

# A Subspace Model-Based Approach to Face Relighting Under Unknown Lighting and Poses

Hyunjung Shim, *Student Member, IEEE*, Jiebo Luo, *Senior Member, IEEE*, and Tsuhan Chen, *Fellow, IEEE*

**Abstract**—We present a new approach to face relighting by jointly estimating the pose, reflectance functions, and lighting from as few as one image of a face. Upon such estimation, we can synthesize the face image under any prescribed new lighting condition. In contrast to commonly used face shape models or shape-dependent models, we neither recover nor assume the 3-D face shape during the estimation process. Instead, we train a pose- and pixel-dependent subspace model of the reflectance function using a face database that contains samples of pose and illumination for a large number of individuals (e.g., the CMU PIE database and the Yale database). Using this subspace model, we can estimate the pose, the reflectance functions, and the lighting condition of any given face image. Our approach lends itself to practical applications thanks to many desirable properties, including the preservation of the non-Lambertian skin reflectance properties and facial hair, as well as reproduction of various shadows on the face. Extensive experiments show that, compared to recent representative face relighting techniques, our method successfully produces better results, in terms of subjective and objective quality, without reconstructing a 3-D shape.

## I. INTRODUCTION

**F**ACE relighting has received increasing attention from image processing, computer vision, and graphics communities due to its various applications such as face recognition, video conferencing, virtual reality, movie special effects, and digital face cloning.

In this paper, we propose a new face relighting approach that synthesizes a face image under any lighting condition given as few as one input image of the same face. The input image can be the image(s) of the same face under the unknown lighting condition(s) in an unknown pose. For realistic face synthesis, the face relighting algorithm needs to hold the following three properties: retaining facial features, reproducing various skin reflectance properties, and preserving the identity of a subject. Although a large number of face relighting approaches have been developed previously, generating a plausible face image is still a challenging task due to the difficulty of modeling the various features of faces. In fact, most of existing face relighting work [1]–[18] aims to improve the face recognition rate instead of the quality of the synthesized face image. In face recognition, the third property, preserving the identity of a subject, is the

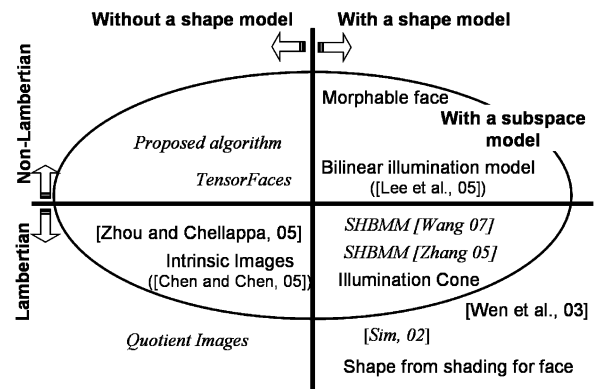


Fig. 1. Related work for face relighting. *Italic*: Experimental comparisons are available in this paper.

primary concern while the other two properties are not critical. Unlike most of the prior work in face relighting, we do not limit the application of the proposed algorithm to the face recognition task. Instead, we develop the algorithm that possesses three desirable properties mentioned earlier such that it can be useful to many practical applications. In face relighting, one must first factorize a single image or multiple images into the lighting and reflectance components. In recent years, many algorithms have been proposed to solve this factorization problem: Morphable faces [20], [21], Quotient Image [22], Illumination Cone [27], [2], [1], Spherical Harmonic Subspace [12], [19], [28], [9], [11], [10], [25], Intrinsic Images [32], [31], [8], [23], and TensorFaces [30], [13], [14]. As shown in Fig. 1, these face relighting methods can be categorized based on whether a (3-D) shape model, the Lambertian assumption, or a subspace model is used or not. With the same criteria, our proposed algorithm is considered as a face relighting method *without a shape model*, *accounting for the non-Lambertian reflectance*, and *using the subspace model* (more precisely the subspace model of the reflectance function). In the following subsections, we review existing approaches in face relighting and highlight the advantages of our approach compared to closely related work.

### A. Shape Models

Morphable face [20] proposed by Blanz and Vetter is known as one of the most powerful shape models for a human face. They constructed a generic subspace model for shape and texture separately from the 3-D face database using principal component analysis (PCA). Then, they used the shape and texture subspace model to constrain the shape and texture estimation, respectively. To recover the reflectance properties, they used the Phong reflectance model. As a result, they rendered realistic face

Manuscript received September 22, 2007; revised March 29, 2008. First published June 24, 2008; last published July 11, 2008 (projected). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Hassan Foroosh.

The authors are with Carnegie Mellon University, Pittsburgh, PA 15213 USA (e-mail: hjs@andrew.cmu.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIP.2008.925390

images under any lighting condition at any pose. However, Morphable model did not handle the facial hair in the input face because both their reflectance model, the Phong model, and their shape model did not account for the facial hair.

Wen *et al.* [34] proposed an efficient face relighting algorithm by using a shape model and a ratio-image-based technique to render the Lambertian faces. Zhao and Chellappa [26] recovered the shape of the face using the shape from shading (SFS) algorithm from a single input image. They imposed an intuitive constraint, a symmetric face for the frontal pose, to improve the shape reconstruction. Smith and Hancock [24] developed a statistical model for the normal map of faces under the SFS framework. Sim [3] used a Lambertian reflectance model to deal with the illumination variation and also included an additional model to encapsulate any residue that was not absorbed by the Lambertian assumption. They trained the statistical model of the albedo-normal map as well as the residue, and used it to re-illuminate the face under a new lighting condition. All these methods, however, did not reproduce the different types of surface reflectance properties and the facial hair in the synthesized face image because they inherently assumed the Lambertian surface for faces.

### B. Quotient Image

Shashua and Riklin-Raviv [22] proposed the Quotient Image technique to extract an illumination invariant signature (e.g., an albedo map) from a single image with the Lambertian assumption. They assumed that the objects in the same class had the same shape but differed only in their albedo functions. Given two objects under the same class, they computed the Quotient Image, which was the ratio of their albedo functions, and used it for face relighting. Due to the underlying assumptions, the same shape and the Lambertian reflectance, their relighting results did not appear convincing.

### C. Subspace Models

The subspace models for the illumination variation have been discussed by many others [27], [12], [11], [19], [28], [15], [5], [25], [30], [8], [4], [14], [23]. Belhumeur and Kriegman [27] have shown that the set of images of an object in the fixed pose, but under all possible lighting directions, form a convex cone, termed the Illumination Cone, in the image space. They have shown that the Illumination Cone of a convex Lambertian object can be constructed from as few as three images, which is analogous to the discussion by Shashua [5]. Basri and Jacobs [19] and Ramamoorthi and Hanrahan [28] independently developed the spherical harmonic representation for the illumination subspace. They have shown that the set of images of a convex Lambertian object under all possible lighting directions can be approximated by a 9-D linear subspace. Later, Zhang *et al.* [11] used the Spherical Harmonic Subspace Model combined with a Morphable Model, referred to as a Spherical Harmonic Basis Morphable Model (SHBMM), to estimate both the shape and the illumination variation of the face. More recently, Wang *et al.* [12] extended the SHBMM from Zhang *et al.* by modeling the statistical distribution and spatial coherence of face texture. As a result, Wang *et al.* reported a better face recognition rate than

Zhang *et al.* under harsh lighting conditions. Zhou and Chellappa [4] proposed to use a linear subspace model for both the illumination and the pose variation of faces under the Lambertian assumption. Chen and Chen [8] used an illumination subspace model to compute the Intrinsic Images [31] (e.g., the reflectance image (albedo) under the Lambertian assumption and the illumination image) from a face image. However, with all of these approaches, the subspace dimension can grow to infinity if both the shadows and the non-Lambertian reflectance are considered. Most faces are not Lambertian and often contain shadows.

TensorFaces [30] provided a multidimensional subspace model to handle various factors in the face image simultaneously (e.g., the illumination, pose, and expression) using a multilinear singular value decomposition (SVD) [33]. In the task of face relighting, there are two factors, illumination and identity, to form a 3-mode tensor. Then, one can derive a bilinear subspace model for the face illumination without the Lambertian assumption and the shape model. Recently, Lee *et al.* [14] proposed a new bilinear subspace model for the illumination and the shape and showed that their model was more efficient to represent the illumination variation on the face images than the Spherical Harmonic Subspace model. While TensorFaces analyzed the most representative multilinear subspace of face images, Yan *et al.* [18] used the most discriminative multilinear subspace of face images for face recognition. Their subspace model can be useful for face recognition but not for face synthesis.

### D. Related Work

A light stage was introduced by Debevec *et al.* [35] to measure the reflectance of a static face. Later, Wegner *et al.* [36] used a high-speed camera and light sources to acquire reflectance of a dynamic face subject. More recently, Peers *et al.* [37] proposed to transfer the lighting of a target face into an input face by warping the ratio image of the target face according to the input face. All these techniques are capable to handle the important facial features including the facial hairs and the non-Lambertian reflectance properties. Furthermore, their relighting performance in general is superior to any other technique mentioned in this paper, including the proposed algorithm. However, their approaches require to capture the input face under the controlled environment. Therefore, the applications of these approaches are limited by this requirement. Note that all other previous work introduced in this paper including our method handles the input face under the unknown lighting condition as well as the uncontrolled environment.

### E. Proposed Algorithm

We construct a subspace model of the reflectance functions from an existing face database (either the PIE database [29] or the Yale database [1]). Note that we derive a subspace model of the reflectance function for *each pixel* and *each pose*. Based on our subspace models, we estimate three unknown factors in the input face image(s), the lighting, the pose, and the reflectance functions, in a sequential manner. Our work is different from all other subspace model-based approaches in the following senses.

*First, we derive the subspace model of the reflectance function (including the non-Lambertian reflectance and shadows)*

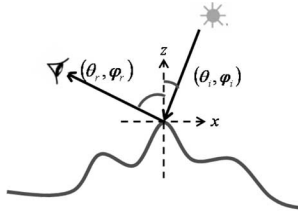


Fig. 2. Illustration of a bidirectional reflectance distribution function (BRDF).

as opposed to the appearance (e.g., the pixel values). Although we do not handle the shape of the face explicitly, our subspace model implicitly incorporates the structural information of general faces as it reflects the illumination caused by the structure of the environment. Since we obtain sampled reflectance functions directly from the database, our model accounts for both *attached shadow*, which is caused by the lighting facing opposite the surface normal [25], and *cast shadow*, which is caused by self-occlusion (e.g., the shadow along the nose). Additionally, our subspace model learns the various reflectance properties of facial features, most notably the facial hair, from the training data, so that we can successfully recover the reflectance of the facial hair as well. Note that virtually all existing face relighting approaches we surveyed, regardless of their models, are incapable to estimate the reflectance properties of the facial hair in the synthesized image since the facial hair does not obey either their reflectance model (e.g., a Lambertian model or a Phong model) or their shape model (e.g., a Morphable model) very well.

Second, we construct a robust subspace model not for an entire face image but for each pixel. Due to this key difference (an image versus a pixel), we have more model parameters and, thus, greater modeling power than all other subspace models. As a result, the proposed subspace model, which uses the basis functions of the reflectance function at each pixel, potentially fits a given face better than other subspace models that use basis images (Illumination Cone, Spherical Harmonics Subspace, and TensorFaces). In fact, it is how we avoid the subspace dimension to grow prohibitively large. While we derive a different subspace model of the reflectance function for each pixel, we take the reflectance functions from the neighboring pixels into account for model learning. This training scheme results that the subspace models in the pixel neighborhood vary smoothly. Indeed, this feature plays an important role in significantly reducing the impact of facial registration error.

As a result, the proposed algorithm is attractive thanks to the following novel features and advantages.

- It can reproduce plausible appearance and fine image details without resorting to a sophisticated shape and texture models.
- It can reproduce different types of skin reflectance properties (e.g., skin specularities and facial hair).
- It is robust to the face registration error.
- It is a flexible model that works well with the population outside of training set.

The proposed approach consists of three stages: model learning, inference, and synthesis. In model learning (Section II), we collect the reflectance functions from a face data-

base and derive a subspace model of the reflectance function for each pixel at each pose. In this paper, we assume the rough alignment of both the training images and the input image. The subspace model is computed just once and used repeatedly for any novel face image. With inference (Section III), we recover the lighting, the pose, and the reflectance functions from input image(s). We use an iterative expectation-maximization (EM) approach to solve for the lighting and the pose, and reflectance functions.

Finally, we synthesize the face under any new lighting condition using the estimated reflectance functions (Section V). Experimental results are presented in Section VI. We will first present some synthesized results to show that we can successfully handle various facial features (Section VI-B). Then, we will compare synthesized face images from the proposed algorithm with the ones from four recent representative face relighting algorithms: Sim's method [3] and TensorFaces [30] (Section VI-C), the quotient image [22] (Section VI-D), and a Spherical Harmonic Basis Morphable Model (SHBMM) [11] (Section VI-E). These algorithms are selected to represent the universe of the face relighting algorithms shown in Fig. 1. In Section VI-F, we present the subjective comparison results between the proposed algorithm and each of the above four face relighting algorithms. Finally, in Section VII, we conclude that the proposed algorithm is robust and effective for the realistic face synthesis, and, therefore, it is useful to many practical applications involving face relighting.

## II. MODEL LEARNING

A bidirectional reflectance distribution function (BRDF),  $F(\theta_i, \phi_i; \theta_r, \phi_r)$ , is defined as the ratio of reflected radiance exiting to a particular view direction  $(\theta_r, \phi_r)$  over the irradiance for a particular incident directional light source at  $(\theta_i, \phi_i)$  (see Fig. 2). By fixing a view direction, the BRDF becomes a 2-D function,  $F(\theta_i, \phi_i; \theta_r = \theta_0, \phi_r = \phi_0)$ , and we refer this 2-D function as to a reflectance function. We discretize  $(\theta_i, \phi_i)$  at  $(\theta_d, \phi_d)$  where  $d = 1, 2, \dots, D$  and concatenate  $D$  discrete samples of this reflectance function to form a column vector  $\mathbf{f}$ . Also, a column vector  $\mathbf{l}$  contains  $D$  discrete samples of  $L(\theta_i, \phi_i)$ , which represents the lighting. According to the definition of the reflectance function, a pixel value,  $i$ , is

$$\begin{aligned} i &= \sum_{d=1}^D F(\theta_d, \phi_d; \theta_r = \theta_0, \phi_r = \phi_0) L(\theta_d, \phi_d) \\ &= \mathbf{f}^T \mathbf{l}. \end{aligned} \quad (1)$$

In general, the lighting changes for every pixel on the object surface, and, therefore, we need a 4-D representation such as  $L(\theta_i, \phi_i; m, n)$  to fully describe spatially varying illumination over the surface (for simplicity, we will use  $n$  instead of  $(m, n)$  to indicate a pixel index). However, in this paper, we ignore such an effect and assume the directionally varying illumination only, same as all other face relighting methods in literature.

To construct the model, we first collect the reflectance functions from a roughly aligned (only according to the positions of eyes and nose) and cropped face database (PIE [29] or Yale [1]). Note that the same alignment scheme will be applied to the input face image. Then, we construct a 3-D tensor as the white

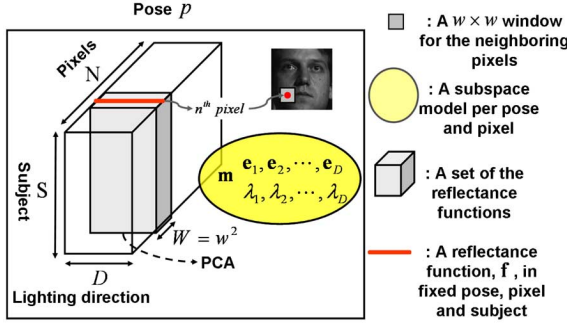


Fig. 3. Illustration of our subspace model for a particular pose.  $S$ : Number of subjects.  $D$ : Number of lighting directions.  $N$ : Number of pixels.  $P$ : Number of poses.  $W$ : Number of neighboring pixels. ( $= w^2$ ),  $w$ : Window size for the neighboring pixels.  $n$ : Index of a pixel.

box shown in Fig. 3. Each face image in the database is illuminated by a single directional light source. Then, we perform PCA on a set of reflectance functions (the gray box) in Fig. 3, where  $W$  is the number of pixels in a neighborhood window centered at a certain pixel location  $n$ , in order to derive the corresponding subspace model. More specifically, we compute the covariance matrix of  $SW$   $D$ -dimensional vectors and derive its mean vector and a set of eigenvectors. The choice of  $W$  will determine the range of the local neighborhood that is used to train the model. By introducing  $W$ , the subspace model at a particular pixel represents the reflectance function for not only the corresponding pixel but also its neighboring pixels. Because of  $W$ , the subspace model can tolerate the mismatch of the pixel registration within a  $w \times w$  window, which eventually can handle face registration error. As a result, we can avoid an elaborate pixel matching process between the model and the input image.

Using the PCA subspace model, the reflectance function can be approximated by the first few eigenvectors corresponding to the largest eigenvalues

$$\mathbf{f} = \sum_{j=1}^D c_j \mathbf{e}_j + \mathbf{m} \approx \sum_{j=1}^J c_j \mathbf{e}_j + \mathbf{m}, \quad (J \ll D). \quad (2)$$

The actual dimension of the tensor is dependent on the database. From the PIE face database [29], we have  $D = 21$ ,  $N = 10\,000$ ,  $S = 65$ , and  $P = 13$ . For the Yale extended B database [1], we have  $D = 64$ ,  $N = 10\,000$ ,  $S = 37$ , and  $P = 9$ . Note that, although the reflectance functions from each database might be sampled sparsely (the true reflectance function is composed of the reflectance values from all possible directional lights located on the hemisphere), we will show that plausible results for face synthesis can be obtained. Also, we use  $J = 4$  compared to  $D = 21$  for the subspace models from the PIE database and  $J = 11$  compared to  $D = 64$  for the models from the Yale database. With the selected  $J$ , we ensure that the first  $J$  eigenvectors capture more than 98% of the variation of the reflectance function in each database. Also, we use  $W = 49$  (pixels within a  $7 \times 7$  window), which is selected by empirical study.

### III. INFERENCE

In this section, we describe the algorithm for estimating the pose, and the reflectance functions of the input face and the lighting condition. We use the subspace model derived in

Section II to constrain the estimation and also use an iterative EM-like approach to sequentially estimate three factors. With the initial guess of the reflectance functions, we compute the pose and the lighting in Step 1. With the estimated lighting and pose from Step 1, we compute the reflectance function for each pixel in Step 2. This constitutes one iteration. **We repeat the iterations until the estimated reflectance function,  $\mathbf{f}$ , over all pixels converge.** From the experiment, we observe that, in general, these three factors converge within three iterations. The detail of each step is demonstrated in Sections III-A and III-B, respectively. Note that  $X$  is the total number of input images and  $x$  is the index of each input.

#### A. Step 1: Estimating the Lighting and the Pose

In the first step, we estimate the lighting  $\mathbf{l}$  for each image and the pose  $p$ . For each image  $x$ , the estimation procedure is as follows.

- 1) For a certain pose, we compute  $\mathbf{l}$  in (3) by setting  $\mathbf{f} = \mathbf{m}$  for initialization. Note that  $\mathbf{f}$  will be updated after the Step 2) in Section III-B

$$\hat{\mathbf{l}} = \arg \min_{\mathbf{l}} \sum_{n=1}^N \|i_n - \mathbf{f}_n^T \mathbf{l}\|^2 \quad \text{s.t. } \min(\mathbf{l}) \geq 0. \quad (3)$$

- 2) Repeat 1) for all possible poses and choose the pose corresponding to the smallest residue. Then, we use corresponding  $\mathbf{l}$  for the estimated lighting condition.

Note that the constraint in (3) implies that the brightness of the lighting should be positive.

#### B. Step 2: Estimating the Reflectance Functions

In the second step (step 2), given  $\mathbf{l}$ ,  $p$  and  $i$  at each pixel, we approximate  $\mathbf{f}$  by computing the objective function in (4) with the subspace constraint. We drop the pixel index  $n$  for the simplicity although  $\mathbf{f}$ ,  $\mathbf{m}$ ,  $i$  and  $\alpha$  depend on  $n$

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \sum_{x=1}^X \|i_x - \mathbf{f}^T \mathbf{l}_x\|^2 + \alpha d(\mathbf{f}, \mathbf{m}) \quad \text{s.t. } \mathbf{f} \in \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_J\}. \quad (4)$$

The first term in (4) finds the best estimate of  $\mathbf{f}$  that minimizes the distance between  $\mathbf{f}$  and a PCA subspace when  $i$  and estimated  $\mathbf{l}$  are given. The second term, a regularization term with  $\alpha$  as a weighting factor, penalizes the deviation of our estimate from the mean  $\mathbf{m}$ . The regularization term is defined by the Mahalanobis distance,  $d(\mathbf{f}, \mathbf{m})$  in (4). Letting  $\mathbf{f} = \sum_{j=1}^J c_j \mathbf{e}_j + \mathbf{m}$ , we rewrite the objective function in terms of  $\{\mathbf{e}_1, \dots, \mathbf{e}_J\}$ ,  $[\lambda_1, \dots, \lambda_J]$ , and  $[c_1, \dots, c_J]$ . That is

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_J \end{bmatrix} = \arg \min_{c_1, \dots, c_J} \sum_{x=1}^X \left\| i_x - \sum_{j=1}^J c_j \mathbf{e}_j^T \mathbf{l} - \mathbf{m}^T \mathbf{l} \right\|^2 + \alpha \sum_{j=1}^J \frac{c_j^2}{\lambda_j} \quad (5)$$

where  $\mathbf{e}_j$ ,  $\lambda_j$ , and  $c_j$  are the  $j$ th eigenvector, eigenvalue, and coefficient, respectively.

We introduce the weight,  $\alpha$ , to balance the two terms in (5). In practice, a proper  $\alpha$  at each pixel  $n$  can suppress the artifacts

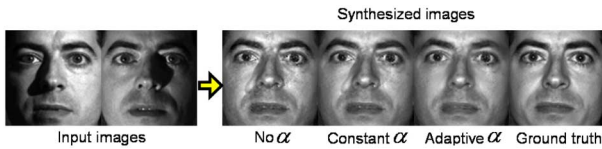


Fig. 4. Synthesized images using different weighting schemes.

caused by the cast shadow. We compose  $\alpha$  for all the pixels into a weighting map. The scheme to compute  $\alpha$  is described in Section IV.

#### IV. IMPLEMENTATION DETAILS

In this section, we first discuss the method to compute the weighting map. Before explaining how to compute the weighting map, we first look at the synthesized image without the regularization term (the first image among the synthesized images in Fig. 4). From this result, we observe that our subspace model does not handle pixels along cast shadow boundaries perfectly, which results in an erroneous estimate of  $\mathbf{f}$ . To control such artifacts, we assign larger weight to the regularization term in (5) wherever it is determined that the pixel belongs to the cast shadow boundaries in the input images. Compared to using no or uniform  $\alpha$  weighting, adaptive weighting reduces cast shadow artifacts, and, thus, it produces synthesized results (column 4) more consistent with the ground truth (column 5).

To detect cast shadow boundaries, we use a *pretrained mask*, which indicates the possible regions for cast shadow boundaries, and the edge map of each input image. As shown in Fig. 5, we find that the edge map of the input image contains cast shadow boundaries as well as intrinsic facial features (e.g., eyes, nose, and lips). Since we aim to detect cast shadow boundaries only, we need a mask that excludes intrinsic facial features and, at the same time, retains cast shadow boundaries. Such a mask can be computed by identifying outliers of our model. To train this mask, we estimate the reflectance functions from one input image in the training database *without the regularization term* and synthesize face images using such estimated reflectance functions. Then, we subtract these synthesized images from the ground truth images. If the error is large at a particular pixel, we consider the pixel in a region that does not follow the subspace model very well. We average the error and threshold them to generate the mask shown in Fig. 5. To compute the edge map, we apply the Canny edge detector to the input image. If multiple input images are available, we add multiple edge maps from all of the input images to generate the final edge map. Multiplying the mask with the final edge map retains only the cast shadow boundaries and excludes intrinsic facial features. Finally, we apply a Gaussian blurring kernel to ensure spatial smoothness in the weighting map. Note that the final weighting map is dependent on the input images (Fig. 5).

In Section III, we implicitly assume that the spatial resolution of the input image matches the resolution of images in the training database. If the resolution of the input image differs from the training images, we perform the nearest neighborhood re-sampling on the PCA subspace model to match the resolution of the PCA models to the input image.

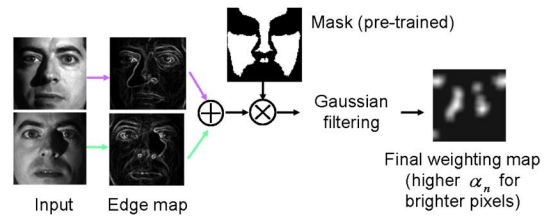


Fig. 5. Example weighting map (alpha) from the two input images in the frontal pose.

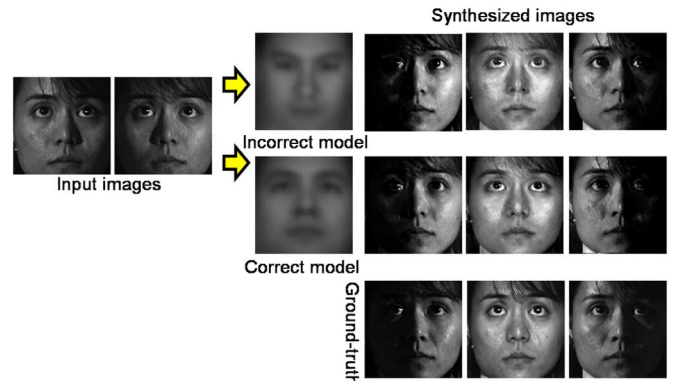


Fig. 6. Synthesized images using the same inputs but with different pose models. First row: Using the incorrect pose model. Second row: Using the correct pose model. Third row: Ground truth. Third column: Mean face from each pose model.

#### V. SYNTHESIS

In Sections III and IV, we discuss our method to estimate the pose, the reflectance functions of the face, and the lighting. For synthesizing the face image under a new lighting condition, we have two schemes.

If the new lighting condition,  $\mathbf{l}$ , is given by the user, we can use (1) to compute a pixel value under this new lighting condition. We call this process the synthesis scheme *A*.

Instead of providing a prescribed lighting condition  $\mathbf{l}$ , users can specify the lighting condition using another face image. That is, users can provide two face images, one for recovering its reflectance (an input image) and another for extracting its lighting condition (a target image). Then, we can re-illuminate the input image under the lighting present in the target image. We call this process the synthesis scheme *B*.

#### VI. EXPERIMENTAL RESULTS

In Sections VI-C–VI-E, we present the synthesized results compared to the results from existing approaches, while in Section VI-F, we are able to provide the subjective comparison results between our approach and others based on visual evaluation by human observers. Except for the results presented in Fig. 12, we use the synthesis scheme *A* to relight the faces. In Fig. 12, we apply the synthesis scheme *B*.

We used two face databases (PIE and Yale) and trained subspace models per the database in Section II. For all of the experiments, we ensure that none of the evaluation subjects is used in training the subspace model. In fact, in Sections VI-C and VI-D, the evaluation subject is either drawn from a different database or an unrelated natural image. In either case, we represent the



Fig. 7. Reproducing the various skin reflectance properties. Last four images in first and second rows: Synthesized faces. Third row: Ground truth. Each column: Same lighting condition.

lighting condition in the input image as a linear combination of  $D$  directional light sources that appeared in the training database. Even with the small mismatch in the lighting configuration between the input image and the trained model, we can still present good results in Sections VI-C and VI-D.

#### A. Pose and Lighting Estimation

In this section, we analyze the performance of the pose and the lighting estimation. We run the experiment 9009 times ( $13$  poses  $\times$   $33$  subjects  $\times$   $21$  lighting conditions) to obtain the results. Also, for the experiment, we use subspace models of the reflectance functions derived from the PIE database and obtain a single input image from the same database (again, not part of the training set).

In this experiment, we obtain 4.53 degrees for the root mean square error (RMSE) in the pose estimation. In fact, we notice that small errors in pose estimation are visually acceptable and do not cause significant problems for the face synthesis. Fig. 6 demonstrates the synthesized results when we deliberately use the model for an incorrect pose (the first row in Fig. 6), instead of the correct pose (the second row in Fig. 6). The first row in Fig. 6 shows that the synthesized images using the subspace model of a neighboring pose (typical for an incorrect pose estimation) are quite acceptable compared to both the ground truth (third row) and the best synthesized results (second row). Based on our experience, if the pose estimation error is below 15 degrees, the results are visually close to the best synthesized results. This property is particularly desirable in practice. From the same experiment, we compute the RMSE for the lighting estimation and the error is only 0.0092 when the intensity of the ground truth lighting is unity.

#### B. Face Synthesis With Various Facial Features

In this section, we compare the synthesized images with their ground truth images in Figs. 7 and 8, respectively. In each row, the first group of images is the input image(s) and the last group of four images is the synthesized images. Each column in the

second group corresponds to the same lighting condition and the image in the third row is the ground truth under the corresponding lighting condition.

Fig. 7 shows that we can recover the various skin reflectance properties such as the non-Lambertian specularities. From Fig. 8, we observe that the facial hair in the original input image is preserved well in the synthesized images. Additionally, the last two columns in Figs. 7 and 8 clearly show the effects of both the attached shadow and the cast shadow. These results justify that, without explicitly modeling the shape of the face, the proposed approach is able to successfully generate the 3-D effects in the synthesized images and handle various skin reflectance properties. As mentioned earlier in Section I, to the best of our knowledge, all existing face relighting methods including the most recent work [12] fail to synthesize the subjects with the facial hair since their models do not account for facial hair [38].

Moreover, with two input images in the second row, we can considerably improve the accuracy of the synthesized images. It is clear that the illumination on the forehead and cheek from the second row is more accurate than the one from the first row in Figs. 7 and 8.

#### C. Comparison With the Sim's Method and TensorFaces

In Figs. 9 and 10, we make a comparison with other work, TensorFaces [30] and a statistical face rendering technique proposed by Sim [3], both qualitatively and quantitatively. In this experiment, all three methods use the PIE database to construct the model as well as to select the input image.

For the TensorFaces, we used all eigenvectors from a 3-mode tensor SVD. Also, we used the ground truth lighting to synthesize the images shown in Figs. 9 and 10, which implies that it is the ideal and most favorable setting for the TensorFaces technique. In contrast, the lighting condition for the proposed algorithm and Sim's method is estimated with its own method.

From the synthesized results presented in Fig. 9, our synthesized face images (column 3) are more realistic and perceptually

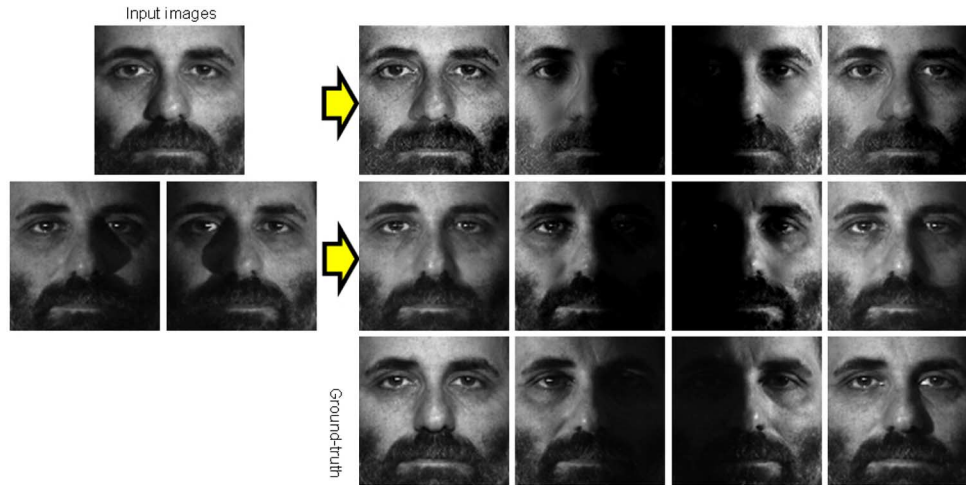


Fig. 8. Preserving the facial hair. Last four images in first and second rows: Synthesized faces. Third row: Ground truth. Each column: Same lighting condition.

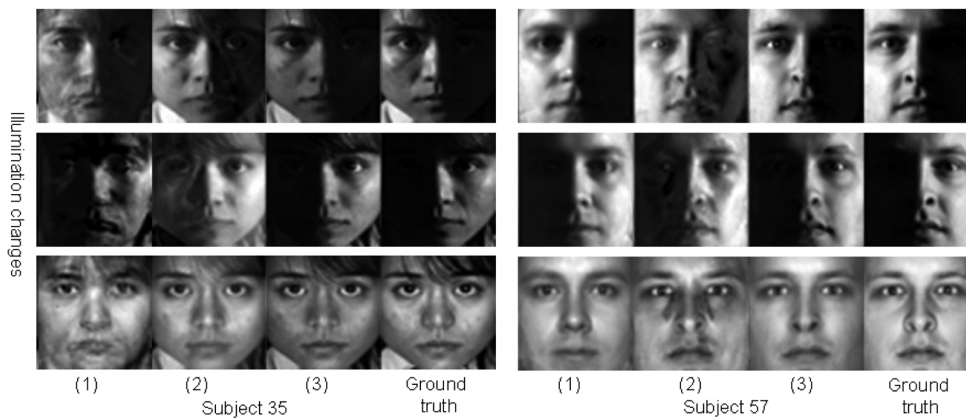


Fig. 9. Qualitative comparison with TensorFaces [30] and Sim's method [3]. (1) TensorFaces. (2) Sim's method. (3) Proposed algorithm and ground truth.

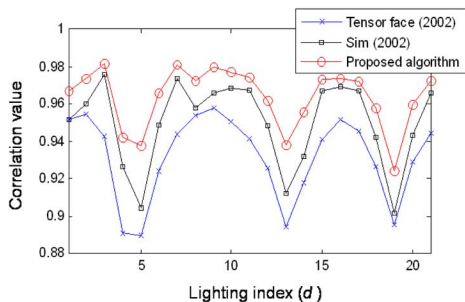


Fig. 10. Quantitative comparison with TensorFaces [30] and Sim's method [3] (correlation).

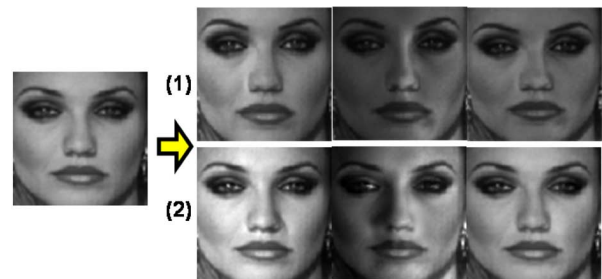


Fig. 11. Comparison with the quotient image [22]. (1) Shushua and Riklin-Raviv [22]. (2) Proposed approach.

ally closer to the ground truth images (column 4). In particular, the synthesized results for subject 57 show that we handle the cast shadow much better than Sim (column 2). The TensorFaces approach often fails to preserve the identity of the subject, i.e., the synthesized face looks like a different person, although it reproduces the illumination reasonably well (subject 57). For quantitative comparison, we compute the correlation value between the synthesized images and the ground truth images. Each curve in Fig. 10 shows the curve of average correlation over 65 subjects with a pair of input images are randomly

drawn ten times. We synthesized face images under 21 different lighting conditions to generate the curve. That means, a total of  $65 \times 10 \times 21 = 13650$  images have been synthesized. The synthesized images using our method are much closer to the ground truth in terms of the correlation. In summary, we outperform Sim [3] and the TensorFaces both qualitatively and quantitatively.

#### D. Comparison With Quotient Image

Fig. 11 compares the Quotient Image with the proposed algorithm. We apply our technique to a single input face image

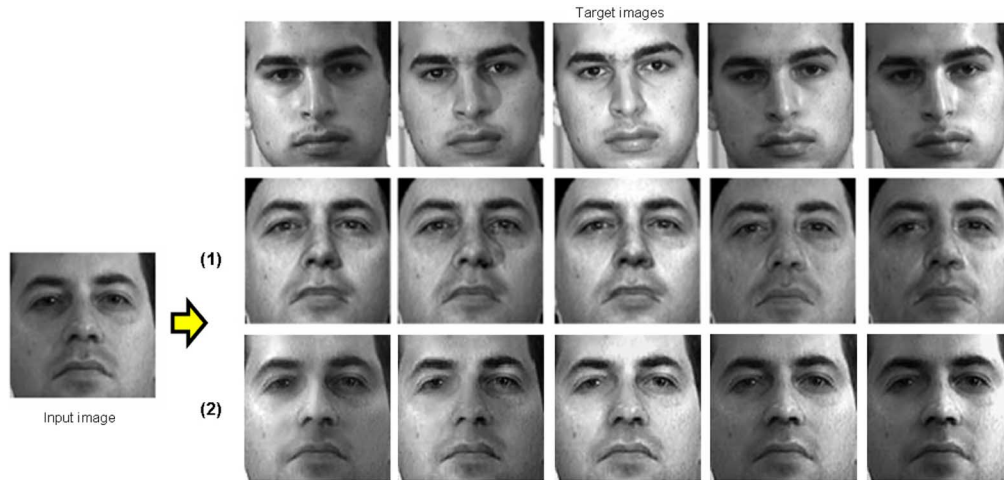


Fig. 12. Comparison with Zhang *et al.*. (1) SHBMM proposed by Zhang *et al.* [11]. (2) Proposed approach.

captured naturally as opposed to in studios. In this experiment, we use the subspace model from the Yale database and obtain the same input images used by Shashua and Riklin-Raviv [22]. Then, we synthesize the face images under various lighting conditions as shown in Fig. 11. In general, the results by [22] do not appear as convincing because the lighting effect is not consistent in different parts of the face. In addition, the skin texture of the synthesized face is not realistic, as its Lambertian assumption allows representation of diffuse texture only. Furthermore, their results do not show the expected cast shadows very well. The same observations hold for other synthesized images.

#### E. Comparison With Spherical Harmonic Basis Morphable Model (SHBMM)

In this section, we compare our synthesized results with the ones from Zhang *et al.* [11]. Note that, in this experiment, we use the subspace models of the reflectance functions from the Yale database while the input image in Fig. 12 is from the PIE database.

The approach introduced by Zhang *et al.* uses the SHBMM to model both the illumination and the shape. In Fig. 12, we compare the face relighting algorithm by Zhang *et al.* [11] with the proposed algorithm. We use the same input image and the target image used by Zhang *et al.* and synthesize the face images using the synthesis scheme  $B$  (see Section V). The synthesis results from Zhang *et al.* present some artifacts along the lip corner. Moreover, the skin reflectance is not realistic as they only preserve the Lambertian property (the synthesized face surface appears plastic). Overall, we observe that our method performs better than Zhang *et al.* [11] in terms of reproducing the exact lighting effects (e.g., the contrast between the left and the right half of the face) and estimating the various skin reflectance properties (e.g., specularly on the bright spot of the forehead and the cheek). Consequently, our method, without an elaborate shape model, is visually pleasing compared to the most recent face relighting algorithm using a shape model and better in producing various skin reflectance properties.

In the last column of Fig. 12, Zhang *et al.* have created a sharp cast shadow along the nose while our synthesized shadow is

somewhat soft. Since the model by Zhang *et al.* does not account for the shadow, they could detect the shadow area as an outlier of the model in the target image and retain the shadows from the target face. Consequently, they were able to copy this shadow to the input face. Although their generated shadow is not realistic either, it may seem plausible to untrained eyes. In the future, we will investigate a method of casting the sharp shadow on the input face by learning from the target image.

#### F. Subjective Evaluation

In Section VI-C, we have shown our quantitative comparison results by presenting the correlation values between the synthesized images and the ground truth images. However, the ground truth images are not often available and, furthermore, the correlation measure does not always correspond to human perception. In this section, we perform the subjective evaluation to obtain the score/rank of our synthesized results as well as the results from other algorithms. The experiment was performed on the examples shown in Sections VI-C–VI-E, and totally 40 participants joined in this evaluation. During the evaluation, the order of the synthesized results was set arbitrarily and the participants spend approximately 10 s to evaluate each image.

**Comparison with Sim and TensorFaces.** We used a total of 18 synthesized images for this subjective study. For each input image, we presented three synthesized images (from the proposed algorithm, Sim, and TensorFaces, respectively) and the ground truth image to participants. Then, we asked the participants to rank three synthesized images between 1 to 3 (1: the best, 2: the second best, 3: the third best).

The average vote for the proposed algorithm, Sim, and TensorFaces were 1.11, 2.21, and 2.63 when their corresponding standard deviation values were 0.13, 0.16, and 0.23, respectively. This is statistically significant and justifies that, in Fig. 10, our synthesized results are superior perceptually to the ones from Sim and TensorFaces.

**Comparison with Quotient Image.** We displayed the input image and eight synthesized images (four images for each algorithm) to participants. Then, the participants were asked to score each synthesized image between 1 to 5 (1: it is a real photograph, 2: it is probably a real photograph, 3: it is hard to judge,

4: it is probably not a real photograph, 5: it is definitely not a real photograph).

The average score for the proposed algorithm was 2.40 with a standard deviation of 0.60, while the average score of the Quotient image was 3.91 with a standard deviation of 0.65. That means the participants, overall, perceive that our synthesized results are more convincing than the results from the Quotient Image.

**Comparison with the SHBMM.** First, to compare our approach with Zhang *et al.*'s method, we provide an input face image and a target image as well as two synthesized images as shown in Fig. 12. Then, the participants rank two synthesized images (1: more likely, 2: less likely) based on which image is more likely to be a real photograph of the input face under the target lighting condition. A total of ten synthesized images were evaluated in this experiment.

As a result, the proposed algorithm scored 1.28 with a standard deviation of 0.20 while Zhang *et al.* [11] achieved 1.72 with the same standard deviation. Our synthesized results performed better in the subjective evaluation task compared to Zhang *et al.*

Based on the subjective evaluation results in this section, we believe that the proposed algorithm is more effective than four recent representative face relighting algorithms for synthesizing perceptually realistic face images.

### G. Summary

We summarize the experimental results presented in Sections VI-A–VI-F, as follows.

- 1) The proposed algorithm achieves high accuracy for the lighting and pose estimation. We also show that the small pose error during the pose estimation is acceptable for face synthesis.
- 2) The proposed algorithm outperforms four existing representative face relighting methods, [3], [11], [22], [30], based on the subjective evaluation results.
- 3) The proposed algorithm preserves the non-Lambertian reflectance property of the skin as well as the facial hair in the synthesized images. Note that virtually all existing face relighting algorithms we surveyed, regardless of their models, could not reproduce the facial hair in the synthesized image very well. Also, without a sophisticated 3-D shape model, we can reproduce both the attached and cast shadow.
- 4) The proposed algorithm can be successfully applied to the practical applications that require retaining facial features, reproducing various skin reflectance properties, and preserving the identity of a subject.

**Limitation of the Proposed Algorithm.** In Sections VI-A–G, we produced plausible synthesized results using the input image(s) under the normal lighting condition. Recently, Wang *et al.* [12] improved the approach by Zhang *et al.* [11] and showed that they could recover the lighting, shape, and albedo from the input image even under harsh lighting condition. Compared to Wang *et al.*, our approach does not produce the good results under the extreme lighting condition. More specifically, the extreme lighting conditions correspond to the lighting of which angles from the camera optical axis is larger than  $50^\circ$ . This is because they use



Fig. 13. Synthesized results using inputs with various image sizes. First column: Input image. First row: Images at  $54 \times 54$ . Second row: Images at  $100 \times 100$ . Third row: Images at  $543 \times 543$ .

the shape model with the spatially dependent texture model to *perceive* the missing components in the input image while we do not impose a spatial constraint on the reflectance functions. However, Wang *et al.* treated other important facial features as outliers of their model and, therefore, could not reproduce various facial features, e.g., non-Lambertian skin reflectance properties and facial hair, in the synthesized image. Despite of the disadvantage of their method, it is definitely useful if we can handle the input face under the harsh lighting condition. We are currently developing a region-based subspace model, which considers the spatial relationship among neighboring pixels, in order to improve the effectiveness of our approach under the harsh lighting condition.

Additionally, our method is currently limited by the occlusion and facial expression changes in the input face image. We believe that a region-based subspace model will be appropriate to solve the occlusion problem since it is capable to recover the missing pixels in the input image based on other pixels in the same region.

**Remark on the Image Resolution.** In Section IV, we briefly mentioned our strategy to handle the input images from various resolutions. The results in Fig. 13 show that we can successfully process the input images which resolution differ from the training images. Particularly, in the third row of the figure, the blurriness did not appeared in the synthesized results although we applied the nearest neighborhood interpolation on our subspace model to increase the resolution of the PCA models up to the input image size. Note that using the same subspace model among the neighboring pixels does not result in the same reflectance functions along the neighboring region. This is how the PCA model is reliable to the resolution of the input images. The same resampling method was applied to obtain the images in the first row.

**Remark on Pose Changes.** To handle the faces under various poses, we constructed the pose-dependent subspace models and applied the pose model corresponding to the pose of the input face for inference and synthesis. From the PIE database, we built a total of 13 pose models, when nine pose models cover approximately  $\pm 75^\circ$  in azimuth and four pose models cover  $\pm 15^\circ$  in altitude. Also, in Section VI-A, we have shown that each pose model can tolerate  $\pm 15^\circ$  of the pose mismatch. Therefore, our subspace model accounts for the poses of  $\pm 90^\circ$  in azimuth and  $\pm 30^\circ$  in altitude. We believe that this range is reasonable to handle most of pose variations in the face images.

## VII. CONCLUSION AND FUTURE WORK

We present a robust subspace model of the reflectance function for face relighting. Using the proposed approach, we can synthesize a face under arbitrary lighting conditions without knowing its original pose and lighting condition.

Our subspace model encapsulates structural information implicitly such that we generate the expected 3-D effects without modeling the 3-D shape explicitly. Moreover, we do not rely on an explicit physical reflectance model (e.g., a Lambertian model or a Phong model), and, thus, we can handle a variety of surface properties, as well as the facial hair. Furthermore, we can handle face registration error because the proposed model of the reflectance function represents the reflectance functions of neighboring pixels as well.

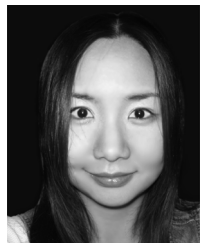
Extensive evaluation and benchmarking with recent representative approaches have demonstrated the robustness of the proposed relighting algorithm. Successful synthesis from naturally acquired images has shown the flexibility of our model for the general face subjects, and, thus, it demonstrates the potential of the proposed algorithm for many other practical applications than face recognition.

Using our pixel-dependent subspace model, we could estimate the lighting for each pixel, which results in a 4-D incident light field for the lighting estimation. In the future, we will develop a robust scheme to handle the 4-D incident light field present in the input image. In addition, we would like to learn the spatial relationship between the pixels on the face and develop a region-based subspace model to further improve the robustness and the effectiveness of our approach.

## REFERENCES

- [1] A. S. Georghiadis, P. N. Belhumeur, and D. J. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 643–660, Jun. 2001.
- [2] A. Georghiadis, D. Kriegman, and P. Belhumeur, "Illumination cones for recognition under variable lighting: Faces," *IEEE Comput. Vis. Pattern Recognit.*, vol. 23, no. 6, pp. 47–54, Jun. 1998.
- [3] T. Sim, "Statistics-based face rendering and its application to face recognition," Ph.D. dissertation, Carnegie Mellon Univ., Pittsburgh, PA, 2002.
- [4] S. K. Zhou and R. Chellappa, "Image-based Face Recognition under Illumination and Pose Variations," *Opt. Soc. Amer. J.*, vol. , pp. 217–229, 2005.
- [5] A. Shashua, "On photometric issues in 3D visual recognition from a single 2D image," *Int. J. Comput. Vis.*, vol. 21(1), no. 99, p. 1229, 1997.
- [6] R. Gross, I. Matthews, and S. Baker, "Fisher light-fields for face recognition across pose and illumination," *Proc. GSPR*, pp. 217–229, 2002.
- [7] R. Gross, S. Baker, I. Matthews, and T. Kanade, "Face recognition across pose and illumination," in *Handbook of Face Recognition*. New York: Springer, 2004.
- [8] C.-P. Chen and C.-S. Chen, "Lighting normalization with generic intrinsic illumination subspace for face recognition," in *Proc. IEEE Int. Conf. Computer Vision*, 2005, pp. 1089–1096.
- [9] L. Zhang and D. Samaras, "Face recognition from a single training image under arbitrary unknown lighting using spherical harmonics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 3, pp. 1914–1930, Mar. 2006.
- [10] L. Zhang and D. Samaras, "Face recognition under variable lighting using harmonic image exemplars," *Comput. Vis. Pattern Recognit.*, vol. I, pp. 9–25, 2003.
- [11] L. Zhang, S. Wang, and D. Samaras, "Face synthesis and recognition under arbitrary unknown lighting using a spherical harmonic basis morphable model," *Comput. Vis. Pattern Recognit.*, vol. II, pp. 209–216, 2005.
- [12] Y. Wang, Z. Liu, G. Hua, Z. Wen, Z. Zhang, and D. Samaras, "Face re-lighting from a single image under harsh lighting conditions," *Comput. Vis. Pattern Recognit.*, pp. 1–8, 2007.
- [13] S. W. Park and M. Savvides, "Estimating mixing factors simultaneously in multilinear tensor decomposition for robust face recognition and synthesis," presented at the CVPR Workshop, 2006.
- [14] J. Lee, B. Moghaddam, H. Pfister, and R. Machiraju, "A bilinear illumination model for robust face recognition," in *Proc. IEEE Int. Conf. Computer Vision*, 2005, pp. 1177–1184.
- [15] K. Lee, J. Ho, and D. Kriegman, "Nine points of light: acquiring subspaces for face recognition under variable lighting," in *Proc. IEEE Int. Conf. Pattern Recognition*, 2001, pp. 519–526.
- [16] W. Y. Zhao and R. Chellappa, "Robust face recognition using symmetric shape-from-shading," Tech. Rep. CARTR-919, Center Autom. Res., College Park, MD, 1999.
- [17] K. C. Lee, J. Ho, and D. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 684–698, Mar. 2005.
- [18] D. Xu, Q. Yang, L. Zhang, X. Tang, and H. J. Zhang, "Multilinear discriminant analysis for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 212–220, Jan. 2007.
- [19] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 2, pp. 218–233, Feb. 2003.
- [20] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," *Proc. ACM SIGGRAPH*, pp. 187–194, 1999.
- [21] V. Blanz, K. Scherbaum, T. Vetter, and H. Seidel, "Exchanging faces in images," presented at the Proc. EuroGraphics, 2004.
- [22] A. Shashua and T. Riklin-Raviv, "The quotient image: Class-based re-rendering and recognition with varying illuminations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 129–139, Feb. 2001.
- [23] H. Barrow and J. Tenenbaum, "Recovering intrinsic scene characteristics from images," *Comput. Vis. Syst.*, pp. 3–26, 1978.
- [24] W. Smith and E. Hancock, "Recovering facial shape using a statistical model of surface normal direction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 2, pp. 351–363, Dec. 2006.
- [25] R. Ramamoorthi, "Analytic PCA construction for theoretical analysis of lighting variability in images of a Lambertian object," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 10, pp. 1322–1333, Oct. 2002.
- [26] W. Zhao and R. Chellappa, "Symmetric shape-from-shading using self-ratio image," in *Proc. IEEE Int. Conf. Computer Vision*, 2001, pp. 55–75.
- [27] P. N. Belhumeur and D. J. Kriegman, "What is the set of images of an object under all possible lighting conditions?," *IEEE Trans. Comput. Vis. Pattern Recognit.*, vol. 18, no. 6, pp. 270–277, Jun. 1996.
- [28] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: Determining the illumination from images of a convex Lambertian object," *J. Opt. Soc. Amer. A*, pp. 2448–2459, 2001.
- [29] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.
- [30] M. A. O. Vasilescu and D. Terzopoulos, "Multilinear analysis of image ensembles: TensorFaces," in *Proc. IEEE ECCV*, 2002, pp. 447–460.
- [31] Y. Weiss, "Deriving intrinsic images from image sequences," presented at the Int. Conf. Computer Vision, 2001.
- [32] H. G. Barrow and J. M. Tenenbaum, "Recovering intrinsic scene characteristics from images," *Comput. Vis. Syst.*, pp. 3–26, 1978.
- [33] L. D. Lathauwer, B. D. Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM J. Matrix Anal. Appl.*, vol. 21, no. 4, pp. 1253–1278, 2000.

- [34] Z. Wen, Z. Liu, and T. S. Huang, "Face relighting with radiance environment maps," *Comput. Vis. Pattern Recognit.*, 2003.
- [35] P. Debevec, T. Hawkins, C. Tchou, H.-P. Duiker, W. Sarokin, and M. Sagar, "Acquiring the reflectance field of a human face," *Proc. SIGGRAPH*, pp. 145–156, 2000.
- [36] A. Wegner, A. Gardner, C. Tchou, J. Unger, T. Hawkins, and P. Debevec, "Performance relighting and reflectance transformation with time-multiplexed illumination," *ACM Trans. Graph.*, pp. 756–764, 2005.
- [37] P. Peers, N. Tamura, W. Matusik, and P. E. Debevec, "Post-production facial performance relighting using reflectance transfer," *ACM Trans. Graph.*, vol. 26, no. 3, p. 52, 2007.
- [38] Y. Wang and Z. Liu, private communication.



**Hyunjung Shim** (S'05) received the B.S. degree in electrical and mechanical engineering from Yonsei University, Korea, in 2002, and the M.S. and Ph.D. degrees in electrical and computer engineering from Carnegie Mellon University, Pittsburgh, PA, under the supervision of T. Chen.

Her research interests span image processing, computer vision, and computer graphics, and she is currently focused on developing image-based modeling, rendering, and relighting techniques, as well as recovering geometry and reflectance information

from images. Note that she initially developed this work during a summer internship with Kodak Research Laboratories (2006), advised by J. Luo.



**Jiebo Luo** (SM'99) received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology of China (USTC) in 1989 and 1992, respectively, and the Ph.D. degree in electrical engineering from the University of Rochester, Rochester, NY, in 1995.

He is a Senior Principal Scientist with Kodak Research Laboratories, Rochester. His research interests include image processing, pattern recognition, computer vision, computational photography, medical imaging, and multimedia communication. He is

the author of more than 120 technical papers and holds more than 40 granted U.S. patents.

Dr. Luo currently serves on the editorial boards of the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, the IEEE TRANSACTIONS ON MULTIMEDIA, *Pattern Recognition*, and the *Journal of Electronic Imaging*. He is a Guest Editor for the Special Issue on Image Understanding for Digital Photos in *Pattern Recognition* (2005), the Special Issue on Real-World Image Annotation and Retrieval in the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE (2008), the Special Issue on Event Analysis in the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY (2008), and the Special Issue on Integration of Content and Context for Multimedia Management in the IEEE TRANSACTIONS ON MULTIMEDIA in 2009. He is a Kodak Distinguished Inventor and a winner of the 2004 Eastman Innovation Award. He has also been an organizer of numerous technical conferences, most notably the general chair of the 2008 ACM International Conference on Image and Video Retrieval (CIVR), an area chair of the 2008 IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), a program cochair of the 2007 SPIE International Symposium on Visual Communication and Image Processing (VCIP), and a special sessions cochair of the 2006 IEEE International Conference on Multimedia and Expo (ICME). He is a fellow of the SPIE.



**Tsuhuan Chen** (F'07) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, R.O.C., in 1987, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, Pasadena, in 1990 and 1993, respectively.

He has been with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA, since October 1997, where he is currently a Professor and Associate Department Head. From August 1993 to October 1997, he was with

AT&T Bell Laboratories, Holmdel, NJ. He co-edited a book titled *Multimedia Systems, Standards, and Networks* (CRC, 2000).

Dr. Chen served as the Editor-in-Chief for the IEEE TRANSACTIONS ON MULTIMEDIA from 2002–2004. He also served on the Editorial Board of the *IEEE Signal Processing Magazine* and as an Associate Editor for the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, the IEEE TRANSACTIONS ON IMAGE PROCESSING, the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and the IEEE TRANSACTIONS ON MULTIMEDIA. He received the Charles Wilts Prize at the California Institute of Technology in 1993. He was a recipient of the National Science Foundation CAREER Award from 2000 to 2003. He received the Benjamin Richard Teare Teaching Award in 2006 and the Eta Kappa Nu Award for Outstanding Faculty Teaching in 2007. He was elected to the Board of Governors, IEEE Signal Processing Society, for 2007–2009. He is a member of the Phi Tau Phi Scholastic Honor Society. He is a Distinguished Lecturer of the Signal Processing Society.