Combination of Projectional and Locational Decompositions for Robust Face Recognition

Fumihiko Sakaue and Takeshi Shakunaga

Department of Computer Science, Okayama University,
Okayama-shi, 700-8530, Japan
{sakaue, shaku}@chino.cs.okayama-u.ac.jp

Abstract. The present paper discusses a method for robust face recognition that works even when only one image is registered and the test image contains a lot of local noises. Two types of facial image decomposition are compared both theoretically and experimentally. That is, we consider both a projectional decomposition, in which images are decomposed into individuality and other components, and a locational decomposition, in which the effects of local noises are suppressed. These two decompositions are simple and powerful and can be applied in collaboration with one another. This collaboration can be realized in a straightforward manner because the decompositions are consistent with one another. They work in a complementary manner and provide better results than when the decompositions are used independently. Finally, we report experimental results obtained using three databases. These results indicate that the combination of projectional and locational decompositions works well, even when only one image is registered and the test images contain significant noise.

1 Introduction

The appearance of the human face changes according to the lighting conditions under which the facial image is captured. However, it is often difficult to control the lighting condition in natural environments. A face recognition algorithm should not be sensitive to the lighting condition in order to realize robust face recognition. Although an eigenface[1,2] can efficiently represent photometric changes, it cannot be constructed appropriately when too few images are available for registration. In this case, photo-insensitive information should be extracted from registered images. In other words, we should decompose the image into individuality and other information. This decomposition is referred to herein as projectional decomposition. This decomposition is a basic and important problem in not only face recognition but also pattern recognition.

Local noises, such as occlusions and shadows, are contained in images and affect the recognition method based on the eigenspace and projection onto the eigenspace. Several algorithms have been proposed[3,4] for robust recognition against various noises. Although these algorithms provide good projection even when an image includes local noises, a great deal of processing time is required.

W. Zhao, S. Gong, and X. Tang (Eds.): AMFG 2005, LNCS 3723, pp. 407–421, 2005.

Alternatively, in another approach, the image is regarded as a set of small components [2,5,6,7]. This approach, referred to herein as locational decomposition, does not spread local noises to the entire image and thus can avoid the above-mentioned problem.

In the present paper, we propose a novel method for robust face recognition by combining projectional and locational decompositions. Since projectional and locational decompositions can be used simultaneously, this combination facilitates the realization of a face recognition algorithm that is robust with respect to noises.

2 Definitions

2.1 Normalized Eigenspace

In this section, we present basic definitions and the notation scheme used herein. Since the proposed method is based on eigenspace, this section deals mainly with the concept of eigenspace.

In the present study, all images are normalized as follows. Let an N-dimensional vector $\mathbf{X}$ denote an original image composed of N pixels, and let $\mathbf{1}$ denote an N-dimensional vector in which each element is 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 $\mathbf{1}^T\mathbf{x} = 1$. An image space constructed by a set of normalized images is called the Normalized Image Space (NIS).

An eigenspace constructed by mean vector $\overline{\mathbf{x}}$ and m-principal eigenvectors Φ_m in NIS is described as $\langle \overline{\mathbf{x}}, \Phi_m \rangle$. In NIS, an image $\mathbf{x}$ is projected onto eigenspace $\langle \overline{\mathbf{x}}, \Phi_m \rangle$ by

$$\tilde{\mathbf{x}}^* = \tilde{\Phi}_m^+ \mathbf{x},$$

where $\tilde{\Phi}_m = [\Phi_m \ \ \overline{\mathbf{x}}]$ and $\tilde{\Phi}_m^+ = (\tilde{\Phi}_m^T \tilde{\Phi}_m)^{-1} \tilde{\Phi}_m^T$.

In order to measure the similarity between an input image $\mathbf{x}$ and the eigenspace $\langle \overline{\mathbf{x}}, \Phi_m \rangle$, we define a normalized correlation in terms of NIS, which can be defined by the cosine of an angle when an image $\mathbf{1}/N$ is regarded as the origin of the NIS. That is, a normalized correlation C_I between $\mathbf{x}$ and $\langle \overline{\mathbf{x}}, \Phi_m \rangle$ is defined as

$$C_I = C(\mathbf{x}, \tilde{\Phi}_m \tilde{\mathbf{x}}^*) \tag{1}$$

where

$$C(\mathbf{x}, \mathbf{y}) = \frac{(\mathbf{x} - \mathbf{1}/N)^T (\mathbf{y} - \mathbf{1}/N)}{||\mathbf{x} - \mathbf{1}/N||^{1/2} ||\mathbf{y} - \mathbf{1}/N)||^{1/2}}. \tag{2}$$

By this definition, a given image $\mathbf{x}$ can be evaluated in terms of NIS without explicit normalization.

2.2 Partial Projection

Let us define an indicator matrix P, which is an $N \times N$ diagonal matrix, each diagonal term of which is 1 or 0, which indicates whether the pixel is effective (1) or ineffective (0) for the projection. Then, $\mathbf{x}$ is partially projected onto $\langle \overline{\mathbf{x}}, \Phi_m \rangle$ with indicator matrix P by

$$\tilde{\mathbf{x}}_P^* = (P\tilde{\Phi}_m)^+ P\mathbf{x}, \tag{3}$$

where $\tilde{\Phi}_m = [\Phi_m \ \overline{\mathbf{x}}]$ and $(P\tilde{\Phi}_m)^+ = (\tilde{\Phi}_m^T P\tilde{\Phi}_m)^{-1}(P\tilde{\Phi}_m)^T$. A partial residual is defined as

$$\tilde{\mathbf{x}}_P^\sharp = P(\mathbf{x} - \tilde{\Phi}_m \tilde{\mathbf{x}}_P^*). \tag{4}$$

The last element of $\tilde{\mathbf{x}}_P^*$ is important and is denoted by β_P. β_P is equivalent to the total pixel values estimated by the partial projection. When the eigenspace cannot be constructed because only one image is available, we can regard the image as a 0-dimensional eigenspace. The normalized correlation C_I can be extended to span the partial projection. A partial correlation C_P between $\mathbf{x}$ and $\langle \overline{\mathbf{x}}, \Phi_m \rangle$ within a pixel set indicated by P is defined as

$$C_P = C(P\mathbf{x}, P\tilde{\Phi}_m \tilde{\mathbf{x}}_P^*). \tag{5}$$

When P is an identity matrix, Eq. (5) is equivalent to Eq. (1).

3 Projectional Decompositions

3.1 Decomposition by Canonical Eigenspace

A facial image contains various types of information, such as head pose, lighting condition, and individuality. In face recognition, it is important to decompose the facial image into the *individuality* and the other information. In the present paper, we refer to this decomposition as a projectional decomposition. In this section, we discuss the projectional decomposition for face recognition.

Principal component analysis (PCA) reduces the dimension of the face space with little loss of representability [1]. Shakunaga and Shigenari[8] proposed an image decomposition by an eigenspace that is constructed from a lot of facial images taken under various lighting conditions. Their method is used as a projectional decomposition in the present paper. We consider an eigenspace constructed from a lot of facial images as the *canonical face space*. The eigenspace is referred to as the canonical space, or CS for short, and the images used for CS construction are referred to as the canonical set. Figure 1 shows examples of the CS. Information that cannot be represented in the CS is regarded as the *individuality*.

The canonical space can be used for decomposing a facial image into the canonical information and the individuality. The former is a projection onto CS, and the latter is the residual of the projection. They are orthogonal to each other.

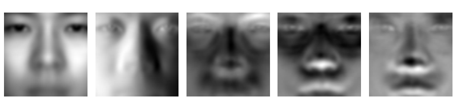


Fig. 1. Example of CS: Mean vector (leftmost image) and four principal bases

Let $\langle \overline{\mathbf{x}}_{\mathrm{cs}}, \Phi_{\mathrm{cs}} \rangle$ denote CS. The projection of an image $\mathbf{x}$ onto CS is given by

$$\tilde{\mathbf{x}}^* = \tilde{\Phi}_{\mathrm{cs}}^{+}\mathbf{x},$$

where $\tilde{\Phi}_{\mathrm{cs}} = [\Phi_{\mathrm{cs}}\ \overline{\mathbf{x}}_{\mathrm{cs}}]$. In the original image space, the projection $\tilde{\mathbf{x}}^*$ is described by

$$\mathbf{x}^{\$} = \tilde{\Phi}_{\mathrm{cs}}\tilde{\mathbf{x}}^*.$$

The residual $\mathbf{x}^{\sharp}$ is then expressed as

$$\mathbf{x}^{\sharp} = \mathbf{x} - \mathbf{x}^{\$}.$$

The decomposition of $\mathbf{x}$ into $\mathbf{x}^{\$}$ and $\mathbf{x}^{\sharp}$ is hereinafter referred to as CS decomposition.

Although the individuality may be represented by only the residual in an ideal environment, it is impossible to completely decompose an input image into the individuality and the other properties in an ordinary environment. Therefore, we simultaneously use both the projection and the residual for face recognition because they are complementary.

A face recognition algorithm is constructed in the conventional way using these two components. In the face registration stage, one eigenspace is constructed from a set of the projections and is denoted by $\langle \overline{\mathbf{x}}^{\$}, \Phi_m^{\$} \rangle$. The other eigenspace is constructed from a set of the residuals and is denoted by $\langle \overline{\mathbf{x}}^{\sharp}, \Phi_m^{\sharp} \rangle$. In the recognition stage, a projection $\mathbf{x}^{\$}$ and a residual $\mathbf{x}^{\sharp}$ are evaluated independently by

$$C^{\$} = C(\mathbf{x}^{\$}, \tilde{\Phi}_m^{\$}\tilde{\Phi}_m^{\$+}\mathbf{x}^{\$}) \tag{6}$$

and

$$C^{\sharp} = C(\mathbf{x}^{\sharp}, \tilde{\Phi}_m^{\sharp}\tilde{\Phi}_m^{\sharp+}\mathbf{x}^{\sharp}). \tag{7}$$

Finally, the image $\mathbf{x}$ is evaluated by adding $C^{\$}$ to $C^{\sharp}$.

The similarity $C^{\$}$, calculated in CS, is a variation of the well known distance-in-feature-space [2]. However, the similarity $C^{\sharp}$ is definitely distinct from the distance-from-feature-space. In the distance-from-feature-space, all of the residual components are simply summed up to L2-norm. In contrast, the similarity

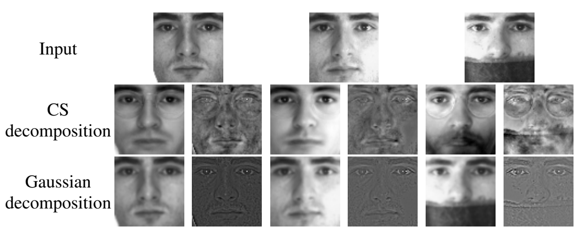


Fig. 2. Examples of CS decomposition and Gaussian decomposition: original images ($\mathbf{x}$) (top row), CS decomposition results ($\mathbf{x}^{\$}$ and $\mathbf{x}^{\sharp}$) (middle row), and Gaussian decomposition results ($\mathbf{x}_G^{\$}$ and $\mathbf{x}_G^{\sharp}$) (bottom row)

$C^\sharp$ is the similarity between the residual $\mathbf{x}^\sharp$ and the eigenspace $\langle \overline{\mathbf{x}}^\sharp, \Phi_m^\sharp \rangle$ in CS. In other words, $C^\sharp$ is the "distance-in-another-feature-space."

Figure 2 shows three examples of the CS decomposition in which input images were not used for constructing CS. The left and center input images, which do not contain an occlusion, are appropriately projected onto CS. Therefore, they are properly decomposed. In the right image, however, an occlusion by a scarf affects both the projection and residual.

3.2 Decomposition by Gaussian Filter

Canonical space decomposition is useful when an appropriate learning set can be prepared for the CS construction. However, often, when a facial image is taken using a different camera under different conditions, CS may not properly decompose the image into the canonical information and the individuality. In addition, when a test image contains numerous noises, such as occlusions, the noises may affect the entire image upon projection onto CS. Furthermore, the test image should be aligned with CS before the CS decomposition. In order to avoid these problems, we consider an alternative method that does not use CS for the projectional decomposition.

Wang et al.[9] proposed a self-quotient image (SQI) that extracts the component that is insensitive to illumination. In their method, a Gaussian filter is used to extract lighting information. The Gaussian filter is used in the proposed method for the projectional decomposition. Let G denote an $N \times N$ matrix that works as the Gaussian filter. Then, the decomposition of image $\mathbf{x}$ into a Gaussian image $\mathbf{x}_G^\$$ and its residual $\mathbf{x}_G^\sharp$ by the Gaussian filter can be formulated as

$$\mathbf{x}_G^\$ = G\mathbf{x} \tag{8}$$

and

$$\mathbf{x}_G^\sharp = \mathbf{x} - \mathbf{x}_G^\$. \tag{9}$$

Since the matrix G can be regarded as a projection matrix, $\mathbf{x}_G^\$$ can be regarded as a component of the Gaussian space. In this formulation, no a priori knowledge is necessary because $\mathbf{x}_G^\$$ can be calculated from only the input image. An input image is simply decomposed into the Gaussian image and its residual. This decomposition is referred to hereinafter as Gaussian decomposition. Figure 2 also shows three examples of Gaussian decomposition. Although the right input image includes an occlusion by a scarf, the effect of this occlusion does not spread to the entire image. [1]

[1] In the self-quotient image, each pixel value of the input image should be divided by the corresponding pixel in the Gaussian image in order to cancel the effect of illumination. In the proposed method, however, the Gaussian image is subtracted from the original image in order to calculate the residual. That is, the Gaussian image and the residual are regarded as approximations of illumination and individuality, respectively. This is an alternative method of calculating the self-quotient image, and the computational cost is lower than the self-quotient image because the residual can be calculated by subtraction rather than division.

4 Locational Decomposition

4.1 Parallel Partial Projections

When an input image contains local noises, such as shadows or occlusions, the noises affect the recognition results. First, in the most commonly used method, although images for face recognition are normalized by some method, when the image contains noises, the image cannot be properly normalized. Second, when we use an eigenspace, the effects of noises is spread to the entire image by the projection onto the eigenspace, affecting the face recognition results.

In order to avoid this problem, we utilize local information independently. In this section, we introduce a locational decomposition algorithm, which can utilize local information independently.

A framework of parallel partial projections (PPP) onto an eigenspace is proposed for face recognition under various lighting conditions[5]. This is one method for implementing the locational decomposition, and so local information is treated independently and the spread of noises is prevented. In the present paper, this method is used as the locational decomposition of the image.

Let us describe the j-th partial projection $\tilde{\mathbf{x}}^*_{P_j}$ onto an individual eigenspace $\langle \overline{\mathbf{x}}, \Phi_m \rangle$. Here, we consider a set of partial projections $\{\tilde{\mathbf{x}}^*_{P_1}, \cdots, \tilde{\mathbf{x}}^*_{P_M}\}$, where M is the number of parts indicated by P_j. This can be represented by the backprojected image, which can be calculated as

$$\mathbf{x}^{\$'} = \sum_{j=1}^{M} P_j \tilde{\Phi}_m \tilde{\mathbf{x}}^*_{P_j}.$$

In the discriminant function for PPP, we use a partial correlation. The (partial) correlation is essentially robust with respect to noises because it represents the cosine of the angle between two vectors. The image $\mathbf{x}$ is evaluated by

$$C' = \sum_{j=1}^{M} C(P_j \mathbf{x}, P_j \tilde{\Phi}_m \tilde{\mathbf{x}}^*_{P_j}), \tag{10}$$

where M is the number of P_j and $C(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y} / (\mathbf{x}^T \mathbf{x} \mathbf{y}^T \mathbf{y})^{1/2}$. Of course, PPP can be used not only for the eigenspace, but also for only one image. When only one image can be registered, the image is regarded as a 0-dimensional eigenspace consisting of the image.

The face recognition algorithm is summarized in Fig 3. Here, local noises are not spread by projection onto the eigenspace.

4.2 Division Scheme

In face recognition using the parallel partial projections, the indicator matrix P can be used to indicate an arbitrary area in the facial image. The image contains some characteristic points such as the eyes, nose and mouth. Several previously

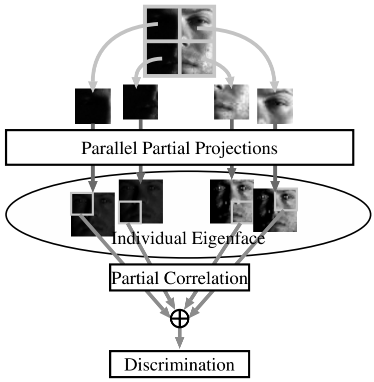


Fig. 3. Parallel partial projections for face recognition

proposed methods have used these characteristic points[2]. Although this method is effective, we do not use characteristic points in the proposed method because correctly determining an effective position for recognition is difficult. Therefore, points that are characteristic points from a human viewpoint may not be characteristic points from the viewpoint of a computer. Furthermore, when the proposed method is applied to the recognition of some other objects, proper characteristic points for the recognition are impossible to conceive of ahead of time.

Therefore, we do not herein consider the optimal placement of P. In the proposed method, images are divided into a set of squares, and experimental results, described later herein, show that the proposed method works well without optimal placement of P.

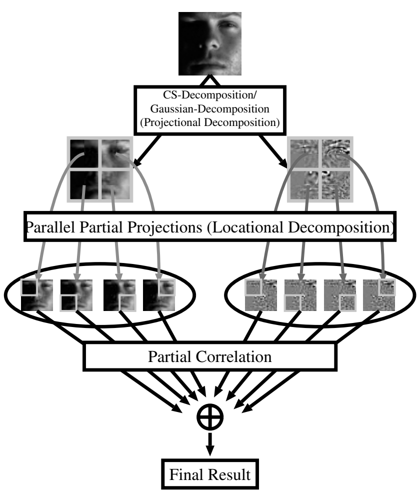


Fig. 4. Combination of projectional and locational decompositions

5 Combination of Locational and Projectional Decompositions

5.1 Combination of CS Decomposition and Parallel Partial Projections

A projectional decomposition and a locational decomposition can be combined in a simple manner. We show two combinations of projectional and locational decompositions. In the combination methods, an input image is projectionally decomposed by either CS decomposition or Gaussian decomposition, and the two decomposed components are evaluated in a framework of the locational decomposition. Figure 4 shows the concept of the combinations.

First, an input image $\mathbf{x}$ is projection decomposed by parallel partial projections onto CS.

$$\tilde{\mathbf{x}}^{*}_{P_j} = (P_j \tilde{\Phi}_{\rm cs})^{+} P_j \mathbf{x} \tag{11}$$

$$\mathbf{x}^{\sharp}_{P_j} = P_j(\mathbf{x} - \tilde{\Phi}_{\rm cs}\tilde{\mathbf{x}}^{*}_{P_j}), \tag{12}$$

where $\tilde{\Phi}_{\rm cs} = [\Phi_{\rm cs} \;\; \overline{\mathbf{x}}_{\rm cs}]$. We define $\mathbf{x}^{\$'}$ and $\mathbf{x}^{\sharp'}$ as

$$\mathbf{x}^{\$'} = \sum_{j=1}^{M} P_j \tilde{\Phi}_{\rm cs}\tilde{\mathbf{x}}^{*}_{P_j} \tag{13}$$

$$\mathbf{x}^{\sharp'} = \sum_{j=1}^{M} P_j \mathbf{x}^{\sharp}_{P_j}, \tag{14}$$

where M is the number of parts. This method realizes the projectional decomposition without any noise expansion because the parallel partial projections onto canonical space do not spread noises. Examples of the decomposition by PPP are shown in Fig 5.

The decomposed images can be locationally decomposed and evaluated in a straightforward manner. In the combination method, the partial correlation should be defined for each component. When an eigenspace constructed from a set of $\mathbf{x}^{\$'}$ is denoted by $\langle \overline{\mathbf{x}}^{\$'}, {\Phi^{\$}_m}' \rangle$, a partial correlation $C^{\$}_{P_j}$ between $\mathbf{x}^{\$'}$ and $\langle \overline{\mathbf{x}}^{\$'}, {\Phi^{\$}_m}' \rangle$ within a pixel set indicated by P_j is calculated by

$$C^{\$}_{P_j} = C(P_j \mathbf{x}^{\$'}, P_j \tilde{\Phi}^{\$'}_m (P_j \tilde{\Phi}^{\$'}_m)^{+} P_j \mathbf{x}^{\$'}). \tag{15}$$

In a similar manner, a partial correlation $C^{\sharp}_{P_j}$ between a residual $\mathbf{x}^{\sharp'}$ and an eigenspace $\langle \overline{\mathbf{x}}^{\sharp'} {\Phi^{\sharp}_m}' \rangle$ is calculated by

$$C^{\sharp}_{P_j} = C(P_j \mathbf{x}^{\sharp'}, P_j \tilde{\Phi}^{\sharp'}_m (P_j \tilde{\Phi}^{\sharp'}_m)^{+} P_j \mathbf{x}^{\sharp'}), \tag{16}$$

where $\langle \overline{\mathbf{x}}^{\sharp'} {\Phi^{\sharp}_m}' \rangle$ is constructed from a set of the residuals defined in Eq. (14). Then, the total correlation ${C_{\rm cs}}'$ is defined as

$${C_{\rm cs}}' = w \sum_{j=1}^{M} C^{\$}_{P_j} + (1-w) \sum_{j=1}^{M} C^{\sharp}_{P_j}, \tag{17}$$

where w is the weight of the projectional components.

Although the locational decomposition provides robustness with respect to local noises, effective information for face recognition does not increase in the entire image. The locational decomposition still requires a sufficient number of registered images for each person because the conventional eigenface method requires a lot of images for the stable recognition. On the other hand, the projectional decomposition often provides stable results even when only a few images are registered. However, the projectional decomposition is sometimes seriously affected by local noises. In the combination method, however, the locational decomposition prevents local noises from spreading to the entire image when the projectional decomposition provides sufficient information for face recognition. Therefore, the combination method works better than the individual decompositions.

5.2 Combination of Gaussian Decomposition and PPP

A combination of the Gaussian decomposition and the parallel partial projections is more straightforward and simpler than the CS decomposition because the Gaussian filter uses only local (independent) information of the input images. In Fig. 5, the right-most images show the results of the Gaussian decomposition, which are similar to the results of the parallel partial projections onto CS using 64(8×8) square subregions as shown in the most upper row. In this method, an input image is decomposed by the Gaussian filter. The Gaussian component and the residual are locationally decomposed and evaluated in a manner similar

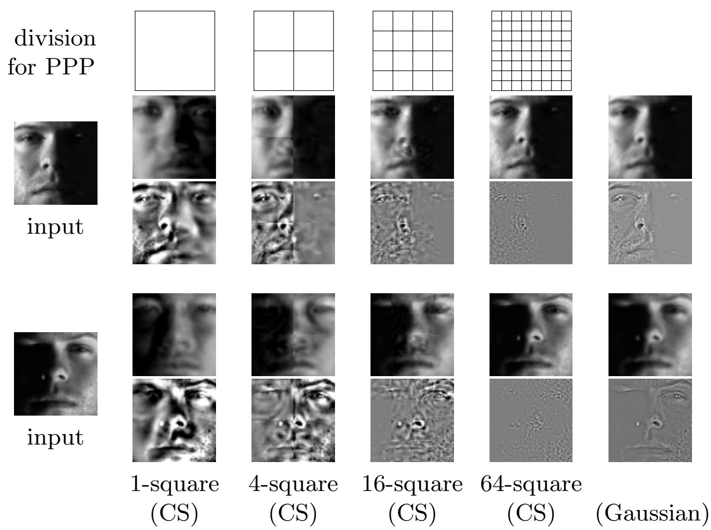


Fig. 5. Examples of the decomposition by parallel partial projections onto CS and the Gaussian filter. The upper row shows the projected images, and the lower row shows the residuals. The first column shows the input images. The second through fifth columns show the images decomposed by the parallel partial projections ($M = 1, 4, 16, 64$) onto CS. The sixth column shows the images decomposed by Gaussian decomposition.

to that described in the previous section. Let $C^{\$}_{GP_J}$ denote the partial correlation between a Gaussian component $\mathbf{x}^{\$}_G$ and an eigenspace constructed in the Gaussian space, and let $C^{\sharp}_{GP_J}$ denote the correlation with the residual. Then, an image $\mathbf{x}$ is evaluated by

$$C_G{}' = w \sum_{j=1}^{M} C^{\$}_{GP_j} + (1-w) \sum_{j=1}^{M} C^{\sharp}_{GP_j}, \tag{18}$$

where w is the weight of the projectional components.

6 Experimental Results

6.1 Results for Yale Face Database B

Data Specifications

We performed discrimination experiments on 640 frontal facial images of 10 people, which were taken from the Yale Face Database B [10]. The database includes 65 frontal facial images of each person. Sixty-four of the images were taken under different lighting conditions, and one special image was taken under ambient light. In order to remove the contribution of ambient light, we prepared 64 images of each person with the ambient image subtracted. At the same time, each image was converted to a 64 × 64 pixel image such that the eyes of all of the images are located at the same coordinates, as shown in Fig. 6.

Discrimination experiments were performed using the segmented data set. Figure 7 shows examples of the five subsets (SS1-5). In the first set of experiments, only the frontal illuminated images in SS1 were used as registered images, and all of the images in SS1 were used in the second set of experiments.

The CS is created from a canonical set from our laboratory, which consists of 1,200 images of 50 people. For each person, images were taken under 24 lighting conditions.

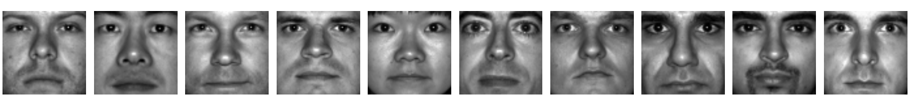


Fig. 6. Segmented facial images in Yale Face Database B

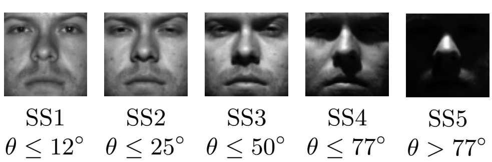


Fig. 7. Example images in subsets 1-5 (SS1-5), where θ is the angle between the light source direction and the camera axis

Discrimination Results

Table 1 shows the discrimination rates for the dataset when only one image is registered from SS1. In the methods that use the PPP, images were divided into sixty-four squares. Among the three single decomposition methods, PPP, the CS decomposition and the Gaussian decomposition, CS decomposition provides the worst results because the method spreads noises by the projection onto CS. Although the PPP provides better results than CS decomposition by preventing the expansion of noises, the results are not sufficient because the method does not include individuality-extraction. Gaussian decomposition provides the best results among the three methods because it can approximately extract individuality without any noise expansion.

The two combination methods, PPP-CS and PPP-Gaussian, work much better than the other methods because they not only extract individuality but also include schemes for avoiding the problems of noises. In addition, in the combination methods, CS decomposition works as well as Gaussian decomposition because CS decomposition does not spread noises by the parallel partial projections.

Table 2 shows discrimination rates when seven images are registered from SS1 for each person. In the experiments, the PPP and the combination methods give the complete discrimination because a sufficient number of images are registered. Two projectional decompositions give slightly worse results than PPP because they cannot sufficiently suppress the noises.

Table 1. Discrimination rates (%) for Yale Face Database B when only one image is registered from SS1. NN denotes the Nearest Neighbor method, PPP denotes the Parallel Partial Projections (locational decomposition), and CS and Gaussian denote the CS and Gaussian projectional decompositions, respectively. In addition, PPP-CS and PPP-Gaussian are combination methods.

Test Class	Method					
	NN	PPP	CS	Gaussian	PPP-CS	PPP-Gaussian
Subset 2	99.2	100	100	100	100	100
Subset 3	74.6	99.2	83.1	99.2	100	100
Subset 4	30.4	78.3	65.9	83.3	98.6	100
Subset 5	12.2	78.3	23.8	44.4	100	100

Table 2. Discrimination rates (%) for Yale Face Database B when seven images are registered from SS1: EF indicates the eigenface method and the other methods are as listed in Table 1

Test Class	Method					
	EF	PPP	CS	Gaussian	PPP-CS	PPP-Gaussian
Subset 2	100	100	100	100	100	100
Subset 3	100	100	100	100	100	100
Subset 4	93.5	100	98.6	98.6	100	100
Subset 5	56.1	100	52.9	74.1	100	100

Table 3. Comparison of the number of parts for each algorithm when seven images are registered from SS1. SS4 and SS5 are used as test sets in the experiment.

# parts	Test Class	Method		
		PPP	PPP-CS	PPP-Gaussian
1 × 1	SS4	93.5	98.6	98.6
	SS5	56.1	52.9	74.1
2 × 2	SS4	96.4	98.6	99.3
	SS5	94.7	90.5	97.4
4 × 4 / 8 × 8	SS4	100	100	100
	SS5	100	100	100
16 × 16	SS4	98.6	98.6	100
	SS5	96.3	99.5	99.5

Table 4. Discrimination rates (%) when one image randomly selected from SS4 is registered

Test Class	Method					
	NN	PPP	CS	Gaussian	PPP-CS	PPP-Gaussian
Subset 1	16.7	41.3	39.2	57.8	92.2	96.7
Subset 2	18.4	41.2	36.0	48.3	90.0	93.8
Subset 3	22.0	37.3	33.4	39.2	71.5	78.3
Subset 5	21.4	37.0	25.4	30.2	83.4	84.3

Table 5. Discrimination rates (%) when seven images randomly selected from SS4 are registered

Test Class	Method					
	EF	PPP	CS	Gaussian	PPP-CS	PPP-Gaussian
Subset 1	86.7	99.5	97.6	100	100	100
Subset 2	91.3	99.9	98.6	99.9	100	100
Subset 3	95.5	97.8	98.5	99.4	100	100
Subset 5	70.9	98.1	75.0	87.0	100	100

Table 3 shows the results when the input image is divided to different numbers of image parts. When the number of image parts is too large, the discrimination rate becomes worse because each part can not provide sufficient information for recognition because it is too small. In the experiments, the best result is provided when the number of parts is 4 × 4 and 8 × 8.

Tables 4 and 5 show results when images classified into SS4 are registered. In these experiments, images for registration are randomly selected from SS4. This process was repeated twenty times and the registered images for each person were varied. Most of the results for these experiments were worse than those shown in Tables 1 and 2 because the images in SS4 include more shadows than SS1. However, the results for the combination methods retained high discrimination rates in the experiments.

Table 6. Discrimination rates (%) using other methods: Illumination cone (IC1), illumination cone with cast shadow (IC2), photometric alignment using RANSAC (PA) and segmented linear subspace method (SLS). Note that only one image is registered for PPP-CS and PPP-Gaussian.

Test Class	Method					
	IC1[11]	IC2[11]	PA[4]	SLS[6]	PPP-CS	PPP-Gaussian
Subset 2	100	100	100	100	100	100
Subset 3	100	100	100	100	100	100
Subset 4	91.4	100	100	100	98.6	100
Subset 5	-	-	81.5	-	100	100

Table 7. Discrimination rates (%) for the AR Database

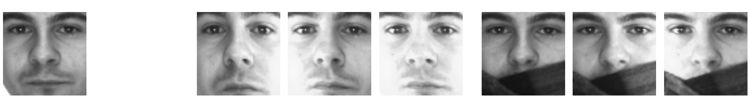
Test Class	Method					
	NN	PPP	CS	Gaussian	PPP-CS	PPP-Gaussian
light	40.5	70.1	89.1	82.2	94.8	95.3
scarf	3.7	45.4	37.0	63.7	83.7	84.7

Table 6 shows a number of results reported in the literature[11,4,6]. This table shows that all of the algorithms provide good results when seven images are registered from SS1. However, the proposed methods, PPP-CS and PPP-Gaussian, can provide almost same results with registering only one image from SS1.

In conclusion, the combination methods work better than the individual decomposition methods. In addition, the combination methods have the advantages of both the projectional and locational decompositions and work well even when only one image is registered and the test images or registered images include a significant number of shadows.

6.2 Results for the AR Database

The AR database[12] contains images of 135 people taken under various conditions for each person. For this experiment, we used database images taken under seven different conditions. The example images are as shown in Fig 8. In this experiment, only one image was registered and the other images were used as the test set from which test images were selected.



Under the normal condition (registered) — Under different lighting conditions (light-set) — Wearing a scarf (scarf-set)

Fig. 8. Examples of segmented images in the AR Database

Table 8. Discrimination rates (%) for our database when one image is registered for each person (a) and when all images classified into Classes 1 and 2 are registered (b)

Registered Class	Test Class	Method					
		EF	PPP	CS	Gaussian	PPP-CS	PPP-Gaussian
(a)	Class 2	54.0	83.6	87.1	81.3	95.9	94.4
	Class 3	20.5	70.2	84.2	72.4	86.6	84.3
(b)	Class 3	93.5	93.2	94.0	90.3	99.8	99.7

Table 7 shows the discrimination rates obtained in the experiments. For the light set, CS decomposition and Gaussian decomposition gave better results than the PPP. However, CS decomposition did not work for the scarf set because the test images included a large occlusion. The combination methods worked better than the other methods for both of the individual sets. The results of this experiment indicate that combination methods work well when only one image is registered and the test images include a large occlusion.

6.3 Results for Another Dataset Under the Same Conditions as the Canonical Set

Finally, experimental results are shown for a database that consists of a set of images taken under the same conditions as the canonical set. The database contains images of 50 people taken under 24 lighting conditions for each person. Each image was converted to a 32×32 pixel image. In the methods that use the PPP, images were divided into sixteen squares. The images are classified into three classes. Images classified into Class 1 are frontal illuminated and were used as registered images. Class 2 images, which contain small shadows, and Class 3 images, which contain large shadows, were used as test sets. Table 8 shows the discrimination rates for the database. In the dataset, the two methods that use CS decomposition work better than those that use Gaussian decomposition, because the illumination conditions of the canonical set are identical to those of the test set. The results suggest that the CS decomposition works better when the lighting conditions are similar between the canonical set and the test set.

7 Conclusions

Combination methods of the two types of decomposition, projectional and locational, has been proposed. The projectional decomposition method can extract the individuality from an image. In particular, the Gaussian decomposition can extract the individuality when the image contains noises. The locational decomposition provides robustness with respect to noises when the eigenspace can be constructed properly. The combination methods have the advantages of both of the decomposition methods. The method of combining projectional and locational decompositions works well even when only one image is registered and test images or registered images contain numerous noises, such as shadows or occlusions. We hope that the concept of the proposed method will be useful in solving other problems in image recognition and computer vision.

Acknowledgments

The authors would like to thank A. Georghiades and A. M. Martinez for granting permission to use the Yale Face Database B [10]. and the AR Database[12]. This work was supported by a Grant-In-Aid for Scientific Research (No.15300062) from the Ministry of Education, Science, Sports, and Culture of Japan.

References

1. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience **3** (1991) 71–86
2. Moghaddam, B., Pentland, A.: Probabilistic visual learning for object representation. IEEE Trans. Pattern Analysis and Machine Intelligence **19** (1997) 696–710
3. Black, M., Jepson, A.: Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. International Journal of Computer Vision **26** (1998) 63–84
4. Okabe, T., Sato, Y.: Object recognition based on photometric alignment using ransac. In: Proc. CVPR2003. Volume 1. (2003) 221–228
5. Sakaue, F., Shakunaga, T.: Face recognition by parallel partial projections. In: Proc. ACCV2004. Volume 1. (2004) 1440–150
6. Batur, A., Hayes III, M.: Linear subspaces for illumination robust face recognition. In: Proc. ICCV2001. Volume II. (2001) 296–301
7. Ohba, K., Ikeuchi, K.: Detectability, uniquensess, and reliability of eigen windows for stable verification of partiallo occluded objects. IEEE Trans. Pattern Analysis and Machine Intelligence **19** (1997) 1043–1048
8. Shakunaga, T., Shigenari, K.: Decomposed eigenface for face recognition under various lighting conditions. In: Proc. CVPR2001. Volume 1. (2001) 864–871
9. Wang, H., Li, S., Wang, Y.: Face recognition under varying lighting conditions using self quotient image. In: Proc. IEEE FG2004. (2004) 819–824
10. Georghiades, A., Belhumeur, P., Kriegman, D.: From few to many: Generative models for recognition under variable pose and illuminations. In: Proc. FG2000. (2000) 277–284
11. Georghiades, A., Belhumeur, P., Kriegman, D.: From few to many: illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Analysis and Machine Intelligence **23** (2001) 643–660
12. Martinez, A., Benavente, R.: The AR face database. Technical Report #24, CVC (1998)