

Lighting Normalization with Generic Intrinsic Illumination Subspace for Face Recognition

Chia-Ping Chen and Chu-Song Chen

Institute of Information Science, Academia Sinica, Taipei, Taiwan

{cpchen, song}@iis.sinica.edu.tw

Abstract

In this paper, we introduce the concept of Intrinsic Illumination Subspace which is based on the intrinsic images. This intrinsic illumination subspace enables an analytic generation of the illumination images under varying lighting conditions. When objects of the same class are concerned, our method allows a class-based generic intrinsic illumination subspace to be constructed in advance. We propose a lighting normalization method based on the generic intrinsic illumination subspace, which is used as a bootstrap subspace for novel images. Face recognition experiments are performed to demonstrate the effectiveness of our method.

1. Introduction

Lighting variation is one of the most difficult problems for vision applications. For face recognition, it has been observed that the variations among images of the same face due to illumination change are larger than image variations due to change in face identity [1]. Recently, many algorithms have been proposed to tackle lighting variations: Illumination Cone [4, 5, 6], Quotient Image [16, 17], Spherical Harmonic Subspace [3, 13], and Intrinsic Images [2, 19].

The Illumination Cone method [4, 5, 6] gave a theoretical explanation that face images under varying lighting directions form a convex polyhedral cone in the image space. Basri et al [3] and Ramamoorthi et al [13] independently developed the spherical harmonic representation, respectively. Their representation explains why low dimensional subspace can describe images of an object under varying lighting conditions.

While the above two approaches assumed that each individual has a different 3D geometry and constructed one subspace for each individual, the Quotient Image [16, 17], i.e. image ratio between a test image and linear combination of three images illuminated by non-coplanar lights, is a simple yet practical algorithm for extracting lighting invariant signatures. The quotient image depends only on the albedo information that is illumination invariant.

Still another approach dealing with lighting problem is from the intrinsic image viewpoint. Barrow and Tenen-

baum suggested that we take every appearance (retinal) image as a composition of a set of latent images, which they referred to as intrinsic images [2]. One type of the intrinsic images, R , represents the reflectance values of the object, while the other type, L , represents the illumination intensities, and their relationship can be described by $I(x, y) = R(x, y)L(x, y)$. See Fig. 1 for example. Barrow and Tenenbaum argued that such a midlevel description, despite not making explicit all the physical causes of image features, can be extremely useful for many visual inferences.

However, recovering two intrinsic images, L and R , from a single input image remains a difficult problem. This is a classic ill-posed problem: the number of unknowns is twice the number of equations. Weiss proposed a maximum-likelihood (ML) estimation method [19] 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 [8] and derived a single reflectance image $R(x, y)$ and T illumination images $L(x, y, t)$ such that

$$I(x, y, t) = R(x, y)L(x, y, t). \quad (1)$$

Nevertheless, this intrinsic image estimation method is offline since this method requires taking the median over the accumulated images to obtain the final intrinsic images. In addition, multiple images of the same object under variant lighting conditions were used, and thus the applicability of this method is restricted.

In this paper, we propose a novel concept called the *intrinsic illumination subspace*, which is defined as the set of illumination images (instead of appearance images used in illumination cone method) of an object under all possible lighting conditions. We analyze the intrinsic illumination subspace in terms of Lambertian model and demonstrate its relationship with the illumination cone and spherical harmonic bases. Based on this intrinsic illumination subspace, a lighting normalization method for a single input image is derived. The effectiveness of our lighting normalization



Figure 1: The intrinsic image decomposition.

method is demonstrated with face recognition experiments.

Our work shares similar motivations with that of Zhou et al [20], but the formulations are quite different. They extended the photometric stereo algorithms to handle all the appearances of all the objects in a class to derive albedo, 3D shape and lighting directions. To solve the highly ill-posed photometric stereo problem, they relied on various constraints, such as single light source, no shadow, symmetric and integrable surfaces. On the contrary, we don't impose these constraints and get significant rank reduction owing to the intrinsic image representation.

This paper is organized as follows: Section 2 depicts the concept of intrinsic illumination subspace. The proposed lighting normalization method is described in Section 3. Section 4 shows experiment results and conclusion is made in Section 5.

2. Intrinsic Illumination Subspace

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, which describes the fraction of the light reflected. We also assume the convex object is illuminated by distant light sources. What we mean by “distant” is that the directions and intensities of light sources are the same for all points of this object.

2.1. Definition of Intrinsic Illumination Subspace

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

$$I(x, y) = R(x, y)n(x, y) \cdot l. \quad (2)$$

When an object is illuminated by k lights instead of only one, the image is given by the sum of the contribution of each light as follows:

$$I(x, y) = R(x, y) \sum_{i=1}^k n(x, y) \cdot l_i, \quad (3)$$

where l_i denotes the intensity and direction of each light source.

Lambertian reflectance function can be thought as a special version of intrinsic images, where R stands for a view-independent reflectance (albedo) value, and L is the shading of a Lambertian surface:

$$L(x, y) = \sum_{i=1}^k n(x, y) \cdot l_i. \quad (4)$$

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

$$L(x, y) = \sum_{i=1}^k \max(n(x, y) \cdot l_i, 0) \quad (5)$$

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

$$\mathcal{L} = \{ L \mid L(x, y) = \sum_{i=1}^k \max(n(x, y) \cdot l_i, 0), \\ \forall l_i \in R^3, \forall k \in Z^+, \} \quad (6)$$

where Z^+ is the set of positive integers.

Previous works [4, 5, 6] have shown that the set of appearance images of an object under all possible lighting conditions forms a convex polyhedral cone in the image space. Therefore a question arises that: 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. Let \mathcal{C} and \mathcal{L} be the illumination cone and intrinsic illumination subspace of the same convex Lambertian object, respectively, then

$$\mathcal{C} = \{R \odot L : L \in \mathcal{L}\} \quad (7)$$

$$\mathcal{L} = \{R^{-1} \odot I : I \in \mathcal{C}\}, \quad (8)$$

where R is the reflectance of the object, $R^{-1}(x, y) = 1/R(x, y)$, and \odot indicates 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, n , of the intrinsic illumination subspace equals to the number of distinct surface normals.

Although the intrinsic illumination subspace, \mathcal{L} , can span n dimensions if there are n distinct surface normals, previous empirical observations and analytical predictions have shown that appearance images of largely diffuse objects actually lie very close to a low dimensional subspace. Their results can also apply to our intrinsic illumination subspace as explained as follows.

According to the spherical harmonic representation in [3, 13], an image I of an object can be represented by

$$I = R \odot (Yl), \quad (9)$$

where R is the reflectance matrix, $Y = [y_1, y_2, \dots, y_n]$ is the spherical harmonic bases, and l is the harmonic light. According to their finding, the first 9 harmonic bases can well describe the images of a diffuse object under different lighting conditions. It follows that the illumination image L can be well represented by the first 9 harmonic bases:

$$L = Yl. \quad (10)$$

Because these bases must be calculated with known 3D geometry, the application range of this representation is limited. According to Ramamoorthi's analysis [14], there is linear relationship between 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:

$$B = YT, \quad (11)$$

where T is a $n \times n$ transformation matrix.

Then equation (10) can be written as

$$L = Yl = Bs, \quad (12)$$

where s is the projected coefficients 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 a densely sampled illumination images under varying lighting conditions, we can get well approximated linearly-transferred spherical harmonic bases, B , without estimating the 3D geometry of this object.

To verify this, Fig. 2a shows 64 illumination images of one person from Yale Face Database B, where the illumination images are derived by Weiss's ML estimation method [19]. The eigen bases and cumulative eigen ratio are shown in Fig. 2b and 2c. In practices, the first 9 eigenvectors can not carried over 95% of image energy for a non-convex face object because there are obvious cast shadows in non-frontally illuminated images. Therefore more eigenvectors are needed for the implementation. In this case, the eigen ratio is 95.27% when 12 eigenvectors are used.

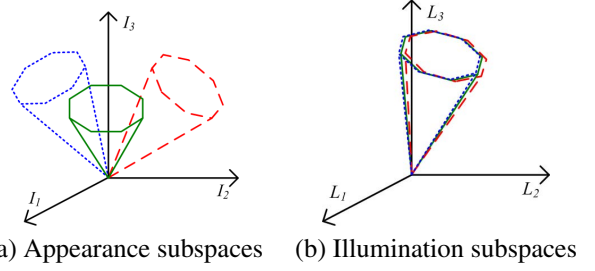


Figure 3: 3D example of generic intrinsic illumination subspace.

2.2. Generic Intrinsic Illumination Subspace

The previous subsection gives the analysis of the intrinsic illumination subspace of one object. We will extend this concept to *generic* intrinsic illumination subspace of multiple objects of the same class, such as human faces.

It is very difficult to model lighting variations for general objects. However, for many vision applications, only objects of the same class are concerned. In this case, some class-based information can be utilized to simplify this problem. Shashua [16, 17] defined an *ideal class* to be a collection of 3D objects that have the same shape but differ in the surface albedo function. So the appearance image space of such a class is represented by:

$$R_i(x, y) \sum_{i=1}^k \max(n(x, y) \cdot l_i, 0), \quad (13)$$

where R_i is the reflectance of object i of this class, and $n(x, y)$ is the surface normal of the object (the same for all objects of this class).

Objects of a class have different albedo functions, so that they span different linear subspaces in the image space (Fig. 3a). However, by removing the reflectance factor from (13), the resulting illumination images only depend on surface normals and lighting conditions. This implies that, given the ideal class assumption, all objects of the same ideal class shares the same *generic* intrinsic illumination subspace (Fig. 3b).

In practice, objects of a class do have shape variations, although the shape is similar at some coarse level or we would not refer to them as a "class." The ideal class could be satisfied if we perform pixel-wise dense correspondence between images. The question is what the degree of sensitivity of our approach to deviations from the ideal class assumption is. Results demonstrate that one can withstand shape changes without noticeable degradation in performance. So there is no need to establish any dense alignment among the images beyond the alignment of the center of mass and scale.

To verify this assumption, we conducted another experiment. We decompose 64 appearance images of each of

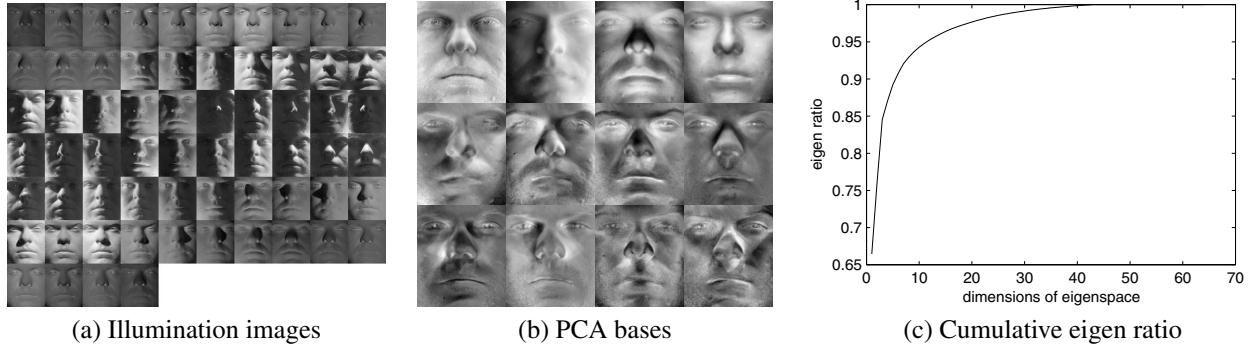


Figure 2: Intrinsic illumination subspace of one person from Yale Face Database B.

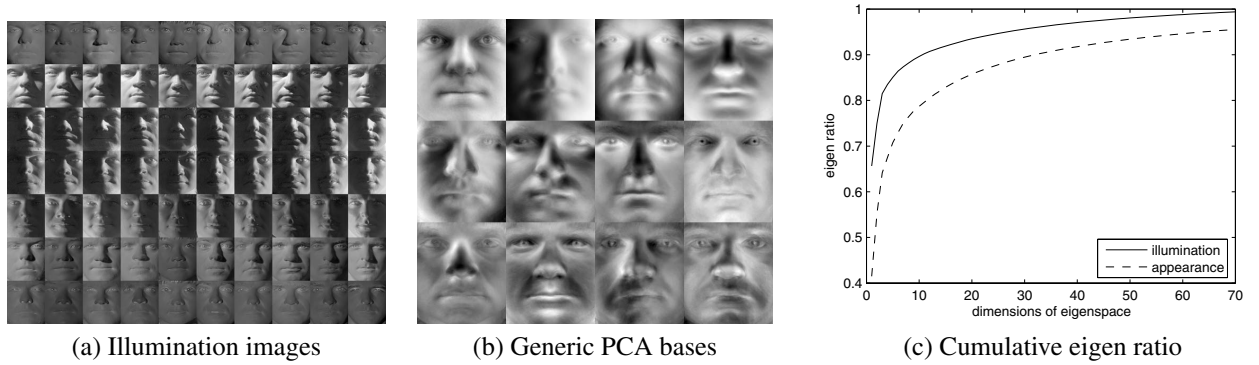


Figure 4: Generic intrinsic illumination subspace of 10 people from Yale Face Database B

the 10 subjects from Yale Face Database B into a single reflectance and 64 illumination images. Then SVD is performed to all 640 appearance images and 640 illumination images respectively. Fig. 4a shows 7 of 64 illumination images of each subject. The resulting generic eigen bases of illumination images are shown in Fig. 4b. Fig. 4c shows the cumulative eigen ratio 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. But 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. So a low dimensional linear subspace is capable of modeling the generic illumination subspace of human faces. On the other hand, a much higher dimensional subspace is required for appearance images.

2.3. Enforcing Non-negative Light with NMF

When we take arbitrary linear combinations of the basis images, such as PCA, we may obtain images that are not phys-

ically realizable. This is because the corresponding linear combination of the basis images may contain negative values. That is, rendering these images may require negative “light,” which is realistically impossible. In our approach, we use *non-negative matrix factorization* (NMF) to enforce the constraint of non-negative light.

NMF is a subspace method proposed by Lee and Seung [11], which has been used for image representation, document analysis and clustering for its parts-based representation property. Given a non-negative $m \times n$ matrix X , the NMF algorithms seek to find non-negative factors B and H of \tilde{X} such that:

$$X \approx \tilde{X} = BH, \quad (14)$$

where $B \in R^{m \times r}$, $H \in R^{r \times n}$.

Intuitively, we think of B as the matrix containing the NMF bases where all values are non-negative, and H as the matrix containing the accompanying coefficients (non-negative weights).

The non-negative B captures representative images under certain lighting conditions and can be used as linear bases to generate illumination images under all possible lighting conditions. The non-negative constraint on B explicitly models that the observed intensities of an image can



Figure 5: Generic NMF Bases.

not be negative. The non-negative restriction on the accompanying coefficients in H results in the additive and non-negative nature of lighting.

Non-negative matrix factorizations can be very difficult to compute. Lee and Seung [12] suggested an approach similar to that used in Expectation-Maximization (EM) algorithms to iteratively update the factorization based on a given objective function. We adopt their method to construct the generic intrinsic illumination subspace. Fig. 5 shows the resulting NMF bases of the generic intrinsic illumination subspace from Yale Face Database B.

3. Lighting Normalization with Intrinsic Illumination Subspace

With the concept of the intrinsic illumination subspace, we propose a lighting normalization method for a single input image in Section 3.1. The basic idea is to estimate the intrinsic illumination image L of the input image I by the intrinsic illumination subspace, which can be a generic one constructed from a pre-collected database containing objects of the same class. Then we can get an estimated reflectance image by dividing the input image with the generated illumination image: $R = I/L$. This estimated reflectance image is much more illumination invariant than the original input image, so it is more suitable for various vision applications.

3.1. The Proposed Lighting Normalization Method

Some previous works [7, 9, 10] 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 to solve this ill-posed problem. Basri [3] also gave a theoretical way of understanding the effects of Lambertian reflectance as that of a low-pass filter on lighting.

In our approach, we use this low-frequency component

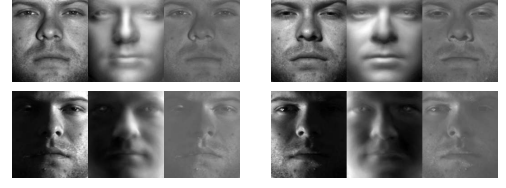


Figure 6: The input, reconstructed illumination, and estimated reflectance images.

as our initialization:

$$L^* = F \star I, \quad (15)$$

where I is the input image, F is the smoothing kernel (a Gaussian smoothing kernel is used in our experiments), \star represents convolution, and L^* is the smoothed version of I .

This initial L^* is used as a query. Best approximation of the corresponding illumination image L is reconstructed from the intrinsic illumination subspace according to the Euclidean distance:

$$\begin{aligned} l &= \arg \min_l \|Bl - L^*\| \\ L &= Bl \end{aligned} \quad (16)$$

where B is the basis matrix of the intrinsic illumination subspace.

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

$$R = I/L. \quad (17)$$

Figure 6 shows the process of the proposed lighting normalization method.

Our lighting normalization method shares some similarities with Quotient Image method [16, 17]. Both methods maintain a bootstrap set to estimate the illumination (lighting direction) of the input image and normalize the input image according to the estimated result. But there are some significant differences between our method and Quotient Image:

1. The bootstrap set of Quotient Image consists of appearance images. Since there are large reflectance variations among different people, the estimation of lighting direction is 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.
2. Quotient Image works under the assumption that face image are illuminated with a single point light source

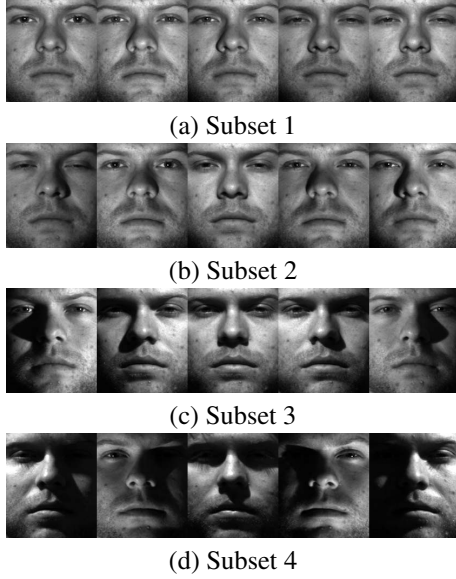


Figure 7: Four Subset of Yale Face Database B.

and there is no shadow, while our method allows multiple light sources and attached shadows are explicitly considered.

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.

3.2. Discussions

Specific or Generic: 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. 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.

Cast Shadows: Cast shadows can be significant in many vision applications. Most algorithms neglect them because nonlocal interactions in non-convex regions make formal analysis difficult. Ramamoorthi et al [15] 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 Heaviside step function. The eigenvalues decay as $1/k$ for Heaviside step function, which is relatively slow comparing to $1/k^2$ for the clamped cosine function (Lambertian reflectance). It suggests that it is possible to develop,



Figure 8: CMU PIE Database.

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

4. Experiments

Experiments were performed to evaluate our method using Yale Face Database B [6] and CMU PIE face database [18]. Frontal face images with lighting variations are selected from these two databases. There are 640 images (10 subjects with 64 images under different lighting conditions for each) from Yale Face Database B. The images are divided into 4 subsets of increasing illumination angles as shown in Fig. 7. There are 68 subjects in CMU PIE and we select frontal face images which are taken under 21 different illuminations without background lighting (Fig. 8). All the images are aligned and cropped roughly by the positions of eyes and mouths.

To construct the generic intrinsic illumination subspace, we randomly select half of 64 images for each subject from Yale Face Database B. These images are first decomposed into reflectance and illumination images by Weiss's method [19]. With all the illumination images, we construct the generic intrinsic illumination subspaces by PCA and NMF, respectively. We only use the first 3 PCA bases to examine the effectiveness of our method. The dimensionality of the NMF intrinsic illumination subspace is 12. The more bases we use to construct the subspace, the better the result is expected to be, but with more computation time.

We also implement the Quotient Image method for comparison. Three appearance images of each subject from Yale Face Database B are selected as the bootstrap set for Quotient Image. A simple recognition scheme, correlation, is used in the following experiments, and only gray-level images are used. Both the template images and test images are normalized by Quotient Image or our method described in Section 3.1. The result of correlation with original appearance images is also shown as the baseline for comparison.

We first perform recognition experiments on the Yale

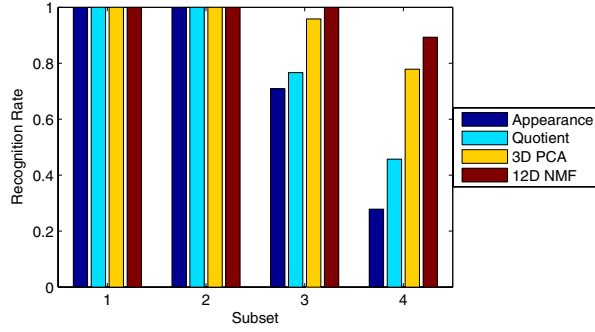


Figure 9: Recognition result for Yale Face Database B.

Method	Recognition Rate
Appearance	55.9%
Quotient Image	65.71%
3D PCA	95.1%
12D NMF	97.0%

Figure 10: Recognition result for CMU PIE.

Face Database B. Only the frontal illuminated images are labeled with the subject's identity and used as templates. Fig. 9 shows the recognition results of different subset. Even with this simple correlation scheme, our lighting normalization method improves the recognition rates significantly. Note that only the first 3 bases of the PCA intrinsic illumination subspace are used. The improvement is impressive. It indicates that our PCA intrinsic illumination subspace is suitable for real-time application because of its simplicity and efficiency.

We then perform recognition experiments on CMU PIE database. Note that the generic intrinsic illumination subspace constructed from Yale Face Database B is used in order to verify its effectiveness for novel images from CMU PIE. 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 are shown in Fig. 10. We can see that although the generic intrinsic illumination subspace is constructed from the 10 subjects of Yale Face Database B, it is also valid for the 68 novel subjects of CMU PIE.

5. Summary and Conclusions

The concept of intrinsic illumination subspace is presented. The illumination images of an object in fixed pose but under all lighting conditions forms a convex polyhedral cone in the image space and can be well described by a low dimensional linear subspace. A single class-based generic intrinsic illumination subspace can be constructed in advance when only objects of the same class are concerned.

This generic intrinsic illumination subspace can be used as a bootstrap set by our lighting normalization method for a single input image. Face recognition experiments were performed to verify the effectiveness of our method. Experimental results show that our lighting normalization method improves the recognition rates significantly even when only 3 bases are used, and the generic intrinsic illumination subspace can be used for novel objects effectively.

Acknowledgments

This work was supported in part under grants NSC 93-2213-E-001-011, NSC 94-2752-E-002-007-PAE, and 94-EC-17-A-02-S1-032.

References

- [1] Y. Adnin, Y. Moses, and S. Ullman, "Face recognition: The problem of compensating for changes in illumination direction," *IEEE Trans. on PAMI*, Vol. 19, No. 7, pp. 712-732, 1997.
- [2] H.G. Barrow and J.M. Tenenbaum, "Recovering Intrinsic scene Characteristics from Images," *Computer Vision System*, pp. 3-26, 1978
- [3] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces," *IEEE Trans. on PAMI*, Vol. 25, No. 2, pp. 218-233, 2003.
- [4] P.N. Belhumeur, D.J. Kriegman, "What is the set of Images of an Object Under All possible Lighting Conditions?" *Proc. of IEEE CVPR*, 1996
- [5] A.S. Georgiades and P.N. Belhumeur, "Illumination cone models for Faces recognition under variable lighting," *Proc. of IEEE CVPR*, 1998
- [6] A.S. Georgiades and P.N. Belhumeur, "From Few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Trans. on PAMI*, Vol. 23, No. 6, pp. 643-660, 2001
- [7] R. Gross, and V. Brajovie, "An Image Propocessing Algorithm for Illumination Invariant Face Recognition," *Proc. of 4th Intl. Conf. on AVBPA*, pp. 10-18, 2003.
- [8] J. Huang and D. Mumford, "Statistics of Natural Images and Models," *Proc. of IEEE CVPR*, 1999.
- [9] D.J. Jobson, S.U. Rahman, and G.A. Woodell, "Properties and Performance of a Center/Surround Retinex," *IEEE Trans. on IP*, Vol. 16, No. 3, pp. 451-462, 1997.
- [10] D.J. Jobson, S.U. Rahman, and G.A. Woodell, "A Multiscale Retinex for Bridging the Gap Between Color Images and the Human Observation of Scenes," *IEEE Trans. on IP*, Vol. 16, No. 7, pp. 965-976, 1997.

- [11] D.D. Lee and H.S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature* 401, pp. 788-791, 1999.
- [12] D.D. Lee and H.S. Seung, "Algorithms for non-negative matrix factorization," *Proc. of NIPS*, 2000.
- [13] R. Ramamoorthi and P. Hanrahan, "On the relationship between radiance and irradiance: determining the illumination from images of a convex Lambertian object," *JOSA A*, Vol. 18, Issue 10, pp.2448V2459, 2001.
- [14] R. Ramamoorthi, "Analytic PCA Construction for Theoretical Analysis of Lighting Variability in Images of a Lambertian Object," *IEEE Trans. on PAMI*, Vol. 24, No. 10, 2002.
- [15] R. Ramamoorthi, M. Koudelka, and P. Belhumeur, "A Fourier Theory for Cast Shadows," *IEEE Trans. PAMI*, Vol. 27, No. 2, 2005.
- [16] T. Riklin-Raviv and A. Shashua, "The Quotient image: Class based recognition and synthesis under varying illumination," *Proc. of IEEE CVPR*, 1999.
- [17] A. Shashua, and T. Riklin-Raviv, "The quotient image: Class-based re-rendering and recognition with varying illuminations," *IEEE Trans. on PAMI*, Vol. 23, No. 2, pp 129-139, 2001.
- [18] T. Sim and S. Baker and M. Bsat, "The CMU Pose, Illumination, and Expression (PIE) Database," *Proc. of IEEE FG*, 2002.
- [19] Y. Weiss, "Deriving Intrinsic Images from Image Sequences," *Proc. of IEEE ICCV*, 2001.
- [20] S.K. Zhou, R. Chellappa, and D.W. Jacobs, "Characterization of human faces under illumination variations using rank, integrability, and symmetry constraints," *Proc. of ECCV*, 2004.