Intrinsic Illumination Subspace for Lighting Insensitive Face Recognition

Chia-Ping Chen and Chu-Song Chen, *Member, IEEE*

Abstract—We introduce the intrinsic illumination subspace and its application for lighting insensitive face recognition in this paper. The intrinsic illumination subspace is constructed from illumination images of intrinsic images, which is a midlevel description of appearance images and can be useful for many visual inferences. This subspace forms a convex polyhedral cone and can be efficiently represented by a low-dimensional linear subspace, which enables an analytic generation of illumination images under varying lighting conditions. When only objects of the same class, such as faces, are concerned, a class-based generic intrinsic illumination subspace can be constructed in advance and used for novel objects of the same class. Based on this class-based generic subspace, we propose a lighting normalization method for lighting insensitive face recognition, where only a single input image is required. The generic subspace is used as a bootstrap subspace for illumination images of novel objects. Face recognition experiments are performed to demonstrate the effectiveness of the proposed lighting normalization method and verify empirically that the class-based generic subspace is applicable to novel objects. Our method is simple and fast, which makes it useful for real-time applications, embedded systems, or mobile devices with limited resources.

Index Terms—Face recognition, intrinsic image, Lambertian reflectance, lighting normalization, spherical harmonics.

I. INTRODUCTION

LIGHTING variation is one of the most challenging problems in vision applications. For face recognition, it has been observed that the variations among images of the same face due to illumination changes are larger than image variations due to changes in face identity [1]. Chen *et al.* [2] illustrated this complexity by showing that there is no discriminative illumination invariant for Lambertian objects under distant point light sources. That is, it is not possible to determine whether two images were created by the same object under two distant point light sources or by different objects. Therefore, given one image of an object, it is difficult, if possible, to predict anything definite about this object or how it will appear under different lighting conditions. Consequently, lighting variations still impede the development of real-world face recognition applications.

At first glance, modeling lighting variations may seem intractable. One way to measure the difficulties presented by any variability is the number of degrees of freedom needed to describe it. For example, to describe the pose of an object relative to the camera, we need six degrees of freedom: three for rotations and three for translations. The illumination, however, can come from any incident direction and can be composed of multiple illuminants. That is, we would need to model the intensity from each of the infinitely many incident lighting directions, meaning that light variations have an infinite number of degrees of freedom.

Despite the daunting nature of this problem, many methods have been proposed in the literature to tackle lighting variations. Histogram equalization is probably the simplest and most common technique to lessen the effects of lighting variations. Discussions about histogram equalization can be found in [3].

Belhumeur and Kriegman [4] and Georghiades *et al.* [5] gave a theoretical explanation for the set of images of an object in fixed pose, but under varying lighting conditions, it forms a convex polyhedral cone called the *illumination cone* in the image space. They also proposed a face recognition method that requires a small number of training images (at least three images) of each face. Each image should be taken under different lighting directions to estimate the shape and albedo, which are used to construct an approximated illumination cone for each face. The identity of a test image is then assigned according to the closest approximated illumination cone.

Basri and Jacobs [6] and Ramamoorthi and Hanrahan [7] independently developed the spherical harmonic representation, respectively. They showed that the Lambertian kernel acts as a low-pass filter and explained why a low-dimensional subspace can describe well the images of a convex object under varying lighting conditions. This led to face recognition with excellent results using harmonic basis images [6]. The basis images can be rendered from a 3-D scan of the face or can be approximated by acquiring a number of images of the same subject under different illuminations [8], [9]. The requirement for a number of training images or 3-D scans of the subjects necessitates special equipments and procedures to capture the training or gallery sets.

While the aforementioned two approaches assumed that each individual has a different 3-D geometry and constructed one subspace for each individual, the *quotient image* method [10] simply takes image ratios between a test image and a linear combination of a set of bootstrap images illuminated by

C.-P. Chen is with the Institute of Information Science, Academia Sinica, Taipei 115, Taiwan, and also with the Department of Computer Science and Information Engineering, National Taiwan University, Taipei 106, Taiwan (e-mail: cpchen@iis.sinica.edu.tw).

C.-S. Chen is with the Institute of Information Science and the Research Center for Information Technology Innovation, Academia Sinica, Taipei 115, Taiwan (e-mail: song@iis.sinica.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.
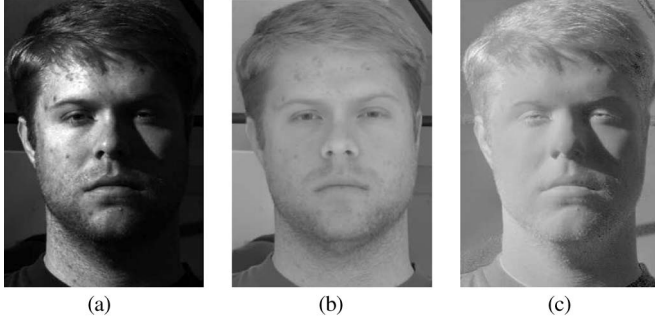
Fig. 1. Intrinsic image decomposition. The appearance image is a composition of two latent images: reflectance and illumination images. The reflectance image contains the lighting-invariant reflectance values of the object, while the illumination image represents the shading and shadowing of the object surface, which accounts for the lighting variations. (a) Appearance image. (b) Reflectance image. (c) Illumination image.

noncoplanar lights. Despite ignoring geometry variations and using the same bootstrap set for all faces, this is a simple yet practical algorithm for extracting lighting-invariant signatures. The extracted quotient image depends only on the albedo information and thus is illumination invariant.

In the 3-D morphable model method [11], each face model is represented by a linear combination of a set of 3-D exemplar faces. Fitting the morphable model to the input image was used successfully in face recognition [12] and face synthesis [11], [13]. The method in [12] is bound to images taken under directional light sources only and requires the knowledge of the light direction, which is not easy to estimate in most cases. Zhang *et al.* [14] further integrated a more general illumination representation into the morphable model approach. In their method, the illumination variations of general lighting conditions are handled by the spherical harmonic bases. Although 3-D morphable model provides excellent performance in both synthesis and recognition applications, it requires a large set of 3-D scans of human faces, which is not easy to collect; the analysis by synthesis strategy demands a lot of computation because both the texture and 3-D shape must be estimated, as well as all other parameters, such as lighting condition, pose, and scale. These make it infeasible in some circumstances.

Still, another approach dealing with lighting variations is from the *intrinsic image* viewpoint. Barrow and Tenenbaum [15] suggested taking every appearance (retinal) image I as a composition of a set of latent images, which they referred to as intrinsic images. One of the intrinsic images is the reflectance image R, which represents the reflectance values of the object, while the other is the illumination image L, which accounts for the shading and shadowing on the object surface. The relationship between the appearance and intrinsic images can be described by

$$I(x,y) = R(x,y)L(x,y) \qquad (1)$$

and is shown in Fig. 1. Barrow and Tenenbaum argued that, despite not making explicit all the physical causes of image appearance, such a midlevel description can be extremely useful for many visual inferences.

However, recovering two intrinsic images R and L from a single appearance image I remains a difficult problem. This

is a classical ill-posed problem: The number of unknowns is twice the number of equations. Weiss proposed a maximum-likelihood (ML) estimation method [16] for a slightly easier version of this problem: A sequence of T images $I(x,y,t)$ is given, in which the reflectance is constant over time and only the illumination changes. He solved this problem based on an assumption that the convolutions of images and derivative filters tend to be sparse [17] and derived a single reflectance image $R(x,y)$ and T illumination images $L(x,y,t)$ such that

$$I(x,y,t) = R(x,y)L(x,y,t). \qquad (2)$$

This intrinsic image estimation method is offline since it requires a sequence of accumulated images under variant lighting conditions to obtain the decomposed intrinsic images.

Yet, there are still a plethora of relevant methods that try to model illumination effects. We just name a few in this paragraph without intending to be complete. Son and Mita [18] adopted edge features and used the mean-field algorithm to reduce noise edges from shadow regions. A kernel-based linear discriminant analysis method, where the input space is mapped to a high-dimensional feature space, was proposed in [19] to handle large illumination variations as well as pose variations. Chen *et al.* [20] proposed the logarithmic total variation model to factorize a single image and obtain illumination invariant facial structure. The merit of this method is that no lighting assumption or any training is required.

In this paper, we propose a novel concept called the *intrinsic illumination subspace*. We analyze the intrinsic illumination subspace in terms of the Lambertian reflectance model and demonstrate its relationship with the illumination cone and spherical harmonic bases. The effects of a single point light source as well as general lighting conditions are captured by the intrinsic illumination subspace, where attached shadows are also explicitly considered. In contrast to the illumination cone method [5] and other low-dimensional subspace methods [21], the advantage of the intrinsic illumination subspace is that a class-based generic subspace can be constructed and used for all objects when only objects of the same class are concerned, instead of one subspace per object.

Based on this class-based generic subspace, a lighting normalization method for a single input image is derived. The generic subspace is used as a bootstrap subspace to normalize images of novel objects of the same class. The effectiveness of our method is demonstrated with face recognition experiments, where the generic subspace is verified empirically to be applicable to novel objects. The recognition accuracy improves significantly with the proposed lighting normalization method even under large lighting variations. Our method is simple and fast, which make it useful for circumstances where available resources are limited, such as real-time applications, embedded systems, and mobile devices.

The remainder of this paper is organized as follows. Essential background knowledge is given in Section II. Section III depicts the concept of intrinsic illumination subspace and the class-based generic subspace. In Section IV, we propose that the non-negative matrix factorization (NMF) can also be used to model the intrinsic illumination subspace. The lighting normalization

method is described in Section V. Section VI shows experimental results, while conclusions are made in Section VII.

II. BACKGROUND

In this section, we briefly introduce essential background knowledge about the quotient image method for lighting normalization and the spherical harmonic bases for the effects of general lighting conditions. More details can be found in [10], [22], [6], and [7], respectively.

A. Quotient Image

The quotient image method adopts the Lambertian reflection model and assumes that the object is illuminated by a single distant point light source and no shadow is present

$$I(x,y) = \rho(x,y) N^{\mathrm{T}}(x,y) \cdot s \qquad (3)$$

where x and y are the image coordinates, ρ and N are the albedo and 3-D surface normal of the face, $\cdot$ denotes the dot product, and s represents the strength and direction of the distant point light source.

The quotient image Q_b of face b against the exemplar face a is defined as the ratio between the albedo of face b and that of face a

$$Q_b(x,y) = \frac{\rho_b(x,y)}{\rho_a(x,y)}. \qquad (4)$$

Since the quotient image Q_b depends only on the albedo information of the two faces, it is illumination invariant and adequate for lighting insensitive face recognition. We briefly introduce the method how Shashua and Riklin-Raviv [10] approximate Q_b in the following.

Let I_1, I_2, and I_3 be three exemplar images of face a illuminated by three noncoplanar light sources s_1, s_2, and s_3, respectively. Then, any point light source s can be taken as the linear combination of s_i with appropriate coefficients: $s = \sum_{i=1}^{3} x_i s_i$. Given the test image I_b of face b illuminated by an unknown light source s, the quotient image Q_b of face b against face a is approximated by assuming that both faces share the same shape (surface normal N)

$$
\begin{aligned}
Q_b(x,y) &= \frac{\rho_b(x,y)}{\rho_a(x,y)} \\
&= \frac{\rho_b(x,y) N^{\mathrm{T}}(x,y) \cdot s}{\rho_a(x,y) N^{\mathrm{T}}(x,y) \cdot s} \\
&= \frac{I_b(x,y)}{\rho_a(x,y) N^{\mathrm{T}}(x,y) \cdot \sum_{i=1}^{3} x_i s_i} \\
&= \frac{I_b(x,y)}{\sum_{i=1}^{3} x_i I_i(x,y)}.
\end{aligned}
\qquad (5)
$$

According to (5), the lighting-invariant quotient image Q_b can be approximated by the ratio between the test image I_b and

a linear combination of the exemplar images I_i with appropriate coefficients x_i, which simulates the light source s in the test image I_b.

Therefore, the key is how to obtain the appropriate coefficients x_i. For that reason, a bootstrap set is needed, where there are exemplar images of multiple faces and each face is illuminated under the same fixed noncoplanar light sources s_1, s_2, and s_3. The shapes of all the faces are assumed the same, which is referred to as the *ideal class assumption* in [10]. Then, the coefficients x_i are estimated by minimizing an energy function which consists of projections of the test image onto the subspaces spanned by the exemplar images of each face in the bootstrap set.

B. Spherical Harmonics

With the distant light source assumption, the intensities of a Lambertian object under general lighting conditions can be represented as follows:

$$I(x,y) = \rho(x,y) \int L(s) \max \left(N^{\mathrm{T}}(x,y) \cdot s, 0 \right) \mathrm{d}s \qquad (6)$$

where ρ and N are the albedo and surface normal, s is a unit vector indicating the direction of incoming light, and $L(s)$ is the radiance intensity from direction s. The integral is over all possible lighting directions.

In [6] and [7], the authors viewed this process of light reflection as a convolution, where the incident illumination signal is filtered by the reflective properties of the surface, which is the Lambertian kernel in our case. They showed that the Lambertian kernel acts like a low-pass filter, preserving only the lowest frequency components of the lighting. According to their results, the effects of general lighting on a Lambertian object can be represented by spherical harmonics

$$I(x,y) = \rho(x,y) \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \sqrt{\frac{4\pi}{2n+1}} k_n l_{nm} Y_{nm}\left(N(x,y) \right) \qquad (7)$$

where Y_{nm} denotes the surface spherical harmonics and k_n and l_{nm} are the coefficients of harmonic expansions of the Lambertian kernel and the lighting condition, respectively.

It has been proved in [6] that, for any distant and isotropic lighting, at least 98% of the reflectance function can be captured by the second-order spherical harmonic approximation. A first-order approximation captures at least 75% of the reflectance function. The accuracy of near lighting is also discussed in [23]. These bounds are not tight. In fact, many common lighting conditions yield significantly better approximations. For example, under a point source illumination, the first- and second-order harmonics approximate the reflectance function to 87.5% and 99.22%, respectively.

Therefore, the following second-order approximation serves the needs of many vision applications well:

$$I(x,y) = \rho(x,y) \sum_{n=0}^{2} \sum_{m=-n}^{n} \sqrt{\frac{4\pi}{2n+1}} k_n l_{nm} Y_{nm}\left(N(x,y) \right). \qquad (8)$$

The first nine harmonics (up to the second order) can be written as

$$y_{00} = \sqrt{\frac{1}{4\pi}} \qquad\qquad y_{10} = \sqrt{\frac{3}{4\pi}} N_z$$

$$y_{11}^e = \sqrt{\frac{3}{4\pi}} N_x \qquad\qquad y_{11}^o = \sqrt{\frac{3}{4\pi}} N_y$$

$$y_{20} = \frac{1}{2}\sqrt{\frac{3}{4\pi}} \left(3N_z^2 - 1\right)$$

$$y_{21}^e = 3\sqrt{\frac{5}{12\pi}} N_x N_z \qquad\qquad y_{21}^o = 3\sqrt{\frac{5}{12\pi}} N_y N_z$$

$$y_{22}^e = 3\sqrt{\frac{5}{12\pi}} \left(N_x^2 - N_y^2\right) \qquad y_{22}^o = 3\sqrt{\frac{5}{12\pi}} N_x N_y \qquad (9)$$

where the superscript e and o denote the even and the odd components of the harmonics, respectively, and (N_x, N_y, N_z) denotes the space coordinates of the unit surface normal N. So, $Y_{nm} = Y_{n|m|}^e \pm Y_{n|m|}^o$, according to the sign of m.

These results help to explain previous empirical results that showed that the set of images produced by an object under varying illuminations lies near a low-dimensional linear subspace [5], [21], and lead to many interesting insights and have practical implications, including [24], [25], and the proposed method.

III. INTRINSIC ILLUMINATION SUBSPACE

We introduce the intrinsic illumination subspace [26] and the class-based generic subspace in this section. We assume that the surface of a convex object has Lambertian reflectance. The only parameter of this model is the albedo at each point on the object surface, which describes the fraction of the light reflected. We also assume that the convex object is illuminated by distant light sources. What we mean by "distant" is that the directions and intensities of light sources can be safely considered the same for all surface points of this object.

A. Intrinsic Illumination Subspace

According to the Lambertian reflectance function, if a distant light source s reaches a surface point with albedo R and surface normal direction N, then the intensity I reflected by the surface point due to this light source is given by

$$I(x,y) = R(x,y)N^{\mathrm{T}}(x,y) \cdot s \qquad (10)$$

where s represents the strength and direction of the light source. When an object is illuminated simultaneously by k light sources instead of only one, the image is given by the summed contributions of all light sources

$$I(x,y) = R(x,y)\sum_{i=1}^{k} N^{\mathrm{T}}(x,y) \cdot s_i \qquad (11)$$

where s_i denotes the strength and direction of the ith light source.

Lambertian reflectance function can be thought of as a special version of intrinsic images, where R stands for the lighting-invariant reflectance (albedo) values and L denotes the shading of a Lambertian surface, which accounts for the lighting variations

$$L(x,y) = \sum_{i=1}^{k} N^{\mathrm{T}}(x,y) \cdot s_i. \qquad (12)$$

Note that, when no part of the object surface is shadowed, it is obvious that L lies in a 3-D linear subspace. When attached shadows are considered, the illumination image L is given by

$$L(x,y) = \sum_{i=1}^{k} \max\left(N^{\mathrm{T}}(x,y) \cdot s_i, 0\right) \qquad (13)$$

where the max function ensures that the contribution of each light source is nonnegative since negative contribution of light is physically not realizable.

We define the *intrinsic illumination subspace* as a set $\mathcal{L}$ of illumination images of a convex Lambertian surface created by varying the directions and intensities of multiple distant light sources

$$\mathcal{L} = \left\{ L|L(x,y) = \sum_{i=1}^{k} \max(N^{\mathrm{T}}(x,y) \cdot s_i, 0) \right.$$
$$\left. \forall\, s_i \in R^3 \qquad \forall\, k \in \mathbb{Z}^+ \right\} \qquad (14)$$

where $\mathbb{Z}^+$ is a set of positive integers.

Previous works [4], [5] have shown that the set of appearance images of an object under all possible lighting conditions forms a convex polyhedral cone, called the illumination cone, in the image space. Therefore, a question arises: What is the shape of $\mathcal{L}$ that consists of illumination images of a convex object under all lighting conditions?

Note that the definition of $\mathcal{L}$ is similar to that of the illumination cone [4]. In fact, there is a linear relationship between the illumination cone and the intrinsic illumination subspace. Letting $\mathcal{C}$ and $\mathcal{L}$ be the illumination cone and intrinsic illumination subspace of the same convex Lambertian object, respectively, then

$$\mathcal{C} = \{R \odot L : L \in \mathcal{L}\} \qquad (15)$$
$$\mathcal{L} = \{R^{-1} \odot I : I \in \mathcal{C}\} \qquad (16)$$

where R is the reflectance of this Lambertian object $R^{-1}(x,y) = 1/R(x,y)$ and $\odot$ indicates the element-by-element multiplication.

With this linear relationship, the intrinsic illumination subspace shares all the properties of the illumination cone. Since the linear transformation of a convex cone is itself a convex cone, the intrinsic illumination subspace also forms a convex polyhedral cone in the image space. The dimensionality of the intrinsic illumination subspace rises as the number of distinct surface normals increases.

Although the intrinsic illumination subspace $\mathcal{L}$ can span many dimensions if there are a large number of distinct surface

normals, previous empirical observations and analytical predictions have shown that appearance images of a largely diffuse object actually lie very close to a low-dimensional subspace. This result can apply to the intrinsic illumination subspace as well.

According to the spherical harmonic representation in [6] and [7], an image I of an object can be represented by

$$I = R \odot (Yl) \tag{17}$$

where R is the reflectance image of the object, Y contains the spherical harmonic bases where each column represents one basis image, and l is the harmonic lighting configuration indicating the contribution of each basis image. According to their finding, the first nine harmonic bases can well describe the images of a diffuse object under general lighting conditions. It follows that the illumination image L can be well represented by the first nine harmonic bases

$$L = Yl. \tag{18}$$

Because these bases must be calculated with known 3-D geometry, the application range of this representation is limited. According to Ramamoorthi's analysis [8], there is a linear relationship between the principal component analysis (PCA) eigenvectors and spherical harmonic bases. We note that the bases B of the *illumination images* of an object under all lighting conditions can also be described as

$$B = YT \tag{19}$$

where T is an $n \times n$ transformation matrix. Then, (18) can be written as

$$L = Yl = Bs \tag{20}$$

where s denotes the projected coefficients of the lighting condition and $s = T^{-1}l$.

Hence, the intrinsic illumination subspace of an object can be well approximated by a low-dimensional linear subspace. If we have densely sampled illumination images under varying lighting conditions, we can get well-approximated linearly transferred spherical harmonic bases B without estimating the 3-D geometry of this object.

B. Generic Intrinsic Illumination Subspace

The previous section introduces the intrinsic illumination subspace of one object and shows that the intrinsic illumination subspace for an object can be constructed by collecting a number of illumination images under different illuminations. However, it is usually not easy to acquire multiple illumination images of each object in many cases. In this section, we extend the intrinsic illumination subspace to a class-based *generic* subspace and show that different objects of the same class can share the same generic subspace, instead of constructing an intrinsic illumination subspace for each object individually. So, this class-based generic subspace can be constructed in advance and used for novel objects of the same class.
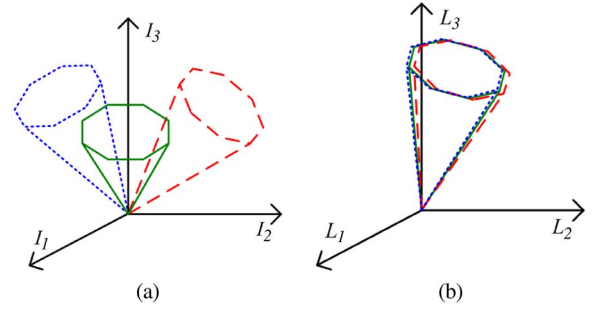

Fig. 2. Illustration for the generic intrinsic illumination subspace. The appearance images of different objects span different subspaces, as illustrated in (a), while the illumination images of different objects in the same class share the same generic intrinsic illumination subspace, as illustrated in (b). (a) Appearance subspaces. (b) Illumination subspaces.

It is very difficult to model lighting variations for general objects. However, for many vision applications, only objects of the same class are concerned, e.g., only human faces are of interest in face recognition applications. In this case, some class-based information can be gathered in advance and utilized to simplify this difficult problem. One common approach is to ignore shape variations among objects of the same class and assume that they all have the same shape. For example, Shashua and Riklin-Raviv [10] defined an *ideal class* to be a collection of 3-D objects that have the same shape but differ in the surface albedo function. Then, the appearance image of an object j in such a class can be represented by

$$I_j(x,y) = R_j(x,y)L_j(x,y)$$
$$= R_j(x,y) \sum_{i=1}^{k} \max\left(N^{\mathrm{T}}(x,y) \cdot l_i, 0\right) \tag{21}$$

where R_j is the reflectance of object j and N contains the shared surface normals of the objects in this class (the same for all objects in this class).

Objects of a class have different albedo functions so that they span different appearance subspaces in the image space, as shown in Fig. 2(a). However, by removing the reflectance function from (21), the resulting illumination image only depends on surface normals and lighting conditions. This implies that, given the ideal class assumption, different objects of the same ideal class have the same illumination image under the same illumination condition because they share identical shape. Then, a useful implication is derived: All objects of the same ideal class share the same generic intrinsic illumination subspace, as shown in Fig. 2(b).

In practice, objects of the same class do have shape variations, although the shape is similar at some coarse level. The question is how sensitive our approach is to deviations from the ideal class assumption. Our experimental results demonstrate that, as long as rough alignment is performed, such as when the center of mass and scale are aligned, the class-based generic subspace can also be well described by a low-dimensional linear subspace. So, we can withstand shape variations among objects of the same class with little degradation in face recognition performance.

The ideal class assumption can be satisfied if we perform pixelwise dense correspondences between images. A gradient-based optic flow algorithm, such as that used in [27], can be utilized to establish correspondences between a pair of images. The computed flow field is expected to improve the performance by eliminating shape variations among objects. However, computing a flow field between a pair of images usually requires salient features to exist in both images, and the results on smooth regions with little texture are sometimes faulty. Our experiments show that dense correspondence is not necessary for some applications, lighting insensitive face recognition in our case, since the recognition accuracy already improves a lot even when shape variations are ignored.

IV. ENFORCING NONNEGATIVE LIGHT WITH NMF

We propose in this section using *NMF* to model the intrinsic illumination subspace. When we take arbitrary linear combinations of the basis images obtained by commonly used PCA or other similar methods [28], [29], we may obtain images that are not physically realizable. This is because the corresponding linear combination of basis images may contain negative values. That is, rendering these images may require negative "light," which is realistically impossible. NMF is an adequate technique to enforce this nonnegative constraint.

In the illumination cone method [4], [5], the nonnegative light constraint was also considered. The set of all images of an object under all possible lighting conditions forms a convex polyhedral cone in the image space. Since the illumination cone is convex, it can be characterized by a set of *extreme rays* (i.e., images illuminated by appropriately choosing single point light sources). All the other images in the cone can then be reconstructed by convex combinations of the extreme rays. However, synthesizing the extreme rays requires estimating the surface albedo and object shape from several images of the object under different illuminations. In our approach, we use NMF to enforce the nonnegative light constraint.

NMF is a subspace method proposed by Lee and Seung [30], which has been used for image representation, document analysis, and clustering for its part-based representation property. Given a nonnegative $m \times n$ matrix X, the NMF algorithms seek to find nonnegative factors B and H such that

$$X \approx \tilde{X} = BH \tag{22}$$

where $B \in R^{m \times r}$ and $H \in R^{r \times n}$. Intuitively, we can think of B as the matrix containing the NMF bases, where all elements are nonnegative, and H as the accompanying coefficient matrix which contains only nonnegative weights.

The nonnegative restriction on the accompanying coefficients in H results in the additive and nonnegative nature of lighting, which was usually achieved by convex optimization. The nonnegative bases in B resemble extreme rays, containing representative images under certain lighting conditions, and can be used as linear bases to generate illumination images under all possible lighting conditions. The nonnegative constraint on B explicitly requires that the intensities of representative images cannot be negative. The primary advantage of nonnegative

bases derived by NMF is that they are image-based analogies to the extreme rays, and therefore, no reconstructed shape or albedo is required.

NMFs can be very difficult to compute. Lee and Seung [31] suggested an approach similar to that used in expectation–maximization algorithms to iteratively update the factorization based on a given objective function. We adopt their multiplicative update algorithm to construct the intrinsic illumination subspace

$$H_{aj} \leftarrow H_{aj} \frac{[B^\mathrm{T} X]_{aj}}{[B^\mathrm{T} W H]_{aj}} \tag{23}$$

$$B_{ia} \leftarrow B_{ia} \frac{[BH^\mathrm{T}]_{ia}}{[BHH^\mathrm{T}]_{ia}}. \tag{24}$$

The aforementioned procedures are performed iteratively, while the 2-norm $\|X - BH\|$ is ensured nonincreasing during the update process. So, the convergence of this iterative algorithm is guaranteed.

V. LIGHTING NORMALIZATION WITH INTRINSIC ILLUMINATION SUBSPACE

Based on the concept of the generic intrinsic illumination subspace, we propose a lighting normalization method for a single input image in this section. The basic idea is to reconstruct the intrinsic illumination image L of the input image I by the generic intrinsic illumination subspace. Then, we can obtain an estimated reflectance image by dividing the input image with the approximated illumination image: $R = I/L$. The estimated reflectance image R is much more illumination invariant than the original input image I, so it is more suitable for various vision applications such as recognition.

To solve the ill-posed problem of decomposing the input image I into two intrinsic images R and L, some previous works [32]–[34] used the low-frequency component of the input image as the estimation of illumination variations and imposed various assumptions or constraints on L or R. Basri and Jacobs [6] also gave a theoretical way of understanding the effects of Lambertian reflectance as that of a low-pass filter on the lighting.

In our approach, we use this low-frequency component as our initialization

$$L^\star = F * I \tag{25}$$

where I is the input image, F is the smoothing kernel, $*$ represents convolution, and $L^\star$ is the smoothed version of I. While (25) can be replaced with other intricate approaches, a simple Gaussian smoothing kernel is used in our experiments to show that our method is insensitive to the initialization.

This initial $L^\star$ is used as a query. The best approximation of the corresponding illumination image L is reconstructed from the intrinsic illumination subspace according to the Euclidean distance

$$l = \arg \min_l \|Bl - L^\star\| \tag{26}$$
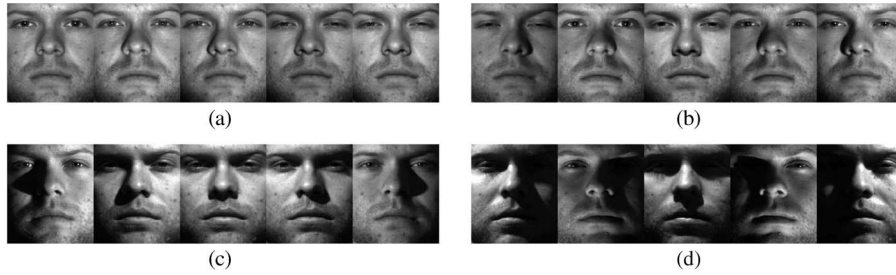
$$L = Bl \tag{27}$$

Fig. 3. Yale Face Database B. There are 64 images under different illuminations for each of the ten subjects. The images are divided into four subsets of increasing illumination angels. (a) Subset 1. (b) Subset 2. (c) Subset 3. (d) Subset 4.

where B is the basis matrix of the intrinsic illumination subspace and l contains the estimated coefficients that represent the lighting condition of the input image. When the intrinsic illumination subspace is constructed by PCA, this minimization can be achieved simply by projecting the query image $L^\star$ onto the subspace since PCA bases are orthonormal to each other. When the intrinsic illumination subspace is constructed by NMF, the minimization is done by using the same iterative procedure in (23), but only the coefficient matrix varies while the basis matrix B is fixed.

Then, lighting normalization is then performed by simply dividing the input image I with the reconstructed illumination image L

$$R = I/L. \qquad (28)$$

The aforementioned lighting normalization method shares some similarities with the quotient image method [10]. Both methods maintain a bootstrap set to estimate the illumination of the input image and normalize the input image according to the estimated result. However, there are some significant differences between our method and the quotient image method.

1) The quotient image method works under the assumptions that face images are illuminated by a single point light source and there is no shadow, while our method allows general light conditions and attached shadows are explicitly considered.

2) The bootstrap set of the quotient image method consists of appearance images of different subjects. Since there are apparent reflectance variations among different subjects, the estimation of lighting direction is inevitably affected by the reflectance variations. Our method maintains illumination images instead of appearance images, so that the estimation of illumination is expected to be more accurate since reflectance variations are removed from the bootstrap set.

3) By collecting illumination images instead of appearance images, our method allows a compact representation of the bootstrap set as a linear subspace for efficient storage and fast computation. Several to tens of basis images, which are constructed from hundreds of illumination images, result in significant improvement in our face recognition experiments.

Our method also shares similar motivations with that of Zhou et al. [35], but the formulations are quite different. They extended the photometric stereo algorithms to handle all the

Fig. 4. CMU PIE database. There are 21 images under different illuminations without background lighting for each of the 68 subjects.

appearances of all the objects in the same class to derive the albedos, 3-D shapes, and lighting directions. To solve this highly ill-posed photometric stereo problem, they relied on various constraints, such as single light source, no shadow, and symmetric and integrable surfaces. On the contrary, we do not impose these constraints, and significant rank reduction is achieved, owing to the intrinsic image representation.

VI. EXPERIMENTS AND DISCUSSION

We performed experiments to evaluate our method with the original Yale Face Database B [5], extended Yale Face Database B [9], and Carnegie Mellon University Pose, Illumination, and Expression (CMU PIE) face database [36]. Frontal face images with lighting variations are selected from these databases. There are 640 images (10 subjects with 64 images under different lighting conditions for each subject) from Yale Face Database B and 1792 images (28 subjects under 64 lighting conditions) from extended Yale Face Database B. The images are divided into four subsets of increasing illumination angles, as shown in Fig. 3. There are 68 subjects in CMU PIE database, and we select frontal face images which are taken under 21 different illuminations without background lighting, as shown in Fig. 4. All the images are aligned and cropped roughly by the positions of eyes and mouths.

A. Intrinsic Illumination Subspace

We first perform experiments to verify properties of the intrinsic illumination subspace. We decompose 64 appearance images of one subject from Yale Face Database B into a single reflectance and 64 illumination images by Weiss's ML estimation method [16]. Fig. 5(a) shows the 64 illumination
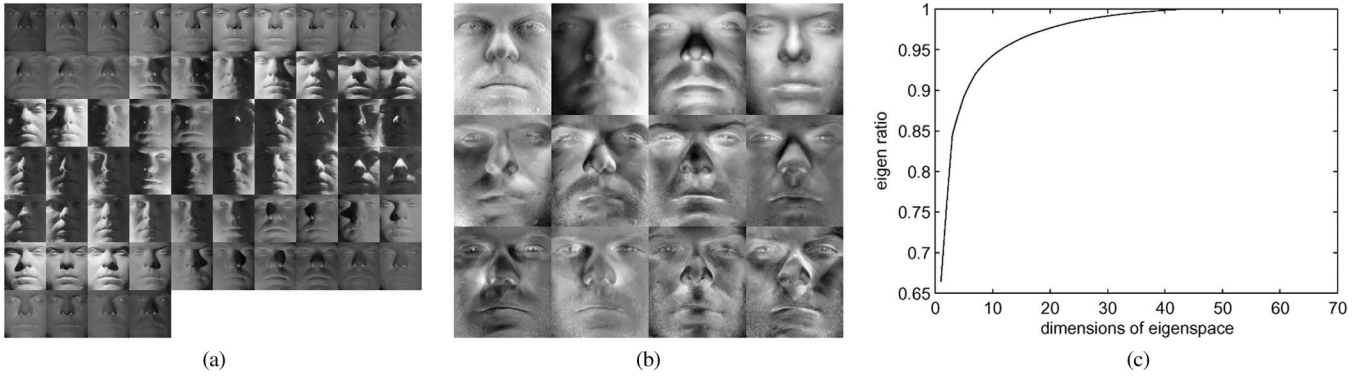
Fig. 5. Intrinsic illumination subspace of one person from Yale Face Database B. (a) All illumination images used to construct the intrinsic illumination subspace. (b) First 12 PCA bases. (c) Cumulative eigen ratios of illumination images of one person.
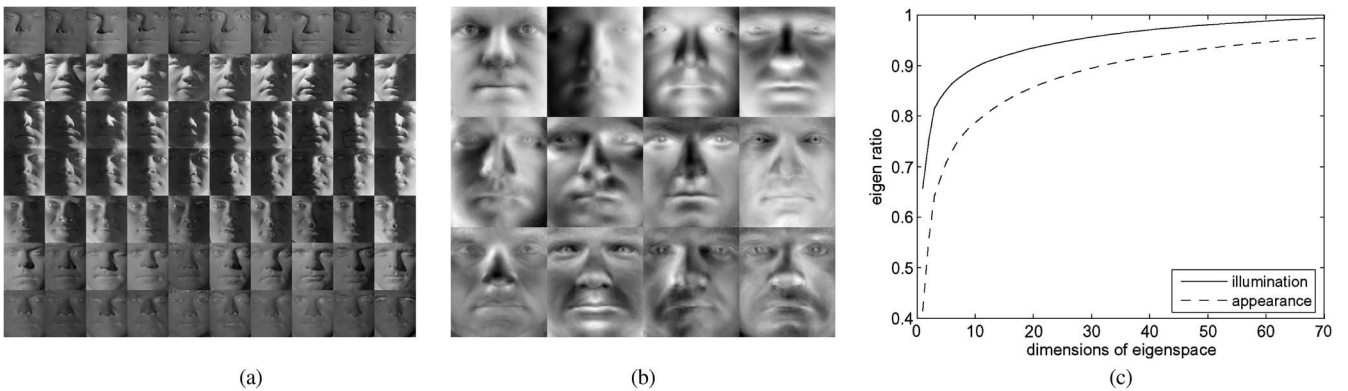

Fig. 6. Generic intrinsic illumination subspace of ten people from Yale Face Database B. (a) Seven of the 64 illumination images of each person, each row contains the illumination images of the ten subjects under the same lighting condition. (b) First 12 generic PCA bases. (c) Cumulative eigen ratios of illumination and appearance images of ten people.

images. The eigen bases and cumulative eigen ratios are shown in Fig. 5(b) and (c), respectively. It is shown that the set of illumination images can be represented by a low-dimensional linear subspace. However, in practice, the first nine eigenvectors cannot carried over 95% of energy for a nonconvex face object because there are obvious cast shadows in nonfrontally illuminated images. More eigenvectors might be needed for the implementation; the eigen ratio is 95.27% when 12 eigenvectors are used in this case.

For generic intrinsic illumination subspace, this time, we decompose the appearance images of each of the ten subjects from Yale Face Database B. Then, SVD is performed to all 640 appearance images and 640 illumination images, respectively. Fig. 6(a) shows 7 of the 64 illumination images of each subject, where each row contains the illumination images of the 10 subjects under the same lighting condition. The resulting generic eigen bases of illumination images are shown in Fig. 6(b). Fig. 6(c) shows the cumulative eigen ratios for both the appearance and illumination image sets.

We note that, even there are shape variations among these 10 subjects, 12 dimensions can capture 90.7% of energy of 640 illumination images without dense correspondence. However, 35 dimensions are needed to achieve the same eigen ratio for appearance images. This result indicates that, by removing reflectance from appearance images, the resulting illumination images exhibit much less variation among different people, which is evident in Fig. 6(a). So, a low-dimensional linear sub-

space is capable of modeling the generic intrinsic illumination subspace of human faces. On the other hand, a much higher dimensional subspace is required for appearance images.

B. Lighting Normalization and Face Recognition

To evaluate our lighting normalization method, we perform face recognition experiments under varying illuminations, where only a single gallery image is required for each subject. A simple recognition scheme, i.e., correlation, is used with the proposed method, and only gray-level images are used.

We randomly select half of the 64 illumination images of each subject from the original Yale Face Database B and reserved the other half for face recognition experiments. With all the selected illumination images, we construct the generic intrinsic illumination subspaces by PCA and NMF, respectively. Fig. 7 shows the first 12 NMF bases of the generic intrinsic illumination subspace.

We then perform recognition experiments on the remaining half of Yale Face Database B. Only the frontal illuminated images are labeled with the subject's identity and used as gallery (template) images, while all the unused images are used as probe (test) images. We also implement the quotient image method for comparison. Three appearance images of each subject are selected as the bootstrap set for the quotient image method. Both the template images and test images are normalized by the quotient image method or our method.

Fig. 7.   First 12 generic NMF bases obtained from Yale Face Database B.
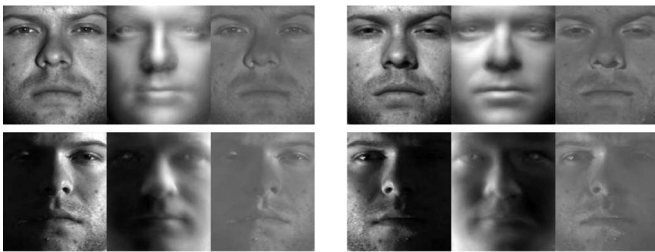

Fig. 8.   Lighting normalization process. The input, reconstructed illumination, and estimated reflectance images.

Fig. 8 shows the process of the proposed lighting normalization method. Then, the identity of a test image is determined by the most correlated template image.

Fig. 9 shows the recognition results when using PCA to construct the intrinsic illumination subspace. Fig. 10 summarizes the comparisons with some lighting insensitive face recognition methods, where the result of correlation with the original appearance images is also shown as the baseline for comparisons. The performance of the baseline algorithm shows that frontally illuminated images (Subset 1&2) are generally easier to recognized, while laterally illuminated images (Subsets 3 and 4) are the main challenges. Note that different methods require different numbers of training images for each of the subjects. For methods that require surface normals (3-D shapes) or train a subspace for each individual, multiple training images for each subject are usually required. Our method requires only one image for each subject once the generic intrinsic illumination subspace is constructed offline.

Even with this simple correlation scheme, our lighting normalization method improves the recognition rates significantly. The more bases we use to construct the subspace, the better the result is expected to be, but with more computation time. Note that the improvement is impressive even when only four PCA bases are used. This indicates that our PCA intrinsic illumination subspace is suitable for real-time applications because of its simplicity and efficiency. Due to its nonnegative constraint, the NMF subspace is comparable with or potentially better than the PCA subspace when the number of bases is large enough. This is not verified for the Subset 4 because the lighting is so extreme that other sources of errors, such as cast shadows and noises, dominate the recognition results.

Two methods, namely, discrete cosine transform (DCT) truncation [37] and harmonic exemplars [38], also require only one image for each subject and obtain good performance as well. However, it is unclear in [37] how to determine the relation between the number of DCT coefficients to truncate, the size of the images, and the cutoff frequency, while the harmonic exemplars method [38] needs a set of 3-D face models to compute a statistical model for each harmonic basis image.

To verify the applicability of the generic intrinsic illumination subspace, we further perform face recognition experiments on the original and extended Yale Face Database B at the same time, 38 subjects under 64 illuminations totally. This merged database is challenging since the number of subjects increases, the lighting conditions may be very extreme (Subset 4), and 28 out of 38 subjects are novel and unseen when constructing the generic illumination subspace. The recognition results are shown in Fig. 11. Despite the difficulties, the proposed method still performs well on this challenging database.

We then perform recognition experiments on CMU PIE database. Note that the generic intrinsic illumination subspace constructed from the ten subjects of the original Yale Face Database B is used again in order to verify its effectiveness for novel subjects and novel lighting conditions in CMU PIE database. Only single frontal illuminated image for every subject is labeled and used as templates. There are 68 templates and 1360 test images. The recognition results of PCA subspaces are shown in Fig. 12, while Fig. 13 shows the comparison results with other methods. We can see that, although the generic intrinsic illumination subspace is constructed from the ten subjects of Yale Face Database B, it is also valid for the 68 novel subjects of CMU PIE database under novel lighting conditions. It can really describe illumination images of novel subjects, so it can serve as a generic subspace. When the lighting conditions are not as extreme as Subject 4 of Yale Face Database B, the NMF subspace outperforms the PCA subspace when enough bases are used (26 bases in this case).

C. Discussions

1) Specific or Generic Subspaces: When multiple illumination images of an object are available, our method can also construct the specific intrinsic illumination subspace of this object. Normalizing input images of this object with its own subspace may yield better results. Performing supervised dimensionality reduction, such as that in [40], to gain more discriminability is also possible. However, multiple illumination images of the same object are not available in many cases. At that time, a generic subspace can be constructed from a database in advance and serves as a bootstrap subspace for a single image of novel objects. We show in the experiments that a generic subspace is valid for novel objects in our face recognition experiments.

2) Cast Shadows: Cast shadows can be significant in many vision applications. Most algorithms neglect them because nonlocal interactions in nonconvex regions make formal analysis difficult. Ramamoorthi *et al.* [41] took a first step toward a formal analysis of cast shadows. They showed that the result of cast shadows is a convolution of the lighting with the Heaviside step function. The eigenvalues decay as $1/k$ for the Heaviside step function, which is relatively slow compared to $1/k^2$ for the clamped cosine function (Lambertian reflectance kernel).
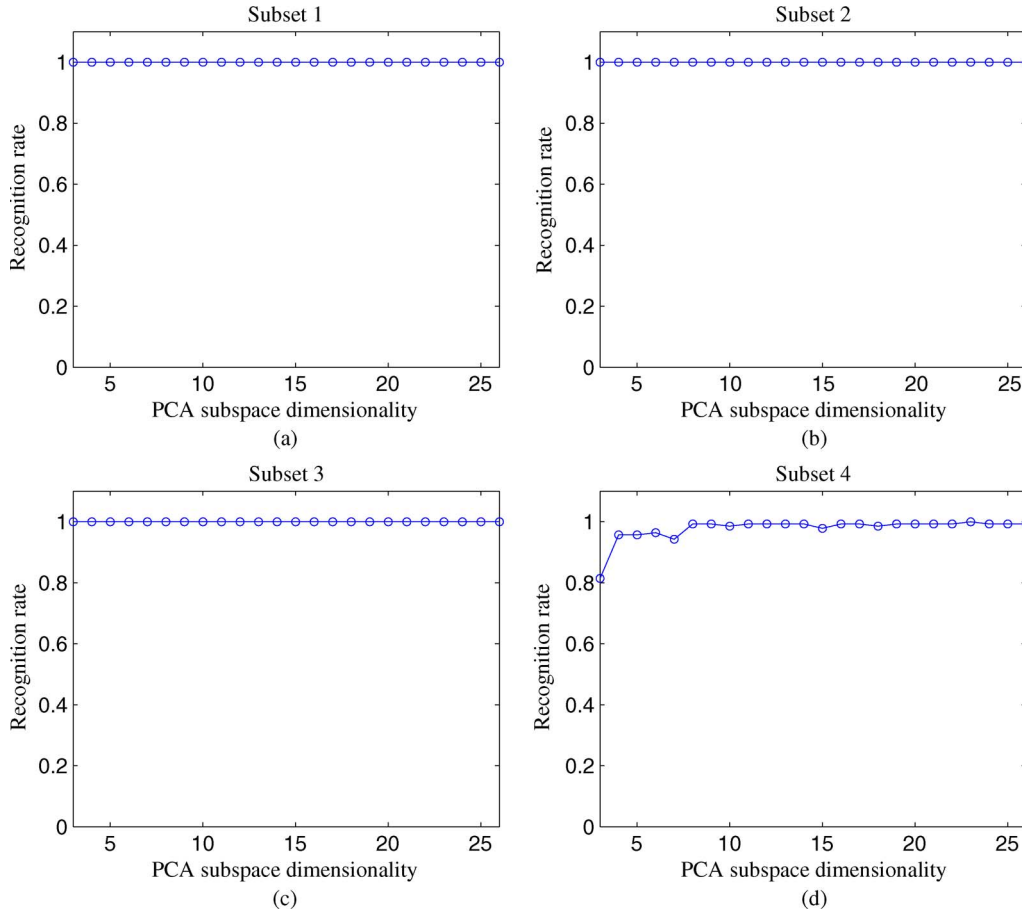
Fig. 9.   Recognition rates for Yale Face Database B when using PCA generic intrinsic illumination subspaces.

Method	Number of Training Images	Estimate Shapes	Error Rate (%)		
			Subset 1&2	Subset 3	Subset 4
Appearance	1	No	0.0	29.17	72.14
Histogram equalization	1	No	0.0	27.33	74.71
Quotient image	1	No	0.0	23.33	54.29
Correlation [5]	6-7	No	0.0	23.3	73.6
Eigenfaces [5]	6-7	No	0.0	25.8	75.7
Eigenfaces w/o 1st 3 [5]	6-7	No	0.0	19.2	66.4
Linear subspace [5]	6-7	Yes	0.0	0.0	15.0
Cones-attached [5]	6-7	Yes	0.0	0.0	8.6
Cones-cast [5]	6-7	Yes	0.0	0.0	0.0
DCT Truncation [37]	1	No	0.0	0.18	1.71
Harmonic Exemplars [38]	1	Yes	0.0	0.3	3.1
5PL [9]	5	No	0.0	0.0	0.0
9PL [9]	9	No	0.0	0.0	0.0
Ours (4D PCA)	1	No	0.0	0.0	4.29
Ours (12D PCA)	1	No	0.0	0.0	0.71
Ours (26D PCA)	1	No	0.0	0.0	0.71
Ours (4D NMF)	1	No	0.0	4.17	28.57
Ours (12D NMF)	1	No	0.0	0.83	7.86
Ours (26D NMF)	1	No	0.0	0.0	5.71

Fig. 10.   Comparison results for Yale Face Database B (ten subjects).

It suggests that it is possible to develop, in a similar fashion, subspaces that encompass the effects of cast shadows *simply by considering more bases*. This result is applicable to our approach.

3) Color Images: The process described so far holds for gray-level images. We describe a simple approach to handle color images while still maintaining a gray-value intrinsic illumination subspace. We make the assumption that varying

Method	Number of Training Images	Estimate Shapes	Error Rate (%)		
			Subset 1&2	Subset 3	Subset 4
Appearance	1	No	1.64	50.15	88.34
Histogram equalization	1	No	1.23	41.14	91.32
Quotient image	1	No	0.45	31.34	75.00
Ours (4D PCA)	1	No	0.0	0.0	13.47
Ours (12D PCA)	1	No	0.0	0.0	12.44
Ours (26D PCA)	1	No	0.0	0.0	9.33
Ours (4D NMF)	1	No	0.0	9.35	36.28
Ours (12D NMF)	1	No	0.0	3.14	12.28
Ours (26D NMF)	1	No	0.0	0.0	9.87

Fig. 11.   Comparison results for the original and the extended Yale Face Database B (38 subjects).
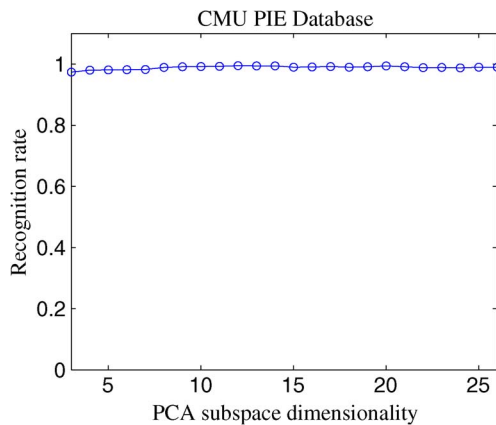

Fig. 12.   Recognition rates for CMU PIE database when using PCA generic intrinsic illumination subspaces.

Method	Number of Training Images	Error Rate (%)
Appearance	1	44.1
Histogram equalization	1	48.27
Quotient Image	3	34.29
Harmonic Relighting [25]	1	13.5
Model- and Examplar-based SFS [39]	1	5
5PL [9]	5	3.5
9PL [9]	9	1.9
DCT Truncation [37]	1	0.36
Ours (4D PCA)	1	1.99
Ours (12D PCA)	1	0.92
Ours (26D PCA)	1	0.71
Ours (4D NMF)	1	27.72
Ours (12D NMF)	1	3.41
Ours (26D NMF)	1	0.57

Fig. 13.   Comparison results for CMU PIE database.

illumination does not affect the saturation and hue composition of the image, only the gray-value distribution (shades of color) of the image. Given this assumption, an input image can be decomposed into H, S, and V channels, and what our method works with is simply V.

VII. CONCLUSION

In this paper, we have presented the concept of intrinsic illumination subspace, which can be well described by a low-dimensional linear subspace. When only objects of the same class are concerned, a single class-based generic intrinsic illumination subspace can be constructed in advance. We proposed

a lighting normalization method for a single input image by using the class-based subspace as a bootstrap set for novel objects. Dense correspondence is not necessary but can also be incorporated for better results if needed. Face recognition experiments showed that the proposed lighting normalization method improved the recognition rates significantly even when only several bases are used.

We assume that the object surfaces are convex and exhibit Lambertian reflectance. One limitation is that the effects of cast shadows and specularities are not accounted for. Applying the proposed method to other object classes will require relaxing the convex Lambertian surface assumption. Another possible future direction is considering shape variations among objects of the same class. When only objects of the same class are concerned, it is possible to reconstruct the object shape from a single input image. The reconstructed shape could significantly improve the lighting normalization results and is itself an important clue for identity recognition as well.

REFERENCES

[1] Y. Adini, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 721–732, Jul. 1997.
[2] H. Chen, P. Belhumeur, and D. Jacobs, "In search of illumination invariants," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2000, vol. 1, pp. 254–261.
[3] P. J. Phillips and Y. Vardi, "Efficient illumination normalization of facial images," *Pattern Recognit. Lett.*, vol. 17, no. 8, pp. 921–927, Jul. 1996.
[4] P. Belhumeur and D. Kriegman, "What is the set of images of an object under all possible lighting conditions?," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 1996, pp. 270–277.
[5] A. Georghiades, P. Belhumeur, and D. 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.
[6] 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.
[7] 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, Opt. Image Sci.*, vol. 18, no. 10, pp. 2448–2459, Oct. 2001.
[8] 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.
[9] 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. 5, pp. 684–698, May 2005.
[10] 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.
[11] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3D faces," in *Proc. 26th SIGGRAPH*, 1999, pp. 187–194.

[12] V. Blanz and T. Vetter, "Face recognition based on fitting a 3D morphable model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1063–1074, Sep. 2003.

[13] V. Blanz, K. Scherbaum, T. Vetter, and H. P. Seidel, "Exchanging faces in images," in *Proc. EUROGRAPHICS*, 2004, pp. 669–676.

[14] L. Zhang, S. Wang, and D. Samaras, "Face synthesis and recognition from a single image under arbitrary unknown lighting using a spherical harmonic basis morphable model," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2005, vol. 2, pp. 209–216.

[15] H. Barrow and J. Tenenbaum, "Recovering intrinsic scene characteristics from images," in *Computer Vision System*. New York: Academic, 1978.

[16] Y. Weiss, "Deriving intrinsic images from image sequences," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2001, vol. 2, pp. 68–75.

[17] J. Huang and D. Mumford, "Statistics of natural images and models," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 1999, vol. 1, pp. 541–547.

[18] T. Son and S. Mita, "Face recognition under variable lighting using the mean-field method and the gray-level pyramid," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2005, vol. 3, pp. 2107–2113.

[19] J. Huang, P. Yuen, W.-S. Chen, and J. H. Lai, "Choosing parameters of kernel subspace LDA for recognition of face images under pose and illumination variations," *IEEE Trans. Image Process.*, vol. 37, no. 4, pp. 847–862, Aug. 2007.

[20] T. Chen, W. Yin, X. S. Zhou, D. Comaniciu, and T. S. Huang, "Total variation models for variable lighting face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 9, pp. 1519–1524, Sep. 2006.

[21] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 1994, pp. 995–999.

[22] S. Shan, W. Gao, B. Cao, and D. Zhao, "Illumination normalization for robust face recognition against varying lighting conditions," in *Proc. IEEE Int. Workshop Anal. Model. Faces Gestures*, 2003, pp. 157–164.

[23] D. Frolova, D. Simakov, and R. Basri, "Accuracy of spherical harmonic approximations for images of Lambertian objects under far and near lighting," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 574–587.

[24] L. Qing, S. Shan, and W. Gao, "Eigen-harmonics faces: Face recognition under generic lighting," in *Proc. IEEE Int. Conf. Autom. Face Gesture Recog.*, 2004, pp. 296–301.

[25] L. Qing, S. Shan, W. Gao, and B. Du, "Face recognition under generic illumination based on harmonic relighting," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 19, no. 4, pp. 513–531, 2005.

[26] C.-P. Chen and C.-S. Chen, "Lighting normalization with generic intrinsic illumination subspace for face recognition," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2005, pp. 1089–1096.

[27] J. Begen and R. Hingorani, "Hierarchical motion-based frame rate conversion," David Sarnoff Res. Center, Princeton, NJ, 1990.

[28] Y. Pang, X. Li, Y. Yuan, D. Tao, and J. Pan, "Fast Haar transform based feature extraction for face representation and recognition," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 3, pp. 441–450, Sep. 2009.

[29] Y. Pang, X. Li, and Y. Yuan, "Robust tensor analysis with L1-norm," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 2, pp. 172–178, Feb. 2010.

[30] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, Oct. 1999.

[31] D. Lee and H. Seung, "Algorithms for non-negative matrix factorization," in *Proc. Adv. Neural Inf. Process. Syst.*, 2001, vol. 13, pp. 556–562.

[32] D. Jobson, Z. Rahman, and G. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997.

[33] R. Gross and V. Brajovic, "An image preprocessing algorithm for illumination invariant face recognition," in *Proc. Int. Conf. Audio Video-Based Biometric Person Authentication*, Jun. 2003, pp. 10–18.

[34] D. Jobson, Z. Rahman, and G. Woodell, "Properties and performance of a center/surround retinex," *IEEE Trans. Image Process.*, vol. 6, no. 3, pp. 451–462, Mar. 1997.

[35] S. Zhou, G. Aggarwal, R. Chellappa, and D. Jacobs, "Appearance characterization of linear Lambertian objects, generalized photometric stereo, and illumination-invariant face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 230–245, Feb. 2007.

[36] 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.

[37] W. Chen, M. J. Er, and S. Wu, "Illumination compensation and normalization for robust face recognition using discrete cosine transform in logarithm domain," *IEEE Trans. Image Process.*, vol. 36, no. 2, pp. 458–466, Apr. 2006.

[38] L. Zhang and D. Samaras, "Face recognition under variable lighting using harmonic image exemplars," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recog.*, 2003, pp. I-19–I-25.

[39] T. Sim and T. Kanade, "Combining models and exemplars for face recognition: An illuminating example," in *Proc. Workshop Models Versus Exemplars Comput. Vis.*, 2001.

[40] F. Nie, S. Xiang, and C. Zhang, "Neighborhood MinMax projections," in *Proc. Int. Joint Conf. Artif. Intell.*, 2007, pp. 993–998.

[41] R. Ramamoorthi, M. Koudelka, and P. Belhumeur, "A Fourier theory for cast shadows," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 288–295, Feb. 2005.
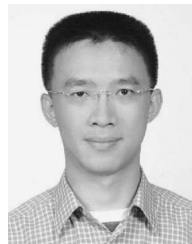
Chia-Ping Chen received the B.S. and M.S. degrees in computer science and information engineering from National Taiwan University, Taipei, Taiwan, in 2001 and 2003, respectively, where he is currently working toward the Ph.D. degree in computer science and information engineering.

His current research interests include computer vision, pattern recognition, and image processing.

Chu-Song Chen (S'94–M'96) received the B.S. degree in control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1989 and the M.S. and Ph.D. degrees from the Department of Computer Science and Information Engineering, National Taiwan University, in 1991 and 1996, respectively.

He is currently with Academia Sinica, Taipei, Taiwan, where he is the Deputy Director of the Research Center for Information Technology Innovation and a Research Fellow of the Institute of Information Science. He is also an Adjunct Professor with the Graduate Institute of Networking and Multimedia, National Taiwan University. He is on the editorial board of the *Journal of Multimedia* (Academy Publisher), *Machine Vision and Applications* (Springer), and *Information Processing Society of Japan Transactions on Computer Vision and Applications*. His research interests include pattern recognition, computer vision, signal/image processing, and multimedia analysis.

Dr. Chen has served as the Program Cochair of the 2005 International Conference on Digital Archive Technologies (ICDAT) and ICDAT 2006, the Theme Chair of the 2009 Pacific-Rim Symposium on Image and Video Technology, and the Area Chair of the 2009 and 2010 Asian Conference on Computer Vision and the 2010 International Conference on Network-Based Information Systems. In 2007-2008, he served as the Secretary-General of the Image Processing and Pattern Recognition Society, Taiwan, which is one of the regional societies of the International Association of Pattern Recognition.