

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

**4,800**

Open access books available

**122,000**

International authors and editors

**135M**

Downloads

Our authors are among the

**154**

Countries delivered to

**TOP 1%**

most cited scientists

**12.2%**

Contributors from top 500 universities



**WEB OF SCIENCE™**

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.

For more information visit [www.intechopen.com](http://www.intechopen.com)



# A Directional-Edge-Based Face Detection Algorithm Adapted to the VLSI Image Recognition System

Yasufumi Suzuki<sup>1</sup> and Tadashi Shibata<sup>2</sup>

<sup>1</sup>*Department of Frontier Informatics, The University of Tokyo*

<sup>2</sup>*Department of Electrical Engineering and Information Systems, The University of Tokyo  
Japan*

## 1. Introduction

There has been a considerable interest in the development of human-like image perception systems. Although a number of algorithms have been proposed so far, it is still a challenging task to achieve a human-like robustness in image perception under various circumstances. For robust image perception, the feature representation algorithm which extracts essential features from an image and represents them in a vector format is of critical importance. Since the discovery by Hubel and Wiesel (Hubel & Wiesel (1959)) in the study of visual cortex of animals, it is well known that the directional edge information in visual inputs plays an essential role in early visual processing. Namely, biological systems rely on the spatial relationship among edges in various directions for image perception. In this regard, local intensity gradients or directional edges have been widely used in computer vision and pattern recognition (Belongie et al. (2002); Dalai et al. (2005); Freeman & Roth (1995); Lowe (2004); Mikolajczyk & Schmid (2005); Shibata et al. (1999)). They are utilized in feature descriptors to represent object images in image and gesture recognition. The present work also relies on the edge information extracted from images.

A feature vector representation algorithm called Projected Principal-Edge Distribution (PPED for short) (Yagi & Shibata (2003)) has been developed to meet the demand of neural-associative-processor-based hardware recognition systems (Shibata (2002)). The original idea of PPED was first disclosed in 1999 (Shibata et al. (1999)) and developed in the work of Yagi & Shibata (2003). The robust nature of the PPED vector representation has been successfully demonstrated through its applications to hand-written pattern recognition and medical radiograph analysis (Yagi et al. (2000); Yagi & Shibata (2002)). The feature representation in PPED is based on the spatial distribution of four directional edges (horizontal,  $+45^\circ$ , vertical, and  $-45^\circ$ ) extracted from an image, in which the local variance of pixel intensities is taken into account for edge detection. Since the processing is computationally very expensive, a dedicated VLSI chip has been developed which can generate PPED vectors at a rate of  $10^6$  vectors/sec (Yamasaki & Shibata (2007)). An architecture that enhances the throughput by a factor of about 30 was also developed (Nakagawa et al. (2009)). VLSI chips specialized for the associative processing (maximum likelihood search) have also been developed in both digital (Ogawa et al. (2002)) and analog (Yamasaki & Shibata (2003)) technologies aiming at real-time performances. The face

detection algorithm in this article has been developed to explore a pattern matching algorithm specifically adapted to the psychologically-inspired VLSI brain model system, a hardware recognition system based on dedicated VLSI chips mimicking the processing in the mind (Shibata (2002; 2007)).

Humans have a specialized skill in face perception because it is directly related to our social life. However, it is still debated among psychologists whether the face perception is an ability specific to humans (Gauthier et al. (1999); Moscovitch et al. (1997)). A widely used technique to localize facial images in video sequences is to use the color information of faces as a clue (Hsu et al. (2002); McKenna et al. (1999)). The human perception of faces, however, does not rely on colors but on features as a clue. A number of face detection algorithms using facial features as a clue have been developed (Yang et al. (2002)) employing principal component analysis (Liu & Wechsler (2002)), neural networks (Feraud et al. (2001); Rowley et al. (1998a;b); Sung & Poggio (1998); Viola & Jones (2004)), support vector machines (Osuna et al. (1997)), and so forth. In these algorithms, however, a large amount of numerical computation is required in general. A fast search algorithm was proposed to apply the detector to real world applications (Feraud et al. (2001)). In Viola & Jones (2004), a real-time face detection system was developed by introducing "Integral Image" representation which is very efficient in computation, and face detection performance comparable to the best prior results was obtained (Rowley et al. (1998a); Sung & Poggio (1998)).

In this article, we focus on the face detection which localizes facial images in an input image without prior information about illumination, scales, numbers of faces, and so on. The PPED representation developed in the previous work (Yagi & Shibata (2003)) is not sufficient for face detection since certain essential information in facial images is lost during the dimensionality reduction in PPED vector formation, making it difficult to discriminate facial images from non-facial images. In order to solve the problem, other directional edge-based feature representations are proposed and the concept of a feature level image fusion called multiple-clue criterion has been introduced. Namely, several feature representations in addition to PPED are employed for template matching and the multiple clues derived from each of the feature representations are utilized for the classification to enhance the face detection performance. Furthermore, the robustness against scale variations and rotations of faces has also been achieved. The robustness of the proposed algorithm has been demonstrated using the CMU test set C (Rowley et al. (1998a)).

Face identification which identifies the localized face images as individuals is another important application of face perception. Developing both face detection and face identification systems realizes an automatic face recognition systems. In order to develop the face identification system, we have adapted the directional edge-based feature representations to a classification level image fusion with pseudo two-dimensional (2D) hidden Markov models (HMM). A dedicated VLSI hardware architecture for pseudo 2D HMM's has also been developed (Suzuki & Shibata (2007)), thus the real-time performance has been performed by combining of the dedicated VLSI chips for the feature generation and the pseudo 2D HMM's. The robustness against illumination variation was demonstrated in the face identification system using Yale face database B (Georghiades et al. (2001)).

The organization of the article is as in the following. Firstly, the directional edge-based feature representations utilized in the present work are described in Section 2. Then, the organization of face detection system and the concept of multiple-clue criterion are presented in Section 3. In Section 4, the proposed feature representations are also applied to multiple-angle-view face detection, which localizes not only frontal face but also posed faces and profiles. In Section 5, the proposed feature representations have been adapted to the hidden Markov model-based

face identification system. Section 6 demonstrates the performance of the algorithm. Section 7 concludes the article.

## 2. Directional edge-based feature representation algorithm

The feature representations utilized in the present work are described in this section. They are based on the spatial distribution of directional edges extracted from an original image. As a region of interests (ROI) for generating a feature vector, we employ a fixed-size window of  $64 \times 64$ -pixel sites, which is compatible with the hardware organization of the VLSI vector generator chip (Yamasaki & Shibata (2007)). A 64-dimension feature vector is generated from the pixel intensities within the window.

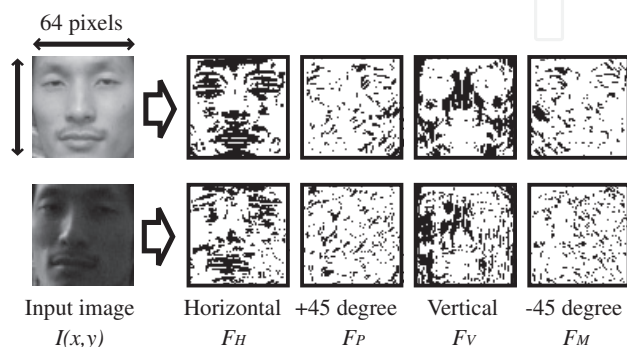


Fig. 1. Directional edge-based feature maps generated from bright and dark facial images. Since edges are detected taking local luminance variance into account, similar edge maps are obtained independent of the illumination condition.

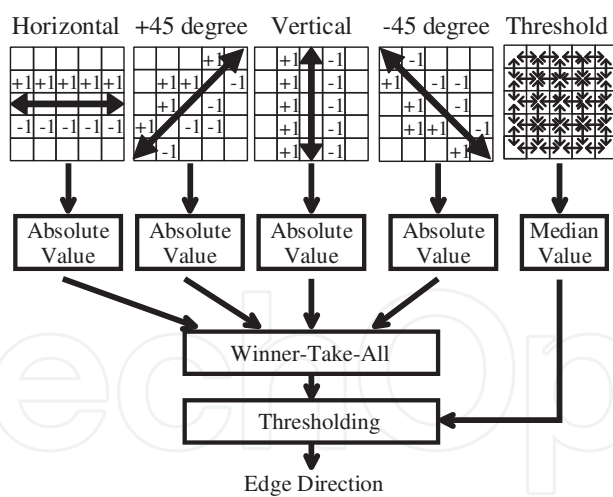


Fig. 2. Procedure of directional edge extraction.

### 2.1 Directional edge-based feature maps

The first step in forming feature vectors is the generation of four feature maps in which edges are extracted from a  $64 \times 64$ -pixel input image in four directions. Fig. 1 shows the relationship between an input image and four feature maps. Each feature map represents the distribution of edge flags corresponding to one of the four directions, i.e. horizontal, +45 degree, vertical, and -45 degree. These four directional-edge-based feature maps are regarded as representing the most fundamental features extracted from the original image

Horizontal	+45 degree	Vertical	-45 degree
0 0 0 0 0	0 0 0 1 0	0 1 0 -1 0	0 -1 0 0 0
1 1 1 1 1	0 1 1 0 -1	0 1 0 -1 0	1 0 -1 -1 0
0 0 0 0 0	0 1 0 -1 0	0 1 0 -1 0	0 1 0 -1 0
-1 -1 -1 -1 -1	1 0 -1 -1 0	0 1 0 -1 0	0 1 1 0 -1
0 0 0 0 0	0 -1 0 0 0	0 1 0 -1 0	0 0 0 1 0
KH	KP	KV	KM

Fig. 3. Filtering kernels utilized for extracting directional edges.

and all feature representations utilized in this work are derived from the feature maps. The procedure of feature-map generation is illustrated in Fig. 2. The input image is firstly subjected to pixel-by-pixel spatial filtering operations using kernels of  $5 \times 5$ -pixel size as in the following:

$$I_d(x, y) = \left| \sum_{i=-2}^2 \sum_{j=-2}^2 K_d(i, j) \cdot I(x + i, y + j) \right| \quad (1)$$

$$d \in \{H, P, V, M\}, \quad (2)$$

where  $I(x, y)$  is the pixel intensity at the location  $(x, y)$ , and  $K_d(i, j)$  is the filtering kernel shown in Fig. 3. The kernel which gives the maximum value of  $I_d(x, y)$  determines the direction of the edge at the pixel site. Namely, the preliminary edge flag  $F_d^*(x, y)$  is determined as follows:

$$F_d^*(x, y) = \begin{cases} 1, & \text{if } I_d(x, y) = \max_{d^* \in \{H, P, V, M\}} I_{d^*}(x, y) \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

This assigns one edge flag at every pixel site. In order to retain only edges of significance in feature maps, thresholding operation is introduced. The threshold value is determined by taking the local variance of luminance data into account. The intensity difference between two neighboring pixels in the horizontal direction  $H_{nm}(x, y)$  and that in the vertical direction  $V_{nm}(x, y)$  are obtained as

$$H_{nm}(x, y) = |I(x + n + 1, y + m) - I(x + n, y + m)| \quad (4)$$

$$V_{nm}(x, y) = |I(x + m, y + n + 1) - I(x + m, y + n)| \quad (5)$$

respectively, where  $n = -2, -1, 0, 1$ , and  $m = -2, -1, 0, 1, 2$ . The threshold value  $TH(x, y)$  is calculated as

$$TH(x, y) = M_{ed}(x, y) \times 5, \quad (6)$$

where  $M_{ed}(x, y)$  is the median of the 40 values of  $H_{ij}(x, y)$  and  $V_{ij}(x, y)$ . For each direction  $d \in \{H, P, V, M\}$ , the directional edge map  $F_d(x, y)$  is obtained as

$$F_d(x, y) = \begin{cases} F_d^*(x, y), & \text{if } I_d(x, y) > TH(x, y) \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

Thanks to such a thresholding operation, essential edges representing facial features are well extracted from both bright and dark facial images as shown in Fig. 1, thus making directional edge-based representations robust against illumination conditions.

**2.2 Projected principal-edge distribution**

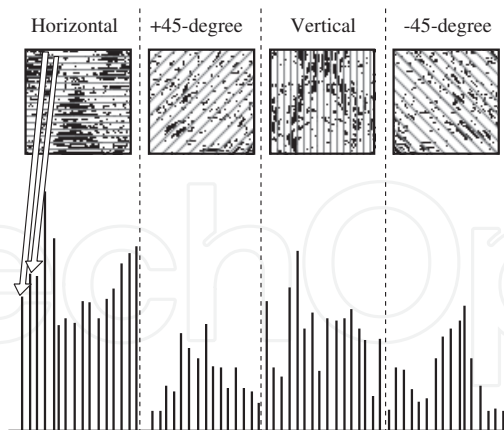


Fig. 4. Partitions of feature maps for vector generation based on projected principal-edge distribution (PPED).

Although the directional edge-based feature maps retain essential feature information in the original image, the amount of data is still massive and dimensionality reduction is essential for efficient processing of classification. Fig. 4 illustrates the procedure of feature-vector generation in the Projected Principal-Edge Distribution (PPED) (Yagi & Shibata (2003)). In the horizontal edge map, for example, edge flags in every four rows are accumulated and spatial distribution of edge flags are represented by a histogram. Similar procedures are applied to other three directions. Finally, a 64-dimension vector is formed by concatenating the four histograms. Details of the PPED feature representation are given in Yagi & Shibata (2003).

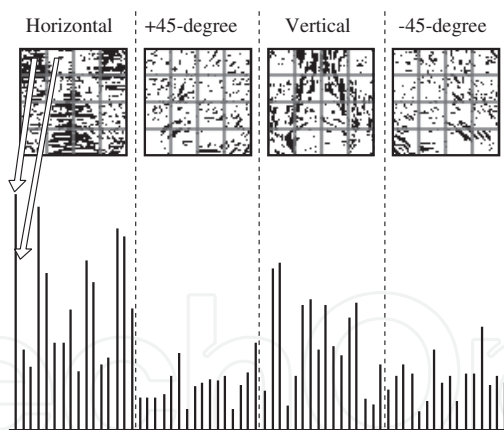


Fig. 5. Partitions of feature maps for vector generation based on averaged principal-edge distribution (APED).

**2.3 Averaged principal-edge distribution**

In the PPED feature representation, the information of edge distribution along the direction identical to the directional edge (e.g. the horizontal distribution of horizontal edge flags) is lost during the accumulation. In order to complement such loss in PPED vectors, other feature-vector generation schemes have been developed (Suzuki & Shibata (2004)). In the Averaged Principal-Edge Distribution scheme (which was originally named Cell Edge Distribution (CED) in Suzuki & Shibata (2004)), every feature map is divided into 4x4 cells

each containing  $16 \times 16$ -pixel sites as shown in Fig. 5. The number of edge flags in each cell is counted and the number constitute a single element in the vector representation by the Averaged Principal-Edge Distribution (APED) scheme. From the four directional edge feature maps  $F_H$ ,  $F_P$ ,  $F_V$ , and  $F_M$ , four 16-dimension vectors are obtained as in the following:

$$\begin{aligned}
 \mathbf{H}(a + 4b) &= \sum_{i=0}^{15} \sum_{j=0}^{15} F_H(16a + i, 16b + j) \\
 \mathbf{P}(a + 4b) &= \sum_{i=0}^{15} \sum_{j=0}^{15} F_P(16a + i, 16b + j) \\
 \mathbf{V}(a + 4b) &= \sum_{i=0}^{15} \sum_{j=0}^{15} F_V(16a + i, 16b + j) \\
 \mathbf{M}(a + 4b) &= \sum_{i=0}^{15} \sum_{j=0}^{15} F_M(16a + i, 16b + j) \\
 a &= 0, 1, 2, 3, \quad b = 0, 1, 2, 3.
 \end{aligned} \tag{8}$$

A 64-dimension feature vector  $\mathbf{X}$  in the Averaged Principal-Edge Distribution scheme is obtained by concatenating these four vectors as

$$\mathbf{X} = [\mathbf{H}, \mathbf{P}, \mathbf{V}, \mathbf{M}]. \tag{9}$$

Two types of 64-dimension feature vectors, the PPED (Projected Principal-Edge Distribution) vector and the APED (Averaged Principal-Edge Distribution) vector, are utilized for face detection in the present work.

### 3. System organization of multiple-clue face detection

In this section, the organization of the face detection system developed in the present work is presented.

#### 3.1 Overview of the system

In order to detect all faces in a target image, a window-scanning technique is employed for face detection as illustrated in Fig. 6. A partial image in the fixed-size window of  $64 \times 64$  pixels is taken from the input image and a 64-dimension feature vector is generated according to the procedure described in Section 2. Then, the feature vector is matched with all template vectors of both face samples and non-face samples stored in the system and classified as a face or a non-face according to the category of the best-matched template vector. Namely, when the best-matched template is found in the group of face samples, the partial image is determined as a face. If the best-matched template is in the group of non-face samples, it is determined as a non-face. In this work, no threshold value is employed in the template matching. The template matching is carried out utilizing the Manhattan distance as the dissimilarity measure, which is given by

$$d(\mathbf{X}, \mathbf{T}) = \sum_{i=1}^{64} |\mathbf{X}(i) - \mathbf{T}(i)|, \tag{10}$$

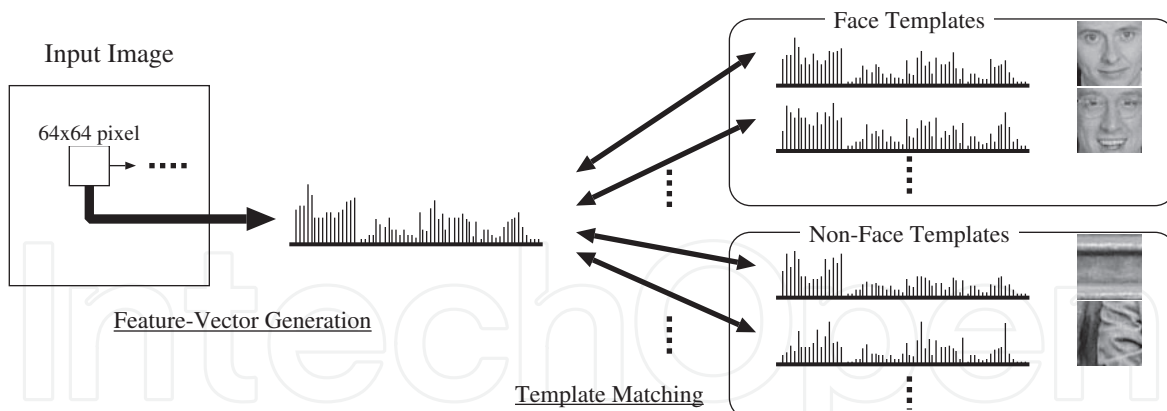


Fig. 6. A partial image in the fixed-size window of 64X64 pixels is taken from the input image and a 64-dimension feature vector is generated. Then, the feature vector is matched with all template vectors of both face samples and non-face samples and classified as a face or a non-face according to the category of the best-matched template vector. The window scans the entire image.

where  $X$  and  $T$  are a feature vector of the partial image and that of a template image, respectively. This classification is carried out by pixel-by-pixel scanning the entire image with the  $64 \times 64$ -pixel window.

### 3.2 Multiple-clue criterion

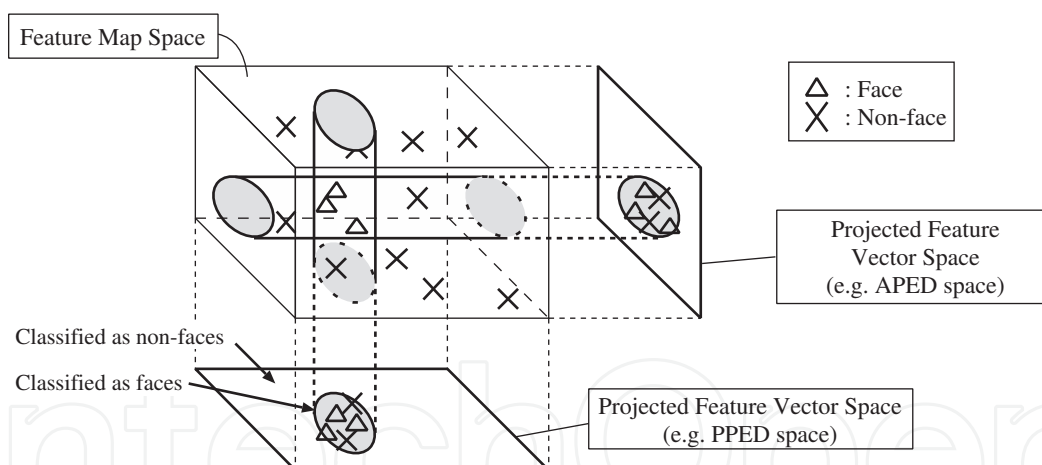


Fig. 7. Concept of multiple-clue criterion.

As we stated earlier, we assume that all essential feature information necessary to carry out correct classification is contained in the four directional edge-based feature maps. Then, the process of feature-vector generation is considered as the dimensionality reduction in the feature map space. In other words, the feature vectors are generated by projecting the data in the high-dimensional space of the four feature maps onto some lower-dimensional spaces as conceptually illustrated in Fig. 7. Then, we anticipate that some important features that separate faces from non-faces are certainly lost during the process of dimensionality reduction. This means, as illustrated in Fig. 7, the mixing of face samples ( $\Delta$ ) and non-face samples ( $\times$ ) necessarily occurs during the projection of sample points from the feature map space onto a reduced-dimension plane, e.g. the PPED space. However, if the sample points in



the feature map space are projected onto another plane, e.g. the APED space, the mixing of face and non-face samples would also occur, but the non-face samples mixed with face samples in the APED space would be different from the non-face samples mixed in the PPED space. Therefore, we can expect that true faces can be detected as sample points classified as faces in both PPED and APED spaces. This is the idea of multiple-clue criterion (Suzuki & Shibata (2004)). If the two low-dimensional spaces (PPED and APED in this example) are well separated from each other, a true face can be detected by taking logical AND of the detection results obtained using two types of vector representations.

In the principal components analysis (PCA), the dimensionality reduction is carried out by transforming the data to proper axes in order to maximize the variance of the data. Although the PCA is expected to retain essential feature information in the feature map space, it is not compatible with the VLSI hardware system. Therefore, in this work, we employed the multiple-clue criterion which is more compatible with the neural-associative processors.

### 3.3 Verification using fine-resolution feature representations

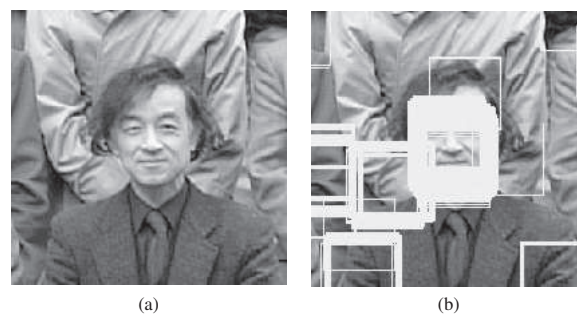


Fig. 8. Result of multiple-clue face detection from input image (a); many candidates are detected around a true face (b).

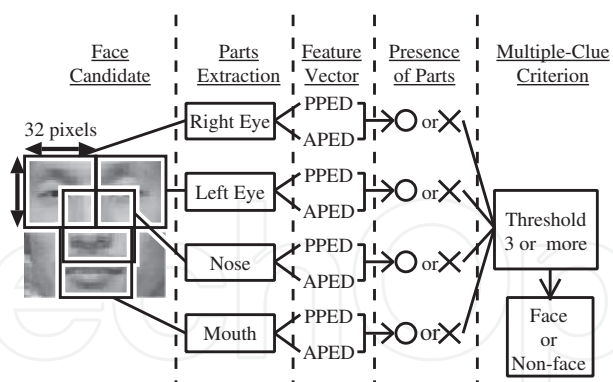


Fig. 9. Face verification by confirming existence of four facial parts by using finer edge information.

Due to the pixel-by-pixel scanning, multiple face candidates are detected around true faces as shown in Fig. 8(b). In order to merge the multiple detection results and determine the correct position of a face, we have utilized the image features extracted from a whole facial image. Confirming the presence of facial parts enhances the accuracy of the classification. Therefore, the face verification algorithm employing the feature vectors generated from local blocks with finer resolutions has been developed. In the verification algorithm, the presence of facial parts is confirmed by using finer edge information. If facial parts cannot be found, the candidate

will be rejected. In the four directional feature maps, as shown in Fig. 1, a lot of edge flags are extracted around the areas of eyes and a mouth in the horizontal edge map and the area of a nose in the vertical edge map. In this work, these four facial parts are utilized in the face verification.

The procedure of the verification is as in the following. Four  $32 \times 32$ -pixel windows are taken from the left-top, the right-top, the center, and the bottom in the candidate window as illustrated in Fig. 9. They correspond to the locations of the right eye, the left eye, the nose and the mouth, respectively. Each  $32 \times 32$ -pixel window is converted to both PPED and APED vectors. In the face verification, finer edge information is utilized as compared to the original PPED and APED vectors representing whole facial features. Namely, the elements representing the horizontal or vertical edge histograms in the PPED representation, for instance, are produced by accumulating the number of edges within two columns or two rows, respectively, instead of four columns or four rows in the face detection. The elements in  $\pm 45^\circ$  histograms are also produced from reduced areas. Similarly, each element of the APED vector is generated as the number of edges in a  $8 \times 8$ -pixel cell instead of a  $16 \times 16$ -pixel cell. Template matching is carried out using the eight feature vectors generated from four facial parts with two options of PPED and APED vectors. The presence of each facial part is determined when both PPED and APED vectors match the respective template vectors. If three out of four parts are confirmed, the candidate is determined as a true face. This threshold value is determined by the experimental results discussed in Section 6 (Fig. 29).

### 3.4 Scale and rotation-invariant detection

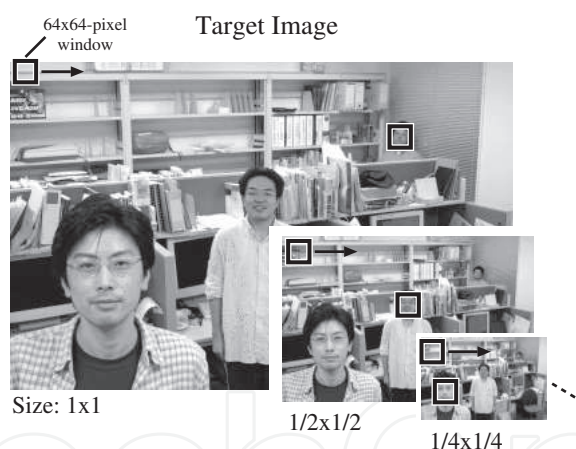


Fig. 10. Scale-invariant face detection is carried out using a fixed-size face detection window ( $64 \times 64$ ) which scans images with full-scale,  $1/2$ ,  $1/4$ , and  $1/8$  scales.

Although we employed the fixed-size window, facial images may appear in various scales in the target image and not fit to the window. In order to accommodate our system to scale variations, the size of the input image is reduced by factors of 1,  $1/2$ ,  $1/4$ , and  $1/8$  as illustrated in Fig. 10, and face detection is carried out for each scale. (Enlargement by factors of 1, 2, 4,  $\dots$  etc. is also possible.) Optional scaled samples are added to the set of face templates to cover the entire range between the original and  $1/2$  scaled images. Due to the robust nature of the directional edge-based representations, the face detection performance does not appreciably degrade for scale variations within  $\pm 15\%$ . Since  $(1 + 0.15 \times 2)^3 = 2.2$  is larger than 2, three types of scaled facial images are required to cover the scale range between 1 and  $1/2$ . Therefore, facial images with reduced sizes of 80% and 60% are used in addition to the original images (see the results in Fig. 21).

Furthermore, facial images do not always appear upright. A similar interpolating scheme has been introduced to the system to accommodate the system to the rotated facial image detection. Namely, face detection is carried out for 90, 180, and 270 degree-rotated images in addition to the upright target image. Tilted face images within the angles between  $-45$  and  $+45$  degrees are added to the set of face template vectors to cover the entire 360 degrees rotation. Due to the symmetry of vector generation algorithm for 90-degree rotations in the PPED and APED vectors, the template vectors for 90, 180, and 270 degree-rotated images can be generated by just swapping elements in the original vector. The PPED and APED vectors have also tolerances for image rotation in a certain range. In this work, 0,  $\pm 18$ , and  $\pm 36$ -degree tilted face images are included as face templates for rotation invariant face segmentation (see Fig. 21(b)).

### 3.5 Noise reduction using Gaussian-blur filtering operation

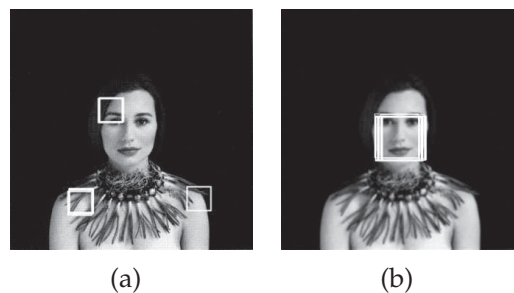


Fig. 11. Face detection results from original image (a) and  $5 \times 5$  Gaussian-blur filtered image (b).

Fig. 11(a) shows an example of face detection and verification result using an image in the CMU test set C (Rowley et al. (1998a)), which missed the true face. In this case, the target image is scanned from printed material, containing a lot of print noises. Such pattern noises degrade the performance of the system since the directional edge extraction is quite sensitive to them. In order to remove the noises, the target image is subjected to  $5 \times 5$  Gaussian-blur filtering operation in advance. The detection result for the blurred image is shown in Fig. 11(b). The true face is detected correctly without false positives. In our system, the face candidates are obtained by taking OR of the candidates from both the original and Gaussian blurred images.

## 4. Multi-view face detection

We have presented the frontal face detection system in Section 3. The system has been adapted to various scales and rotations of facial images. However, posed faces such as profiles cannot be detected since their appearance is different from frontal faces. The simple solution to this problem is to include posed faces into the template sets. Fig. 12 shows the directional edge-based feature maps generated from frontal face, half-way posed face, and profile images. Facial parts such as the eyes, the nose and the mouth can be easily recognized in the horizontal and vertical edge maps generated from frontal face images. However, it is difficult to identify the patterns of facial parts from the feature maps of the profile. The left-hand side of the feature map represents the rear half of the profile face, which varies a lot depending on the hairstyle of each person. Therefore, it degrades the performance of profile detection. Since the important facial parts are on the right-hand side of the feature map, the front side of the profile is utilized as a focus-of-attention (FOA) area for profile face detection. Fig. 13 illustrates the

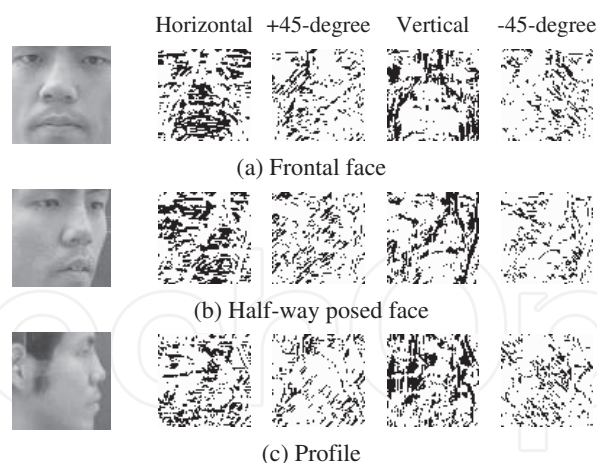


Fig. 12. Directional edge-based feature maps generated from frontal face (a), half-way posed face (b), and profile (c) images.

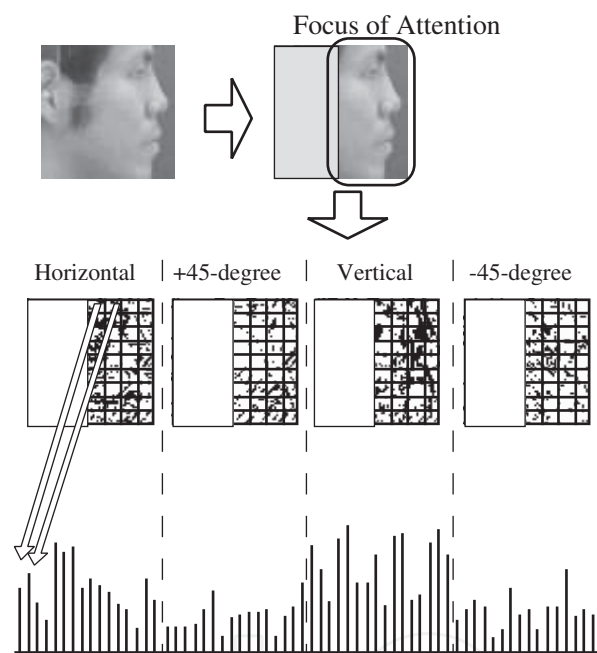


Fig. 13. Profile-specific feature-vector generation based on Averaged Principal-Edge Distribution from focus-of-attention (FOA) area.

profile-specific feature-vector generation from the focus-of-attention (FOA) area. The feature vectors are formed by similar procedures as in the APED vector generation. Namely, the FOA area is divided into  $4 \times 4$  cells each of which contains  $8 \times 16$ -pixel sites, and each element of the 64-dimension vector represents the number of edge flags within the corresponding cell. In this system, the face verification algorithm described in Section 3.3 was not carried out.

### 5. Face identification employing pseudo two-dimensional hidden Markov models

In this section, we focused on a face identification system where the localized facial images, which are already detected and localized in the scene, are identified as individuals registered in the system. In order to realize a face identification system, a classification level image fusion with pseudo two-dimensional (2D) hidden Markov models (HMM) has been employed.

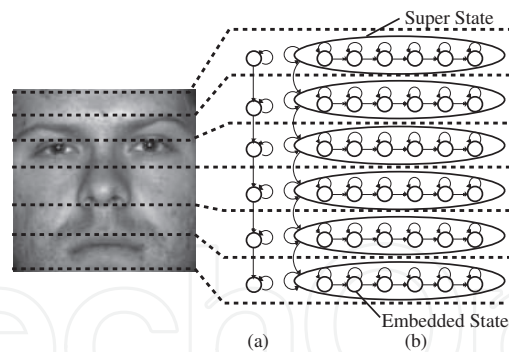


Fig. 14. 1D Hidden Markov Model (a) and pseudo-2D Hidden Markov Model (b).

HMM's are one of the statistical classifiers successfully applied to the speech recognition. The original form of HMM's is based on a simple one dimensional (1D) Markov chain. In order to process images as 2D data, pseudo-2D HMM's have been introduced (Kuo & Agazzi (1994)) and applied to face recognition (Nefian & Hayes (1998)). The pseudo-2D HMM consists of a set of super states which contain a 1D HMM within themselves. The pseudo-2D HMM is utilized for modeling facial images in a hierarchical manners as in the following. Several super states correspond to the vertical facial features, such as forehead, eyes, nose, and mouth as illustrated in Fig. 14 (b). Each state within the super state is utilized for modeling the horizontal sequence of the localized feature. The pseudo-2D HMM-based face identification has often utilized the coefficients of 2D discrete cosine transform (DCT) for feature representation (Eickeler et al. (2000); Nefian & Hayes (1999)). However, the representation is sensitive to the change in illumination conditions. In order to develop an illumination-invariant system, the directional edge-based feature representation is employed for the pseudo-2D Hidden-Markov-Model-based face identification.

### 5.1 Feature representations

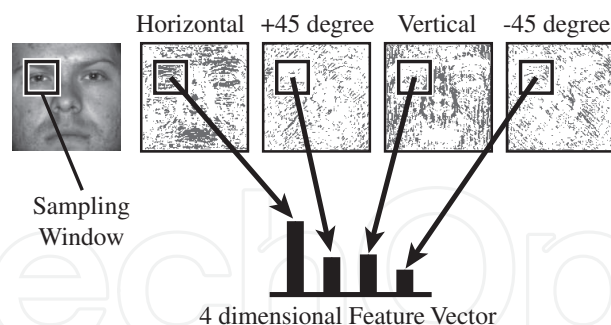


Fig. 15. Feature-vector-generation scheme from edge-based feature maps: each element represents the number of edge flags within sampling window.

The proposed feature representation vectors are generated from the edge-based feature maps described in Section 2.1 by counting the number of edge flags within the sampling window as illustrated in Fig. 15. Each element of the vector represents the number of edge flags within the window of the corresponding edge direction. Since four directional edges are available in the feature maps, four elements of the feature representation vector are generated from a sampling window. In order to represent the 2D structural information within the sampling window, another type of feature-vector-generation scheme illustrated in Fig. 16 was also employed. In this scheme, a larger size sampling window is employed and then divided into  $2 \times 2$  cells. A 16-dimension feature vector is generated by counting the number of edge flags in each cell.

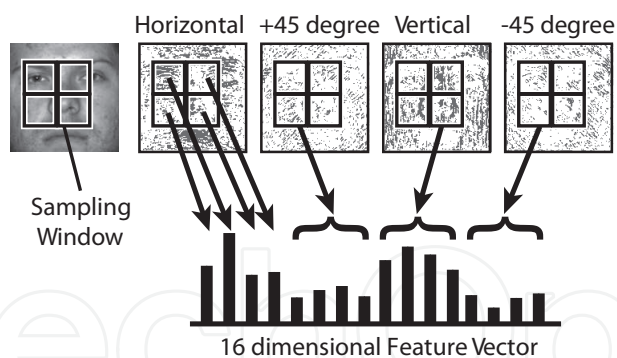


Fig. 16. Another feature-vector-generation scheme: four times larger size sampling window is divided into  $2 \times 2$  cells and each element represents the number of edge flags within a cell.

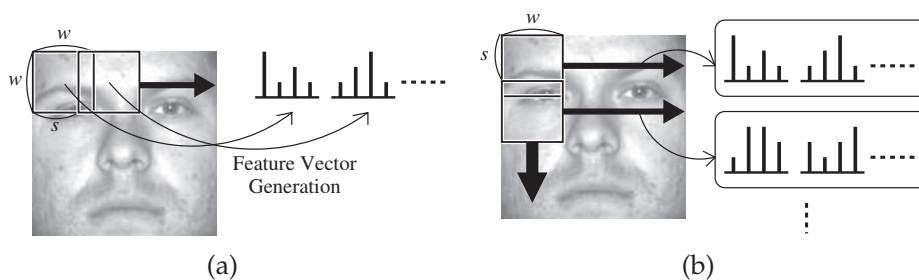


Fig. 17. Sampling window scans face image; horizontal scanning (a) and vertical scrolling (b).

The sequence of observation vectors is acquired as follows. The  $w \times w$ -pixel window scans the image from left to right and makes the sequence of feature vectors as illustrated in Fig. 17(a). For each sampling, the sampling window shifts  $s$  pixels to right. The horizontal scanning is repeated with the vertical pitch of  $s$  pixels from the top to bottom as shown in Fig. 17 (b). In this manner, a series of feature vectors is generated. These feature vectors are utilized as observation vectors for the pseudo-2D Hidden Markov Models on both training and identification stages.

### 5.2 Face identification using pseudo-2D HMM's

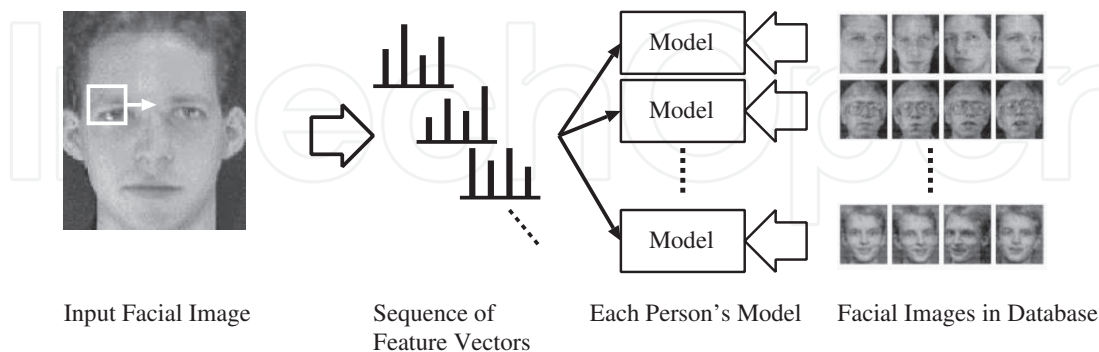


Fig. 18. Pseudo-2D HMM-based face identification system. A sequence of feature vectors is generated from the partial image in the scanning window, which is then matched with each person's HMM using the Viterbi algorithm. Then, the maximum-likelihood model yields the identity of the face.

The identification of the target facial image is carried out as in the following. The sequence of edge-based feature-vectors is classified by the pseudo-2D Hidden Markov Models. The  $6 \times 6$ -state left-right model illustrated in Fig. 14 (b) is utilized in this work. The model consists of six super states each of which contains an one-dimension Hidden Markov Model with six embedded states. The sum of three Gaussian mixtures is employed for the probability function of the embedded state. These parameters of the HMM are optimized by the experimental results on the AT&T face database (Samaria & Harter (1994)).

In the training, one HMM is generated for each person using the Baum-Welch algorithm. For avoiding false local minima, the common initial face model (Eickeler et al. (2000)) is utilized. Namely, the common initial face model is firstly trained on all faces in the training set. The face model for each person is obtained from the common model by refining it on the training faces of the person with the Baum-Welch algorithm. In the identification, the target facial image is evaluated by each face model using the Viterbi algorithm as illustrated in Fig. 18. The face model which gives the maximum probability determines the identity of the input image.

### 5.3 Dedicated VLSI architecture for pseudo-2D HMM classifier

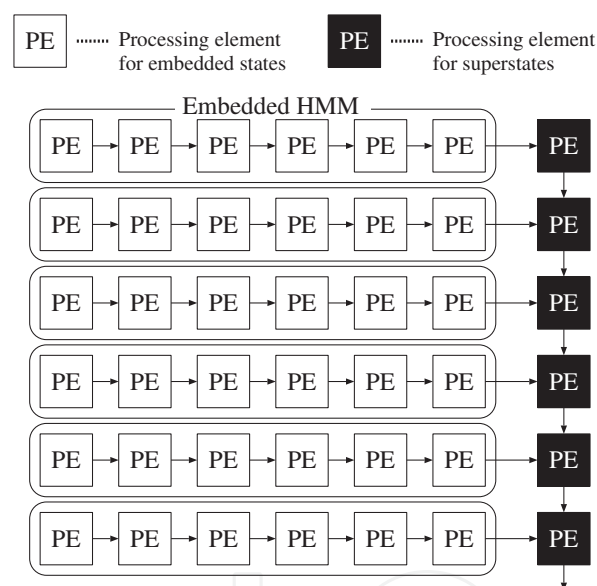


Fig. 19. Block diagram of pseudo-2D HMM VLSI processor. Each PE corresponds to an embedded state or a superstate in pseudo-2D HMM.

In order to realize the real-time face identification system, a dedicated VLSI hardware architecture for pseudo-2D HMM classifier has been developed. The proposed architecture consists of processing elements (PE's) each of which corresponds to an embedded state or a super state as illustrated in Fig. 19. The processing element for the embedded state is composed of the Viterbi decoder and the observation probability calculator which identifies the maximum-likelihood path among all possible state transitions within the HMM. The details of the proposed architecture are described in Suzuki & Shibata (2007).

The proposed architecture has been implemented in a field programmable gate array (FPGA). The processing time for identifying a  $92 \times 112$  pixel grayscale facial image from 40 people was 4.42 ms at 100 MHz clock frequency. Therefore, the processor is capable of identifying more than 220 facial images in a second.

## 6. Experimental results and discussion

### 6.1 Frontal face detection



Fig. 20. Group photo in which 2,000 non-face images were randomly taken from the background scenery as non-face templates.

In this work, the upright frontal facial images of 300 Japanese people from HOIP facial image database were utilized as face templates in the experiment of frontal face detection. As the non-face templates, 2,000 non-face images which were randomly taken from the background scenery of only a group photo shown in Fig. 20 were employed. Updating of the non-face templates by feeding back the results of false positives, or other sorts of maintenance were not conducted in this work. The same non face templates were used throughout the experiments. The performance of the face detection system has been evaluated on the CMU test set (Rowley et al. (1998a)).

The system was evaluated by using two types of measures which are detection rates and false positive rates. They were defined as in the followings. Because of the pixel-by-pixel scanning, multiple windows of the detection results are overlapped as shown in Fig. 8(b). If two windows were overlapped in area more than 50% each other, they were merged into one window. Only when the merged window covered the entire face including the eyes, the nose, and the mouth, we determined that the face was detected correctly. The detection rate is the number of faces detected against faces in the test images. The false positive rate is the number of false detections against all  $64 \times 64$ -pixel windows examined in the evaluation. The system decides whether an image within the window is a face or not at each pixel site independently. Therefore, the false positive rate described above is employed for the evaluation.

#### 6.1.1 Optimization for scale-invariant detection

As described in Section 3.4, it is required to append the optional scaled samples to the set of face templates to cover the range between the original and 1/2 scaled images. In order to optimize the sets of templates for scale variations, face detection was carried out with the template vectors derived from different combinations of scaled face samples. Fig. 21(a) demonstrates the results of face detection rate. The template set of 100%+70%, for example, means that 70% scaled 300 face samples were appended to the original 300 face samples for the face templates of reduced scales. As illustrated in Fig. 21(a), when three types of scaled face samples (100%, 80%, and 60%) were utilized, over 99% detection rate was shown. Therefore, the three types of face samples (100%, 80%, and 60%) were utilized in the following experiments.



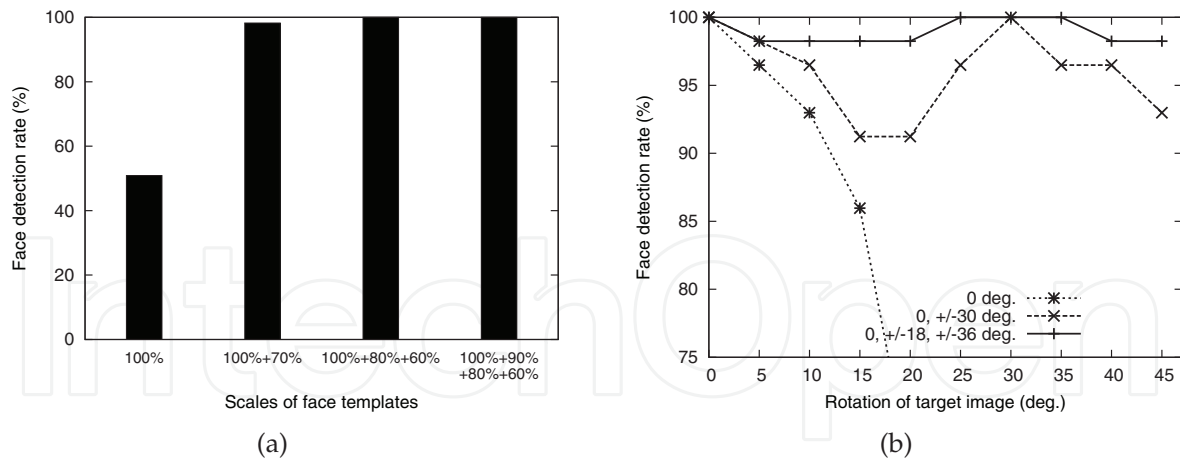


Fig. 21. Face detection rates with template vectors derived from different combinations of scaled face samples (a); detection rates from rotated target images with template vectors derived from different combinations of tilted face samples.

### 6.1.2 Optimization for rotation-invariant detection

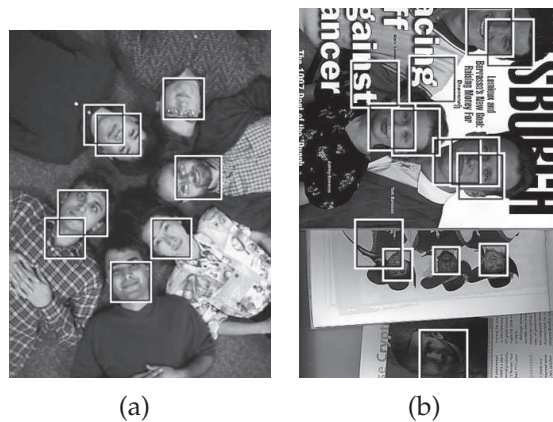


Fig. 22. Results of scale and rotation-invariant face detection using images in Rowley et al. (1998b). Verification by facial parts were not carried out in the experiment.

As described in Section 3.4, it is required to prepare the tilted face samples as the face templates to cover the range between  $-45$  and  $+45$  degrees. To optimize the face template sets, face detection was carried out on rotated target images between  $0$  and  $+45$  degrees using different combinations of tilted template vectors. The detection rates are presented in Fig. 21(b). Using only upright templates, namely only  $0$ -degree template vectors, the detection rate falls rapidly at  $10$  degrees of tilting. Although the set of  $0$  and  $\pm 30$ -degree template vectors covers the range, the detection rate fluctuates depending on the tilt angle. The results obtained with the template set generated from  $0$ ,  $\pm 18$ , and  $\pm 36$ -degree tilted facial images show the detection rate over  $95\%$  all over the range. Fig. 22 demonstrates the results of scale and rotation-invariant face detection using images that contain difference size and angle of faces in the CMU rotation test set (Rowley et al. (1998b)).

### 6.1.3 Tests of face detection and verification

In order to eliminate false positives, face verification algorithm described in Section 3.3 was employed here. Here, if at least three facial parts were confirmed out of four parts, the



Fig. 23. Result of face detection (a) and face verification (b).

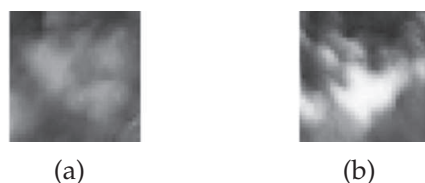


Fig. 24. Examples of face candidates which were detected in Fig. 23(a) but rejected by verification in Fig. 23(b); (a) and (b) are on top left and top right corner in original image, respectively.

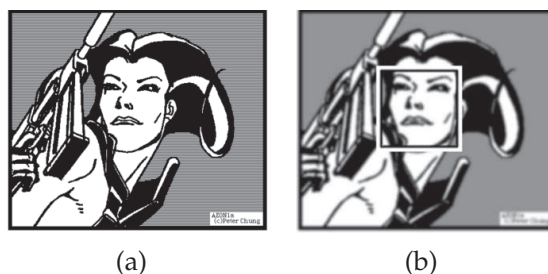


Fig. 25. Face detection results on line drawing face using original image (a) and  $5 \times 5$  Gaussian-blur filtered image (b).

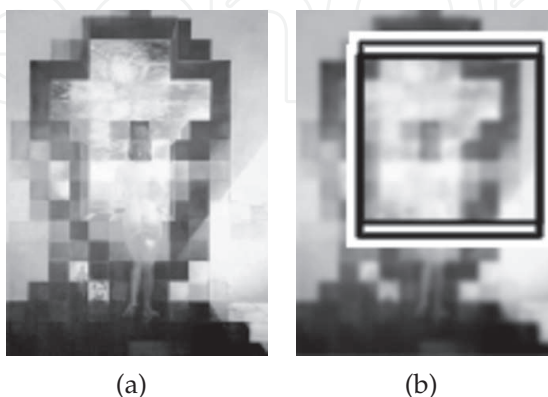


Fig. 26. Results of face detection on Dali's "Gala Contemplating the Mediterranean Sea" from original image (a) and  $5 \times 5$  Gaussian-blur image (b).

candidate was accepted as a true face. Fig. 23(a) demonstrates the result of face detection using a picture in the CMU test set B. All faces in the picture are detected. However, a lot of false positives occur. These detected faces were screened using the face verification algorithm and the result of verification is demonstrated in Fig. 23(b). All false positives except one were rejected by the verification without missing the true faces. Fig. 24 shows examples of false positives in face detection shown in Fig. 23. These candidates, which look like faces at a glance by human perception, were rejected by verification algorithm since facial parts could not be confirmed.

As described in Section 3.5, the face candidates are obtained from both original and  $5 \times 5$  Gaussian-blur filtered images. This Gaussian-blur operation makes another effect to face detection. A line drawing facial image shown in Fig. 25(a) was not detected from the original image but detected and verified from the blurred one as illustrated in Fig. 25(b). Fig. 26 demonstrates the face verification results for the drawing by Salvador Dali, "Gala Contemplating the Mediterranean Sea" which is also called "Lincoln in Dalivision". Although no faces are detected from the original image in a reduced size shown in Fig. 26(a), Lincoln is detected and verified correctly in the blurred image as illustrated in Fig. 26(b).

## 6.2 Performance evaluation on CMU test set

In order to evaluate the performance of the system, the proposed algorithm has been evaluated on the CMU test set C which consists of 65 images containing 183 faces. Fig. 27 presents the detection rates and false positive rates using the PPED and APED feature representations. The detection rate is the number of faces detected correctly against all 183 faces in the test set. The false positive rate is the number of false detections against 80,146,631 which is the number of  $64 \times 64$ -pixel windows examined in this evaluation. Generally, there are trade-offs between the detection rates and the false positive rates, and the closer the data points approach the left-top corner, the better the performance is. When PPED vectors are utilized, 93% of faces are detected from the original images and 97% detection rate is achieved for  $5 \times 5$ -pixel Gaussian-blur filtered images. APED vectors show better performance than PPED, and 97% and 99% detection rates are obtained for the original and Gaussian blurred images, respectively. Only one facial image shown in Fig. 28 was not detected from either original or blurred images. This image is composed of black-and-white dots and the high frequency components in the image were not removed by the Gaussian-blur filter, which contributed edge flags to the feature maps. When the image was subjected to Gaussian-blur filtering operation one more time, however, the face was detected. On the other hand, the false positive rates are about 0.01 (1%) and a huge number of false positives occur. Fig. 27 also shows the detection results using the multiple-clue criterion using both PPED and APED, which are 88% and 95% for the original and Gaussian-blur filtered images, respectively. Although the face detection rate is a little degraded as compared to the case using only PPED vectors or APED vectors, the false positive rate is improved. In order to improve the detection rate, we merged the detection results from both the original images and Gaussian blurred images. This OR operation achieves 95% detection rate when the multiple-clue criterion is employed as shown in Fig. 27.

In order to further reduce the number of false positives, the face verification algorithm explained in Section 3.3 was introduced and the number of facial parts to confirm in the face verification algorithm was varied. The face detection rates for the verification algorithm in which the existence of no less than one, two, three, and four facial parts out of four parts are given in Fig. 29. If all of four facial parts are required for verification, the detection rate is degraded to about 50% while the number of false positive is drastically reduced. From

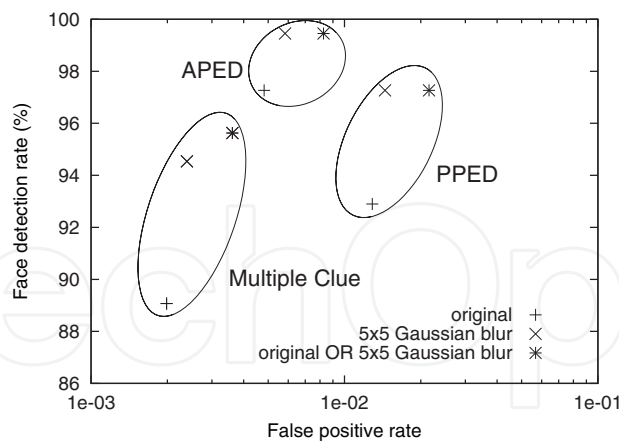


Fig. 27. Detection rates and false positive rates using PPED, APED, and multiple-clue criterion of them.

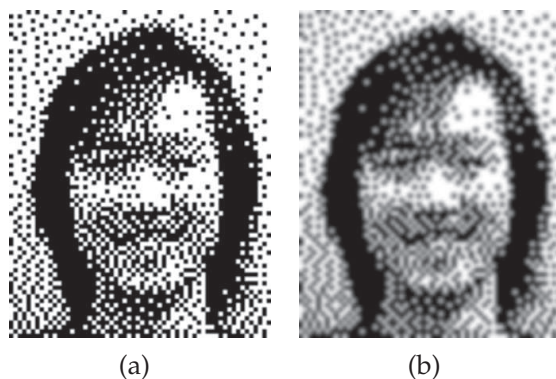


Fig. 28. Facial image could not be detected using APED feature vector from either original image (a) or blur image (b).

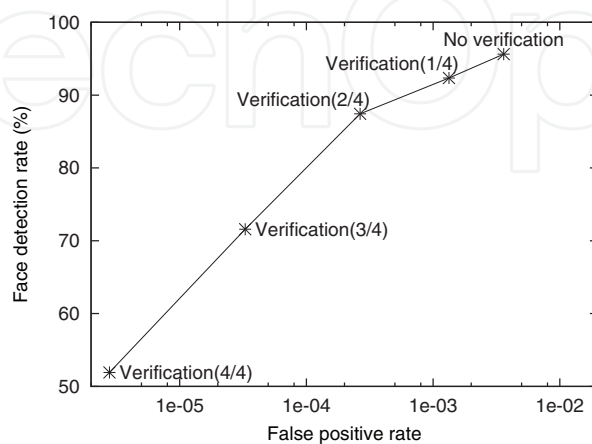


Fig. 29. Detection rates and false positive rates using face verification.

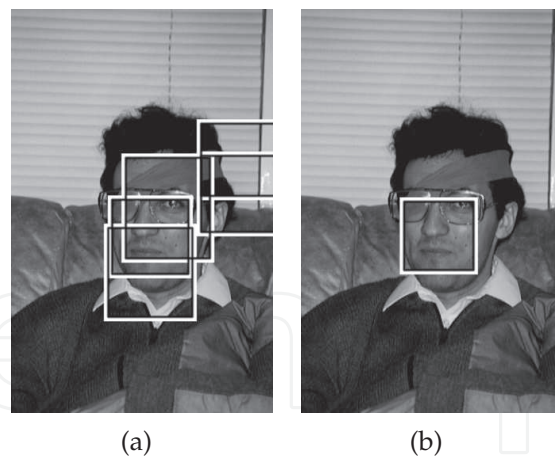


Fig. 30. Results of face detection (a) and verification (b) from facial image whose right eye is occluded.

these results, we adopted three or more facial parts as the threshold for verification. Fig. 30 illustrates the result of detection and verification on an occluded face. In this case, the right eye failed to be confirmed but the other parts were validated and the true face was correctly verified.

As demonstrated in Fig 29, there are trade-offs between the detection rates and the false positive rates, thus the system yielding less false positives demonstrates the better performance at the same face detection rate. In the face detection system employing two facial parts for the threshold value in face verification, for instance, the detection rate using the CMU test set C was 87.4% while 689 false positives occurred. When the system was evaluated on all images in the CMU test set A, B, and C, the detection rate was degraded to 71.5% while 1,055 false positives occurred. The systems developed by Viola & Jones (2004) and by Rowley et al. (1998a) achieved the higher detection rate (81.1% and 83.2%) with 10 false positives, respectively. These results indicate that the performance of our system is not comparable to other systems. In the learning, Viola and Jones employed the AdaBoost algorithm which increases the effective number of non-face templates. Rowley et al. updated the non face-templates by feeding back the results of false positives. These techniques have not been studied in this work yet. Introducing such techniques to the learning algorithm of false positives would improve the performance of our system in reducing the false positive detection rate, which would be the subject of future study.

The present face detection algorithm has been developed as a typical pattern matching algorithm specifically tuned for the VLSI brain mimicking system under intensive development (Shibata (2002; 2007)). Therefore, it involves a lot of computationally demanding operations. However, they are very efficiently processed using dedicated VLSI chips developed for the system. For instance, the directional-edge based vector generation from every pixel site in a VGA-size image is carried out at a rate of 6.1 frames/s (Yamasaki & Shibata (2007)). Further enhancement in the processing speed by a factor of 30 has been achieved by an advancement in the chip architecture (Nakagawa et al. (2009)). In order to accelerate the scale-invariant detection, a CMOS image sensor chip capable of performing multiple-scale filtering processing for a  $64 \times 64$ -pixel image has been developed and operation at a 680 frames/s was demonstrated (Takahashi et al. (2009)). Therefore, the processing in the present algorithm is all very efficiently conducted in the VLSI brain mimicking system. However, regarding the recognition performance, the present results are not sufficient when compared to those of advanced software systems. If the false positive rate is defined as the

number of false detections against all the detection results, our results are in the range of several 10%, which are far inferior to the results in Rowley et al. (1998a) and Viola & Jones (2004). PPED-based face detection was also studied using a support vector machine (SVM) as a classifier instead of the simple nearest-neighbor search method employed in this work, which showed much better results. In this regard, the basic VLSI circuits for SVM have also been developed for use in a certain kind of classification problems (Kang & Shibata (2009)).

### 6.3 Experiment on multi-view face detection

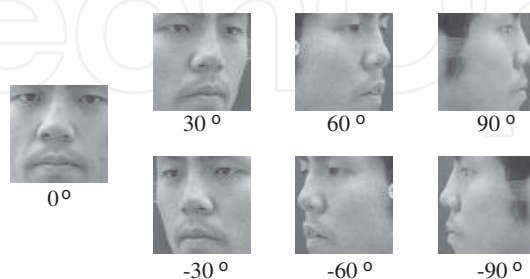


Fig. 31. Template images for multi-view face detection.

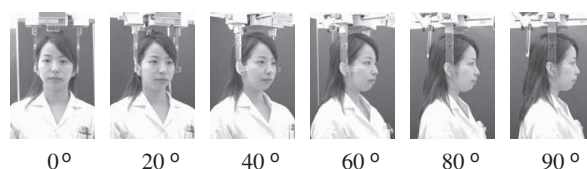


Fig. 32. Target images for multi-view face detection.

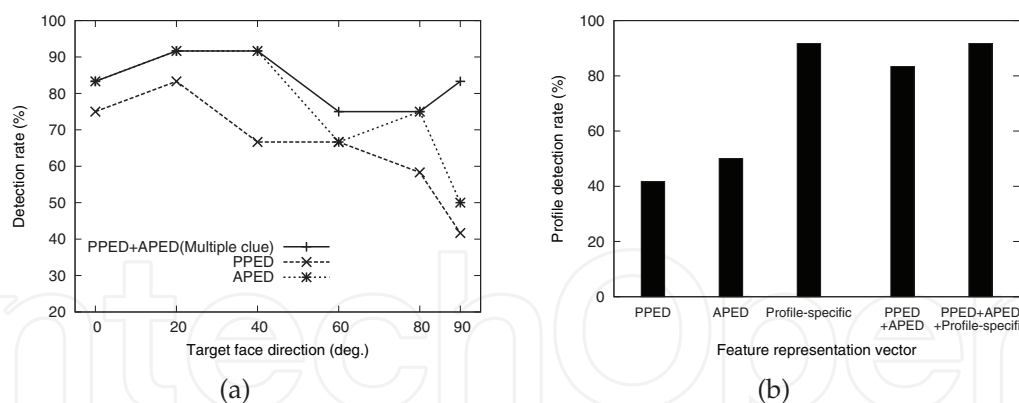


Fig. 33. Results of multi-view face detection rate (a); results of profile detection rate (b).

In order to evaluate the multi-view face detection, 130 face images of 13 people taken from 10 directions between the range of 0 to 90 degrees with an increment of 10 degrees were utilized. These sample images were prepared as a preliminary database to use in the multi-dimensional facial soft tissue analysis for the treatment planning in orthodontics. Four direction posed faces at 0, 30, 60, and 90 degrees were utilized as templates. In addition, -30, -60, and -90-degree direction faces were generated by taking the mirror images of original samples as shown in Fig. 31. Face detection was carried out on face images angled at 0, 20, 40, 60, 80, and 90 degrees for 13 people as shown in Fig. 32. The detection rate was evaluated by the cross validation. Namely, all face images except for one person were utilized as templates

and the face detection was carried out for the face images of the person excluded from the templates. This procedure was repeated for all images in the database. In this experiments, based on the prior knowledge that only one face image is existing in the target image, the location where the feature vector gives the minimum distance to the best matched template was determined as a face. And the direction of the template which gives the minimum distance among the all templates determines the pose of the target face. Fig. 33(a) shows the detection rates of six direction faces using the PPED and APED feature vectors, respectively. Although almost over 80% detection rates are obtained in frontal and 20-degree faces for both feature vectors, the detection rates degrade to less than 60% in profile images. In order to improve the performance, the multiple-clue criterion have been utilized. Fig. 33(a) illustrates the correct detection rates using the multiple-clue criterion. The detection rates all of angled faces were improved and approximate 80% detection rates were obtained. Fig. 33(b) shows the detection rates of profile images for three types of feature vectors: the PPED, APED, and profile-specific APED as discussed in Section 4. In addition, the detection rates using the multiple clue derived from all three feature vectors are also shown in Fig. 33(b). Only about 50% of the target profile images were detected with the original feature vectors (PPED and APED). However, over 90% detection rate was obtained with the proposed profile-specific APED feature vector.

#### 6.4 Experiment on face identification using pseudo-2D HMM's

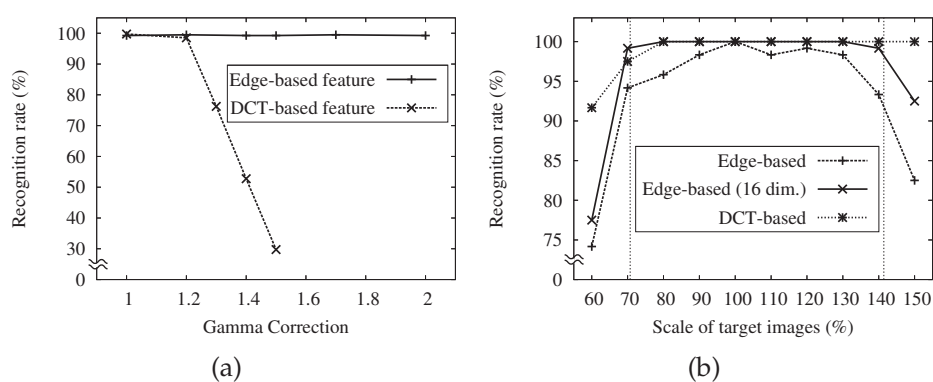


Fig. 34. Recognition rate on AT&T database with different illumination conditions (a) and with different scales of target facial images (b).



Fig. 35. Viterbi segmentation of face images under different illumination conditions using edge-based feature vectors (a) and DCT-based feature vectors (b); each block corresponds to each state of pseudo-2D HMM.

In order to evaluate the performance of the proposed face identification system, the AT&T face database (Samaria & Harter (1994)) was used for both training and test sets. This database

contains 10 different images for each of 40 people. These face images are taken at similar illumination conditions. As a preliminary evaluation of the robustness against illumination variations, various illumination conditions were emulated by the gamma correction. The recognition rate was evaluated by the cross validation. Namely, each person's model was trained on all face images except for one image and the classification was carried out for the image excluded in the training. This procedure was repeated for all images in the database. The recognition rates on various illumination conditions employing the edge-based and the DCT-based feature representations are presented in Fig. 34(a). In this experiment, first nine coefficients of discrete cosine transform on the sampling window are utilized for the DCT-based feature vectors, and the four dimensional feature vector of Fig. 15 was used as the edge-based feature vectors.  $w = 8$  and  $s = 4$  were employed for both feature vectors as the parameters of window scanning. Figure 34(a) shows that the recognition rate of the DCT-based feature vectors falls rapidly at  $\gamma = 1.3$  while the edge-based feature vectors performed over 99% recognition rate at the entire range. Figure 35 shows the results of the Viterbi segmentation on the two test images ( $\gamma = 1.0$  and  $\gamma = 2.0$ ). Each block separated by the white lines corresponds to each state of the pseudo-2D HMM illustrated in Fig. 14. The segmentation results using the edge-based features are almost the same independent of illumination conditions as shown in Fig. 35(a). On the other hand, Fig. 35(b) shows that the states to which the nose and the mouth belong do not match between the dark and bright images in case of the DCT-based feature vectors.

Figure 34(b) shows the recognition rate on the various size of the target faces. Although the recognition rate using the edge-based feature vectors (4 dimensions) is a little degraded as compared to that using the DCT-based ones, the recognition rate using the 16 dimensional edge-based feature vectors is almost same with that using the DCT-based ones. Over 99% recognition rate was obtained with the 16 dimensional edge-based vectors on the scale range between 70% ( $\approx \frac{1}{\sqrt{2}}$ ) and 140% ( $\approx \sqrt{2}$ ). The scale-invariant face detection system described in Section 3 is capable of enclosing facial images of any sizes in a frame in which the scale variation of the face image is limited within the range between  $\frac{1}{\sqrt{2}}$  and  $\sqrt{2}$ . Therefore, it would be possible to build a scale-invariant face identification system using our already-developed face detection system as a preprocessing stage in the present identification system.

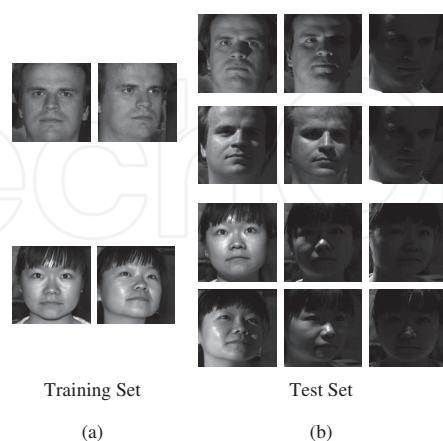


Fig. 36. Examples of facial images from Yale face database B used for training (a) and test (b).

The proposed system demonstrates a good performance. However, the results are obtained from the AT&T face database which consists of relatively simple images for identification. We also evaluated the performance on the Yale face database B (Georghades et al. (2001)). The



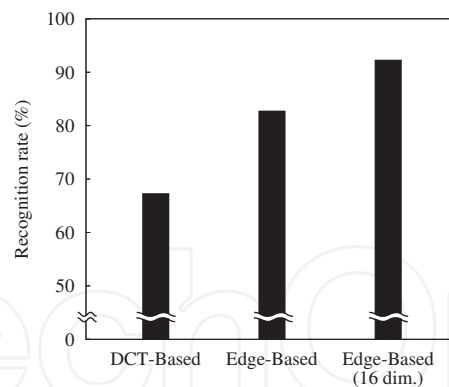


Fig. 37. Recognition rate on Yale face database B.

Yale face database B contains 5760 single light source images of 10 people each seen under 576 viewing conditions (9 poses  $\times$  64 illumination conditions). The face model of each person was learned from the nine poses with frontal illumination, and then the face identification is carried out for the images under various illumination conditions. Examples of facial images are shown in Fig. 36. Figure 37 illustrates the results of face identification. The recognition rate obtained with the proposed 16 dimensional edge-based feature vectors was 92% and better than that of the DCT-based feature vectors.

## 7. Summary

A face detection system has been developed aiming at exploring a pattern matching algorithm specifically adapted to the psychologically-inspired VLSI brain model system (Shibata (2002; 2007)). Namely, the computations involved in the algorithm are all very efficiently carried out by the dedicated VLSI chips developed for the system. Four directional edges are extracted from a  $64 \times 64$ -pixel image and two kinds of image feature vectors (PPED and APED) are generated by forming spatial distribution histograms from the edge maps. By the introduction of multiple-clue criterion using these two edge vectors, a face detection robust against scale variation and rotation has been developed. Although the total performance including the false positive rate has not yet reached those of advanced software systems, the algorithm has shown a lot of potentiality and issues for use in the VLSI hardware recognition system. The system has also been applied to the detection of posed faces and it has been shown that the performance for the profile face detection can be enhanced by generating edge-based vectors only from the region of focus-of-attention. Furthermore, the directional edge-based feature representations has been adapted to the HMM-based face identification system and the robust nature of the representations has been proven.

## 8. Acknowledgment

The authors would like to thank Daisuke Moriya at The University of Tokyo for his contribution to the work of posed face detection, and Prof. Kenji Takada and Dr. Masakazu Yagi at Osaka University for their support throughout the work.

## 9. References

Belongie, S., Malik, J. & Puzicha, J. (2002). Shape matching and object recognition using shape contexts, *IEEE Trans. Pattern Anal. Machine Intell.* 24(4): 509–522.

- Dalai, N., Triggs, B., Rhone-Alps, I. & Montbonnot, F. (2005). Histograms of oriented gradients for human detection, *Proc. of Computer Vision and Pattern Recognition (CVPR 2005)*, Vol. 1, San Diego, CA.
- Eickeler, S., Muller, S. & Rigoll, G. (2000). Recognition of JPEG compressed face images based on statistical methods, *Image and Vision Computing* 18: 279–287.
- Feraud, R., Bernier, O. J., Viallet, J. E. & Collobert, M. (2001). Fast and accurate face detector based on neural networks, *IEEE Trans. Pattern Anal. Machine Intell.* 23(1): 42–53.
- Freeman, W. & Roth, M. (1995). Orientation histograms for hand gesture recognition, *Proc. International Workshop on Automatic Face-and Gesture-Recognition*, Zurich, Switzerland, pp. 296–301.
- Gauthier, I., Behrmann, M. & Tarr, M. (1999). Can face recognition really be dissociated from object recognition?, *Journal of Cognitive Neuroscience* 11(4): 349–370.
- Georghiades, A., Belhumeur, P. & Kriegman, D. (2001). From few to many: Illumination cone models for face recognition under variable lighting and pose, *IEEE Trans. Pattern Anal. Machine Intell.* 23(6): 643–660.
- Hsu, R. L., Abdel-Mottaleb, M. & Jain, A. K. (2002). Face detection in color images, *IEEE Trans. Pattern Anal. Machine Intell.* 24(5): 696–706.
- Hubel, D. & Wiesel, T. (1959). Receptive fields of single neurons in the cat's striate cortex, *Journal of Physiology* 148(3): 574–591.
- Kang, K. & Shibata, T. (2009). An on-chip-trainable Gaussian-kernel analog support vector machine, *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2009)*, Taipei, Taiwan, pp. 2661–2664.
- Kuo, S. & Agazzi, O. (1994). Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models, *IEEE Trans. Pattern Anal. Machine Intell.* 16(8): 842–848.
- Liu, C. & Wechsler, H. (2002). Gabor feature base classification using the enhanced fisher linear discriminant model for face recognition, *IEEE Trans. Image Processing* 11(4): 467–476.
- Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision* 60(2): 91–110.
- McKenna, S. J., Raja, Y. & Gong, S. (1999). Tracking colour objects using adaptive mixture models, *Image and Vision Computing* 17: 225–231.
- Mikolajczyk, K. & Schmid, C. (2005). A performance evaluation of local descriptors, *IEEE Trans. Pattern Anal. Machine Intell.* 27(10): 1615–1630.
- Moscovitch, M., Winocur, G. & Behrmann, M. (1997). What is special about face recognition? nineteen experiments on a person with visual object agnosia and dyslexia but normal face recognition, *Journal of Cognitive Neuroscience* 9(5): 555–604.
- Nakagawa, T., Fujita, K. & Shibata, T. (2009). A real-time image feature vector generator employing functional cache memory for edge flags, *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS 2009)*, Taipei, Taiwan, pp. 3026–3029.
- Nefian, A. & Hayes, M. (1998). Hidden Markov models for face recognition, *Proc. Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP 1998)*, Vol. 5, pp. 2721–2724.
- Nefian, A. & Hayes, M. (1999). An embedded HMM-based approach for face detection and recognition, *Proc. Int'l Conf. Acoustics, Speech and Signal Processing (ICASSP 1999)*, Vol. 6, pp. 3553–3556.
- Ogawa, M., Ito, K. & Shibata, T. (2002). A general-purpose vector-quantization processor employing two-dimensional bit-propagating winner-take-all, *Digest of Technical Papers of 2002 Symposium on VLSI Circuits*, Honolulu, HI, pp. 244–247.

- Osuna, E., Freund, R. & Girosi, F. (1997). Training support vector machines: an application to face detection, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition 1997*, San Juan, Puerto Rico, pp. 130–136.
- Rowley, H., Baluja, S. & Kanade, T. (1998a). Neural network-based face detection, *IEEE Trans. Pattern Anal. Machine Intell.* 20(1): 23–38.
- Rowley, H., Baluja, S. & Kanade, T. (1998b). Rotation invariant neural network-based face detection, *Proc. IEEE International Conference on Computer Vision and Pattern Recognition 1998*, Santa Barbara, CA, pp. 38–44.
- Samaria, F. & Harter, A. (1994). Parameterisation of a stochastic model for human face identification, *Proc. 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, FL, pp. 138–142.
- Shibata, T. (2002). Intelligent signal processing based on a psychologically-inspired VLSI brain model, *IEICE Trans. Fundamentals* E85-A(3): 600–609.
- Shibata, T. (2007). A VLSI brain system mimicking the processing in the mind, *Proc. COE Symp. Advanced Electronics for Future Generations*, Tokyo, Japan, pp. 183–191.
- Shibata, T., Yagi, M. & Adachi, M. (1999). Soft-computing integrated circuits for intelligent information processing, *Proc. 2nd International Conference on Information Fusion*, Vol. 1, Sunnyvale, CA, pp. 648–656.
- Sung, K. & Poggio, T. (1998). Example-based learning for view-based human face detection, *IEEE Trans. Pattern Anal. Machine Intell.* 20(1): 39–51.
- Suzuki, Y. & Shibata, T. (2004). Multiple-clue face detection algorithm using edge-based feature vectors, *Proc. IEEE International Conference on Acoustic Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, Vol. 5, Montreal, Canada, pp. 737–740.
- Suzuki, Y. & Shibata, T. (2007). A hardware architecture for pseudo-2d hidden-Markov-model-based face recognition systems employing Laplace distribution functions, *Japanese Journal of Applied Physics* 46(4B).
- Takahashi, N., Fujita, K. & Shibata, T. (2009). A pixel-parallel self-similitude processing for multiple-resolution edge-filtering analog image sensors, *IEEE Trans. Circuits and Systems-I* 56(11): 2384–2392.
- Viola, P. & Jones, M. J. (2004). Robust real-time face detection, *International Journal of Computer Vision* 57(2): 137–154.
- Yagi, M., Adachi, M. & Shibata, T. (2000). A hardware-friendly soft-computing algorithm for image recognition, *Proc. European Signal Processing Conference (EUSIPCO 2000)*, Tampere, Finland, pp. 729–732.
- Yagi, M. & Shibata, T. (2002). Human-perception-like image recognition system based on the associative processor architecture, *Proc. European Signal Processing Conference (EUSIPCO 2002)*, Toulouse, France, pp. 103–106.
- Yagi, M. & Shibata, T. (2003). An image representation algorithm compatible with neural-associative-processor-based hardware recognition systems, *IEEE Trans. Neural Networks* 14(5): 1144–1161.
- Yamasaki, H. & Shibata, T. (2007). A real-time image-feature-extraction and vector-generation VLSI employing arrayed-shift-register architecture, *IEEE J. Solid-State Circuits* 42(9): 2046–2053.
- Yamasaki, T. & Shibata, T. (2003). Analog soft-pattern-matching classifier using floating-gate MOS technology, *IEEE Trans. Neural Networks* 14(5): 1257–1265.
- Yang, M. H., Kriegman, D. J. & Ahuja, N. (2002). Detecting faces in images: a survey, *IEEE Trans. Pattern Anal. Machine Intell.* 24(1): 34–58.



## **Image Fusion and Its Applications**

Edited by Dr. Yufeng Zheng

ISBN 978-953-307-182-4

Hard cover, 242 pages

**Publisher** InTech

**Published online** 24, June, 2011

**Published in print edition** June, 2011

The purpose of this book is to provide an overview of basic image fusion techniques and serve as an introduction to image fusion applications in variant fields. It is anticipated that it will be useful for research scientists to capture recent developments and to spark new ideas within the image fusion domain. With an emphasis on both the basic and advanced applications of image fusion, this 12-chapter book covers a number of unique concepts that have been graphically represented throughout to enhance readability, such as the wavelet-based image fusion introduced in chapter 2 and the 3D fusion that is proposed in Chapter 5. The remainder of the book focuses on the area application-orientated image fusions, which cover the areas of medical applications, remote sensing and GIS, material analysis, face detection, and plant water stress analysis.

### **How to reference**

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Yasufumi Suzuki and Tadashi Shibata (2011). A Directional-Edge-Based Face Detection Algorithm Adapted to the VLSI Image Recognition System, Image Fusion and Its Applications, Dr. Yufeng Zheng (Ed.), ISBN: 978-953-307-182-4, InTech, Available from: <http://www.intechopen.com/books/image-fusion-and-its-applications/a-directional-edge-based-face-detection-algorithm-adapted-to-the-vlsi-image-recognition-system>

**INTECH**  
open science | open minds

### **InTech Europe**

University Campus STeP Ri  
Slavka Krautzeka 83/A  
51000 Rijeka, Croatia  
Phone: +385 (51) 770 447  
Fax: +385 (51) 686 166  
[www.intechopen.com](http://www.intechopen.com)

### **InTech China**

Unit 405, Office Block, Hotel Equatorial Shanghai  
No.65, Yan An Road (West), Shanghai, 200040, China  
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元  
Phone: +86-21-62489820  
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](#), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen