Robust face recognition by combining projection-based image correction and decomposed eigenface

Takeshi Shakunaga
Okayama University

Fumihiko Sakaue
Okayama University

Kazuma Shigenari
Okayama University

Robust Face Recognition by Combining Projection-Based Image Correction and Decomposed Eigenface

Takeshi Shakunaga Fumihiko Sakaue Kazuma Shigenari

Department of Information Technology, Faculty of Engineering
Okayama University
Okayama-shi, Okayama 700-8530, Japan
E-mail: {shaku,sakaue}@chino.it.okayama-u.ac.jp

Abstract

This paper presents a robust face recognition method which can work even when an insufficient number of images are registered for each person. The method is composed of image correction and image decomposition, both of which are specified in the normalized image space (NIS). The image correction[1, 2] is realized by iterative projections of an image to an eigenspace in NIS. It works well for natural images having various kinds of noise, including shadows, reflections, and occlusions. We have proposed decomposition of an eigenface into two orthogonal eigenspaces[3], and have shown that the decomposition is effective for realizing robust face recognition under various lighting conditions. This paper shows that the decomposed eigenface method can be refined by projection-based image correction.

1. Introduction

A human face changes in appearance with lighting conditions, and difficulty is encountered in controlling lighting conditions in natural environments where face images are taken. These facts suggest that robust face recognition requires construction of a face recognition algorithm that is insensitive to lighting conditions. Meanwhile, appearance-based face recognition can be resolved into the eigenface method [4], which in many cases is identical with the subspace method [5, 6]. Eigenfaces are widely used for both personal identification and detection of (unknown) faces in an image. When intended for detection of faces in an image, eigenfaces are constructed from many persons and when intended for personal identification, each eigenface should be constructed from face images of the individual. In the present paper we focus on the second purpose, and as used herein an eigenface (EF) always means an eigenspace constructed from face images of the individual for the purpose of personal identification. If many face images can be collected in the registration stage, the EF can be constructed by Principal Component Analysis (PCA). However, an EF for an individual cannot be stably constructed when an insufficient number of sample images are available or when the sample images have been taken under very similar lighting conditions. In these situations, realizing illumination-insensitive identification requires some refinements of the eigenface method. In our previous paper [3], we analyzed the eigenface approach and proposed concepts of virtual eigenfaces and the decomposed eigenface for the foregoing purpose. In this paper, we combine these concepts with a projection-based image correction method [2] in order to refine the face recognition methodology.

2. Normalized Image Space and Normalized Eigenspace

2.1. Normalized Image Space

In object recognition, eigenspaces are often constructed in a least-squares sense, faithful to the original images [4, 6]. Eigenspaces are also effective for image-based rendering under changing lighting conditions [7, 8]. They can also be constructed from original images, by use of the photometric SVD (Singular Value Decomposition) algorithm [9, 10]. These methods are commonly discussed in the original image space.

Although the original image space is effective for some purposes, it often fails to work when illumination falls out of range. In such a case, we empirically utilize some types of image normalization.

In this paper, normalization is based on L1-norm. Let an N-dimensional vector $\mathbf{X}$ denote an image whose elements are all non-negative, and let $\mathbf{1}$ denote an N-dimensional vector whose elements are all 1. The normalized image $\mathbf{x}$ of an original image $\mathbf{X}$ is defined as $\mathbf{x} = \mathbf{X}/(\mathbf{1}^T\mathbf{X})$. After the normalization, $\mathbf{x}$ is normalized in the sense that $\mathbf{1}^T\mathbf{x} = 1$.

By this normalization, any nonzero image $\mathbf{X}(\neq \mathbf{0})$ is

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE

mapped to a point in the Normalized Image Space (NIS). The NIS is closed to any averaging operation, and the real variance is encoded up to a scale factor.

2.2. Normalized Eigenspace

When an image class is given, a k-dimensional eigenspace is constructed in NIS by the conventional PCA from the mean vector $\overline{\mathbf{x}}$ and covariance matrix Σ

$$\overline{\mathbf{x}} = \frac{1}{K}\sum_{j=1}^{K}\mathbf{x}_j \quad \text{and} \quad \Sigma = \frac{1}{K}\sum_{j=1}^{K}(\mathbf{x}_j - \overline{\mathbf{x}})(\mathbf{x}_j - \overline{\mathbf{x}})^T,$$

where K is the number of images in the class.

Let Λ denote a diagonal matrix in which diagonal terms are eigenvalues of Σ in descending order, and Φ a matrix in which the i-th column is the i-th eigenvector of Σ. Then PCA implies $\Lambda = \Phi^T\Sigma\Phi$. Using a submatrix Φ_k of Φ, which consists of k principal eigenvectors, the projection, $\mathbf{x}^*$, of $\mathbf{x}$ onto the eigenspace and the residual, $\mathbf{x}^\sharp$, of the projection are given by

$$\mathbf{x}^* = \Phi_k^T(\mathbf{x} - \overline{\mathbf{x}}), \tag{1}$$

$$\mathbf{x}^\sharp = \mathbf{x} - \overline{\mathbf{x}} - \Phi_k\mathbf{x}^*. \tag{2}$$

In our problems, k is a small number because human faces are almost Lambertian.

Let us call the k-dimensional eigenspace the Normalized Eigenspace (NES). We also use another notation, $\langle\overline{\mathbf{x}}, \Phi_k\rangle$, which explicitly specify $\overline{\mathbf{x}}$ and Φ_k.

2.3. Canonical Space

In this paper, a face space is defined as a space composed from a set of frontal faces, which includes images of numerous persons taken under a wide variety of lighting conditions. To simplify the problem, we assume that good segmentation is readily accomplished as shown in Fig. 10. Eigenspace analysis on the face space reduces the dimension of the face space, with little loss of representability [6, 4].

Let $\langle\overline{\mathbf{x}}_c, \Phi_c\rangle$ denote a NES constructed over a canonical set. We call this the canonical space (CS). In our experiments, a 45d CS is constructed from a canonical image set that consists of face images of 50 persons taken under 24 lighting conditions.

3. Noise Detection and Image Correction

3.1. Effect of Noise in NIS

Let us analyze the effect of noise on the projection onto NES $\langle\overline{\mathbf{x}}, \Phi_k\rangle$ and the residual. Suppose that an object view is completely encoded to the NES by

$$\mathbf{x}^* = {\Phi_k}^T(\mathbf{x} - \overline{\mathbf{x}}).$$

Let $\mathbf{n}$ and $\mathbf{y}$ denote the normalized images of image $\mathbf{N}$ and $\mathbf{Y} = \mathbf{X} + \mathbf{N}$, where $\mathbf{X}$ is a signal, $\mathbf{N}$ is a noise and $\mathbf{Y}$ is an input image. Let us define $\mathbf{y} = (1-\alpha)\mathbf{x} + \alpha\mathbf{n}$, where $\alpha = \mathbf{1}^T\mathbf{N}/\mathbf{1}^T\mathbf{Y}$. Then the projection, $\mathbf{y}^*$, and the residual, $\mathbf{y}^\sharp$, are respectively represented by

$$\mathbf{y}^* = (1-\alpha)\mathbf{x}^* + \alpha{\Phi_k}^T(\mathbf{n} - \overline{\mathbf{x}}) \tag{3}$$

$$\mathbf{y}^\sharp = \alpha\mathbf{n}^\sharp = \alpha(\mathbf{n} - \overline{\mathbf{x}}) - \alpha\Phi_k{\Phi_k}^T(\mathbf{n} - \overline{\mathbf{x}}). \tag{4}$$

In Eq. (4), the first term indicates the existence of $\mathbf{n}$ itself. The second term indicates that the noise affects the whole image with weight $-\alpha\Phi_k{\Phi_k}^T$. Even if $\mathbf{n}$ is very localized, the effect spreads to the whole image. Since $\Phi_k{\Phi_k}^T$ is positive semidefinite, the second term yields a counteraction to the noise. A negative reaction is generated from a positive noise, whereas a positive reaction is generated from a negative noise. Refer to [1] on the estimations of α and $\mathbf{x}^*$ as well as a detection of noise region in image correction.

3.2. Noise Detection by Relative Residual

Let us define a relative residual r_i for the i-th pixel of $\mathbf{x}$ by

$$r_i = \frac{\mathbf{e}_i^T\mathbf{x}^\sharp}{\mathbf{e}_i^T(\overline{\mathbf{x}} + \Phi_k\mathbf{x}^*)},$$

where $\mathbf{x}^\sharp$ is an absolute residual given by Eq.(2), and $\mathbf{e}_i$ is a unit vector which consists of 1 in the i-th element and 0 in the other elements. We use the relative residual instead of absolute residual for the i-th pixel, $\mathbf{e}_i^T\mathbf{x}^\sharp$, because we would like to suppress noise not in the absolute scale but in the relative scale. For example, low noise in a dark area should be suppressed when the relative residual is sufficiently large.

Noise detection is basically performed by $|r_i|$. However, Eq. (4) suggests that when a considerable amount of noise is included, the zero level of the relative residual may shift in response to the amount of noise and Φ_k. In order to compensate the possible shift, we use $|r_i - \widehat{r}|$ instead of $|r_i|$, where $\widehat{r}$ is the median of the whole r_i. We don't use the average, because we would like to neglect the direct noise factors in Eq. (4). Consequently, noise can be detected when $|r_i - \widehat{r}| \geq r_\theta$, where r_θ is a threshold.

3.3. Image Correction by Projection

The noise correction algorithm can be created on the basis of the noise detection, as follows: When $|r_i - \widehat{r}| \geq r_\theta$, the i-th pixel of $\mathbf{x}$ should be replaced to $(1-\alpha)\mathbf{e}_i^T(\overline{\mathbf{x}} + \Phi_k\mathbf{x}^*)$, where $\mathbf{x}^*$ and α are provided simultaneously by calculating a partial projection[1] when the partial region excludes noise regions. The image correction makes an intensity value consistent with the projection. For example, shadows and reflection regions are corrected when the NES is

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE

Figure 1: Example of eigenplane for Lambertian surface: Average and 2 principal bases.

Figure 2: Two examples of image correction: In each triplet from left to right, original $\mathbf{x}$, detected noise region, and final result.

constructed over an image set including a small amount of shadows and reflections.

Note that the normality of the image doesn't hold after the correction. Therefore, the corrected image should be re-normalized when all the pixels are checked and corrected.

The projection-based correction just changes outliers to inliers. When more than a few pixels are corrected, $\mathbf{x}^*$ also changes to some extent. Therefore, a few iterations of correction should be performed for better noise suppression. After iterations, $\mathbf{x}$ converges to an image having little noise.

3.4. Experiments of Noise Detection/Correction

(1) Eigenplane for Lambertian Object
A 2d eigenspace (an eigenplane) is constructed for a Lambertian object. Figure 1 shows an example eigenplane which is constructed in NIS from ten images of a Napoleon statue made of plaster. In Fig. 1, the left image shows the average image, and the other images show the orthonormal bases of the eigenplane.

Figure 2 shows two examples of image correction. In both examples, the corrections are made around the nose, where the cast shadow regions are detected and corrected by the projection.

(2) Image Correction with Individual Eigenface
For a non-Lambertian object, the image correction also works well. Figure 3 shows an example eigenface which is constructed in NIS from 6 individual faces. In Fig. 3, the left image shows the average image and the others show the 3 principal bases of the eigenface.

Figure 4 shows 3 examples of image correction. In the left and center pairs, the corrections are made around the nose, where the cast shadow regions are detected and corrected by the projection. In the right-most example, the correction is made for an artificial occlusion. These examples show that the projection-based image correction is very robust to both shadows and occlusions.

Figure 3: Individual Eigenface: Average and 3 principal bases.

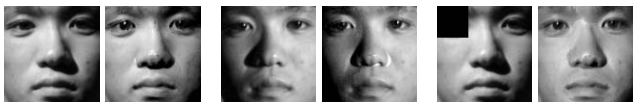

Figure 4: Examples of image correction: In each pair, $\mathbf{x}$ is shown on the left and the final result on the right.

(3) Image Correction with Universal Eigenface
For a class of human faces, the projection-based image correction still works. A universal eigenface, as shown in Fig. 5, is constructed in NIS from images of 50 faces, each taken under 20 lighting conditions. In Fig. 5, the left image shows the average image and the others show 5 principal bases of the eigenface.

In Fig. 6, we compare the results between 3d and 45d eigenfaces. In these two examples, similar results are obtained with 3d and with 45d eigenfaces, although $\mathbf{x}^*$ is less similar to $\mathbf{x}$ when the 3d eigenface is used. The right example also shows that the image correction works for the mirror reflection on his eyeglasses as well as for half-transparency. In this case, his glasses are removed in the results, because the universal eigenface doesn't include persons wearing glasses.

4. Decomposed Eigenface Method

4.1. Concept of Decomposed Eigenface

If numerous face images can be collected in the registration stage, the EF can be constructed by PCA. Yuille et al. [9] have shown that an EF can also be constructed by photometric SVD when a lot of images of an individual are taken under various illumination conditions. However, an EF for an individual cannot be stably constructed when an insufficient number of sample images are available or when the sample images have been taken under very similar lighting conditions. In these situations, illumination-insensitive identification requires some refinements to the eigenface method. In this section, we introduce a decomposition of EF for the foregoing purpose.

Figure 5: Universal eigenface constructed from images of 50 persons: Average and 5 principal bases.

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE

Figure 6: Two examples of image correction: In each triplet from left to right, an original $\mathbf{x}$, an image corrected with the 3d eigenface and an image corrected with the 45d eigenface.

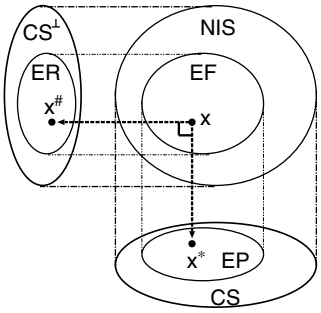

Figure 7: Decomposition of EF to EP and ER.

4.2. Eigen-Projection and Eigen-Residual

A projection of $\mathbf{x}$ to CS, $\langle \overline{\mathbf{x}}_c, \Phi_c \rangle$, and the residual are respectively represented by

$$\mathbf{x}^* = {\Phi_c}^T(\mathbf{x} - \overline{\mathbf{x}}_c),$$

$$\mathbf{x}^\sharp = \mathbf{x} - \overline{\mathbf{x}}_c - \Phi_c \mathbf{x}^*.$$

Thus, a normalized image $\mathbf{x}$ can be decomposed into the canonical component $\mathbf{x}^*$ and the residual component $\mathbf{x}^\sharp$. Because of the above definitions, they are orthogonal.

The orthogonal components $\mathbf{x}^*$ and $\mathbf{x}^\sharp$ enable us to decompose the eigenface (EF) in NIS. That is, as shown in Fig. 7, two eigenspaces can be constructed independently in CS and in the orthogonal complement CS$^\perp$. The first eigenspace, called an eigen-projection (EP), is constructed from the canonical components in CS. The second eigenspace, called an eigen-residual (ER), is constructed from the residual components in CS$^\perp$. EP and ER are constructed by PCA in CS and CS$^\perp$, respectively.

In the construction of EP, the mean vector $\overline{\mathbf{x}_p^*}$ and the covariance matrix Σ_p^* are respectively calculated by

$$\overline{\mathbf{x}_p^*} = \frac{1}{L}\sum_{l=1}^{L} \mathbf{x}_{pl}^* \tag{5}$$

$$\Sigma_p^* = \frac{1}{L}\sum_{l=1}^{L} (\mathbf{x}_{pl}^* - \overline{\mathbf{x}_p^*})(\mathbf{x}_{pl}^* - \overline{\mathbf{x}_p^*})^T, \tag{6}$$

where L is the number of images for each person. Let Φ_p^* and Λ_p^* denote the eigenvectors and the diagonal matrix, respectively. Then PCA implies $\Lambda_p^* = {\Phi_p^*}^T \Sigma_p^* \Phi_p^*$. Using a submatrix Φ_{pm}^* of Φ_p^*, which consists of m principal eigenvectors, the projection of $\mathbf{x}^*$ to the p-th EP is given by

$$\tilde{\mathbf{x}}_p^* = {\Phi_{pm}^*}^T(\mathbf{x}^* - \overline{\mathbf{x}_p^*}). \tag{7}$$

ER can also be constructed in the same way as described for EP. We can define $\overline{\mathbf{x}_p^\sharp}$, $\Sigma_p^\sharp$, $\Phi_p^\sharp$, $\Lambda_p^\sharp$, and $\Phi_{pm}^\sharp$ in the same way. Consequently, the projection of $\mathbf{x}^\sharp$ to the p-th ER is given by

$$\tilde{\mathbf{x}}_p^\sharp = {\Phi_{pm}^\sharp}^T(\mathbf{x}^\sharp - \overline{\mathbf{x}_p^\sharp}).$$

4.3. Registration and Recognition Schemes

In the registration stage, EP and ER are created independently in CS and in CS$^\perp$. In CS$^\perp$, statistical noise reduction is effective for the construction of the Eigen-residual (ER) because most of noise appears in residuals.

In the recognition stage, we can realize face identification by combining the two eigenspaces. Given an unknown face $\mathbf{x}$, two similarity measures are defined by normalized correlations in CS and CS$^\perp$, where $C(\mathbf{x}, \mathbf{y})$ shows a normalized correlation between $\mathbf{x}$ and $\mathbf{y}$:

(1) Similarity between $\mathbf{x}^*$ and EP $\langle \overline{\mathbf{x}_p^*}, \Phi_{pm}^* \rangle$ in CS:

$$C1_p(\mathbf{x}) = C(\Phi_{pm}^* \tilde{\mathbf{x}}_p^* + \overline{\mathbf{x}_p^*}, \mathbf{x}^*).$$

(2) Similarity between $\mathbf{x}^\sharp$ and ER $\langle \overline{\mathbf{x}_p^\sharp}, \Phi_{pm}^\sharp \rangle$ in CS$^\perp$:

$$C2_p(\mathbf{x}) = C(\Phi_{pm}^\sharp \tilde{\mathbf{x}}_p^\sharp + \overline{\mathbf{x}_p^\sharp}, \mathbf{x}^\sharp).$$

(3) Combined similarity of C1 and C2:

Because $C1$ and $C2$ are calculated independently in CS and CS$^\perp$, they can be combined as

$$C3_p(\mathbf{x}) = \frac{C1_p(\mathbf{x})}{C1_{\widehat{p_1}}(\mathbf{x})} + \frac{C2_p(\mathbf{x})}{C2_{\widehat{p_2}}(\mathbf{x})},$$

$$\text{where} \quad \widehat{p_i} = \arg\max_{1 \le p \le P} Ci_p(\mathbf{x}).$$

A simple discrimination is then made for $Ci(i = 1, 2, 3)$, by selecting a person

$$\arg\max_{1 \le p \le P} Ci_p(\mathbf{x}).$$

5. Refinement of Decomposed Eigenface Method

5.1. Image Correction and Refinement of CS

In this section, we refine the original decomposed eigenface method by applying the image correction prior to the CS construction. In the image correction, CS is used as a universal eigenface in 3.4(3). Since the image correction method described in 3.2 and 3.3 can be applied to any image for any purpose, it is also applied to the canonical images when a preliminary CS is made up without use of the

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE

Figure 8: Example of iterative image correction: From left to right, original $\mathbf{x}$, and after first, second, and third corrections.

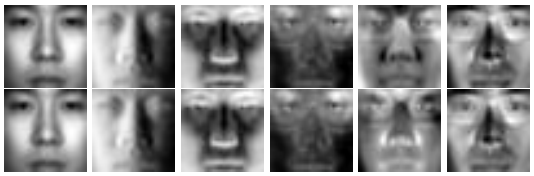

Figure 9: Comparison of average faces and principal bases: Upper and lower rows show them before and after the image correction, respectively.

image correction. We can redefine the canonical space by making an eigenspace over the canonical images after the image correction on the preliminary CS.

Figure 8 shows an example of image correction. The original image on the left is gradually corrected. Especially, the corrections suppress reflections in her glasses.

Figure 9 shows the average faces and 5 principal bases of CS before and after image correction. The image correction refines the canonical spaces, because many reflections and shadows are removed from the canonical image set.

5.2. Refined Registration and Recognition Schemes

Both the registration and recognition schemes can be refined. In the registration scheme, after the image correction, EP is made up in the same manner as described in 4.3. Because both the CS and the registered image are corrected, each EP includes much less noise than the original method. Construction of ER does not require additional noise reduction, because the image correction sufficiently eliminates the noise.

In the recognition scheme, the subspace method is applied to an image after the image correction. Because EP and ER include less noise, the recognition scheme works better than the original method.

6. Face Discrimination Experiments

6.1. Data Specification

Data specification is summarized in Table 1. Facial images were taken from a fixed camera located in our laboratory. Each of 100 persons were looking forward while sitting on a chair located in a fixed distance from the camera. The chair was fixed in order to obtain a frontal facial image of each person.

Table 1: Data specification

	Canonical set	Test images
# of persons	50	50
# of lighting conditions	24	24
Image size	32×32	32×32
persons wearing glasses	9	15

Figure 10: Averages of canonical images taken under 24 lighting conditions.

As shown in Section 2.3, CS is created from the canonical image set, which consists of 1200 images of 50 persons. For each person, images are taken under 24 lighting conditions, which are controlled by changing the position of a light. In the canonical set, 9 persons wear glasses. Figure 10 shows averages of the canonical images taken under the 24 lighting conditions. The remaining 50 persons are used for the test data, and 15 of these persons wear glasses. Figures 11 (a) and (b) show 4 examples of canonical images and 4 examples of the test images, respectively, taken under a fixed lighting condition.

6.2. Comparison

For personal registration, K images were randomly sampled from 24 images of each person in the test data. There-

(a) canonical images (b) test images

Figure 11: Examples of canonical/test image sets.

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE

Table 2: Discrimination rates[%] for 50 persons

Method	Number of samples(K)/person				
	1	2	3	4	5
conventional NN	35.8	50.2	63.1	63.6	68.7
conventional EF	44.8	73.2	86.5	91.9	94.3
$C1$	81.2	88.3	91.1	91.8	93.3
$C2$	89.0	96.2	98.2	99.1	99.3
$C3$	92.1	97.3	98.6	99.3	99.3
$C1'$	83.1	91.4	93.1	94.9	94.6
$C2'$	91.0	97.0	98.9	99.6	99.8
$C3'$	95.1	98.5	99.4	99.8	99.9

fore, the discrimination experiment was carried out from the remaining $24 - K$ images of each registered person. This process was repeated one hundred times while registered images for each person were varied.

Table 2 shows average discrimination rates. Eight methods are compared for the same canonical and test sets. The first 2 rows show the results of conventional NN (nearest-neighbor) and EF (eigenface) methods. The second 3 rows ($C1, C2, C3$) show the results of the decomposed eigenface method without prior image correction, and the remaining 3 rows ($C1', C2', C3'$) show the results of the refined method. In all the methods, face symmetry is used in the registration stage. The dimension of EP is $K + 3$, and the dimension of ER is $\min(2K - 1, K + 2)$.

The table shows that the image correction improves the discrimination rates for all K and for all similarity measures. The best result is provided with $C3'$, with recognition reaching 95.1% when only one image is registered for each person. This shows an improvement of 3 points over the $C3$. When five images are registered for each person, the result obtained with $C3'$ reaches 99.9%.

6.3. Experiments in Public Databases

In order to confirm the effectiveness of the refinement, we also conducted experiments in Yale Face Database B[11] and AR Database[12]. These experiments employ the same CS as used in the above experiments.

(1) Experiments in Yale Face Database B[11]

The database includes 10 individuals taken under 64 different lighting conditions. Images were classified to five subsets(SS1-SS5) by angle between the light source direction and the camera axis. Each image was segmented and converted to 32 by 32 pixels so that all of the faces have eyes in the same coordinates. Our discrimination experiments were carried out over the segmented data set. SS1 which includes seven images was used as a registered set and the other subsets were used as test sets. Table 3 shows the experimental results.

$C3$ and $C3'$ worked well for test sets SS2 and SS3 even if too few images were registered. For test sets SS4 and SS5, $C3'$ resulted in slightly worse results than $C3$. This is because the image correction in $C3'$ could not work when input images include excessive noise. You also should know that the CS was constructed from only Japanese facial images. They look very different from images in the Yale Database. If a CS is constructed from international data set, the image correction works much better and $C3'$ is expected to provide better results.

Table 3: Discrimination rates[%] for Yale Database B

		Number of samples(K)/person						
		1	2	3	4	5	6	7
SS2	$C3$	98.0	99.7	100	100	100	100	100
	$C3'$	98.0	99.7	100	100	100	100	100
SS3	$C3$	89.3	95.3	99.0	100	100	100	100
	$C3'$	88.3	95.8	99.7	100	100	100	100
SS4	$C3$	59.7	78.9	87.5	92.9	94.6	95.1	95.7
	$C3'$	56.2	75.3	83.3	88.0	90.8	91.6	91.3
SS5	$C3$	23.2	36.0	42.0	46.3	48.7	49.4	53.5
	$C3'$	23.9	29.8	33.1	33.9	34.6	35.2	34.4

Table 4: Comparison of discrimination rates: $C3$ and $C3'$ are proposed in this paper, PPP indicates a linear subspace method using parallel partial projections[13], Cones-attached indicates the basic illumination cone[14], Cones-cast indicates the illumination cone with cast shadow representation[14], 9PL indicates the nine points light method[15], Segm. LS indicates the segmented linear subspace method[16], and PA indicates the photometric alignment[17].

method	SS2	SS3	SS4	SS5
$C3$	100	100	95.7	50.8
$C3'$	100	100	91.3	34.4
PPP[13]	100	100	100	100
Cones-attached[14]	100	100	91.4	-
Cones-cast[14]	100	100	100	-
9PL[15]	100	100	97.2	-
Segm. LS[16]	100	100	100	-
PA[17]	100	100	100	81.5

In Table 4, we compare results on Yale Face Database B reported in [14, 15, 16, 17, 13]. In the experiments, seven images in SS1 are registered. The results for SS5 are not reported in three papers. We should note that we cannot simply compare the results since the cropped regions and image resolutions are different. Although our method works slightly worse than the other methods, it can be improved if an appropriate CS is used. Furthermore, the comparison result indicates the parallel partial projections (PPP) provides the best scores. This suggests that a better method may be constructed by a combination of the parallel partial projections and the decomposed eigenface.

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE

Table 5: Results for 2 databases [%]

Method		$C3$		$C3'$	
Database	P	K=1	K=2	K=1	K=2
Ours	50	92.1	97.3	95.1	98.5
AR	50	91.6	98.1	92.3	99.4

(2) Experiments in AR Database[12]
The database contains images of 126 persons taken under 3 lighting conditions for each person. In order to compare with the results in our database, 50 persons were randomly selected from the database, and the discrimination experiment was carried out using $C3$ and $C3'$. In the experiment, one or two images were registered for each person and the remaining image(s) were used for the test. Table 5 shows that the refinement can stably discriminate persons in the AR database, as well as persons in our database.

7. Conclusions

We have integrated several concepts on the normalized eigenspaces for realizing a robust face recognition system which can work even when an insufficient number of images can be registered for each person. The normalized image space (NIS) is very useful when combined with eigenspaces. Normalized Eigenspaces (NES) enable us to realize object recognition, photometric analysis, and image correction in the same domain. The projection-based image correction on NIS is very effective for a variety of eigenspaces. The image correction method has wide applications, including recognition of face and object images including shadows, reflections, and occlusions. The decomposed eigenface provides a high capacity of face discrimination, which is realized simply by the projection of an image onto the canonical space. Projection-based image correction is employed to refine the decomposed eigenface method. Experimental results show that the integrated method improves the discrimination rates on our own and the AR databases as well as on the Yale database. The proposed method can be applied to face recognition under natural lighting conditions, even if the lighting condition is unknown or changes with time.

Acknowledgment: We would like to thank Athinodoros Georghiades and Aleix M. Martinez for allowing us to use with the Yale Face Database B [11] and the AR Database[12]. This work has been supported in part by Grant-In-Aid for Science Research under Grant No.15300062 from the Ministry of Education, Science, Sports, and Culture, Japanese Government.

References

[1] F. Sakaue and T. Shakunaga, "Robust projection onto normalized eigenspace using relative residual analysis and optimal partial projection," *IEICE Trans. Inf. & Syst.*, vol. E87-D, no. 1, 2004.

[2] T. Shakunaga and F. Sakaue, "Natural image correction by iterative projections to eigenspace constructed in normalized image space," in *Proc. ICPR2002*, vol. 1, pp. 648–651, 2002.

[3] T. Shakunaga and K. Shigenari, "Decomposed eigenface for face recognition under various lighting conditions," in *Proc. CVPR2001*, vol. 1, pp. 864–871, 2001.

[4] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991.

[5] E. Oja, *Subspace Methods of Pattern Recognition*. Research Studies Press Ltd., 1983.

[6] M. Kirby and L. Sirovich, "Application of the karhunen-loeve procedure for the characterization of human faces," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, 1990.

[7] K. Nishino, Y. Sato, and K. Ikeuchi, "Eigen-texture method: appearance compression and synthesis based on a 3d model," *IEEE Trans. PAMI*, vol. 23, no. 11, pp. 1257–1265, 2001.

[8] Y. Mukaigawa, H. Miyaki, S. Mihashi, and T. Shakunaga, "Photometric image-based rendering for image generation in arbitrary illumination," in *Proc. ICCV2001*, pp. 652–659, 2001.

[9] A. Yuille, S. Epstein, and P. Belhumeur, "Determining generative models of objects under varying illumination: Shape and albedo from multiple images using svd and integrability," *International Journal of Computer Vision*, vol. 35, no. 3, pp. 203–222, 1999.

[10] K. Hayakawa, "Shape, surface reflectance, illuminant direction and illuminant intensity from shading images," *Trans. IEICE*, vol. J77-D-II, no. 4, pp. 700–710, 1994 (in Japanese).

[11] A. Georghiades, P. Belhumeur, and D. Kriegman, "From few to many: Generative models for recognition under variable pose and illuminations," in *Proc. FG2000*, pp. 277–284, 2000.

[12] A. Martinez and R. Benavente, "The ar face database," Tech. Rep. #24, CVC, 1998.

[13] F. Sakaue and T. Shakunaga, "Face recognition by parallel partial projections," in *Proc. ACCV2004*, vol. 1, pp. 1440–150, 2004.

[14] 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 Analysis and Machine Intelligence*, vol. 23, no. 6, pp. 643–660, 2001.

[15] K. Lee, J. Ho, and D. Kriegman, "Nine points of light: acquiring subspaces for face recognition under variable lighting," in *Proc. IEEE CVPR2001*, vol. 1, pp. 519–526, 2001.

[16] A. Batur and M. Hayes III, "Linear subspaces for illumination robust face recognition," in *Proc. ICCV2001*, vol. II, pp. 296–301, 2001.

[17] T. Okabe and Y. Sato, "Object recognition based on photometric alignment using ransac," in *Proc. CVPR2003*, vol. 1, pp. 221–228, 2003.

Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04)
0-7695-2122-3/04 $ 20.00 © 2004 IEEE