PERGAMON

Pattern Recognition 32 (1999) 1707–1718

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

Automatic human face detection and recognition under non-uniform illumination

Toshiaki Kondo*,[1], Hong Yan

Department of Electrical Engineering, University of Sydney, NSW 2006, Australia

Received 17 February 1997; received in revised form 1 December 1998; accepted 1 December 1998

Abstract

A system for automatic human face detection and recognition is presented. The procedure consists of five steps: (1) the Haar wavelet transform, (2) facial edge detection, (3) symmetry axis detection, (4) face detection and (5) face recognition. Step 1 decomposes an input image, reducing image redundancy. Step 2 excludes non-facial areas using edge information, whereas Step 3 narrows down face areas further using gradient orientation. Step 4 restricts face-like areas by template matching. Finally, Step 5 determines the best face location in the face-like areas and identifies the face based on principal component analysis (PCA). The system shows a remarkably robust performance under non-uniform lighting conditions.

Keywords: Wavelet transform; Edge detection; Symmetry detection; Face detection; Face recognition; Template matching; Correlation; Principal component analysis; K–L expansion

1. Introduction

Human face detection and recognition have drawn considerable interest and attention from many researchers for decades [1, 2]. A system that performs face detection or recognition will find many applications such as surveillance cameras and security control systems. However, it is still a challenging research topic since the human face may change its appearance because of facial expressions, beards, mustaches, hair styles, make-up, glasses, aging, surgery, etc. In addition to these internal variations, we have to consider the external distortions such as the scale, lighting, position, tilt and orientation of the face. It should be also noted that a complex background in an image makes it far more difficult to locate faces and facial features. Therefore, it is always necessary for an automatic face recognition system to set some constraints on input images. Our goal is to relax these constraints. In this paper, we focus on images captured with cluttered backgrounds under non-uniform illumination as shown in Fig. 1. These images are a part of the face database produced by the MIT Media Lab. Each image has a resolution of 256 by 240 pixels with 256 gray levels.

2. Face detection

The overall flow diagram of our system is given in Fig. 2. We describe Steps 1–4 in this section. These

*Corresponding author. Digital Imaging Technology Development, Dept. 24, Digital Imaging Products Development Center, Canon Inc., 30-2, Shimomaruko 3-Chome, Ohta-Ku, Tokyo, 146-8501 Japan. Tel.: + 81-337-579-885; fax: + 81-337-578-841.

E-mail address: tkondo@cam.canon.co.jp (T. Kondo)

0031-3203/99/$20.00
PII: S0031-3203(98)00176-9

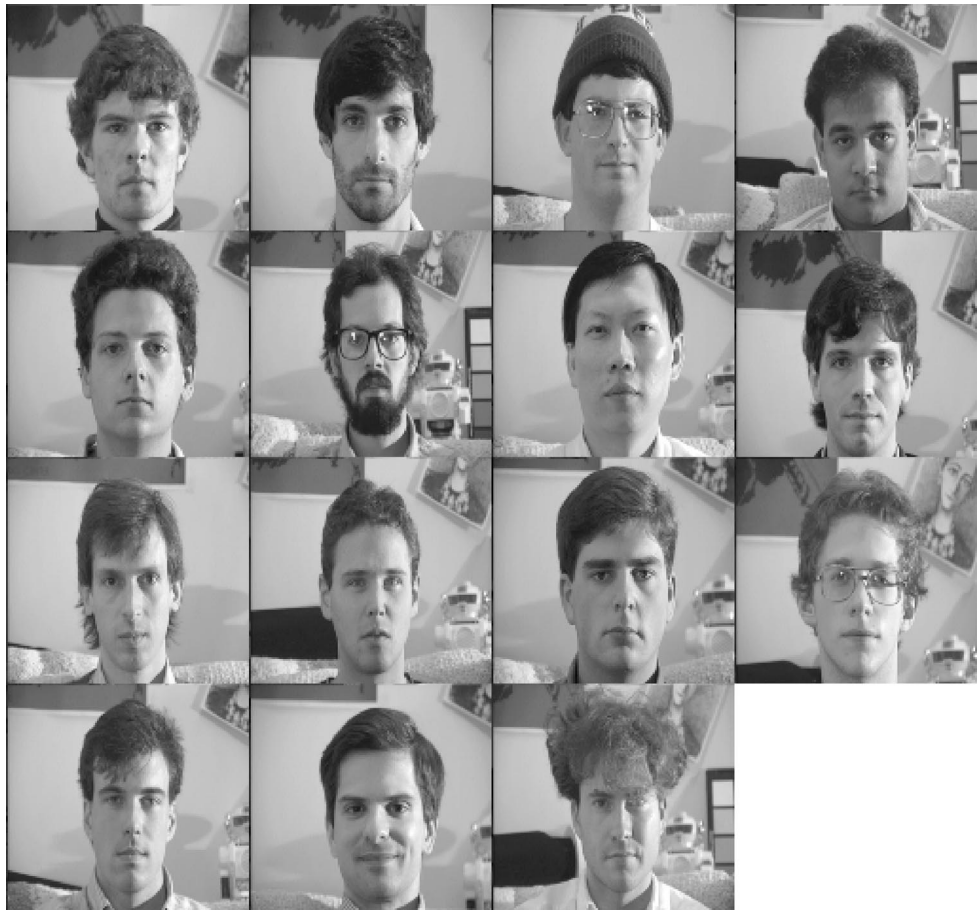

Fig. 1. Face images captured under non-uniform illumination.

procedures aim to realize a system which allows examinees to stand rather freely in front of a camera with a cluttered background.

2.1. Haar wavelet transform

A multifrequency channel analysis is employed for an efficient search for a face. One advantage is that redundancy in an image is reduced. It is reported that there is enough information in images of size 32×32 pixels with 16 gray levels for human eyes to detect and identify faces [3]. Even mosaic images at a lower resolution of 12×12 were successfully used for face location and identification in Refs. [4,5]. Another advantage is that a multifrequency channel analysis enables us to use a different technique suitable for each channel, which makes the subsequent procedure more efficient. In addition, the wavelet used in our system is the simplest of all finite impulse response (FIR) filters and highly suited to hardware implementation.

In Step 1, the wavelet transform decomposes an input image $f(x,y)$ with a resolution of 256×240 pixels with 256 gray levels into four subimages of size 128×120 pixels. The Haar function is used as a wavelet since it is the simplest linear phase filter of all wavelets [6]. In the Haar wavelet, low-pass filtering is conducted by averaging two adjacent pixel values, while high-pass filtering is performed by subtracting two adjacent pixel values. The four subimages are a low-pass filtered image (f_{LL}), a horizontally low-pass filtered and vertically high-pass filtered image (f_{LH}), a horizontally high-pass filtered and vertically low-pass filtered image (f_{HL}) and a diagonally high-pass filtered image (f_{HH}). As shown in Fig. 2, subimage f_{LL} is used for face detection (Step 4), while both f_{LH} and f_{HL} are used in facial edge detection (Step 2) and symmetry axis detection (Step 3). An input image and all its subimages except for f_{HH} are shown in Fig. 3 (No. 0, 1, 2 and 3). f_{HH} is not currently utilized as it does not contain much image information.

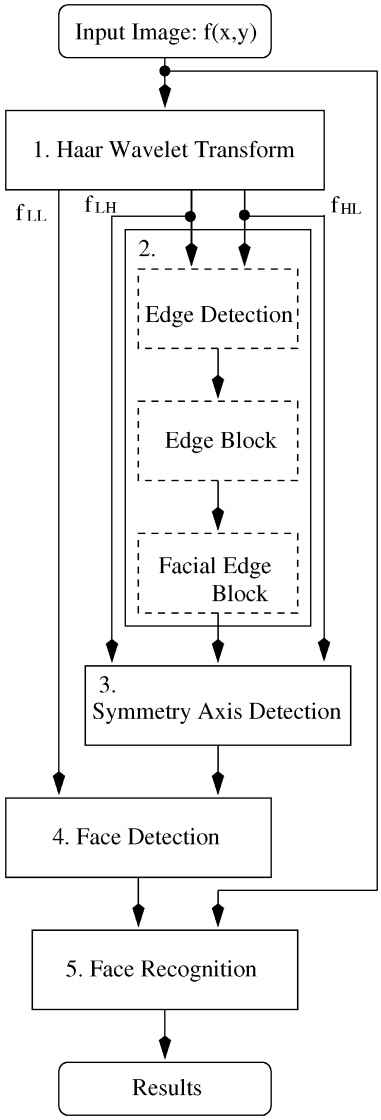

Fig. 2. Overall flow diagram.

2.2. Facial edge detection

Step 2 (facial edge detection) can be divided into the three parts as shown in the dashed boxes in Fig. 2. Firstly, oriented edges are effectively detected from subimages f_{LH} and f_{HL} by thresholding. The horizontal edges surrounding the eyes and mouth are extracted from f_{LH}, whereas the vertical edges of the nose are detected in f_{HL}. Then the binary edge image is obtained by the logical OR operation between the two subimages f_{LH} and f_{HL} after they are converted to binary images at a prespecified threshold ($\geqslant |\pm 4|$). The threshold is set very low to avoid rejecting face areas mistakenly.

Next, the edge image is partitioned into small blocks of size 8×8 pixels (see No. 4 in Fig. 3) and the number of edges in each block is determined. A block containing a sufficient number of edges ($\geqslant 8/64$) is selected as an edge block (shown as a black block in No. 5 of Fig. 3), while a block with few edges is ignored in the following process as a plain background (shown as a white block in No. 5 of Fig. 3).

Finally, we separate facial edge blocks from the edge blocks by examining connections with their neighboring blocks. We regard a block as valid which has adjoining edge blocks on both sides and in its downward direction. This T-shaped neighbor arrangement indicates the area encompassing the eyes and nose in a face as illustrated in Fig. 4. If a block under test has five or six edge blocks in its T-shaped neighbors, we select it as a facial edge block (see No. 6 in Fig. 3). The block under test is not counted because there may not be a sufficient number of edges between two eyes. Seven passable patterns of the facial edge block are depicted in Fig. 5. One non-edge block is allowed to be included in the T-shaped arrangement so that poor lighting is tolerated. All blocks in the image are tested in the same manner.

Although it is simple, Step 2 is fairly effective for reducing target areas for face detection. In our experiment using 45 images from 15 persons taken under three different lighting conditions, the target areas were successfully reduced to less than 20% of the entire area of the edge image on average.

2.3. Symmetry axis detection

We restrict face locations further by introducing another constraint, bilateral symmetry of the human face, in Step 3. Fig. 6 illustrates the gradient vectors. Assume the image in the right half is brighter than that in the left half. Since the length of each vector changes according to the illumination on it, the gradient vectors in the right half are longer than those in the left. In this way, all image gradient or intensity-based methods are influenced by changes in lighting conditions.

Thus, instead of image gradient or intensity, we use gradient orientation [6] to detect the symmetry axis. The gradient orientation is defined as

$$\theta(x,y) = \arctan\left(\frac{\partial f(x,y)}{\partial y} \Big/ \frac{\partial f(x,y)}{\partial x}\right). \tag{1}$$

The above image gradients $\partial f(x,y)/\partial$y and $\partial f(x,y)/\partial$x can be approximately described by subimages f_{LH} and f_{HL}, respectively. Therefore, Eq. (1) can be rewritten as

$$\theta(x,y) = \arctan\left(\frac{f_{LH}(x,y)}{f_{HL}(x,y)}\right), \tag{2}$$

where f_{HL} is not equal to zero. When f_{HL} is zero, we assign $\pi/2$ to θ if $f_{LH} \geqslant 0$, and assign $-\pi/2$ to θ if $f_{LH} < 0$. Consequently, the range of θ is $(-\pi, +\pi]$.

Input image $f(x,y)$ 0		*Subimage* f_{LL} 1	*Subimage* f_{LH} 2
		Subimage f_{HL} 3	*Partitioned edge image* 4
Edge blocks 5	*Facial edge blocks* 6	*Symmetry index map* 7	*Facial edge blocks* AND *Symmetry index map* 8
High symmetry index area 9	*Face-like area* 10	*Best matching point* 11	*Synthesized image with a face template* 12

Labeling of face detection procedures shown below

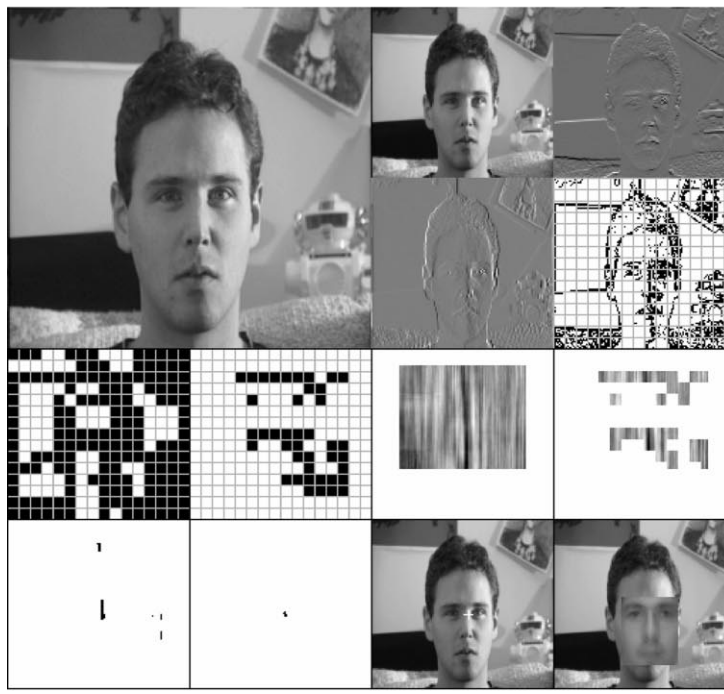

Fig. 3. Procedures of face detection (Step 1 to 4).

Eq. (2) can be computed efficiently provided a lookup table is used. Fig. 7 illustrates the lookup table. With the lookup table, we only have to read subimages, f_{LH} and f_{HL}, and convert their image values to the gradient orientation θ. The range of the subimages, f_{LH} and f_{HL}, is $[-255, +255]$ since they are obtained by subtraction between two adjacent pixel values of the original image vertically and horizontally. Hence, Step 3 has two favorable structural aspects: (1) the gradient information is directly fed from Step 1, and (2) the calculation of trigonometry can be replaced with a lookup table.

As illustrated in Fig. 8, a block assumed to have roughly the same size, $h \times v$, as that of the expected input face is scanned across the facial edge area for computing

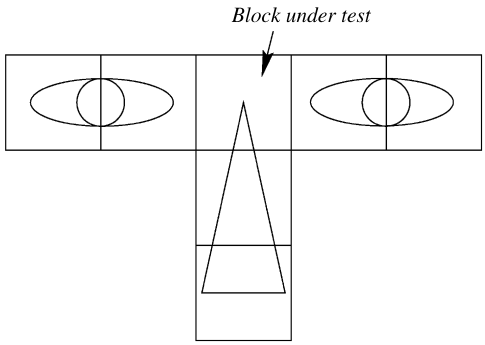

Fig. 4. T-shaped distributed blocks.

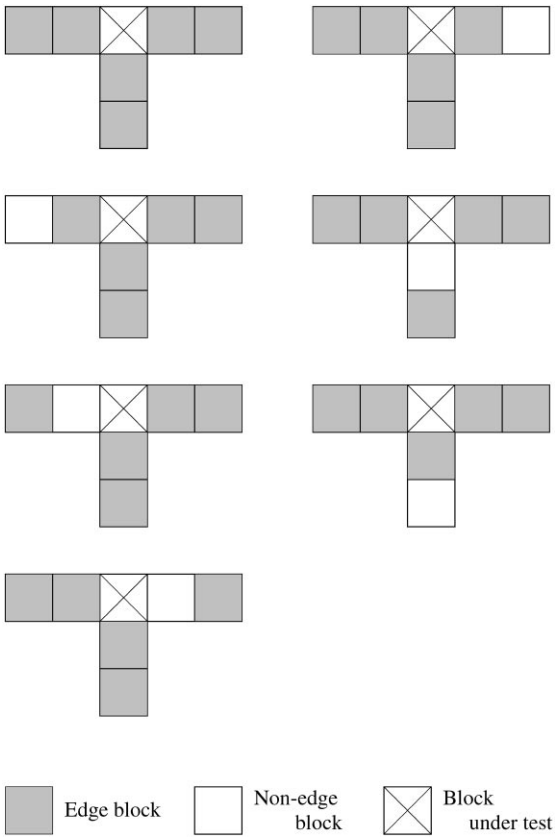

Fig. 5. Seven passable patterns of the facial block.

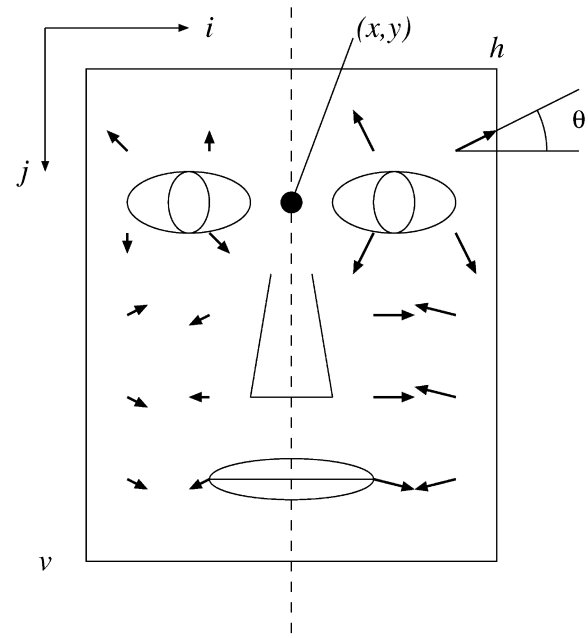

Fig. 6. Gradient vectors.

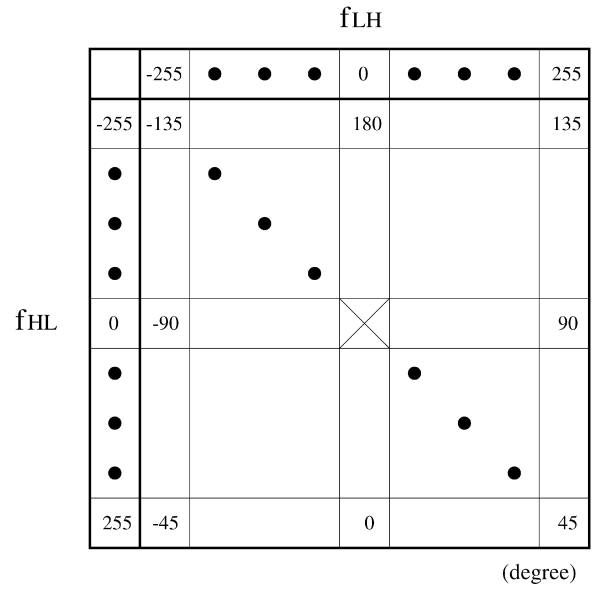

Fig. 7. A lookup table.

the symmetry index s_{xy}. The distribution of the symmetry index s_{xy} forms a symmetry index map. The definition of s_{xy} is given by

$$s_{xy} = \sum_j \sum_i \{|\cos\theta(i,j) + \cos\theta(2x-i,j)|^2 + |\sin\theta(i,j) - \sin\theta(2x-i,j)|^2\}, \quad (3)$$

where $(y - v/4) \leqslant j \leqslant (y + 3v/4)$ and $(x - h/2) \leqslant i \leqslant x$. Note that the position of (x, y) does not coincide with the central point of the scanning block. This is because we have defined the midpoint between two eyes as a face location. When the image in the block is bilaterally symmetric, each term of Eq. (3) takes a small value because of cancellation. Conversely, when it is asymmetric,

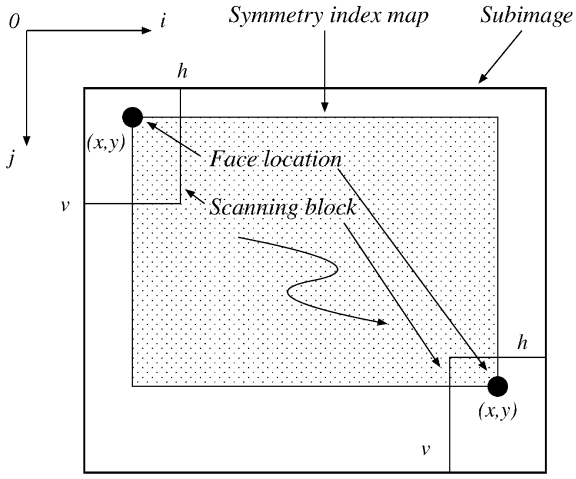

Fig. 8. Process of computing symmetry indexes.

Fig. 9. Gradient orientation and symmetric index map.

Fig. 10. A face template (40×48 pixels).

the index gives a larger value. Usage of a lookup table can be extended to $\cos\theta$ and $\sin\theta$ in Eq. (3).

The distributions of $\cos\theta$, $\sin\theta$ and s_{xy} are visualized by being scaled in the range 0–255 in Fig. 9. The upper left gives a subimage f_{LL}. The upper right shows $\cos\theta$, whereas the lower right shows $\sin\theta$. Since the gradient orientation θ on the face is bilaterally symmetric, $\cos\theta$ image shows an inverted pattern with respect to the vertical axis centered on the face, while $\sin\theta$ image shows a symmetric pattern. The symmetry index s_{xy} is shown in the lower left of Fig. 9. The dark vertical stripe shows the symmetry axis centered on the face. Another symmetry index map is given in Fig. 3 (No. 7). Although s_{xy} is computed in the entire area for viewing purposes, in practice, it is computed only in the facial edge area as shown in Fig. 3 (No. 8). The following process is to be carried out over the high symmetry index area ($\geqslant 90$) shown in Fig. 3 (No. 9).

We have found that gradient orientation is illumination invariant to a great extent. Even if the face is shadowed because of non-uniform illumination, the gradient orientation is preserved except for the border of the shadow. In a comparison test, the gradient orientation-based symmetry detection technique has outperformed a normalized cross-correlation-based method. This result is worth noting because normalized cross correlation is frequently used to overcome changes in the amplitude of images caused by various lighting conditions. There are two clear advantages of our approach over a correlation-based method. Firstly, our method can exclude a plain background unless it is completely flat. On the contrary, a correlation-based method regards a plain background as a highly symmetric area because it only deals with the image intensity. Secondly, our scheme has a hardware-oriented structure as mentioned before, whereas a correlation based method is computationally rather expensive.

2.4. Template matching

In Step 4, normalized cross correlation is computed between the input image $f(x, y)$ and a face template $w(x, y)$ shown in Fig. 10. Assuming $f(x, y)$ and $w(x, y)$ exist in the region $0 \leqslant x \leqslant X-1$, $0 \leqslant y \leqslant Y-1$, we define the cross correlation as

$$\gamma_{xy} = \frac{\sum_{j=y}^{Y-1}\sum_{i=x}^{X-1}[f(i,j)-\bar{f}][w(i-x,j-y)-\bar{w}]}{\sqrt{\sum_{j=y}^{Y-1}\sum_{i=x}^{X-1}[f(i,j)-\bar{f}]^2\sum_{j=y}^{Y-1}\sum_{i=x}^{X-1}[w(i-x,j-y)-\bar{w}]^2}}, \quad (4)$$

where $\bar{f}$ is the average of $f(x,y)$ over the overlapped region with $w(x,y)$, and $\bar{w}$ is the average of $w(x,y)$. The correlation coefficient γ_{xy} is scaled in the range -1–1. This

Table 1
Face detection results

Illumination angle	Detection rates	
Head-on	15/15	100.00%
45°	14/15	93.33%
90°	15/15	100.00%

computation is carried out separately in the left and right half of the template to reduce the effect of the non-uniform illumination. The face template is obtained from all 15 subimages, f_{LL}, produced in Step 1 from the original images captured under uniform illumination. Firstly, the face areas are cropped manually from the subimages of size 128×120 pixels. The size of the face template is determined as 40×48 pixels so that the images cropped do not contain the backgrounds. Secondly, all the cropped images are simply averaged to generate the face template. Again, by eliminating the locations with low correlation coefficients ($\gamma_{xy} \leqslant 0.18$), the face-like area is extracted from the high symmetry index area (see No. 10 in Fig. 3).

2.5. Detection results

We have tested our face detection scheme under three different lighting conditions using 45 images from 15 persons as mentioned before. The best matching point marked ' + ' in Fig. 3 (No. 11) shows the face location detected by our method, and the face template is superposed on f_{LL} in No. 12.

The face detection rates are given in Table 1. We have assumed that the face detection is successful when the best matching point is located within 1 pixel from the midpoint between two eyes. Our algorithm shows a robust performance under non-uniform illumination. The only detection error occurred when the only oriental face in the MIT database was used. The face template needs to be sufficiently similar to each face in the database. Hence, we should have several different face templates when the number of face images increases. To increase tolerance to image deformation, it would be effective to adopt a deformable template matching strategy [8]. Another approach to improve detection accuracy is to adopt neural network based techniques [9,10].

3. Face recognition

Our face recognition method is based on principal component analysis (PCA) since the principal component of the K–L expansion is the optimal filter that maximizes the signal-to-noise ratio of the output correlation when the input signal is distorted [11].

3.1. Principal component analysis

At the recognition stage, we use the original images instead of the subimages to achieve higher accuracy. The face areas of size 80×96 pixels are cropped manually from the original images of size 256×240 pixels in the same manner as the face template of size 40×48 pixels was made from the subimages of size 128×120 pixels.

3.1.1. Computation of eigenfaces

Let us summarize the computation of eigenfaces [12,13]. Assume a set of face images to be $\{\Gamma_1, \Gamma_2, \ldots, \Gamma_M\}$. The average face of the set is defined by $\Psi = 1/M \sum_{k=1}^{M} \Gamma_k$, which is identical to the face template mentioned before. Each face differs from the average by the vector $\Phi_k = \Gamma_k - \Psi$. We define the matrix A as $[\Phi_1 \Phi_2 \ldots \Phi_M]$. The matrix A has M columns and N rows. The number N is equal to the number of pixels in an image. Consider the eigenvectors u_i of AA^{T} and v_i of $A^{\mathrm{T}}A$ such that

$$(AA^{\mathrm{T}})u_i = \mu_i u_i, \tag{5}$$

$$(A^{\mathrm{T}}A)v_i = \mu_i v_i, \tag{6}$$

where μ_i are the eigenvalues. It is always possible to diagonalize the matrix AA^{T} and $A^{\mathrm{T}}A$ because they are real and symmetric. However, it is laborious to solve Eq. (5) directly since N, the size of the matrix, is huge. A fast technique is used here to solve the problem. Multiplying both sides of Eq. (6) by A, we obtain

$$(AA^{\mathrm{T}})(Av_i) = \mu_i(Av_i). \tag{7}$$

Comparing Eq. (5) with Eq. (7), we have

$$u_i = Av_i. \tag{8}$$

Thus, we only have to solve Eq. (6) to obtain u_i. The eigenvectors u_i are referred to as eigenfaces or eigenpictures as they appear face like. Fig. 11 shows the first eigenface of size 80×96 pixels.

It is known that the original image Φ_k can be represented as a weighted sum of eigenvectors u_i as follows:

$$\Phi_k = \sum_{i=1}^{M} u_i(u_i^{\mathrm{T}}\Phi_k). \tag{9}$$

The scalars $u_i^{\mathrm{T}}\Phi_k$ determine the weights of eigenvectors u_i. The eigenfaces and weights are computed when the database is established and they are made available during the recognition process. The weights of the input image Φ_n are also obtained by the inner product $u_i^{\mathrm{T}}\Phi_n$. We discriminate individuals using the distribution of the weights $u_i^{\mathrm{T}}\Phi_n$.

3.1.2. Recognition process

PCA-based recognition is fairly simple and fast in comparison with featured-based methods. The input face

Fig. 11. First eigenface u_1 (80×96 pixels).

image Γ is projected into the eigenface space after the average face Ψ is subtracted. The projection is simply an inner product between the input face and each eigenface:

$$\omega_i = u_i^{\mathrm{T}}(\Gamma - \Psi) \tag{10}$$

for $i = 1, 2, \ldots, M$. The weights ω_i form a vector $\Omega^{\mathrm{T}} = [\omega_1, \omega_2, \ldots, \omega_M]$ that shows the contribution of each eigenface. The input face is identified by finding the minimum Euclidean distance ε defined as

$$\varepsilon_i = ||(\Omega - \Omega_i)||^2, \tag{11}$$

where Ω_i is the precalculated vector of the ith face in the database.

3.2. Normalized PCA

In this section, we discuss the relation between PCA-based face recognition and illumination. The MIT research group claimed that the PCA-based algorithm was robust to variations in lighting conditions [13]. In fact, the face recognition ratio was not strongly affected by non-uniform illumination in their report. However, there is a crucial problem in their experiment. Although they multiplied the input face image by a two-dimensional Gaussian window centered on the face, the background is not completely excluded. Furthermore, each face has its peculiar background which is fixed under the different lighting conditions as shown in Fig. 12. Since each background is a fixed pattern attached to each face, it helps to discriminate individuals. This problem is also mentioned in Ref. [14].

We conducted a similar simulation to confirm this positive effect of the background, especially under different illumination conditions. To make it clear, we did not employ any filter for suppressing the background. All 80×96 face areas are clipped manually from the original images of size 256×128 pixels. Table 2 shows the recognition results of PCA using the original images, whereas Table 3 shows the results when only the face areas are used.

The stark contrast between Tables 2 and 3 shows that the background positively contributes to the recognition process. As shown in Table 3, we should recognize that PCA-based recognition is rather sensitive to changes in lighting conditions and preprocessing should be considered.

In our method, we conduct intensity normalization before PCA. The normalization of an image intensity is done by subtracting the mean intensity and dividing by the standard deviation in the region of interest. We have normalized all face areas cropped from the original images in the database before computing eigenfaces. The normalization is carried out over the entire area of each face since the face in the database should have been illuminated uniformly. In the recognition process, the input face image is also normalized before being projected into the eigenspace. This normalization is performed separately in the left and right half of the face to reduce the effect of the non-uniform illumination.

3.3. Recognition results

The result of face recognition by normalized PCA is shown in Table 4. When an input face image has the minimum Euclidean distance to his/her own face in a database, we say the input image is recognized. The recognition rates have been successfully improved under oblique illumination (45°). On the other hand, we could not achieve higher recognition rates under side illumination (90°). However, there is a significant improvement in reliability in spite of the same recognition rates. The reliability index is defined as

$$R = (\varepsilon_{2nd} - \varepsilon_{min})/\varepsilon_{min}, \tag{12}$$

where ε_{min} and ε_{2nd} are the Euclidean distances of the best and the second best matching. The larger the index R, the more reliable the recognition result. The reliability indexes are given in Table 5. The reliability leaps by 3.5 times after illumination normalization in both lighting conditions.

3.4. System integration

In this section, we discuss how face detection and recognition are integrated. System integration is an essential part because we carry out face detection at a lower resolution, while we perform face recognition at the original resolution. Table 6 shows how the recognition rates vary when the face is shifted vertically from -5 to 5 pixels at the original resolution. It is apparent that PCA is sensitive to the displacement. The best recognition rates shown in Table 4 are maintained only between -1 and 1 pixel with respect to the best matching point under

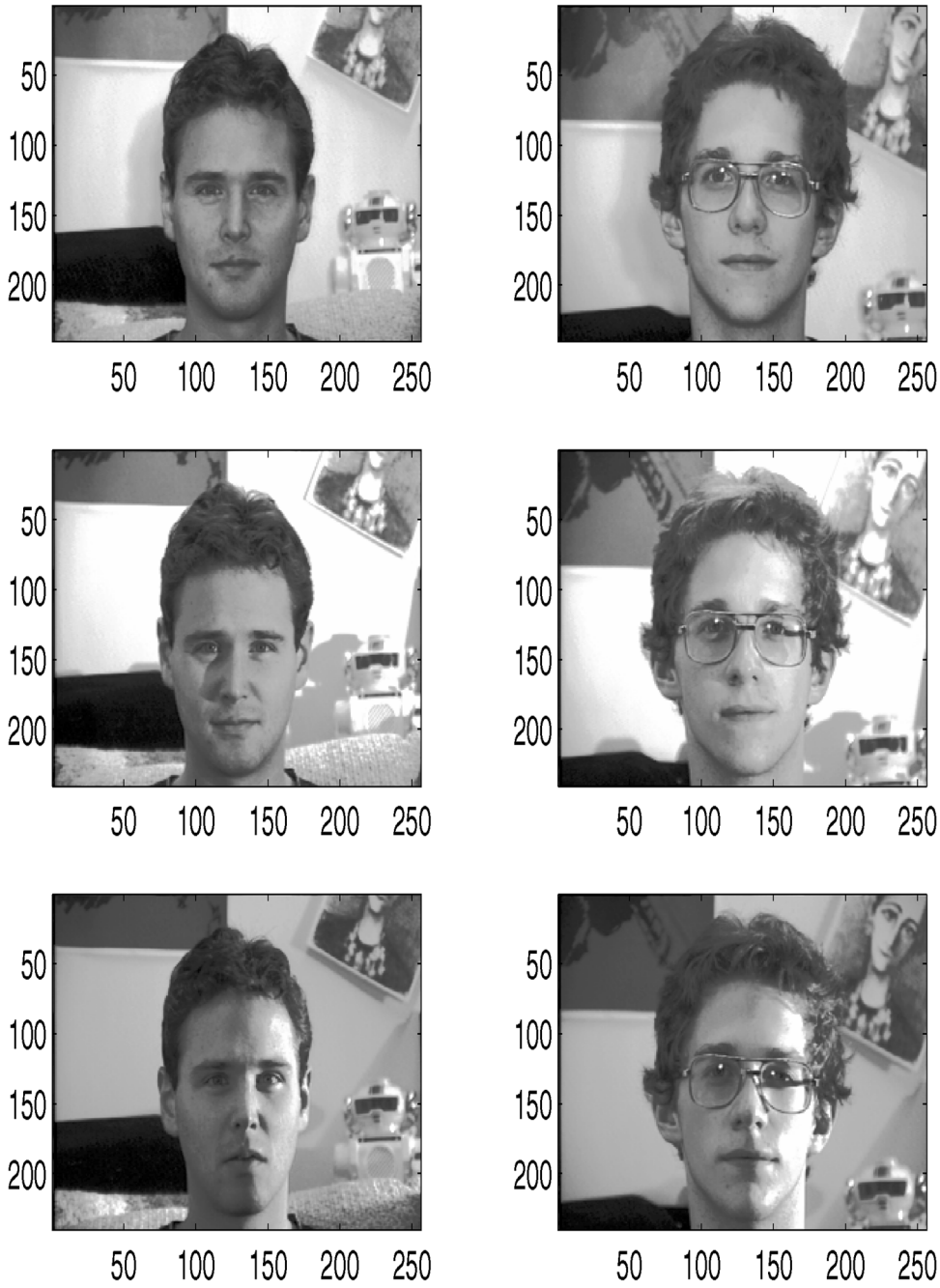

Fig. 12. Fixed backgrounds under three lighting conditions (top: head on lighting, middle: 45° oblique lighting, bottom: 90° side lighting).

three lighting conditions. The 2-pixel displacement in Table 6 corresponds to a 1-pixel displacement in the face detection stage. Consequently, it should be noted that the requirement for face detection accuracy is quite demanding. Hence, it is necessary to search for the precise face location at the recognition stage by applying PCA not only to the best matching point but also to its neighboring pixels.

Fig. 13 depicts the flow of the integration. First of all, the best matching point, determined by template

matching, is projected to the original image of size 256×240 as illustrated in Fig. 14. Normalized PCA is applied to the projected pixel and its eight adjacent pixels, namely nine pixels altogether. The recognition result is determined by R_{max}. If $R_{max} \geqslant 1.0$, the result is regarded as reliable and the recognition process is terminated. If $R_{max} < 1.0$, we apply normalized PCA to the other face candidates which satisfy $\gamma_{xy} > (0.85 \times \gamma_{max})$. In this integration step, the only face detection error was corrected. The integrated system achieved the same results as Table 4 automatically.

Table 2
PCA using 256×240 original images

Illumination angle	Recognition rates	
Head-on	15/15	100.00%
45°	15/15	100.00%
90°	14/15	93.33%

Table 3
PCA using 80×96 face areas

Illumination angle	Recognition rates	
Head-on	15/15	100.00%
45°	6/15	40.00%
90°	10/15	66.67%

Table 4
Normalized PCA using 80×96 face areas

Illumination angle	Recognition rates	
Head-on	15/15	100.00%
45°	14/15	93.33%
90°	10/15	66.67%

Table 5
Average reliability of correct recognition results

Illumination angle	Reliability of PCA	Reliability of Normalized PCA
45°	1.43	5.07
90°	0.48	1.70

3.5. System evaluation

The important issue regarding face recognition is how to suppress a false acceptance rate (FAR) and false rejection rate (FRR). This is often more important than face recognition accuracy. We attempt to reduce these rates by setting a threshold against the reliability R. Tables 7 and 8 show recognition rates, FARs and FRRs at two

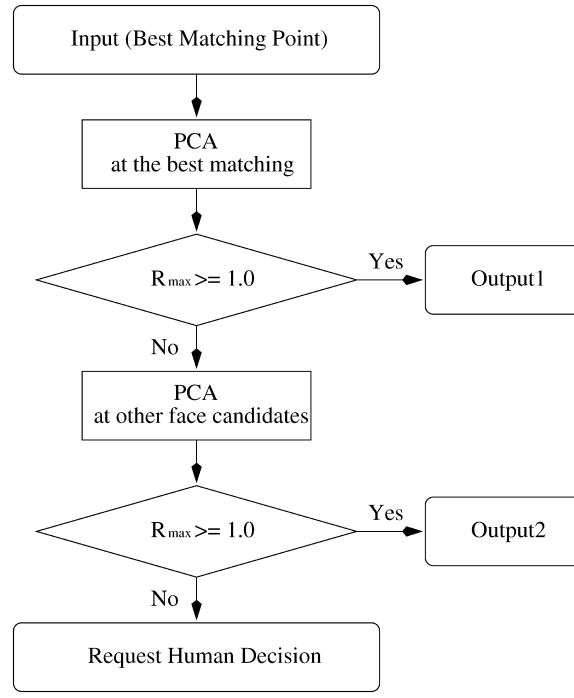

Fig. 13. Integration of face detection and recognition.

Table 6
Face displacement vs. recognition rates for 15 faces

Illumination angle	Vertical displacement (pixel)										
	−5	−4	−3	−2	−1	0	1	2	3	4	5
Head-on	9	11	11	14	15	15	15	15	14	12	10
45°	8	11	11	12	14	14	14	14	11	9	7
90°	5	5	6	10	10	10	10	9	8	8	7

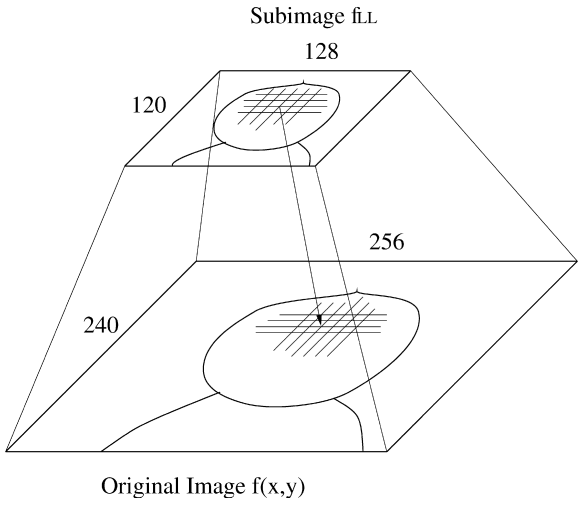

Fig. 14. Projection of the best matching point.

Table 7
System performance (threshold: $R = 1.0$)

Illumination angle	Recognition rates	FARs	FRRs
Head on	100% (15/15)	0% (0/15)	0% (0/0)
45°	100% (14/14)	0% (0/14)	0% (0/1)
90°	83% (5/6)	17% (1/6)	56% (5/9)

Table 8
System performance (threshold: $R = 0.5$)

Illumination angle	Recognition rates	FARs	FRRs
Head on	100% (15/15)	0% (0/15)	0% (0/0)
45°	93% (14/15)	7% (1/15)	0% (0/0)
90°	77% (10/13)	23% (3/13)	0% (0/2)

different threshold levels. At the higher threshold 1.0, it increases the recognition rates resulting in low FARs. However, FRRs are too high under side illumination. At the lower threshold 0.5, the recognition rates fall slightly and FRRs become 0% under any illumination conditions. An appropriate threshold level is determined by an application of the system.

To acquire higher recognition accuracy, PCA would need to be applied to each facial feature [15]. Alternatively, PCA could be combined with a feature-based approach [16] or a neural-network-based method [2,4].

4. Conclusion

We have described a system that automatically detects and recognizes a human face in an image captured with a cluttered background under non-uniform illumination. The system has been tested under three different lighting conditions using 45 images from 15 persons. Our face detection scheme has achieved a fairly robust performance under non-uniform lighting conditions, whereas our face recognition method has achieved a significant improvement over the conventional PCA-based approach.

One of the features of our system is to employ a symmetry axis detection method based on gradient orientation. Most of the existing image processing techniques are based on image intensity or image gradient. Owing to this, they are destined to be affected by the effects of non-uniform illumination. However, we have shown that gradient orientation is illumination-invariant by nature. Human vision is also quite stable under changing lighting conditions. This can be interpreted to mean that the local variations of the image intensity are more important than the image intensity itself. Gradient orientation is one way of representing some of the local variations of the image intensity. Thus, it would be possible to improve many of the existing image processing methods by using gradient orientation based techniques. Finally, in this paper, we have tackled a non-uniform illumination problem in face detection and recognition. Further work will be conducted to solove scale and rotation problems.

References

[1] A. Samal, P.A. Iyengar, Automatic recognition and analysis of human faces and facial expressions: a survey, Pattern Recognition 25 (1) (1992) 65–77.
[2] D. Valentin, H. Abdi, A.J. O'Toole, G.W. Cottrell, Connectionist models of face processing: a survey, Pattern Recognition 27 (1994) 1209–1230.
[3] A. Samal, Minimum resolution for human face detection and identification, SPIE: Human Vision, Visual Process. Digital Display II 1453 (1991) 81–89.
[4] M. Kosugi, Robust identification of human face using mosaic pattern and BPN, in: Proc. IEEE Workshop on Neural Networks for Signal Processing II, 1992.
[5] M. Kosugi, Human-face search and location in a scene by multipyramid architecture for personal identification, Systems Comput. Japan 26 (6) (1995) 27–38 (Translated from Denshi Joho Tsushin Gakkai Ronbunshi J77-D-II (1994) 672–681.
[6] M. Antonini, M. Barlaud, P. Mathieu, I. Daubechies, Image coding using wavelet transform, IEEE Trans. Image Process. 1 (2) (1992) 205–220.
[7] C. Sun, Symmetry detection using gradient information, Pattern Recognition Lett. 16 (1995) 987–996.

[8] A.L. Yuille, Deformable templates for face recognition, J. Cognitive Neurosci. 3 (1) (1991) 59–70.
[9] K. Sung, T. Poggio, Example-based learning for view-based human face detection, A.I. Memo No. 1521, CBCL. Paper No. 112, 1994. (CBCL stands for the Center for Biological and Computational Learning in the Department of Brain and Cognitive Sciences at the MIT.)
[10] H.A. Rowley, S. Baluja, T. Kanade, Human face detection in visual scenes, CMU-CS-96-158R, 1995.
[11] B.V.K. Vijaya Kumar, D. Casasent, H. Murakami, Principal-component imagery for statistical pattern recognition correlators, Opt. Engng 21 (1) (1982) 43–47.
[12] L. Sirovich, M. Kirby, Low-dimensional procedure for the characterization of human faces, J. Opt. Soc. Am. A 4 (3) (1987) 519–524.
[13] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–86.
[14] N. Intrator, D. Reisfeld, Y. Yeshurun, Face recognition using a hybrid supervised/unsupervised neural network, Pattern Recognition Lett. 17 (1996) 67–76.
[15] B. Moghaddam, A. Pentland, Face recognition using view-based and modular eigenspaces, SPIE: Automat. Systems Ident. Inspect. Humans 2277 (1994).
[16] R. Brunelli, T. Poggio, Face recognition: features versus template, IEEE PAMI 15 (10) (1993) 1042–1052.

About the Author—TOSHIAKI KONDO received his B.E. degree in Optics and his M.E. degree in Information Processing from the Tokyo Institute of Technology, Japan in 1986 and 1988, respectively.

Since 1988, he has been employed by Canon Inc. He has worked in projects dealing with image flow detection, digital filter design, image segmentation, digital zoom circuit and stereo vision. Currently, he is working on human face detection and recognition at the University of Sydney on leave from Canon Inc. His research interests include image processing, pattern recognition, computer vision and human visual perception.

About the Author—HONG YAN received his B.E. degree from Nanking Institute of Posts and Telecommunications in 1982, M.S.E. degree from the University of Michigan in 1984, and Ph.D. degree from Yale University in 1989, all in Electrical Engineering. From 1986 to 1989 he was a research scientist at General Network Corporation, New Haven, CT, USA, where he worked on developing a CAD system for optimizing telecommunication systems.

Since 1989 he has been with the University of Sydney where he is currently a Professor in Electrical Engineering. His research interests include medical imaging, signal and image processing, neural networks and pattern recognition. He is an author/co-author of one book and more than 150 technical papers in these areas. Dr. Yan is a fellow of the Institution of Engineers, Australia (IEAust), a senior member of the IEEE, and a member of the SPIE, the International Neural Network Society, the Pattern Recognition Society, and the International Society for Magnetic Resonance in Medicine.