Object Recognition Based on Photometric Alignment Using RANSAC

Takahiro OKABE and Yoichi SATO
Institute of Industrial Science, University of Tokyo
4-6-1, Komaba, Meguro-ku, Tokyo 153-8505, Japan
{*takahiro, ysato*}*@iis.u-tokyo.ac.jp*

Abstract

For object recognition under varying illumination conditions, we propose a method based on photometric alignment. The photometric alignment is known as a technique that models both diffuse reflection components and attached shadows under a distant point light source by using three basis images. However, in order to reliably reproduce these components in a test image, we have to take into account outliers such as specular reflection components and shadows in the test image. Accordingly, our proposed method utilizes RANdom SAmple Consensus (RANSAC), which has been used successfully for estimating basis images. In the present study, we have conducted experiments using the Yale Face Database B and confirmed that a combination of the photometric alignment and RANSAC provides a simple but effective method for object recognition under varying illumination conditions.

1. Introduction

The appearance of an object depends upon poses of the object, viewpoints, and illumination conditions. Therefore, object recognition based on computer vision suffers from three problems: variations due to changes in pose, viewpoint, and illumination. In this work, we focus on the problem of object recognition under varying illumination conditions, provided that both the pose and the viewpoint are fixed.

When we consider only primary reflection from a point light source and disregard secondary reflection such as interreflection, we can classify image pixels into four components: diffuse reflection, specular reflection, attached shadow, and cast shadow [26]. For object recognition under varying illumination conditions, *generative methods* have been developed [26, 4, 11, 10, 1, 15, 13, 2] in which diffuse reflection components (and shadows) under varying illumination conditions are synthesized on the assumption of the Lambertian model.

Assuming the Lambertian model, diffuse reflection components under a distant point light source can be represented by a linear combination of three basis images. Here, pixels where the diffuse reflection components represented by the linear combination become negative correspond to attached shadows. Therefore, attached shadows can be also synthesized by replacing negative pixel values with zeros. A technique proposed by Shashua [26] that models both diffuse reflection components and attached shadows under a distant point light source by using three basis images is called the *photometric alignment*.

Object recognition based on the photometric alignment has two processes: estimation of three basis images (training process) and reproduction of a test image (recognition process). If images consist of only diffuse reflection components, the two processes are self-evident. Namely, we can use three images under distant point light sources placed in linearly independent directions as the basis images. The coefficients of linear combination for reproducing the test image can be obtained by projecting the test image onto each of the basis images. However, images generally contain outliers such as specular reflection components and shadows. Therefore, we have to take into account the outliers in both the training and recognition processes.

The training process is closely related to photometric stereo [29] because three basis images of an object are determined by the albedo and normal of the object surface. In the context of photometric stereo, object recognition, and image synthesis, various robust methods have been proposed by Georghiades et al. [11], Yuille et al. [30], Mukaigawa et al. [18] and Ishii et al. [14] for estimating the basis images from training images containing outliers. These methods are based on Singular Value Decomposition with Missing Data (SVDMD) [27], SVDMD with the integrability constraint, voting, and RANSAC [9] respectively. In addition, Nakashima et al. [20] have proposed a method for estimating the basis images from a sequence of images of an object in motion.

On the other hand, the recognition process has not been sufficiently examined until now. Although robust estima-

Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'03)
1063-6919/03 $17.00 © 2003 IEEE

tion techniques have already been used by Black and Jepson [6] and Leonardis and Bischof [16], they deal with other types of outliers, typically occlusions.

In the present paper, we discuss a recognition process that is robust against outliers such as specular reflection components and shadows in test images. In particular, in order to estimate the coefficients of linear combination reliably, we utilize RANSAC that has been used successfully for estimating the basis images [14, 20]. Then, we compare synthesized images with a test image, and classify the test image based on the number of correctly reproduced pixels.

To demonstrate the effectiveness of our proposed method, we have conducted experiments using the Yale Face Database B [10]. The main contribution of the present work is to demonstrate experimentally that a combination of the photometric alignment and RANSAC provides a simple but effective method for object recognition under varying illumination conditions.

The rest of the paper is organized as follows. In Section 2, we summarize related works and describe the advantages of our method. Then, in Section 3, we introduce the photometric alignment in detail and clarify the problem of object recognition based on the photometric alignment. In Section 4, we propose a recognition method utilizing RANSAC. We report the results of experiments conducted by using the Yale Face Database B in Section 5. Finally, in Section 6, we present concluding remarks.

2. Previous work

With regard to the problem of object recognition under varying illumination conditions, three distinct approaches have been developed: *feature-based methods*, *appearance-based methods*, and generative methods.

The first approach is based on features that are insensitive to changes in illumination. However, features such as edges and corners cannot always be extracted stably. Moreover, information essential for recognition may be lost by utilizing a part of images [7]. Recently, Chen et al. [8] have shown that the direction of image gradient is insensitive to changes in illumination direction.

The second approach applies algorithms of pattern recognition in which all pixel values are used as inputs. It has been shown that these appearance-based methods are effective for object recognition under varying poses and/or illumination conditions [28, 19, 3, 22, 25]. However, in order to recognize an object under a certain imaging condition, they need training images of the object taken under similar imaging conditions.

In contrast to feature-based methods and appearance-based methods, generative methods [26, 4, 11, 10, 1, 15, 13, 2] model diffuse reflection components (and shadows) under varying illumination conditions on the assumption of the Lambertian model. Therefore, they are effective even when illumination conditions of the test images differ greatly from those of the training images. In the rest of this section, we review each of the generative methods briefly.

In the case where images consist of only diffuse reflection components, it is known that images under varying illumination conditions can be represented by linear combinations of three basis images [26]. The recognition method based on the distance between a test image and three-dimensional subspace spanned by the basis images is called the linear subspace method [3]. It has been shown that this method is effective only when diffuse reflection components are dominant in the test image [11, 10].

On the other hand, in the case where images consist of diffuse reflection components and attached shadows, that is, in the case of a convex Lambertian object, Belhumeur and Kriegman [4] have shown that a set of images of the object under arbitrary illumination conditions forms a convex cone, called the illumination cone, in the image space. The illumination cone is represented by convex combinations of the extreme images that are generated from three basis images. It has been shown that the recognition method based on the distance between a test image and the illumination cone is effective even when attached shadows are dominant [11, 10]. Moreover, the model has been extended to consider cast shadows [11, 10] by restoring the shape and albedo of the object [5, 30].

The illumination cone model is superior to the linear subspace method in that the model can generate not only diffuse reflection components but also attached shadows. However, it is difficult to apply the model to object recognition in practice because a large number of extreme images are necessary to represent the illumination cone. To cope with this problem, two approaches have been proposed.

One approach is to represent the illumination cone approximately by using a small number of basis images. It has been shown experimentally that face images under varying illumination conditions are approximately represented by linear combinations of a small number of basis images [12]. Based on this observation, the illumination cone is represented approximately by applying Principal Component Analysis (PCA) [11, 10].

Moreover, Basri and Jacobs [1] and Ramamoorthi and Hanrahan [23, 24] have shown recently that, based on the analysis in the frequency domain using spherical harmonics, the illumination cone of a convex Lambertian object can be represented approximately by using 4 to 9 basis images. It has been shown that the recognition method based on the distance between a test image and the subspace spanned by the basis images performs as well as the illumination cone model [15, 13]. However, the basis images correspond to images of the object under illuminations whose distributions are the same as spherical harmonics with low frequencies. Therefore, it is necessary to synthesize the basis images by restoring the shape and albedo of the object [1] or to approximate them by using images under point light sources placed in special directions [15, 13].

The other approach is to improve the linear subspace method. Supposing point light sources distributed randomly around a convex object, regions of the object surface with similar normals are illuminated by similar point light sources. Based on this observation, Batur and Hayes [2] have proposed the segmented linear subspace method in which images are segmented into regions whose surface normals are close to one another and then the linear subspace method is applied to each of the regions. Although it has been reported that the performance of the method is nearly equal to that of the illumination cone model, it is necessary to examine the way of determining the number of segmented regions appropriately.

Our proposed method based on the photometric alignment has several distinct advantages over the previously proposed generative methods. For instance, our method is superior to the linear subspace method [26, 3] in the point that our method can represent not only diffuse reflection components but also attached shadows. Unlike the illumination cone model [4, 11, 10], our method does not require a large number of extreme images. In addition, as opposed to the method based on the analysis in the frequency domain [1, 15, 13], our method does not need the shape and albedo of objects or training images taken under special lighting directions. Finally, our method does not require appropriate segmentation of images, unlike the segmented linear subspace method [2]. Another important merit of our method is that our method can generate diffuse reflection components and attached shadows accurately while the other methods [1, 15, 13, 2] consider only the approximation of these components.

Unfortunately, our method also has limitations. Because the photometric alignment assumes a point light source, our method cannot be applied to images taken under complex illumination conditions. However, our method should be extended to the scene under complex illumination by using two images: one taken under a complex illumination, and the other taken under the complex illumination and a distant point light source prepared in advance. In this case, we can regard the difference of the two images as an input image.

3. Photometric alignment

In this section, we introduce representation of diffuse reflection components and attached shadows based on the photometric alignment in detail. Then, we clarify the problem of the photometric alignment when it is applied to object recognition under varying illumination conditions.

3.1. Representation of diffuse reflection components

First, we consider the case where images of an object consist of only diffuse reflection components.

Let us assume the Lambertian model and a distant point light source. Then, the i-th pixel value in an image with n pixels denoted by $x_i^{(D)}$ is given by an inner product as

$$x_i^{(D)} = \rho_i \boldsymbol{n}_i^T \boldsymbol{s} \equiv \boldsymbol{b}_i^T \boldsymbol{s},\ (i = 1, 2, \cdots, n), \tag{1}$$

where ρ_i and $\boldsymbol{n}_i$ are the albedo and normal of the object surface corresponding to the i-th pixel respectively, and $\boldsymbol{s}$ is a vector whose direction and norm are equal to the direction and strength of the light source[1]. Therefore, the image denoted by $\boldsymbol{x}^{(D)}$ is represented by

$$\boldsymbol{x}^{(D)} = B\boldsymbol{s}, \tag{2}$$

where B is a $n \times 3$ matrix whose i-th row is given by $\boldsymbol{b}_i^T$.

Equation (2) means that arbitrary images lie in a linear subspace defined by

$$\mathcal{L} = \{\boldsymbol{x} | \boldsymbol{x} = B\boldsymbol{s}, \forall \boldsymbol{s} \in \mathrm{R}^3\} \tag{3}$$

in the image space. The dimension of the subspace $\mathcal{L}$ is equal to the rank of the matrix B, that is, three except for special cases such as planer and cylindrical objects. Hence, an image consisting of only diffuse reflection components is represented by a linear combination of three basis images $\boldsymbol{e}^{(j)}$ $(j = 1, 2, 3)$ taken under point light sources placed in linearly independent directions as

$$\boldsymbol{x}^{(D)} = c_1 \boldsymbol{e}^{(1)} + c_2 \boldsymbol{e}^{(2)} + c_3 \boldsymbol{e}^{(3)}, \tag{4}$$

where c_j $(j = 1, 2, 3)$ are coefficients of linear combination.

A technique that models diffuse reflection components as linear combinations of three basis images is called the linear subspace method [3]. We should note that the linear subspace method simplifies the photometric alignment by neglecting attached shadows.

3.2. Representation of attached shadows

Secondly, we consider the case where images of an object consist of diffuse reflection components and attached shadows.

A point on the object is in attached shadow if the angle between the surface normal and the direction of a point light source is obtuse. Then, pixels where inner products of equation (1) become negative $(\boldsymbol{b}_i^T \boldsymbol{s} < 0)$ correspond to

[1] We focus on variations due to changes in illumination direction. Thus, we do not consider the dependence on the spectral response of a camera and the spectral distribution of light sources.

attached shadows. Therefore, an image $\boldsymbol{x}^{(D+AS)}$ consisting of diffuse reflection components and attached shadows is represented by

$$\boldsymbol{x}^{(D+AS)} = \max\left(B\boldsymbol{s}, \boldsymbol{0}\right), \tag{5}$$

where $\max(\boldsymbol{z}, \boldsymbol{0})$ replaces negative components of a vector $\boldsymbol{z}$ with zeros.

In the same way, the image is represented by using the three basis images as

$$\boldsymbol{x}^{(D+AS)} = \max\left(\sum_{j=1}^{3} c_j \boldsymbol{e}^{(j)}, \boldsymbol{0}\right). \tag{6}$$

A technique that models diffuse reflection components and attached shadows by using three basis images is called the photometric alignment [26].

3.3. Application to object recognition

The photometric alignment shows that diffuse reflection components and attached shadows of an object under a point light source placed in a novel direction can be synthesized if three basis images of the object are prepared in advance. Shashua [26] pointed out that the photometric alignment is applicable to object recognition under varying illumination conditions.

However, training and test images generally contain outliers such as specular reflection components and shadows. Then, when we estimate three basis images by, for instance, applying SVD to three or more training images, the estimated basis images are affected by the outliers. Moreover, if we calculate coefficients of linear combination for synthesizing a test image by, for instance, projecting the test image onto each of the basis images, we cannot obtain the correct coefficients. Therefore, we have to deal with the outliers carefully in both the training and recognition processes. In addition, it is necessary to examine the classification method after synthesizing images from the basis images.

4. Proposed method

In the context of photometric stereo, object recognition, and image synthesis, various robust methods have already been proposed for estimating three basis images [11, 30, 18, 14, 20]. Accordingly, in this section, we propose the rest of process for object recognition under varying illumination conditions. Namely, we propose a recognition process that is robust against outliers such as specular reflection components and shadows in test images.

4.1. Outline of our proposed method

Let us consider the case where a test image and three basis images belong to the same class, that is, they are images of the same object. Generally, the test image contains not only diffuse reflection components but also specular reflection components, attached shadows, and cast shadows. Among these components, diffuse reflection components are represented by a linear combination of the basis images as equation (4). Accordingly, we can compute the coefficients of linear combination from pixels in the test image that correspond to diffuse reflection components. Then, we can synthesize an image consisting of diffuse reflection components and attached shadows from the estimated coefficients and the basis images as equation (6). Here, if a pixel in the test image corresponds to a diffuse reflection component or an attached shadow, its pixel value is equal to that in the synthesized image. In particular, in the case of a convex Lambertian object, the test image is identical with the synthesized image.

On the other hand, when a test image and three basis images belong to different classes, that is, different objects, it is expected that we will encounter difficulty in reproducing the test image from the basis images.

Motivated by the above observations, we propose the following recognition method. For each class, we estimate the coefficients of linear combination by using RANSAC so that diffuse reflection components and attached shadows in a test image are synthesized. Then, we classify the test image based on the number of correctly reproduced pixels.

4.2. Recognition process

Here, we describe in detail our proposed method utilizing RANSAC.

For each class, we estimate the number of correctly reproduced pixels as follows.

1. **Compute coefficients of linear combination**
   Suppose that three pixels randomly chosen in a test image correspond to diffuse reflection components, compute coefficients of linear combination $\hat{c}_j$ $(j = 1, 2, 3)$.

2. **Synthesize image and label each pixel**
   Synthesize an image consisting of diffuse reflection components and attached shadows from the computed coefficients and three basis images as

$$\hat{x}_i = \max\left(\sum_{j=1}^{3} \hat{c}_j e_i^{(j)}, 0\right), \ (i = 1, 2, \cdots, n), \tag{7}$$

   where $e_i^{(j)}$ is the i-th pixel value in the j-th basis images. Then, assuming that a pixel is correctly reproduced if the difference between pixel values in the test image and the synthesized image is less than a predefined threshold t, label each pixel as

$$\xi_i = \begin{cases} 1 & (|x_i^{(test)} - \hat{x}_i| < t : inlier) \\ 0 & (|x_i^{(test)} - \hat{x}_i| \geq t : outlier) \end{cases} \quad (8)$$

where $x_i^{(test)}$ is the i-th pixel value in the test image. The threshold t should depend on deviation of the actual diffuse reflection components from the ideal Lambertian model (see Oren and Nayar [21] for instance).

3. **Count the number of inliers**
   Count the number of inliers and define it as

$$support = \sum_{i=1}^{n} \xi_i. \quad (9)$$

4. **Repeat procedures 1 through 3**
   Repeat procedures 1 through 3, and regard the coefficients that maximize $support$ as the correct ones.

5. **Improve coefficients by least squares**
   Improve the estimated coefficients by applying least-squares method to pixels that are considered to correspond to diffuse reflection components. Namely, remove a pixel from the set of inliers if the pixel is considered to be in an attached shadow or saturated according to the following criteria with threshold t, and assign weights as [2]

$$w_i = \begin{cases} 1 & (\xi_i = 1, t < \hat{x}_i < 1 - t) \\ 0 & (others) \end{cases} . \quad (10)$$

   Then, compute the coefficients $\hat{c}_j$ $(j = 1, 2, 3)$ that minimize the following function

$$\mathcal{C} = \sum_{i=1}^{n} w_i \left(x_i^{(test)} - \sum_{j=1}^{3} \hat{c}_j e_i^{(j)} \right)^2 . \quad (11)$$

6. **Synthesize image and label each pixel**
   By using the improved coefficients, synthesize an image and label each pixel in the synthesized image in the same way as in procedure 2.

7. **Repeat procedures 5 and 6**
   Repeat procedures 5 and 6 until the label $\xi_i, (i = 1, 2, \cdots, n)$ converges.

8. **Count the number of reproduced pixels**
   Count up $support$ again.

[2] We assume that the pixel value $\hat{x}_i$ is normalized to take value between 0 and 1, for instance, by dividing it by 255 when it is represented as 8 bit integer.

Figure 1. Cropped images of 10 individuals.

Finally, we compare $supports$ obtained from basis images of all classes and classify the test image into the class that maximizes $support$.

As a part of the process for estimating three basis images, Ishii et al. [14] calculate the coefficients of linear combination by using RANSAC. Unlike their method, our method effectively uses attached shadows in a test image and estimates the optimal coefficients in the framework of least squares method.

5. Experimental results

To demonstrate the effectiveness of our proposed method, we have conducted experiments using the Yale Face Database B [10]. It is known that face recognition under varying illumination conditions is a very difficult task because variations due to changes in illumination are usually larger than those due to changes in face identity [17].

5.1. Face image database

The database consists of images of 10 individuals in 9 poses acquired under 64 different point light sources and an ambient light: 5850 images in total. The coordinates of the left eye, right eye, and mouth are appended for images in the frontal pose. Each image is assigned to one of 5 subsets according to the angle θ between the direction of the light source and the optical axis of a camera.

In our experiments, we used 650 images in the frontal pose. These images were cropped and down-sampled to 40×40 pixels by averaging. In order to remove any bias due to the scale and position of a face in each image from the recognition performance, they were aligned so that the locations of the eyes were the same. In addition, the image taken under the ambient light was subtracted from them.

In Figure 1, we show cropped images of 10 individuals. We also show images of an individual belonging to each subset in Figure 2. One can confirm that images vary significantly depending on the direction of the light source.

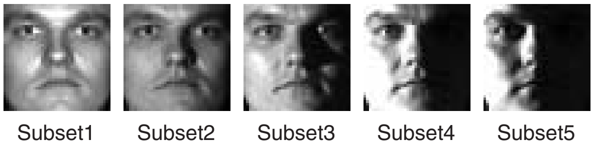

Figure 2. Images of an individual belonging to each subset: the angle θ between the light source direction and the optical axis lies $[0°, 12°)$, $[20°, 25°]$, $[35°, 52°)$, $(60°, 77°)$, and $(85°, 128°)$ respectively.

Figure 3. Three basis images of an individual.

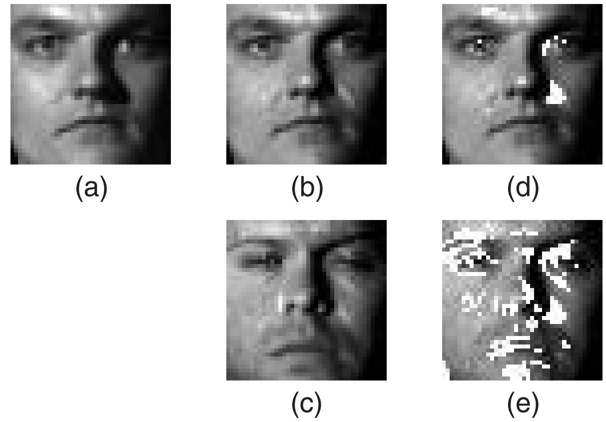

Figure 4. Recognition process of our method: (a) a test image, (b) an image synthesized from three basis images of the same person, (c) an image synthesized from those of different person. White pixels in images (d) and (e) represent locations where the difference between the test image and the synthesized image is larger than the threshold t.

5.2. Training

We used 7 images belonging to subset 1 as training images for each person. For estimating the basis images, we adopted a method using SVDMD [11]. The method identifies pixels corresponding to highlights and shadows based on their pixel values, and regards these pixels as missing data, and then applies SVD to the rest of pixels. In Figure 3, we show the estimated basis images of an individual.

Furthermore, we also determined the threshold t in equation (8) by synthesizing the training images from the estimated basis images. We computed the standard deviation σ of the difference between synthesized and original pixel values with respect to pixels used for estimating the basis images. We set the threshold t to be a few times the standard deviation σ.

5.3. Recognition

We compared the following three methods.

- **Linear subspace method (LS)**
  Compute the distance between a test image and three-dimensional linear subspace spanned by three basis images of each class, and then classify the test image into the class that minimizes the distance.

- **Photometric alignment using projection (PA1)**
  Compute the distance between a test image and a synthesized image from three basis images of each class, and then classify the test image into the class that minimizes the distance. The coefficients of linear combination are estimated by projecting the test image onto each of the basis images.

- **Photometric alignment using RANSAC (PA2)**
  As proposed in Section 4, estimate the coefficients of linear combination for synthesizing a test image from three basis images of each class, and then classify the test image based on the number of correctly reproduced pixels.

First, we show the process of our method (PA2). In Figure 4, we show (a) a test image, (b) an image synthesized from three basis images of the same person, (c) an image synthesized from those of different person. White pixels in images (d) and (e) represent locations where the difference between the test image and the synthesized image is larger than the threshold t. In the case of the same person, one can find that the test image is correctly reproduced except that specular reflection components appeared in the forehead and cast shadows appeared around the nose and eye. On the other hand, in the case of different person, one can find that our method cannot reproduce the test image in relatively large regions.

Secondly, we show recognition error rates of the three methods for each subset in Table 1. In our method, we show the results when the threshold t was set to be 4σ. We confirmed that the error rates varied about 1% at most for subset 4 and 2% for subset 5 when the threshold varied between

Table 1. Recognition error rates (%): linear subspace method (LS), photometric alignment using projection (PA1), photometric alignment using RANSAC (PA2), illumination cone model (IC), nine points of light (9PL and 9PL'), and segmented linear subspace method (SLS).

Method	Subset2	Subset3	Subset4	Subset5
LS	0	0	5.8	55.6
PA1	0	0	0.7	39.7
PA2	0	0	0	18.5
IC [10]	0	0	8.6	—
9PL [15]	0	0	2.8	—
9PL' [13]	0	0.7	1.4	—
SLS [2]	0	0	0	—

3.5σ and 5.5σ. This means that our method is not sensitive to the value of the threshold.

One can find that the methods based on the photometric alignment (PA1 and PA2) are superior to the linear subspace method (LS). This shows the effectiveness of the photometric alignment that can generate not only diffuse reflection components but also attached shadows. Figure 2 shows that, as the angle θ between the light source and the optical axis becomes large, the contribution of attached shadows become more dominant in general. Therefore, the difference of the recognition performance between the linear subspace method and the methods based on the photometric alignment becomes larger as the angle θ becomes larger. Moreover, one can find that our method using RANSAC (PA2) is superior to the method using projection (PA1). This means that the method using RANSAC is robust against outliers in test images. However, our method cannot generate specular reflection components and cast shadows while it can generate both diffuse reflection components and attached shadows accurately. The mis-classifications for subset 5 seem to be caused by these components that cannot be synthesized in the framework of the photometric alignment.

In Table 1, we also show recognition error rates on the Yale Face Database B reported in other papers [10, 15, 13, 2] in order to compare the performance of our method to that of the other methods. The error rates for subset 5 are not reported in the other papers. We should note that we cannot necessarily compare the performance simply since the cropped regions and image resolution are different in each experiment. However, our method based on only three basis images seems to have better performance than the illumination cone model that requires a large number of extreme images [10]. Moreover, our method using seven training images under unknown lighting directions seems to perform better than the methods based on the analysis in the frequency domain that use nine training images under special lighting directions [15, 13]. Furthermore, our method performs as well as the segmented linear subspace method that requires appropriate segmentation of images [2].

6. Conclusions and future work

In this paper, we have proposed a method for object recognition under varying illumination conditions; our proposed method is based on the photometric alignment that models both diffuse reflection components and attached shadows under a distant point light source by using three basis images. In order to deal with outliers such as specular reflection components and shadows in test images, our method utilizes RANSAC. We conducted experiments using the Yale Face Database B and confirmed that a combination of the photometric alignment and RANSAC provides a simple but effective method for object recognition under varying illumination conditions.

In the present study, we consider specular reflection components and cast shadows in test images as outliers based on the photometric alignment. However, these components seem to be dominant in images under extreme lighting directions and cause mis-classifications. Moreover, generative methods cannot be applied to non Lambertian objects because the methods assume the Lambertian model. Therefore, in the future work, we will also examine how to utilize specular reflection components and cast shadows by modeling them efficiently.

Acknowledgements

We would like to thank Athinodoros Georghiades for providing us with the Yale Face Database B. A part of this work was supported by Grants-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan (No. 13224051).

References

[1] R. Basri and D. Jacobs, "Lambertian reflectance and linear subspaces", In *Proc. IEEE ICCV 2001*, pp.383–390, 2001.

[2] A. Batur, and M. Hayes, "Linear subspaces for illumination robust face recognition", In *Proc. IEEE CVPR 2001*, 2, pp.296–301, 2001.

[3] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", *IEEE Trans. PAMI*, 19(7), pp.711–720, 1997.

[4] P. Belhumeur and D. Kriegman, "What is the set of images of an object under all possible lighting conditions?", *Int'l. J. Computer Vision*, 28(3), pp.245–260, 1998.

[5] P. Belhumeur, D. Kriegman, and A. Yuille, "The bas-relief ambiguity", In *Proc. IEEE CVPR '97*, pp.1060–1066, 1997.

[6] M. Black and A. Jepson, "Eigen tracking: robust matching and tracking of articulated objects using a view-based representation", In *Proc. ECCV '96*, pp.329–342, 1996.

[7] R. Brunelli and T. Poggio, "Face recognition: features versus templates", *IEEE Trans. PAMI*, 15(10), pp.1042–1052, 1993.

[8] H. Chen, P. Belhumeur, and D. Jacobs, "In search of illumination invariants", In *Proc. IEEE CVPR 2000*, pp.254–261, 2000.

[9] M. Fischler and R. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", *Communications of the ACM*, 24(6), pp.381–395, 1981.

[10] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: illumination cone models for face recognition under variable lighting and pose", *IEEE Trans. PAMI*, 23(6), pp.643–660, 2001.

[11] A. Georghiades, D. Kriegman, and P. Belhumeur, "Illumination cones for recognition under variable lighting: faces", In *Proc. IEEE CVPR '98*, pp.52–58, 1998.

[12] P. Hallinan, "A low-dimensional representation of human faces for arbitrary lighting conditions", In *Proc. IEEE CVPR '94*, pp.995–999, 1994.

[13] J. Ho, K.-C. Lee, and D. Kriegman, "On reducing the complexity of illumination cones for face recognition", In *Proc. CVPR Workshop on Idetifying Objects Accross Variations in Lighting*, 2001.

[14] Y. Ishii, K. Fukui, Y. Mukaigawa, and T. Shakunaga, "Photometric linearization based on classification of photometric factors", In *Proc. MIRU 2002*, 2, pp.167–176, 2002. (in Japanese)

[15] K.-C. Lee, J. Ho, and D. Kriegman, "Nine points of light: acquiring subspaces for face recognition under variable lighting", In *Proc. IEEE CVPR 2001*, 1, pp.519–526, 2001.

[16] A. Leonardis and H. Bischof, "Dealing with occlusions in the eigenspace approach", In *Proc. IEEE CVPR '96*, pp.453–458, 1996.

[17] Y. Moses, Y. Adini, and S. Ullman, "Face recognition: the problem of compensating for changes in illumination direction", In *Proc. ECCV '94*, pp.286–296, 1994.

[18] Y. Mukaigawa, H. Miyaki, S. Mihashi, and T. Shakunaga, "Photometric image-based rendering for image generation in arbitrary illumination", In *Proc. IEEE ICCV 2001*, pp.652–659, 2001.

[19] H. Murase and S. Nayar, "Visual learning and recognition of 3-D objects from appearance", *Int'l. J. Computer Vision*, 14(1), pp.5–24, 1995.

[20] A. Nakashima, A. Maki, and K. Fukui, "Constructing illumination image basis from object motion", In *Proc. ECCV 2002* (LNCS 2352), pp.195–209, 2002.

[21] M. Oren and S. Nayar, "Generalization of the Lambertian model and implications for machine vision", *Int'l. J. Computer Vision*, 14(3), pp.227-251, 1995.

[22] M. Pontil and A. Verri, "Support vector machines for 3D object recognition", *IEEE Trans. PAMI*, 20(6), pp.637–646, 1998.

[23] 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. Am. A*, 18(10), pp.2448–2459, 2001.

[24] R. Ramamoorthi and P. Hanrahan, "A signal-processing framework for inverse rendering", In *Proc. ACM SIGGRAPH 2001*, pp.117–128, 2001.

[25] T. Shakunaga and K. Shigenari, "Decomposed eigenface for face recognition under various lighting conditions", In *Proc. IEEE CVPR 2001*, 1, pp.864–871, 2001.

[26] A. Shashua, "On photometric issues in 3D visual recognition from a single 2D image", *Int'l. J. Computer Vision*, 21(1/2), pp.99–122, 1997.

[27] H.-Y. Shum, K. Ikeuchi, and R. Reddy, "Principal component analysis with missing data and its application to polyhedral object modeling", *IEEE Trans. PAMI*, 17(9), pp.854–867, 1995.

[28] M. Turk and A. Pentland, "Face recognition using eigenfaces", In *Proc. IEEE CVPR '91*, pp.586–591, 1991.

[29] R. Woodham, "Photometric method for determining surface orientation from multiple images", *Optical Engineering*, 19(1), pp.139–144, 1980.

[30] A. Yuille, D. Snow, R. Epstein, and P. Belhumeur, "Determining generative models of objects under varying illumination: shape and albedo from multiple images using SVD and integrability", *Int'l. J. Computer Vision*, 35(3), pp.203–222, 1999.