Scale Invariant Face Recognition Method Using Spectral Features of Log-Polar Image

Kazuhiro HOTTA[a], Takio KURITA[b], and Taketoshi MISHIMA[a]

[a]Saitama University, 255, Shimo-Okubo, Urawa City, Saitama, 338-8570 JAPAN
[b]Electrotechnical Laboratory, 1-1-4, Umezono, Tsukuba City, Ibaraki, 305-8568 JAPAN

ABSTRACT

This paper presents scale invariant face detection and classification methods which use spectral features extracted from Log-Polar image. Scale changes of a face in a image are represented as shift along the vertical axis in Log-Polar image. In order to make them robust to the scale changes of faces, spectral features are extracted from the each row of the Log-Polar image. Autocorrelations, Fourier power spectrum, and PARCOR coefficients are used as spectral features. Then these features are combined with simple classification methods based on the Linear Discriminant Analysis to realize scale invariant face detection and classification. The effectiveness of the proposed face detection method is confirmed by the experiment using the face images which are captured under the different scales, backgrounds, illuminations, and dates. We have also performed the experiments to evaluate the proposed face classification method using 2800 face images with 7 scales under 2 different backgrounds.

Keywords: Scale Invariant, Face Detection, Face Classification, Spectral Features, and Log-Polar Image

1. INTRODUCTION

Face Recognition has many applications such as access control to building etc.,[1] man-machine interface,[2] vision system on the mobile robot,[3] the search and indexing from the video library using face as key,[4–6] identification of the passport holders at airport and users at ATM,[7] and so on. Because of these applications, many face recognition methods have been proposed.[8,9] Some methods can achieve high recognition rates if the scale and the position of a face in a given image is normalized in advance.[10,11] For many of practical applications, however, scale and position of a face in a given image will change because the relative position between the person and the camera is different. The robustness to the scale and the position changes is very important for practical applications, especially for online applications.

To cope with the position changes of faces, face recognition task is usually divided into two stages; face detection and face classification. A position of a face candidate is searched in face detection stage and the face centered at that position is classified by face classification.

Face detection should be robust to the changes of scale of the faces because the sizes of faces are not identical in all the images. In the face detection methods based on pattern matching,[12–15] the scale changes of the face is coped with by changing the size of the image itself. However the normalization of the scale is not easy and requires much computation. The color informations of faces are often utilized to segment the face region.[16] Although the color information of faces is effective to narrow down the candidates of faces in the image frame, it is difficult to detect the correct position of the face. The position of facial features are also utilized for face detection,[17] but there is the same difficulty to find the position of the facial features.

Many of the typical face classification methods are not robust to the scale changes of the face. This means that the normalization of the scale after the face detection are assumed in these methods. Therefore the performance of face recognition heavily depends on the accuracy of the normalization.

From these consideration, we have already proposed a face recognition method which uses the scale and rotation invariant features extracted from Log-Polar image.[18,19] Both in face detection[19] and face classification,[18] we used

Further author information: (Send correspondence to K.Hotta)
K.Hotta: E-mail: hotta@me.ics.saitama-u.ac.jp
T.Kurita: E-mail: kurita@etl.go.jp
T.Mishima: Email: mishima@me.ics.saitama-u.ac.jp

Higher-order Local AutoCorrelation (HLAC) features.[20] However the robustness to 2D rotations is not important for face recognition tasks because (1) the face direction in 2D is usually fixed such that the hair is up and the chin is down, and (2) the robustness to 2D rotations of the face may induces the miss classification because it increases the possibility that the features extracted from different faces become similar.[14] It is expected that the face detection and classification can be improved by using the features which are robust to the only scale changes of the face but not to the rotations. In this paper we propose a scale invariant face detection and classification method which uses spectral features extracted from Log-Polar image. Autocorrelations, Fourier power spectrum, and PARCOR coefficients are used as spectral features. Then these features are combined with simple classification methods based on the Linear Discriminant Analysis to realize scale invariant face detection and classification.

In section 2, Log-Polar image and spectral features are explained. Face detection method and face classification method are explained in section 3 and 4, respectively. Both methods use simple classification method based on Linear Discriminant Analysis (LDA). The experimental results of the scale invariant face detection are shown in section 5. In section 6 the effectiveness of the face classification method is described.

2. SCALE INVARIANT FEATURES

It is well known that the density of photo-receptors in the retina is space-variant. One of the simplest models of space variant sensor is Log-Polar transformation.[21,22] Higher weights are given for the central region of the input image than the peripheral region in the Log-Polar images which are obtained through Log-Polar transformation. This property of the Log-Polar images is good for target recognition because the peripheral region often includes unnecessary information such as background. Another important property of Log-Polar image is that the scale changes and rotations of a target on an input image are represented as shifts in the Log-Polar image if the center of the target is fixed at the center of the image. This means that we can obtain scale and rotation invariant features by extracting shift invariant features from Log-Polar image.

By using these property we have already proposed scale and rotation invariant features which are based on Higher-order Local AutoCorrelation (HLAC).[20] These features are used for face detection[19] and face classification.[18] However the robustness to 2D rotations is not important for face recognition tasks because the face direction in 2D is usually fixed and the robustness to 2D rotations of the face may induces the miss classification. Thus this paper proposes features which are robust to the only scale changes but not to the 2D rotation. Since scale changes of the target are represented as the shift along the horizontal axis on the Log-Polar image, we can obtain scale invariant features by taking shift invariant features along the horizontal axis on the Log-Polar image. It is well known that spectral features are shift invariant. Autocorrelation features, Fourier power spectrum features, and PARCOR features are used as the spectral features.

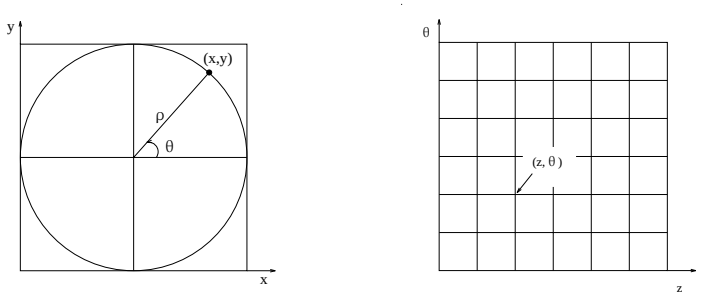
2.1. Log-Polar Transformation

Input image is generally represented as a collection of pixel points on the Cartesian coordinate. Here we take the origin at the middle pixel in the width and the height of a image. Then Log-Polar image can be constructed by the following transformations of the coordinates. At first, the point (x, y) on the Cartesian coordinate is transformed into the point $(\rho = \sqrt{(x^2 + y^2)}, \theta = \arctan(\frac{y}{x}))$ on the Polar coordinate. The point on the Polar coordinate is transformed into the point $(z = \log(\rho), \theta)$ on the Log-Polar coordinate by taking the logarithm of the scale ρ. Figure 1 shows Cartesian coordinate and Log-Polar coordinate.

We use the re-sampling method by the inverse transformation to obtain the Log-Polar image from the input image. To obtain the pixel value at the point (z_i, θ_j) on the Log-Polar image (coordinate), the point is inversely transformed into the point $(\exp(z_i)\cos(\theta_j), \exp(z_i)\sin(\theta_j))$ on the Cartesian coordinate in which the center of the image represent $(0, 0)$. Then the value of the point (z_i, θ_j) is estimated as the mean intensity value of the neighboring points of the back-projected point $(\exp(z_i)\cos(\theta_j), \exp(z_i)\sin(\theta_j))$ on the input image. We can obtain a Log-Polar image by performing this estimation for all points on the Log-Polar coordinate. Figure 2 shows the example of an input image, the sampling points used to construct the Log-Polar image, and its Log-Polar image. It is noticed that the sampling density increases toward the center of the image and decreases from the center to the periphery. This means that the extracted features contain much information of the target on the center region than that on the periphery regions such as background etc.

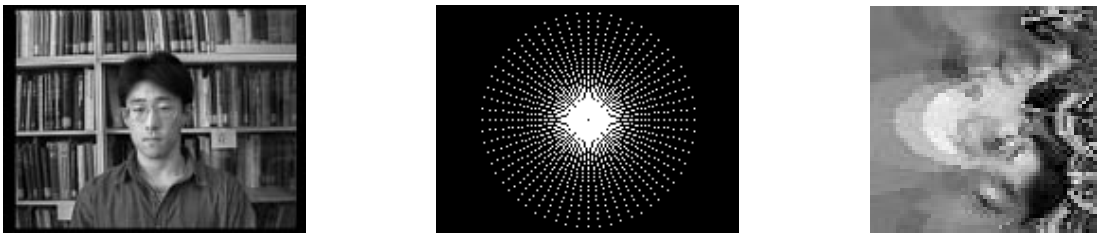
Log-Polar image has a good property for scale invariant feature extraction. Scalings of a target are represented as shifts along z $(= \log(\rho))$ axis on the Log-Polar image (coordinate). Rotations of a target are also represented as

shifts along θ axis. Figure 3 shows Log-Polar image of a simple 2D shape with different scales and rotations. It is noticed that both scalings and rotations of the target are represented as shifts in Log-Polar image (coordinate).



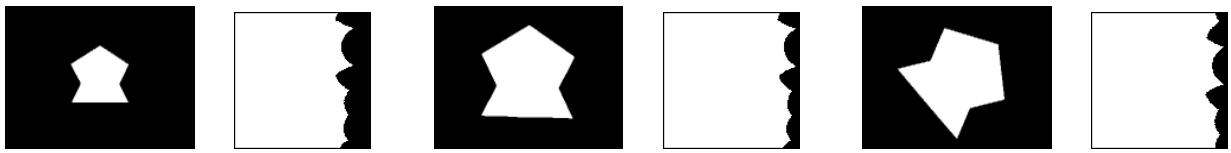
(a) Cartesian coordinate (b) Log-Polar coordinate

Figure 1. Cartesian coordinate and Log-Polar coordinate.



(a) (b) (c)

Figure 2. Log-Polar transformation. (a) Input image (160×120 pixels). (b) Sampling points used to construct the Log-Polar image. (c) Log-Polar transformed image (60×60 pixels).



(a) (b) (c) (d) (e) (f)

Figure 3. Examples of Log-Polar image of 2D shapes. (a) Small size. (b) Log-Polar image of (a). (c) Normal size. (d) Log-Polar image of (c). (e) 45° rotated image of (c). (f) Log-Polar image of (e).

2.2. Spectral Features

The scale changes of a target are represented as shifts along the vertical axis on the Log-Polar image. To make the features robust to the scale changes of the face, spectral features are extracted from each row on the Log-Polar image. Figure 4 shows how to extract the spectral features from Log-Polar image.

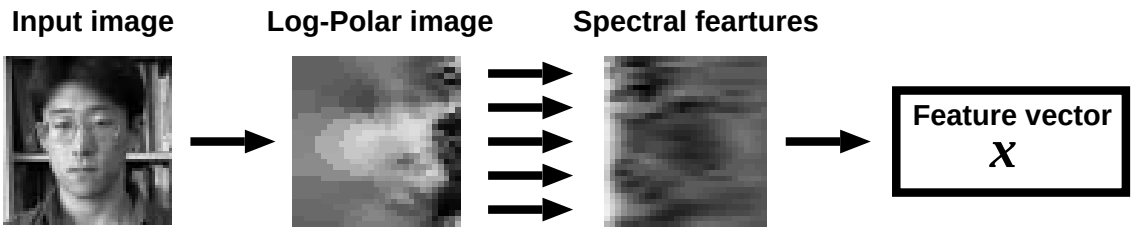


Figure 4. How to extract the spectral features from the Log-Polar image.

2.2.1. Autocorrelation Features

Let an row on the Log-Polar image be $x(t)$. The auto-covariance of the input signal $x(t)$ is defined by

$$R(s) = \frac{1}{N}\sum_{t=0}^{N-1}(x(t)-\bar{x})(x(t+s)-\bar{x}) \tag{1}$$

Then the auto-covariance of $x(t)$ depends on only the difference.

The autocorrelation $\rho(s)$ of signal $x(t)$ is defined by

$$-1 \leq \rho(s) = \frac{R(s)}{R(0)} \leq 1 \tag{2}$$

The autocorrelation has maximum at $s = 0$ and becomes robust to the scaling of the signal because it is normalized by the variance $R(0)$. This means that the features based on the autocorrelation are robust to the changes of scaling of the intensities, namely lighting conditions.

2.2.2. Fourier Power Spectrum Features

The Fourier transform $X(w)$ of the signal $x(t)$ is defined by

$$X(w) = \sum_{\tau=0}^{N-1} x(\tau)e^{\frac{-j2\pi w\tau}{N}}. \tag{3}$$

Then the Fourier power spectrum is given as $|X(w)|$. This is also sift invariant.

There is a close relation between the autocorrelation and Fourier power spectrum. Fourier power spectrum of the auto-covariance is called periodogram and is define by

$$S(w) = \sum_{\tau=0}^{N-1} R(\tau)\ e^{\frac{-j2\pi w\tau}{N}}\ . \tag{4}$$

The periodogram is related to the Fourier power spectrum $|X(w)|$ of the signal $x(t)$ as

$$S(w) = \frac{1}{N}|X(w)|^2\ . \tag{5}$$

2.2.3. PARCOR Features

The PARtial autoCORrelation (PARCOR) coefficients (features) are often used in speech signal processing to extract dominant information from a time signal. Let the forward autoregressive (AR) model of order τ be

$$x(n) = \sum_{i=1}^{\tau}\alpha^{\tau}(i)\ x(n-i) + \varepsilon_f^{\tau}(n), \tag{6}$$

where $\{\alpha^\tau(i)|i=1,\cdots,\tau\}$ and $\varepsilon_f^\tau(n)$ are forward AR coefficients and forward prediction error, respectively. Similarly backward AR model of order τ is given by

$$x(n-\tau-1) = \sum_{i=1}^{\tau} \beta^\tau(i)\ x(n-i) + \varepsilon_b^\tau(n), \tag{7}$$

where $\{\beta^\tau(i)|i=1,\cdots,\tau\}$ and $\varepsilon_b^\tau(n)$ are backward AR coefficients and backward prediction error, respectively. Then PARCOR coefficients are defined as the correlations between prediction errors obtained by the forward and the backward autoregressive models of order $\tau-1$. Namely, the PARCOR coefficients are defined by

$$k_\tau = \frac{\sum_{n=\tau}^{N-1}\ \varepsilon_f^{\tau-1}(n)\ \varepsilon_b^{\tau-1}(n)}{\sqrt{\sum_{n=\tau}^{N-1}\{\varepsilon_f^{\tau-1}(n)\}^2 \sum_{n=\tau}^{N-1}\{\varepsilon_b^{\tau-1}(n)\}^2}}\ . \tag{8}$$

where the prediction errors are defined by

$$\begin{aligned} \varepsilon_f^{\tau-1}(n) &= x(n) - \sum_{i=1}^{\tau-1} \alpha^{\tau-1}(i)\ x(n-i), \\ \varepsilon_b^{\tau-1}(n) &= x(n-\tau) - \sum_{i=1}^{\tau-1} \beta^{\tau-1}(i)\ x(n-i). \end{aligned} \tag{9}$$

It is well known that the PARCOR coefficients k_τ is the same as the AR coefficients $\alpha^\tau(\tau)$ or $\beta^\tau(\tau)$ of AR models of order τ. There is a fast recursive algorithm for calculating AR and PARCOR coefficients.[23] The algorithm can compute all AR and PARCOR coefficients of orders 1 through τ with $O(\tau^2)$. The optimal dimension of the order of the PARCOR coefficient can be determined by An Information Theoretical Criterion (AIC).

3. FACE DETECTION METHOD

The spectral features extracted from a Log-Polar image are general and primitive and are independent on the recognition task. It is expected that these features have enough information to distinguish faces. To get new effective features for the given recognition task, it is necessary to combine these features. For this purpose, we use Linear Discriminant Analysis (LDA).

For face detection, we have to design a classifier which can classify "face" and "not face". It is expected that "face" class includes only face images but "not face" class includes many kinds of images except face images. It is difficult to recognize "not face" class as a single cluster in the feature space. Thus we modified the discriminant criterion such that the covariance of "face" class is minimized while the covariance between "face" class and each of the learning samples in "not face" class is maximized.

Let "face" class and "not face" class (samples) are represented as

$$C_F = \{\boldsymbol{x}_{Fi} \mid i=1,\cdots,N_F\}, \qquad C_{NF} = \{\boldsymbol{x}_{NFk} \mid k=1,\cdots,N_{NF}\}, \tag{10}$$

where N_F is the number of "face" samples and N_{NF} is the number of "not face" samples. Then the mean vector of "face" class is given by

$$\bar{\boldsymbol{x}}_F = \frac{1}{N_F}\sum_{i=1}^{N_F} \boldsymbol{x}_{Fi}\ . \tag{11}$$

The covariance matrix (Σ_F) of "face" class and the covariance matrix (Σ_C) between the mean vector of "face" class and each samples of "not face" class are given by

$$\begin{aligned} \Sigma_F &= \frac{1}{N_F}\boldsymbol{x}_{Fi}\boldsymbol{x}_{Fi}^T - \bar{\boldsymbol{x}}_F\bar{\boldsymbol{x}}_F^T\ , \\ \Sigma_C &= \frac{1}{N_{NF}}\sum_{k=1}^{N_{NF}} (\boldsymbol{x}_{NFk} - \bar{\boldsymbol{x}}_F)(\boldsymbol{x}_{NFk} - \bar{\boldsymbol{x}}_F)^T\ , \end{aligned} \tag{12}$$

where the symbol ^{T} denotes the transpose.

New features $\boldsymbol{y} = (y_1, \cdots, y_L)^T$ are obtained by linear combination of primitive features $\boldsymbol{x} = (x_1, \cdots, x_M)^T$ as

$$\boldsymbol{y} = A^T \boldsymbol{x}, \tag{13}$$

where $A = [a_{ij}]$ is a coefficients matrix, L is the number of new features, and M is the number of primitive features.

To construct the discriminant space in which the covariance of "face" class is minimized and the covariance between the mean vector of "face" class and each samples of "not face" class is maximized, we use the discriminant criterion

$$J = \boldsymbol{tr}(\hat{\Sigma}_F^{-1} \hat{\Sigma}_C), \tag{14}$$

where $\hat{\Sigma}_F$ and $\hat{\Sigma}_C$ are the covariance matrix of "face" class and the covariance matrix between the mean vector of "face" class and each samples of "not face" class in the discriminant space, respectively. The optimal coefficient A, which maximizes this discriminant criterion J, is obtained by solving the eigen-value problem

$$\Sigma_C A = \Sigma_F A \Lambda \qquad (A^T \Sigma_F A = I). \tag{15}$$

The discriminant space for "face" and "not face" classification is constructed by using learning samples. Then all the positions are checked whether there is a "face" or not. If the distance from the mean vector of "face" class is less than a given threshold, then that position is classified as "face". In this classification, the performance of face detection depends on the value of the threshold. The optimal value of the threshold is experimentally determined by using the following two probabilities. First probability is $P_F = 1 - \frac{n_F}{N_F}$ in which the samples of "face" class are miss-classified as "not face", where n_F is the number of the samples of "face" class which has a value less than the threshold and N_F is the total number of samples of "face" class. Second probability is $P_{NF} = \frac{n_{NF}}{N_{NF}}$ in which the samples of "not face" class are miss-classified as "face", where n_{NF} is the number of the sample of "not face" class which has a value less than the threshold and N_{NF} is the total number of samples of the "not face" class. When the threshold is increased from zero to infinity, two probabilities may change depending on the threshold. Since these two probabilities are error probabilities, we would like to minimize both of these probabilities. Thus we can select the optimal threshold in which the sum of the two probabilities is minimized.

By combining spectral features of Log-Polar image in section 2 with "face" and "not face" classification method in this section, scale invariant face detection can be realized. Although the spectral features of Log-Polar image are robust to the scale changes of a face, these features are heavily influenced by the position of the face in the image. However this property is good for face detection because the only position in which a face is centered gives similar features with "face" class. Namely the sensitivity of the face detection is equivalent to the sensitivity of the position changes.

The proposed face detection method consists of the following three steps.

(1) The center of the Log-Polar transformation is set to a certain position in the input image and a Log-Polar image is constructed from the input image by resampling.

(2) The spectral features are extracted from the Log-Polar image.

(3) The extracted features are projected into the discriminant space and classified as "face" of "not face".

By applying this process to all positions in the input image, the face detector can find "face" in the image. Figure 5 shows the flow of the proposed face detection method.

4. FACE CLASSIFICATION METHOD

For face classification, we can use the usual LDA. New features $\boldsymbol{y} = (y_1, \cdots, y_L)^T$ are also computed by linear combinations of the spectral features extracted from Log-Polar image $\boldsymbol{x} = (x_1, \cdots, x_M)^T$ as

$$\boldsymbol{y} = A^T \boldsymbol{x}, \tag{16}$$

where $A = [a_{ij}]$ is a coefficients matrix.

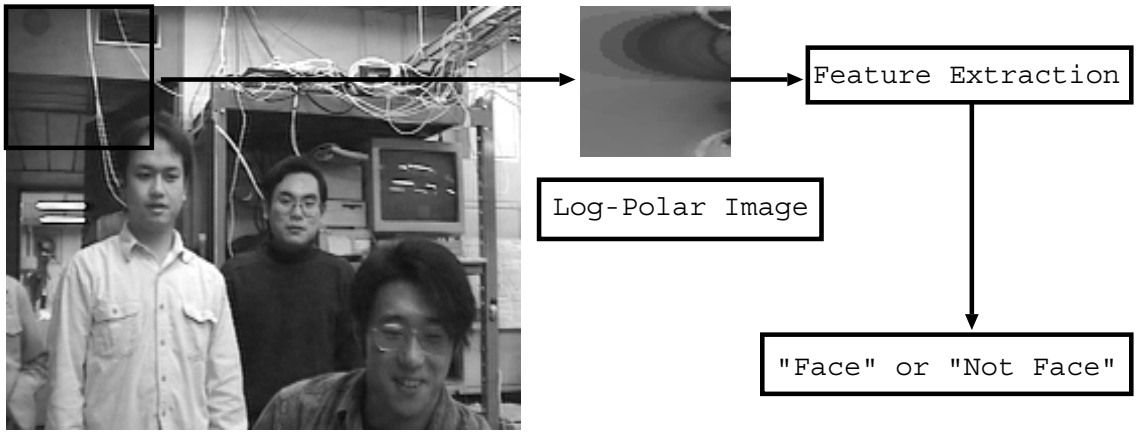


Figure 5. The flow of proposed face detection method.

Suppose that we have K classes $\{C_k\}_{k=1}^{K}$. Then the within-class and the between-class covariance matrices of the spectral features extracted from Log-Polar image are computed from the training samples as

$$\Sigma_W = \frac{1}{N}\sum_{k=1}^{K}\sum_{i=1}^{N_k}(\boldsymbol{x}_{ki}-\bar{\boldsymbol{x}}_k)(\boldsymbol{x}_{ki}-\bar{\boldsymbol{x}}_k)^T, \quad \Sigma_B = \frac{1}{K}\sum_{k=1}^{K}(\bar{\boldsymbol{x}}_k-\bar{\boldsymbol{x}}_T)(\bar{\boldsymbol{x}}_k-\bar{\boldsymbol{x}}_T)^T, \tag{17}$$

where $N=\sum_{k=1}^{K}N_k$, $\bar{\boldsymbol{x}}_k$, and $\bar{\boldsymbol{x}}_T$ denote the mean vector of C_k and the total mean vector, respectively. Similarly, the within-class and between-class covariance matrices of new features in the discriminant space are defined as

$$\hat{\Sigma}_W = A^T\Sigma_W A, \qquad \hat{\Sigma}_B = A^T\Sigma_B A. \tag{18}$$

The discriminant criterion

$$J = \mathrm{tr}(\hat{\Sigma}_W^{-1}\hat{\Sigma}_B) \tag{19}$$

is used to evaluate the performance of the discrimination of the new features $\boldsymbol{y}$. The optimal coefficient matrix A to maximize the discriminant criterion J is also given by solving the following eigen-value problem

$$\Sigma_B A = \Sigma_W A\Lambda \qquad (A^T\Sigma_W A = I), \tag{20}$$

where $\Lambda = diag(\lambda_1 \geq \lambda_2 \geq \cdots \lambda_L \geq 0)$ is a diagonal matrix of eigenvalues and I denotes the unit matrix. The j-th column of A is the eigenvector corresponding to the j-th largest eigenvalue.

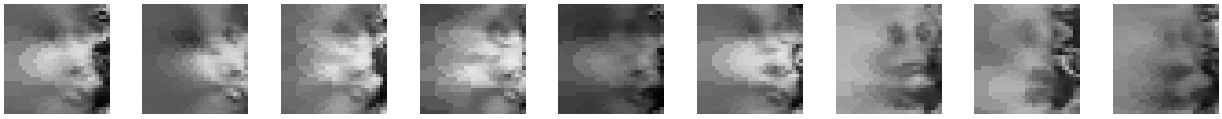
To identify the person from the given unknown image, we can use a simple classifier which compares the distances between the unknown input $\boldsymbol{y}$ and each class mean vectors $\bar{\boldsymbol{y}_k}$ in the discriminant space and classifies the unknown input image to the class C_k which gives the shortest distance.

5. EXPERIMENTS ON FACE DETECTION

In following experiments, the gray-scale image is used. The number of the "face" images are over 3000 and they are taken from over 70 persons. Face images from MIT face database[24] are also included in the learning samples. They include different sizes of the faces. The number of the "not face" images are over 1000 images. They include many kinds of images such as books, chair, hand, clothes, and so on. Some of the "face" and "not face" images and their Log-Polar image are shown in Figure 6.

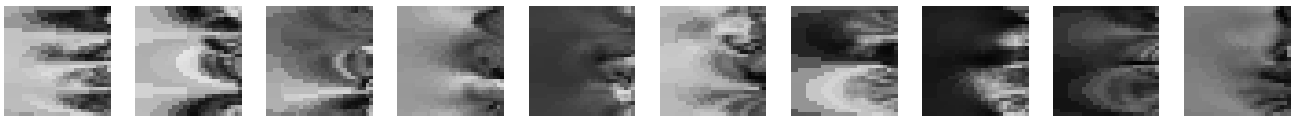
We have performed the experiment to evaluate the performance of the face detection method using autocorrelation features, Fourier power spectrum features, PARCOR features, and HLAC features. The 200 images are selected at random from the learning images of "face" class. The correct position of the "face" is measured in advance for all of the selected 200 images. These selected 200 images are used to evaluate the performance of the face detection. The

(a) The examples of "face". (59×59 pixels)



(b) Log-Polar image of (a). (30×29 pixels)

(c) The examples of "not face". (59×59 pixels)



(d) Log-Polar image of (c). (30×29 pixels)

Figure 6. Examples of "face" and "not face" samples

discriminant space is constructed by using both the remaining "face" images and "not face" images. The dimension of the discriminant space is set to 20 for all the proposed spectral features. The optimal dimension of PARCOR coefficients is determined by AIC. The images used in test are shown in Figure 7. The face image used in this experiment are taken with different places, dates, illuminations, scales of a face.

Figure 7. The examples of the images used for test.

Face detection method is applied to the selected images and the distance between the correct position and the center position of the region which gives the minimum distance is measured. If the measured distance is less than 5.0 (Euclidean distance), then it is considered that the detection is correct. The results are shown in Table 1. The '"face" images' in Table 1 shows the results obtained by using the 200 "face" images. For comparison, the results obtained by using HLAC features are also shown in Table 1. For the "face" images captured under the many different places, dates, illuminations, and scales, the proposed methods give very high accuracy of face detection. It is noticed that accuracy is much higher than HLAC. This means that the proposed methods are improved by dropping the 2D rotation invariance.

To investigate the precision of the face detection, the correct rate is evaluated by changing the permission distance between the detected position and the correct position. Figure 8 shows the correct detection rates when the permission distance is changed from 0.0 to 5.0. From this graph, it is noticed that Fourier power spectrum features

	"face" images
Autocorrelation	190/200 (0.95)
Fourier power spectrum	195/200 (0.975)
PARCOR	168/200 (0.84)
HLAC	84/200 (0.42)

Table 1. The performance of face detection.

give highest detection rates and PARCOR coefficients are lower than the others. Both autocorrelation and Fourier power spectrum can detect face at the correct position with more than 80% of the correct detection rate.

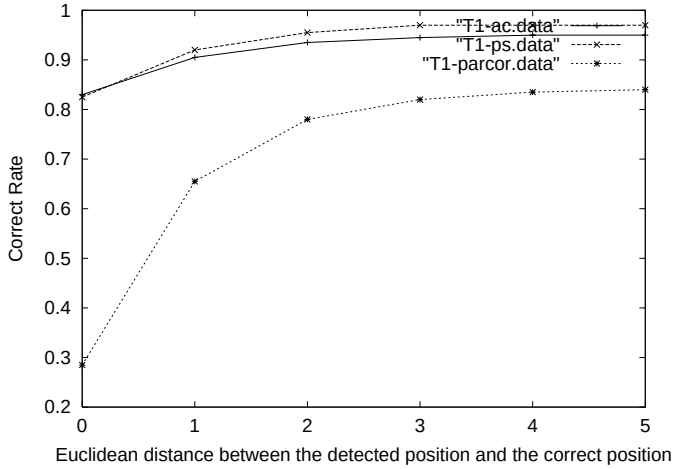


Figure 8. The correct detection rates when the permission distance is changed.

6. EXPERIMENTS ON FACE CLASSIFICATION

We have performed the experiments to evaluate the robustness to the scale changes of the face. The face images with 7 different scales are gathered under the 2 different background. The scale of the face are changed by changing zoom parameter of a camera. The total number of images is 2800 images (20 images $\times$ 7 scales $\times$ 2 background $\times$ 5 people $\times$ 2 times). The examples of the face images and their Log-Polar images are shown in Figure 9. It is noticed that Log-Polar image is robust to the scale changes of face and background changes. The robustness to the change of background come from the space variant sampling of Log-Polar image whose resolution is sparse in the periphery region such as background.

The 400 images (20 images $\times$ 2 scales $\times$ 2 background $\times$ 5 people) are used for learning and the 1400 images (20 images $\times$ 7 scales $\times$ 2 background $\times$ 5 people) are used to evaluate the performance of the proposed face classification method. The proposed spectral features are compressed by Principal Component Analysis (PCA) because the dimension of the primitive features (the spectral features of Log-Polar image) is very high. Then LDA is applied to the compressed features. The dimension of the discriminant space is set to 4 for fair comparison. Then the test samples are classified by a simple classifier which compares the distances between the input $\boldsymbol{y}$ and each class means $\bar{\boldsymbol{y}_{\boldsymbol{k}}}$ and classifies it to the class C_k which gives the shortest distance. The results are shown in Table 2. The resolution (the number of sampling) of the rotation axis on the Log-Polar image is changed in the experiment because they probably influence to the recognition rates. The top row in Table 2 represents the size of Log-Polar image (rotations : θ $\times$ scale : $\log(\rho)$) . For comparison, the recognition rates with HLAC features are also shown in Table 2.

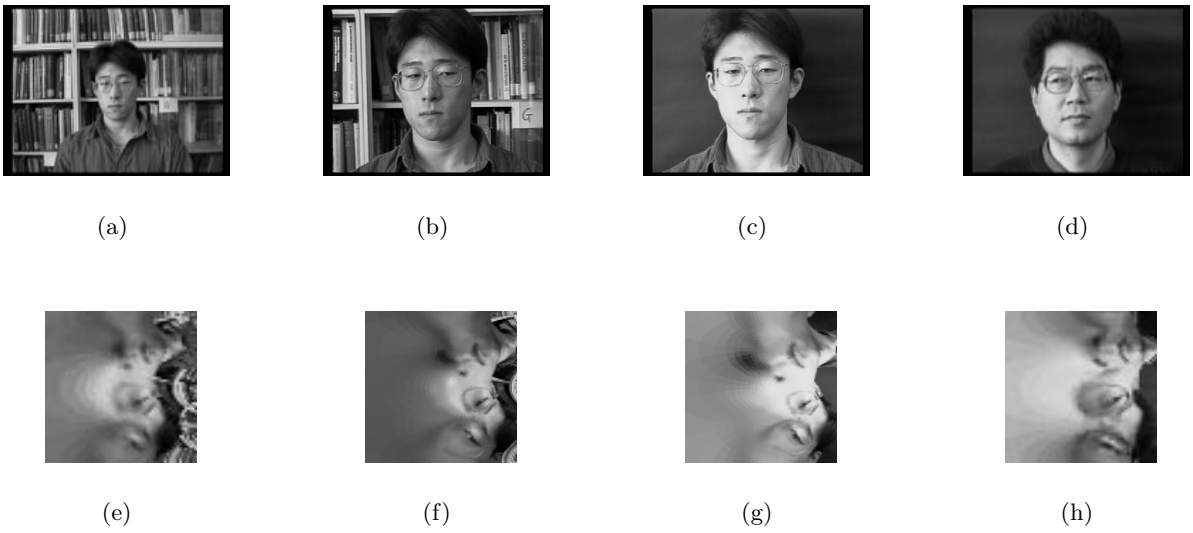


Figure 9. The examples of face images and their Log-Polar images. (a) The example contained a small face. (b) The example contained a large face. (c) The example contained a large face with different background. (d) The example contained an another face. (e) The Log-Polar image of (a). (f) The Log-Polar image of (b). (g) The Log-Polar image of (c). (h) The Log-Polar image of (d).

It is notices that the recognition rates of the proposed methods are higher than the previous method based on HLAC features. This means that the proposed methods are improved by dropping the 2D rotation invariance. The changes of the resolution axis is less influenced to the recognition rate.

Table 2. Face Classification Experiment 1 (Input image : 80×60)

	30×30	60×30	90×30	120×30
Autocorrelation	97.64%	97.79%	97.29%	96.64%
Fourier power spectrum	98.93%	99.50%	99.14%	98.29%
PARCOR	91.79%	93.93%	89.07%	95.93%
HLAC	82.21%	77.36%	82.79%	85.93%

ACKNOWLEDGMENTS

We would like to thank Dr.Nakashima, the head of Information Science Division, Electrotechnical Laboratory, for giving the opportunity of this work and also thank to the members of the Information Dynamics Lab. for helpful discussions. The first author was supported in part by Japan Society for the Promotion of Science (JSPS).

REFERENCES

1. M.Doi, K.Sato, and K.Chihara, "A robust face identification against lighting fluctuation for lock control," in *Proc. of the third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 42–47, 1998.
2. O.Hasegawa, K.Itou, T.Kurita, S.Hayamizu, K.Tanaka, K.Yamamoto, and N.Otsu, "Active gent oriented mutimodal interface system," in *Proc. of International Joint Conference on Artificial Intelligence*, pp. 82–87, 1995.
3. I.Hara, T.Matui, H.Asoh, T.Kurita, M.Tanaka, K.Hotta, T.Mishima, and A.Zelinsky, "Communicative functions to support human robot cooperation," in *Proc. of IEEE/RSJ International Conference on Intelligent Robots and Systems (to appear)*, 1999.

4. S.Satoh and T.Kanade, "Name-it: Association of face and name in video," tech. rep., CMU-CS-96-205, 1996.
5. H.D.Wactlar, T.Kanade, M.A.Smith, and S.M.Stevens, "Intelligent access to digital video: The informedia project," *IEEE Computer* **29**(5), pp. 46–52, 1996.
6. Y.Ariki, Y.Sugiyama, and N.Ishikawa, "Face indexing on video data – extraction, recognition, tracking and modeling –," in *Proc. of third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 62–69, 1998.
7. S.A.Rizvi, P.J.Phillips, and H.Moon, "The feret verification testing protocol for face recognition algorithm," in *Proc. of third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 48–53, 1998.
8. A.Samal and P.A.Iyenger, "Automatic recognition and analysis of human faces and facial expression: A survey," *Pattern Recognition* **25**(1), pp. 65–77, 1992.
9. R.Chellappa, C.L.Wilson, and S.Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE* **83**(5), pp. 705–740, 1995.
10. M.Turk and A.Pentland, "Face recognition using eigenfaces," in *Proc. of the third IEEE Conference on Computer Vision and Pattern Recognition*, pp. 586–591, 1991.
11. A.Pentland, B.Moghaddam, and T.Starner, "View-based and modular eigenspaces for face recognition," in *Proc. of the third IEEE Conference on Computer Vision and Pattern Recognition*, pp. 84–91, 1994.
12. K.Sung and T.Poggio, "Example-based learning for view-based human face detection," tech. rep., A.I. Memo 1521, CBCL Paper 112, 1994.
13. H.A.Rowley, S.Baluja, and T.Kanade, "Human face detection in visual scenes," tech. rep., CMU-CS-95-158R, 1995.
14. H.A.Rowley, S.Baluja, and T.Kanade, "Rotation invariant neural network-based face detection," tech. rep., CMU-CS-97-201, 1997.
15. B.Moghaddam and A.Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **19**(7), 1997.
16. G.Xu and T.Sugimoto, "Rits eye: A software-based system for realtime face detection and tracking using pan-tilt-zoom controllable camera," in *Proc. of 14th International Conference on Pattern Recognition*, pp. 1194–1197, 1998.
17. K.C.Yow and R.Cipolla, "Feature-based human face detection," tech. rep., Cambridge University Engineering Department CUED/F-INFENG/TR249, 1996.
18. T.Kurita, K.Hotta, and T.Mishima, "Scale and rotation ivariant recognition method using higher-order local autocorrelation features of log-polar image," in *Proc. of the third Asian Conference on Computer Vision*, vol. II, pp. 89–96, 1998.
19. K.Hotta, T.Kurita, and T.Mishima, "Scale invariant face detection method using higher-order local autocorrelation features extracted from log-polar image," in *Proc. of the third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 70–75, 1998.
20. N.Otsu and T.Kurita, "A new scheme for practical flexible and intelligent vision systems," in *Proc. of IAPR Workshop on Computer Vision*, pp. 431–435, 1988.
21. G.Sandini and V.Tagliasco, "An anthropomorphic retina-like structure for scene analysis," *Computer Graphics and Image Processing* **14**, pp. 365–372, 1980.
22. L.Massone, G.Sandini, and V.Tagliasco, "Form-invariant: Topological mapping strategy for 2d shape recognition," *Computer Vision, Graphics and Image Processing* **30**, pp. 169–188, 1985.
23. J.Durbin, "The fitting of time-series models," *Rev. Inst. de Stat.* **28**, pp. 233–244, 1960.
24. *MIT face database*. ftp://whitechapel.media.mit.edu/pub/images/.