors respectively.

Figure 3. The images in the first row are the input images under different illumination conditions. Second row shows the corresponding de-lit images. The rightmost column shows a failure example where the input image is saturated.

According to Eq. 7, by mapping the input image to our SHBMM, an error image E can be computed: $E = I_s - (\bar{b} + B\beta_s)\ell_s$. As described in Section 3.2, to achieve better parameter recovery, the errors caused by cast shadows were filtered out and thus not being minimized. Hence, such errors should contribute significantly in the error image E and could be easily identified and removed. As shown in Figure 2, our experiments verify this hypothesis. With E_{cast} representing the error caused by cast shadows, the input face I_s can be de-lit as:

$$I_d = \frac{(I_s - E_{cast}) \times \sqrt{4\pi} b_{00}}{(\bar{b} + B\beta_s)\ell_s} \tag{13}$$

Figure 5. Face re-lighting results: the images in the first row are the target images and those in the first column are input images. Images with remarkable quality are synthesized even if only one input image is available.

The error image is computed in 3D space, thus in the relighting step, the shadow errors will be added to the appropriate voxels. By assuming human faces have similar cast shadows under the same illumination conditions, we can relighting a face by:

$$I_r = \frac{(\bar{b} + B\beta_s)\ell_t \times (I_s - E^s_{cast})}{(\bar{b} + B\beta_s)\ell_s} + E^t_{cast} \qquad (14)$$

where E^s_{cast}, E^t_{cast} represent the shadow errors of face I_s, I_t respectively,

The third column of Figure 5 exhibits an example of Eq. 14 where the cast shadows in the target image are added to the re-lit images.

Given an input image, a set of SHBMM parameters $\{\alpha, \beta, \ell\}$ can be recovered and a de-lit image can be computed. Based on this, two illumination invariant methods for face recognition from a single training image are proposed in the following section.

4.2. Face Recognition

Recognition based on SHBMM parameters: For each image I_i in the training set and a testing image I_t, we recover SHBMM parameters $\{\alpha_i, \beta_i, \ell_i\}$ and $\{\alpha_t, \beta_t, \ell_t\}$. Since the identity of a face is represented by $\{\alpha, \beta\}$, we recognize the face of subject i whose recovered parameters $\{\alpha_i, \beta_i\}$ are the closest to $\{\alpha_t, \beta_t\}$. In this method, the training image and the testing image can be acquired under different arbitrary illumination conditions and poses. In our implementation, the shape recovery was performed in a semi-automatic way, thus for images of one face under the same pose, the shape parameters recovered were almost the same. To avoid using this subject identity information captured from human interaction and to examine our recognition method unbiasedly, we performed experiments by just using the spherical harmonic basis parameters β. In a complete application, shape would be recovered automatically [6][9], so both shape and texture parameters would be used for recognition.

Recognition based on de-lit images: For each image I_i in the training set and a testing image I_t, we compute de-lit images I^i_d and I^t_d. We recognize the face whose de-lit image is closest to that of the testing image. In this method, images should be aligned and taken under same view-point.

5. Experiments and Results

We used the CMU-PIE data set [19] which provides images of both pose and illumination variations for face synthesis and recognition experiments. The CMU-PIE database contains 68 individuals, none of which is also in the USF set used to compute the Spherical Harmonic Basis Morphable Model.

5.1. Face Synthesis Experiments

The results of the first face synthesis experiment are shown in Fig. 3. where the four images in the first row are the input images under different unknown illuminations and

	Frontal			Side			Profile		
	Frontal	Side	Profile	Frontal	Side	Profile	Frontal	Side	Profile
based on SHB parameters	96.5%	94.58%	78.75%	93.91%	96.75%	78.58%	81.83%	81.5%	90.58%
based on De-lit image	99.25%	-	-	-	99.08%	-	-	-	98.75%

Table 1. Recognition results of our methods: the first row reports the recognition results based on the Spherical Harmonic Basis parameters β and the second row reports the results of using de-lit images.

Figure 4. The first two images in the top row are the input images to be de-lit and re-lit as the illumination condition of the third image in the top row (frontal lighting). The left two columns are de-lit images and the right two columns are re-lit images. The images shown in the middle row are computed without any cast shadows processing while the images shown in the last row are computed with cast shadows processing.

the images in the second row shows the corresponding de-lit images. We found that the de-lit images we computed exhibit much greater invariance to illumination effects. Quantitatively, for one subject, we computed the de-lit images of 40 images under different illumination conditions. The variance of the 40 de-lit images was 6.34 intensity levels per pixel while the variance of the original images was 24.55. In the rightmost column of the Fig. 3., we show a de-lighting example where the input image is taken under extreme illumination condition and part of the face is saturated.

Fig. 4 shows examples of face de-lighting and re-lighting when the input images have cast shadows. In the first row, the first two images are input images to be de-lit and re-lit while the third image in the first row is the target image of re-lighting (used as groundtruth). By removing the cast shadow E_{cast}, we generate more photo-realistic images.

Fig. 5 shows a series of face re-lighting experiments where the input images in the first column are re-lit "driven" by the target images in the first row. Images with remarkable quality can be synthesized even if only one single input image under arbitrary unknown lighting is available (Please refer to the accompanying video for higher resolution images). Fig. 6 shows re-lit images of the same input image"driven" by target images of different subjects under similar illumination conditions. We see that those re-lit images are rather similar given that the target images only have approximately same lighting condition. This suggests that our method extracts illumination information consistently across different subjects.

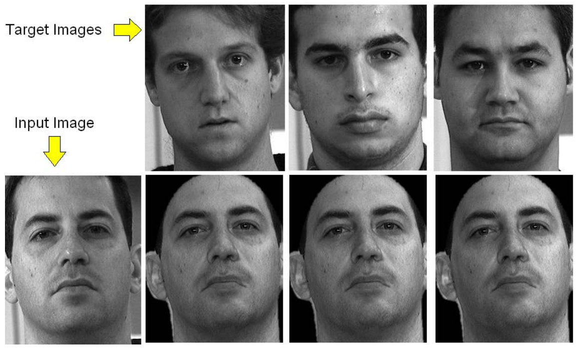

Figure 6. Comparison of re-lit images from the same input image "driven" by target images of different subjects under similar illumination conditions. The illumination information is preserved to a large extent, across different subjects.

5.2. Recognition Experiments

In our face recognition experiments, we used an image set of 3600 images which contains 30 subjects, 3 poses for each subject and 40 different illuminations for each pose. In the CMU-PIE database [19], for each subject, there are 22 images under single directional illuminant with an ambient light. To study the performance of our method on images taken under multiple directional illumination sources, we synthesized images by combining face images under different illumination conditions. For each subject, we randomly selected 2-4 images from the training data set and combined them together with random weights to simulate face images under multiple directional illumination sources(18 images per subject). Thus, for each subject, we have a total 40 images under different illuminations per pose. We didn't do the experiments on all 68 subjects in CMU-PIE data set due to time limitations yet the experimental results on our data set demonstrate the performance of our recognition methods sufficiently well.

Recognition methods proposed in Section 4.2 were tested on the set of 3600 images and the recognition results are reported in Table 1. From the experimental results, we found that we achieved high recognition rates for images under a wide range of illumination conditions, including multiple sources of illumination. As described in Section 4.2, in this paper, we used only the Spherical Harmonic Basis parameters for recognition. We expect to achieve better recognition results by combining our SHBMM with an automatic shape recovery method[6][9] and using both SHB parameters β and shape parameters α for recognition.

6. Conclusions and Future Work

We have shown that by using a Spherical Harmonic Basis Morphable Model, any image can be represented by three low dimensional vectors (SHBMM parameters). We also demonstrated that we can recover the set of SHBMM parameters given a single image of the subject under arbitrary unknown lighting. We provided a face synthesis technique by combining our SHBMM and the Ratio image technique. We showed that, with our model, given an input image, we could remove the illumination effects from the image (de-lighting) and synthesize new images (re-lighting) "driven" by a second image. Furthermore, by detecting the shadow errors from the image difference between the input image and the rendered image, we remove the cast shadows from the input image and add cast shadows to the re-lit image to generate more photo-realistic images. We also proposed two illumination invariant face recognition methods which achieve high recognition rates for images under a wide range of illumination conditions.

To add cast shadows to the re-lit face, we made an assumption that the input face and target face had similar cast shadows under the same illumination which in some cases may not be true. Thus, one of our future research directions is to relate the cast shadows with spherical harmonic basis parameters for better synthesis results. In our recognition experiments, we tested both images under single- and synthesized multiple- directional illuminations. At this time, there exist relatively few publicly available sets of images of faces under arbitrary illumination conditions, so we plan to continue validation of our method with databases with more types of light sources, e.g. area sources, when they become available.

References

[1] Y. Adini, Y. Moses, and S. Ullman. Face recognition: The problem of compensating for changes in illumination direction. *PAMI*, 19(7):721–732, July 1997.

[2] R. Basri and D.W. Jacobs. Lambertian reflectance and linear subspaces. *PAMI*, 25(2):218–233, February 2003.

[3] P.N. Belhumeur and D.J. Kriegman. What is the set of images of an object under all possible illumination conditions. *IJCV*, 28(3):245–260, July 1998.

[4] V. Blanz, K. Scherbaum, T. Vetter, and H. Seidel. Exchanging faces in images. In *EuroGraphics*, 2004.

[5] V. Blanz and T. Vetter. A morphable model for the synthesis of 3d-faces. In *SIGGRAPH*, 1999.

[6] V. Blanz and T. Vetter. Face recognition based on fitting a 3d morphable model. *PAMI*, 25(9):1063–1074, Sept. 2003.

[7] P.E. Debevec. Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. In *SIGGRAPH*, pages 189–198, 1998.

[8] P.E. Debevec, T. Hawkins, C. Tchou, H Duiker, W. Sarokin, and M. Sagar. Acquiring the reflectance field of a human face. In *SIGGRAPH*, pages 145–156, 1998.

[9] M. Dimitrijevic, S. Ilic, and P Fua. Accurate face models from uncalibrated and ill-lit video sequences. In *CVPR04*.

[10] A.S. Georghiades, P.N. Belhumeur, and D.J. Kriegman. Illumination-based image synthesis: Creating novel images of human faces under differing pose and lighting. In *IEEE Workshop on Multi-View Modeling and Analysis of Visual Scenes*, pages 47–54, 1999.

[11] 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):643–660, June 2001.

[12] P.W. Hallinan. A low-dimensional representation of human faces for arbitrary lighting conditions. *CVPR*, 1994.

[13] K.C. Lee, J. Ho, and D.J. Kriegman. Nine points of light: Acquiring subspaces for face recognition under variable lighting. In *CVPR01*, pages I:519–526, 2001.

[14] S.R. Marschner and D.P. Greenberg. Inverse lighting for photography. In *IST/SID Fifth Color Imaging Conference*, 1997.

[15] R. Ramamoorthi. Analytic pca construction for theoretical analysis of lighting variability in images of a lambertian object. *PAMI*, 24(10), Oct. 2002.

[16] R. Ramamoorthi and P Hanrahan. An efficient representation for irradiance environment maps. In *SIGGRAPH*, 2001.

[17] E. Sali and S. Ullman. Recognizing novel 3-d objects under new illumination and viewing position using a small number of examples. In *ICCV98*, pages 153–161, 1998.

[18] A. Shashua. On photometric issues in 3d visual recognition from a single 2d image. *IJCV*, pages 21:99–122, 1999.

[19] T. Sim, S. Baker, and M. Bsat. The cmu pose, illumination, and expression database. *PAMI*, pages 1615–1618, 2003.

[20] T. Sim and T. Kanade. Combining models and exemplars for face recognition: An illuminating example. In *Proc. CVPR Workshop on Models versus Exemplars in Comp. Vision*, 2001.

[21] Riklin-Raviv T. and Shashua A. The quotient image: Class based re-rendering and recognition with varying illuminations. In *CVPR*, pages 566–571, 1999.

[22] Z. Wen, Z. Liu, and T. S. Huang. Face relighting with radiance environment maps. In *CVPR*, 2003.

[23] L. Zhang and D. Samaras. Face recognition under variable lighting using harmonic image exemplars. In *CVPR*, volume I, pages 19–25, 2003.

[24] L. Zhang and D. Samaras. Pose invariant face recognition under arbitrary unknown lighting using spherical harmonics. In *Biometric Authentication*, pages 10–23, 2004.

[25] W.Y. Zhao and R. Chellappa. Illumination-insensitive face recognition using symmetric shape-from-shading. In *CVPR00*, pages I: 286–293, 2000.

[26] S.K. Zhou, R. Chellappa, and D.W. Jacobs. Characterization of human faces under illumination variations using rank, integrability, and symmetry constraints. In *ECCV*, 2004.