Fast Face Detection via Morphology-based Pre-processing[1]

Chin-Chuan Han†, Hong-Yuan Mark Liao†[2], Kuo-Chung Yu‡, and Liang-Hua Chen§

†Institute of Information Science,

Academia Sinica, Nankang, Taipei, Taiwan

‡Institute of Computer Science and Information Engineering,

National Central University, Chung-Li, Taiwan

§Department of Computer Science and Information Engineering,

Fu Jen University, Taiwan

Abstract

An efficient face detection algorithm which can detect multiple faces in cluttered environment is proposed. The proposed system consists of three main steps. In the first step, a morphology-based technique is devised to perform eye-analogue segmentation. Morphological operations are applied to locate eye-analogue pixels in the original image. Then, a labeling process is executed to generate the eye-analogue segments. In the second step, the previously located eye-analogue segments are used as guides to search for potential face regions. The last step of the proposed system is to perform face verification. In this step, every face candidate obtained from the previous step is normalized to a standard size. Then, each of these normalized potential face images is fed into a trained backpropagation neural network for identification. After all the true faces are identified, their corresponding poses are located based on the guidance of optimizing a cost function. The proposed face detection technique may locate multiple faces oriented in any directions. Besides, the morphology-based eye-analogue segmentation process is able to reduce the background part of a cluttered image up to 95%. This process significantly speeds up the subsequent face detection procedure because only 5-10% regions of the original image are left for further processing. Experiments demonstrate that an approximately 94% success rate is reached and the relative false detection rate is very low.

[1]This work is supported by National Science Council of Taiwan under grant no. NSC86-2213-E-001-023.

[2]To whom correspondence should be addressed.

Key Words: Face Detection, Backpropagation Neural Network, Morphological Opening/Closing Operation

1 Introduction

Human face detection and recognition have long being a difficult research topic. In the last two decades, researchers have devoted much effort to these two problems and have received some satisfactory results. Part of these previous efforts were focused on face recognition[1, 2, 3, 4]. Turk and Pentland[4] have successfully employed the eigenface approach to recognize human face. However, an accurate and efficient method for human face detection is still deficient. In 1989, Govindaraju *et al.*[5, 6] presented a system which could locate human faces in photographs of newspapers, but the approximate size and expected number of faces must be known in advance. Sirohey[7] utilized an elliptical structure to segment the human heads from cluttered images. Yang and Huang[8, 9] utilized a three-level hierarchical knowledge-based method to locate the human faces in complex backgrounds. Sung and Poggio[10, 11, 12] applied two distance metrics to measure the distance between the input image and the cluster center. Twelve clusters including six face and six non-face clusters were trained by a modified *k-mean* clustering technique. A feature vector consisting of 12 values was inputted into a multi-layer perceptron network for the verification task. Pentland *et al.*[13, 14, 15, 16, 17, 18] applied principal component analysis(PCA) to describe face patterns with lower dimensional features. They designed a distance function called *distance from feature space*(dffs) as a metric to evaluate the difference between the input face image and the reconstructed face image. The system can be considered a special case of Sung and Poggio's system.

Face detection can also be achieved by detecting geometrical relationships among facial components such as nose, eyes, and mouth. Juell and Marsh[19] proposed a hierarchical neural network to detect the human face in gray scale images. An edge enhancing preprocessor and four backpropagation neural networks arranged in a hierarchical structure were devised to find multiple faces in the scene. Leung *et al.*[20, 21, 22] coupled a set of local feature detectors via a statistical model to find the facial components for face locating. Their apprach was invariant with respect to translation, rotation, and scale. Besides, they could also handle partial occlusion of faces. Instead of the gray scale images, Sobottka and Pitas[23], and Chen *et al.*[24, 25, 26, 27, 28] located the poses of human faces and facial features from color images. In [23], the oval shape of a face could be approximated by an ellipse in *Hue-Saturation-Value*(HSV) color space. Chen *et al.*[24, 25, 26, 27, 28] proposed a skin color distribution function on *perceptually uniform color space* to detect the face like region. The skin color regions in color images were modeled as several 2-D patterns and verified with the built face model by a fuzzy pattern matching approach.

Most of the above mentioned systems limit themselves to deal with human faces in frontal view. Namely, the orientation problem which is a potential trouble in this type of problem was not seriously considered. Besides, some previous approaches adopted a small window(20×20) to slide over all portions of an image at various scales. This brute force search is no doubt a time-consuming procedure. Jeng *et al.*[29] proposed a geometrical face model to solve the above mentioned problem. In their method, the geometrical face template was used to precisely locate the faces in a cluttered image. The average execution time for locating a face using a SPARC-20 machine was less than 5 seconds. However, the major drawbacks of their approach are: (1) the size of a face must be larger than 80×80 and,(2) the accuracy of using a geometrical face model may be significantly influenced due to the change of lighting conditions.

In this paper, an efficient face detection system is proposed. The proposed system consists of three main steps. In the first step, a morphology-based technique is devised to perform eye-analogue segmentation. Morphological operations are applied to locate eye-analogue pixels in the original image. Then, a labeling process is executed to generate the eye-analogue segments. In the second step, the previously located eye-analogue segments are used as guides to search for potential face regions. A face region is temporarily marked if a possible geometrical combination of eyes, noise, eyebrows, and mouth exists. Of course this step is a relatively loose process because we do not intend to miss any candidate faces. Therefore, this step may result in numerous number of face candidates, including both true faces and fake ones. The last step of the proposed system is to perform face verification. Face verification here includes identifying faces and locating their corresponding poses. In this step, every face candidate obtained from the previous step is normalized to a 20×20 image. Then, each of these normalized potential face images is fed into a trained backpropagation neural network for identification. After all the true faces are identified, their corresponding poses are located based on the guidance of optimizing a cost function. The proposed face detection technique may locate multiple faces oriented in any directions. Besides, the morphology-based eye-analogue segmentation process is able to reduce the background part of a cluttered image up to 95%. This process significantly speeds up the subsequent face detection procedure because only 5-10% regions of the original image are left for further processing. In the experiments, we use 122 cluttered images which contain totally 130 faces to test the effectiveness of the proposed method. 122 faces(approximately 94%) are successfully located and the false detection rate is very low. In average, our algorithm requires 20 seconds to finish the detection task on a 512×339 test image. Therefore, the proposed approach is indeed an efficient technique for face detection.

The rest of this paper is organized as follows. The extraction of eye-analogue region using morphological operations is described in Section 2. In Section 3, eye-pairs and normalized images are generated based on some geometric rules. A neural network-based verifier and a $dffs$ function are presented to verify the face region in Section 4. Some experimental results to show the validity of our detection system are demonstrated in Section 5. Finally, concluding remarks are given in Section 6.

2 Eye-Analogue Segmentation

As mentioned in [29, 30], eyebrows, eyes, nostrils and mouth always look darker than the rest part of a face. Among these facial features, the degree of darkness of eyebrows is dependent on the density or color, and that of the nostrils relies on how they are occluded. As to the shape of a mouth, it is frequently displayed in various appearances and sizes due to facial expressions. In comparison with the above mentioned unstable facial features, the eyes can be considered a salient and relatively stable feature on a face. Therefore, in this section we propose a morphology-based method to locate the eye-analogue segments as the first step of our system. The eye-analogue pixels in a cluttered image are segmented first. Then, the small segments of eye-analogue regions are grouped together by conditional morphological dilation operations and labeled by a traditional labeling process.

In 1993, Chow and Li[30] employed a morphological *opening residue* operation to extract the intensity valleys as potential eye-analogue pixels via a 5×5 circle structuring element. In our approach, we apply the morphological closing and clipped different operations to find the candidate pixels of eye-analogue. Let X be the original image and $X_{1/2}$ be the half image with scale 0.5. A horizontal structuring element S_h with size 1×7 and a vertical structuring element S_v with size 7×1 are, respectively, operated on X and $X_{1/2}$. It is known that[30] the eye-analogue pixels are located at the intensity valleys whose size are smaller than a preset value. Therefore, by mixing the closing ($\bullet$), clipped different ($\biguplus$), thresholding (T), and OR ($\vee$) into different operations, the eye-analogue pixels in an image can be located. The set of operations which are able to identify the eye-analogue pixels are expressed as follows:

$$\begin{aligned}
E_1 &= T_1\left(X \biguplus (X \bullet S_h)\right), \\
E_2 &= T_2\left(X_{1/2} \biguplus (X_{1/2} \bullet S_h)\right), \\
E_3 &= T_3\left(X \biguplus (X \bullet S_v)\right), \\
E_4 &= T_4\left(X_{1/2} \biguplus (X_{1/2} \bullet S_v)\right), \\
E &= E_1 \vee E_2^2 \vee E_3 \vee E_4^2, \qquad (1)
\end{aligned}$$

where the superscript, 2, of E_2 and E_4 are used to enlarge them into twice of the original size. T_1, T_2, T_3, and T_4 are four threshold functions whose values are the average values of the images E_1, E_2, E_3 and E_4, respectively.

As mentioned in the previous paragraph, the eye-analogue pixels are located at the intensity valleys. However, when the scene is complex, some pixels of this kind may not be the eye-analogue pixels. For instance, text characters in a notice board, frames of windows, edges among a set of text books, ..., etc., are frequently segmented as the eye-analogue pixels via operations in Eq. (1). Basically, these pixels can be considered noises in facial images. Therefore, we apply the conditional morphological dilation operation at this stage to remove these unwanted pixels in the background. Then, a conventional labeling process is performed to locate the eye-analogue segments. The eye-analogue detection algorithm is described in detail as follows.

Eye-analogue Detection Algorithm:

Step 1 : Perform labeling process on image E and compute a set of geometrical data from each segment including the lengths of the major and minor axes, the orientation, the center point, and the minimal bounding rectangle.

Step 2 : If the length of the major axis of segment i is larger than $0.6N$(where N is the smallest width of a face region), terminate the conditional dilation operation for segment i and eliminate segment i from image E. Otherwise, go to next step.

Step 3 : Perform conditional dilation operation on every segment i by the structuring element $SE = \{1_{(x,y)} | (x, y) \in Z\}$, where

1. if segment i is a horizontal segment, i.e., the orientation of segment i is located at $[-\frac{\pi}{8}$ to $\frac{\pi}{8}]$, the structuring element SE is assigned as $\{1_{(-1,0)}, 1_{(0,0)}, 1_{(1,0)}\}$,
2. if segment i is a left slant segment, i.e., the orientation of segment i is located at $(-\frac{\pi}{8}$ to $-\frac{3\pi}{8}]$, choose the element $\{1_{(1,-1)}, 1_{(0,0)}, 1_{(-1,1)}, 1_{(-1,0)}, 1_{(1,0)}\}$ as the structuring element SE,
3. if segment i is a right slant segment, i.e., the orientation of segment i is located at $(\frac{\pi}{8}$ to $\frac{3\pi}{8}]$, the structuring element SE is defined as $\{1_{(-1,-1)}, 1_{(0,0)}, 1_{(1,1)}, 1_{(-1,0)}, 1_{(1,0)}\}$,
4. if segment i is a vertical segment, i.e., the orientation of segment i is located at $(-\frac{3\pi}{8}$ to $-\frac{4\pi}{8}]$ or $(\frac{3\pi}{8}$ to $\frac{4\pi}{8}]$, choose the element $\{1_{(0,-1)}, 1_{(0,0)}, 1_{(0,1)}\}$ as the structuring element SE.

Step 4 : Repeat steps 1 to 3 for $N/5$ times.

An example demonstrating the eye-analogue segmentation process is shown in Fig. 1. Fig. 1(a) is a gray scale image. After the morphological closing and clipped different operations, the potential eye-analogue pixels and regions are shown in Fig. 1(b). After performing the conditional dilation and a labeling process on the potential eye-analogue segments in Fig. 1(b), Fig. 1(c) shows the located eye-analogue segments with bounding rectangles. One thing to be noted is that the text characters in the notice board(Fig. 1(b)) are merged together to form fewer segments. The main advantage of the two above mentioned processes is to combine those non-eye analogue segments whose sizes are small. These processes will reduce the number of potential eye-analogue segments.

3 Finding Potential Eye-pairs from Eye-analogue Segments

As stated in Section 2, the eyes are considered the primary features in a face image. After performing the segmentation process, each eye-analogue segment can be considered a candidate of one of the eyes. In this section, we propose four matching rules to guide the merging of potential eye-analogue segments into pairs. Then, based on these potential eye pairs, the potential poses of the face regions are determined.

Suppose that there are M eye-analogue segments in the image E, there are at most M^2 potential combinations of face regions generated via the matched eye pairs. In order to reduce the execution time, the geometrical constraint between a real eye pair is introduced to screen some impossible pairs. In [19], Juell and Marsh use a 19×11 window to find the location of eyes. They assume the ratio between the width and the height of an eye is roughly about 2. In what follows, a set of rules is proposed to merge those previously desired eye-analogue segments into potential eye pairs.

Matching Rules

(a) *The length ratio between the major and the minor axes of segment i must be smaller than 10.* Here, the threshold value of the length ratio is set to 10 and should be larger than 2 for tolerating segmentation errors.

(b) *The distance between the center points of two eye-analogue segments must be larger than $0.6N$.* Here, the value 0.6 is a rough estimate of the ratio between the distance between the two pupils of eyes and the width of a face, and N is the shortest width of a face. This value can be derived from the training samples.

(c) *Each eye must be located by extending a small range from the other eye* as shown in Fig. 2(a). Since two eyes can be located through the assistance of their base line, they will be in co-linear direction even the face is rotated.

(d) *The area of segment i should be larger than 10 pixels.*

Once the potential eye pairs are determined, their corresponding face regions can be easily extracted and each of these faces is normalized into the 20×20 standard size. As shown in Fig. 2(b), (x_i, y_i) and (x_j, y_j) are two center points of segment i and j, respectively. $(x_1, y_1), (x_2, y_2), (x_3, y_3)$, and (x_4, y_4) are four corner points of a normalized face region. Let $x_i + x_j = A$, $x_i - x_j = B$, and $y_i - y_j = C$. The the coordinates of the four corner points can be calculated as follows.

$$
\begin{aligned}
x_1 &= \frac{1}{2}A - \frac{c_2}{c_1}B + \frac{c_3}{c_1}C, \\
y_1 &= \frac{1}{2}A - \frac{c_2}{c_1}C - \frac{c_3}{c_1}B, \\
x_2 &= \frac{1}{2}A + \frac{c_2}{c_1}B + \frac{c_3}{c_1}C, \\
y_2 &= \frac{1}{2}A + \frac{c_2}{c_1}C - \frac{c_3}{c_1}B, \\
x_3 &= \frac{1}{2}A - \frac{c_2}{c_1}B + \frac{c_4}{c_1}C, \\
y_3 &= \frac{1}{2}A - \frac{c_2}{c_1}C - \frac{c_4}{c_1}B, \\
x_4 &= \frac{1}{2}A + \frac{c_2}{c_1}B + \frac{c_4}{c_1}C, \\
y_4 &= \frac{1}{2}A + \frac{c_2}{c_1}C - \frac{c_4}{c_1}B, \qquad (2)
\end{aligned}
$$

where c_1 is the average distance between two pupils of the eyes and c_2 is one half of the width of a face region. Besides, c_3 and c_4 are, respectively, the distances from the base line of two eyes to the top and bottom boundaries of a face region(see Fig. 2(b).) From the training samples used in the experiments, c_1, c_2, c_3, and c_4 are, respectively, 12.5, 10, 4, and 16 in a normalized 20×20 face image. Based on these geometrical relations, a face region can be easily rotated and normalized. In Fig. 2(c), the potential eye pairs which satisfy the matching rules are linked via solid line segments. According to Eq. (2), the potential face regions which cover the potential eye pairs in Fig. 2(c) are extracted and normalized to 20×20 standard size. These normalized potential face regions are shown in Fig. 2(d).

4 Face Verification

In the previous section, we have selected a set of potential face regions in an image. These potential face regions are allowed to have different sizes and orientations. However, after the normalization process, all potential faces are normalized to a fixed size, i.e., 20×20, and rotated into frontal position. In this section, we propose a coarse-to-fine face verification process to locate the real positions of faces in an image. In coarse verification, a trained backpropagation neural network[31] is applied to decide whether a potential region contains a face. Before the neural network is applied to execute coarse verification, a preprocessing step which is similar to the work of Sung and Poggio[10], and Rowley *et al.*[31] has to be performed first. The preprocessing step consists of masking, illumination gradient correction, and histogram equalization. The trained backpropagation net is then applied to filter out those non-face regions. If the neural network outputs a positive response, the potential region may contain a face. Otherwise, this region will be eliminated from the potential candidate face list. Basically, this step can filter out part of the non-face regions and we thus call it coarse verification. In the fine verification process, we apply a cost function, $dffs$, proposed by Moghaddam and Pentland[16, 17] to perform final selection. In what follows, the detailed procedure, from coarse to fine, will be described.

To train the neural networks, a technique similar to the one reported in [31] is adopted. A set of 11200 face images generated from 700 face samples are collected as the positive samples by randomly, slightly rotating(up to 10°), scaling (90% to 110%), translating(up to half a pixel), and mirroring. As mentioned by Rowley *et al.*[31], it is difficult to characterize the prototype of non-face images because of the huge size of non-face images. Therefore, the network training task on non-face images is basically a difficult one. Sung and Poggio's work[10] conformed this point. They collected 4000 positive(face) samples and 47500 negative(non-face) samples to train their network. It is obvious that the latter is much larger than the former. For reducing the number of non-face training samples, Rowley *et al.*[31] used the bootstrap algorithm to train the network. In their experiment, 16000 positive and only 9000 negative face samples were used. For the non-face training samples, they repeated a bootstrap algorithm to collect the data. Therefore, it requires significant computation time spent on training. Here, we modify their bootstrap algorithm as follows.

1. Create an initial set of non-face images by generating 1000 images with random intensities. Apply the pre-processing steps to each of these images. Initially, the network's weights are randomly initialized. Train a neural network to produce an output of 1 for the face samples, and -1 for the non-face samples.

2. Run the system on 20 scenery images which contain no faces. Collect the subimages that the network incorrectly identifies as faces. Select 250 of these subimages as non-face samples.

3. Apply the preprocessing steps on the collected face and non-face samples. Then, retrain the network to obtain a new version of face verifier.

The major difference between the training process of ours and Rowley's is the non-face sample collection process. In our system, 5000 non-face images are collected from 20 scenery images as negative samples. Therefore, only one training phase is needed in our scheme. In the recall process, if the trained neural network outputs a positive response, it means that there exists a face in the region under checking. If a negative response is received, the region under checking is considered a non-face region. The coarse verification process using neural networks filters out significant number of non-face regions. However, part of the non-face regions are retained due to the loose constraints used in the coarse verification process. Besides, some overlapping candidate faces are retained simultaneously and are therefore required to be further checked. In the fine verification process, an evaluation function proposed by Moghaddam and Pentland[16, 17] is applied to eliminate the above mentioned overlapping detections as well as those previously retained non-face regions. The evaluation function is called the residual reconstruction error which is defined as follows:

$$\epsilon^2 = \sum_{i=M+1}^{N} y_i^2 = ||x - \overline{x}||^2 - \sum_{i=1}^{M} y_i^2, \tag{3}$$

where N is the size of an image x which is to be checked, M is the number of principal components used to reconstruct the original images. Since the $dffs$ value denotes the difference between the input image and the reconstructed image via PCA, one has to choose the region with the positive response and the local smallest $dffs$ value as a face region. Once a face region is confirmed, the last step is to eliminate those regions that overlap with the chosen face region.

5 Experimental Results and Discussion

A set of experimental results are demonstrated to show the effectiveness and efficiency of the proposed system. 122 test images containing totally 130 faces were used to test the validity of our system. All test images were of size 512×339. In these test images, all human faces were oriented into various directions and positioned arbitrarily in cluttered background. In this research, the minimum size of a face which could be detected was 50×50.

Figs. 3(a)-(j) show 10 test images which have correct detection results. The bounding rectangle that bounds a face region is used to justify whether the detection is correct or not. Among the successful cases, Fig. 3(b) demonstrated that the proposed method could detect faces with facial expression. The case shown in Fig. 3(c) reflected that the illumination problem could also be solved by our approach. One thing worth noticing is that the test image shown in Fig. 3(f) contained two faces with a nearly 180 degree difference, it is obvious that our system worked perfectly in dealing with this kind of problem. For an overall evaluation, 122 faces were detected successfully out of the total of 130 faces. Therefore, the success rate for detecting face was roughly about 94%. On the other hand, the proposed system also detected totally 25 fake faces from the cluttered backgrounds of the test images. Fig. 4 shows some unsuccessful detections including fake faces(the left bounding rectangle of Fig. 4(a)), missed faces(the right-hand side face of Fig. 4(b)), and inclinedly detected faces(the right bounding rectangle of Fig. 4(a)). There are three possible causes to generate the mis-detection problem. First, the size of a face to be detected has to be larger than 50×50. If the size of a face region is smaller than 50×50 as shown in Fig. 4(a), the system failed to detect it. The second cause that leads to mis-detection is missing eye pairs. If a potential eye pair which corresponds to a true face is not found, it is impossible to further locate the exact pose of this face. Fortunately, the above mentioned situation does not happen very often. In our experiments, only the mis-detection of the face in the right-hand side of Fig. 4(b) was due to this reason. There is another reason to generate mis-detection problem. As shown in Fig. 4(c), a partially occluded face was mis-detected. In the future, we shall extend the capability of the system to deal with partially occluded faces as well as to handle faces with sizes smaller than 50×50.

As to the execution time problem, the time required to locate the precise locations of the faces in the test image set is dependent upon the size and complexity of images. For example, a 512×339 image shown in Fig. 3(e) needed less than 18 seconds to locate the correct face position using a SPARC 20 SUN workstation. As to the case shown in Fig. 3(f), the execution time under the same environment was about 23 seconds.

6 Concluding Remarks

In this paper, we have proposed an efficient face detection algorithm to find multiple faces in cluttered images. In the first stage, morphological operations and labeling process were performed to obtain the eye-analogue segments. Based on some matching rules and the geometrical relationship on a face, eye-analogue segments were grouped into pairs and used to locate potential face regions. Finally, the potential face regions were verified via a trained neural network and the true faces were determined by optimizing a distance function. Since the morphology-based eye-analogue segmentation process can efficiently locate the potential eye-analogue regions, the subsequent processing only has to deal with 5-10% area of the original image. Therefore, the execution time is much less than most of the existing systems. Besides, the proposed system can detect faces in arbitrary orientations and scales as long as they are larger than 50×50.

References

[1] M. A. Shackleton and W. J. Welsh, "Classification of facial features for recognition", in *Proc. Computer Vision and Pattern Recognition*, Lahaina, Maui, Hawaii, Jun. 1991, pp. 573–579.

[2] B. S. Manjunath, "A feature based approach to face recognition", in *Proc. Computer Vision and Pattern Recognition*, Champaign, Illinois, Jun. 1992, pp. 373–378.

[3] R. Brunelli and T. Poggio, "Face recognition: Features verus templates", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 1042–1052, 1993.

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

[5] V. Govindaraju, D. B. Sher, and R. K. Srihari, "Locating human face in newspaper photographs", in *Proc. Computer Vision and Pattern Recognition*, San Diego, California, Jun. 1989, pp. 549–554.

[6] V. Govindaraju, S. N. Srihari, and D. B. Sher, "A computational model for face location", in *Proc. Computer Vision and Pattern Recognition*, 1990, pp. 718–721.

[7] S. A. Sirohey, "Human face segmentation and identification", Master's thesis, University of Maryland, 1993.

[8] G. Yang and T. S. Huang, "Human face detection in a scene", in *Proc. Computer Vision and Pattern Recognition.* New York, Jun. 1993, pp. 453–458.

[9] G. Yang and T. S. Huang, "Human face detection in a complex background", *Pattern Recognition*, vol. 27, no. 1, pp. 53–63, 1994.

[10] K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection", Tech. Rep., Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Dec. 1994.

[11] K. K. Sung and T. Poggio, "Example-based learning for view-based human face detection", in *Proc. Image Understanding Workshop*, Monterey, Calif., Nov. 1994, pp. 843–850.

[12] K. K. Sung, *Learning and Example Selection for Object and Pattern Detection*, PhD thesis, MIT, 1996, anonymous ftp from publications.ai.mit.edu.

[13] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modular eigenspaces for face recognition", in *IEEE Conference on Computer Vision and Pattern Recognition*, Seattle, Washington, Jun. 1994, pp. 84–91.

[14] B. Moghaddam and A. Pentland, "Face recognition using view-based and modular eigenspaces", in *Proc. Automatic Systems for the Identification and Inspection of Humans, SPIE*, Jul. 1994.

[15] B. Moghaddam and A. Pentland, "Probabilistic visual learning for object detection", in *Proc. 5th International Conference on Computer Vision*, Cambridge, Mass., Jun. 1995, pp. 786–793.

[16] B. Moghaddam and A. Pentland, "An automatic system for model-based coding of faces", in *Proc. IEEE Data Compression Conference*, Snowbird, Utah, Mar. 1995.

[17] B. Moghaddam and A. Pentland, "A subspace method for maximum likelihood target detection", in *Proc. IEEE International Conference on Image Processing*, Washington DC, Oct. 1995.

[18] T. Darrel, B. Moghaddam, and A. P. Pentland, "Active face tracking and pose estimation in an interactive room", in *Proc. Computer Vision and Pattern Recognition*, San Francisco, California, Jun. 1996, pp. 67–72.

[19] P. Juell and R. Marsh, "A hierarchical neural network for human face detection", *Pattern Recognition*, vol. 29, no. 5, pp. 781–787, 1996.

[20] T. K. Leung, M. C. Burl, and P. Perona, "Finding faces in clustered scenes using random labeled graph matching", in *Proc. Computer Vision and Pattern Recognition*, Cambridge, Mass., Jun. 1995, pp. 637–644.

[21] M. C. Burl, T. K. Leung, and P. Perona, "Face localization via shape statistics", in *Proc. Internal Workshop on Automatic Face and Gesture Recognition*, 1995.

[22] M. C. Burl and P. Perona, "Recognition of planar object classes", in *Proc. Computer Vision and Pattern Recognition*, San Francisco, California, Jun. 1996, pp. 223–230.

[23] K. Sobottka and I. Pitas, "Extraction of facial regions and features using color and shape information", in *Proc. 13th International Conference on Pattern Recognition*, Vienna, Austria, Aug. 1996, pp. 421–425.

[24] H. Wu, Q. Chen, and M. Yachida, "An application of fuzzy theory: Face detection", in *Proc. International Workshop on Automatic Face and Gesture Recognition*, Zurich, Switzerland, Jun. 1995, pp. 26–28.

[25] Q. Chen, H. Wu, and M. Yachida, "Face detection by fuzzy pattern matching", in *Proc. Computer Vision and Pattern Recognition*, Cambridge, Mass., Jun. 1995, pp. 591–595.

[26] Q. Chen, H. Wu, and K. Chihara, "Real-time face detection", in *Proc. ACCV'95 Second Asian Conference on Computer Vision*, Singapore, Dec. 1995, pp. 479–483.

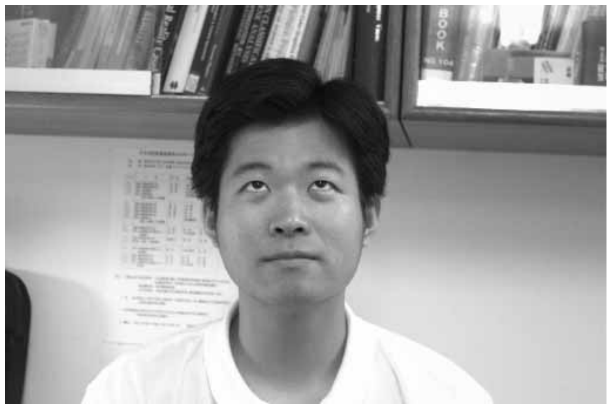
[27] H. Wu, Q. Chen, and M. Yachida, "A fuzzy-theory-based face detector", in *Proc. 13th International Conference on Pattern Recognition*, Vienna, Austria, Aug. 1996, pp. 406–410.

[28] H. Wu, Q. Chen, and M. Yachida, "Facial feature extraction and face recognition", in *Proc. 13th International Conference on Pattern Recognition*, Vienna, Austria, Aug. 1996, pp. 484–488.

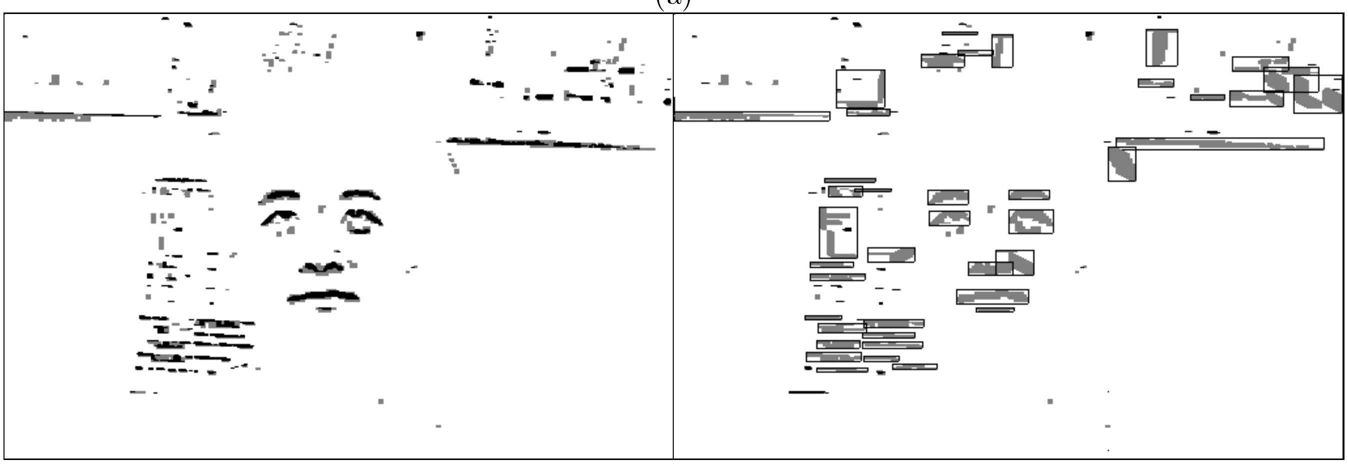
[29] S. H. Jeng, H. Y. Mark Liao, Y. T. Liu, and M. Y. Chern, "An efficient approach for facial feature detection using geometrical face model", in *Proc. 13th International Conference on Pattern Recognition*, Vienna, Austria, Aug. 1996, pp. 426–430.

[30] G. Chow and X. Li, "Towards a system for automatic facial feature detection", *Pattern Recognition*, vol. 26, no. 12, pp. 1739–1755, 1993.

[31] H. A. Rowley, S. Baluja, and T. Kanade, "Human face detection in visual scenes", Tech. Rep., Carnegie Mellon University, 1995.



(a)



(b) (c)

Figure 1: The eye-analogue segmentation process. (a) The original image, (b) the eye-analogue pixels and segments after the closing and clipped different operations, (c) the eye-analogue segments after the conditional dilation and labeling processes.

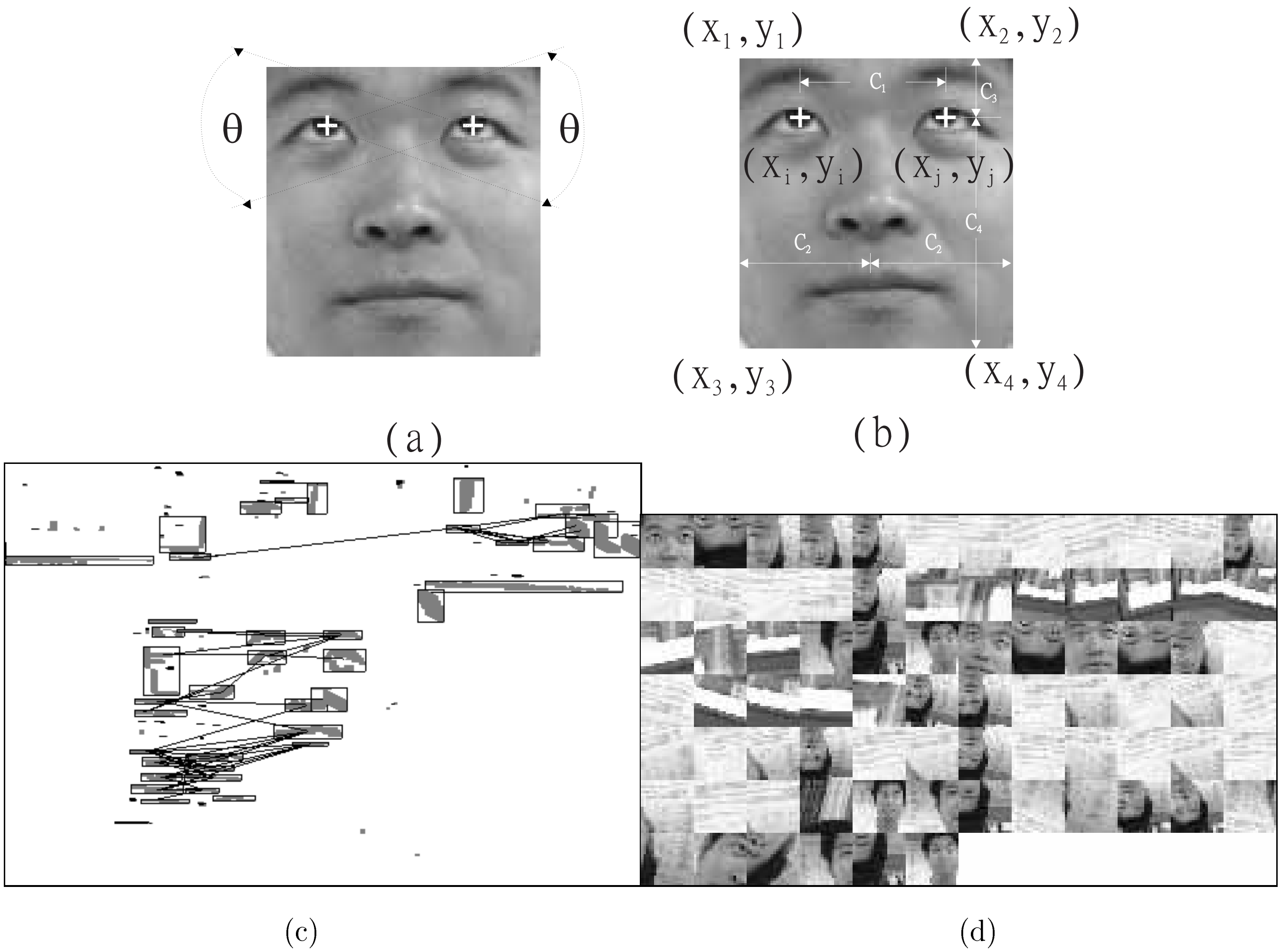
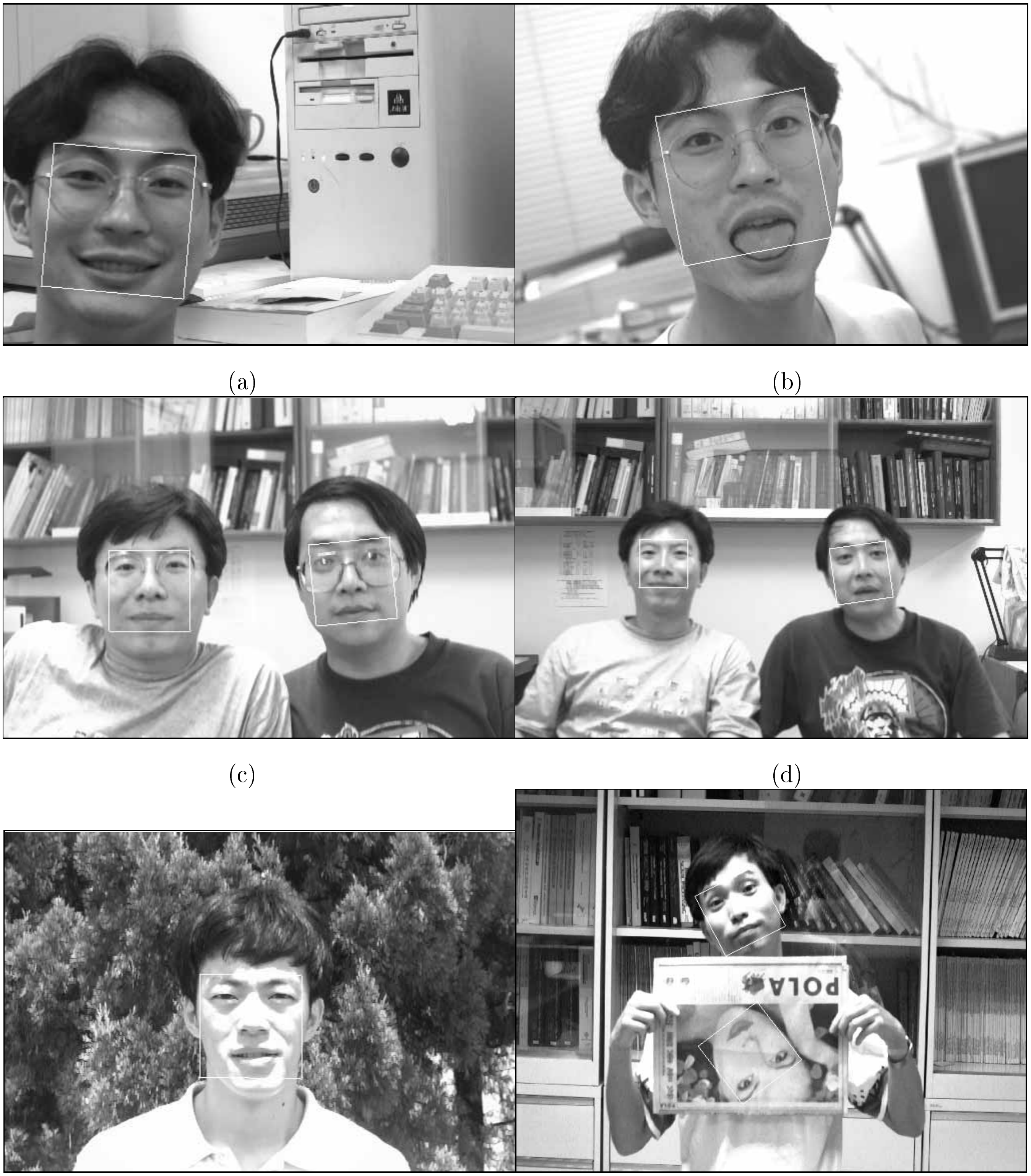


Figure 2: The extraction of face region. (a) The matching rule (c), (b) the geometrical relationship of face region, (c) the potential eye pairs of Fig. 1(c), (d) the potential face regions.

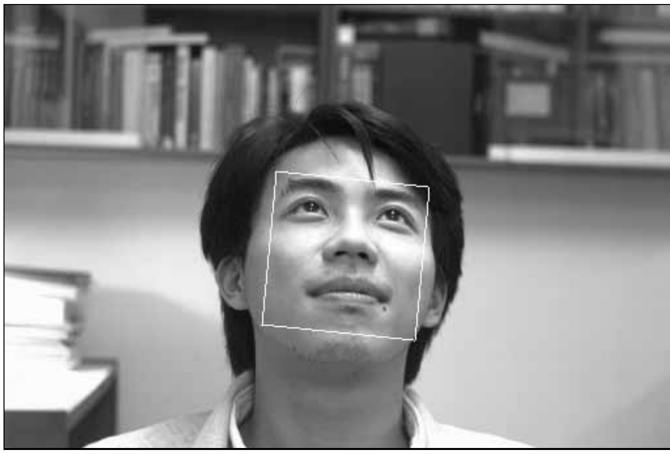


(a) (b)

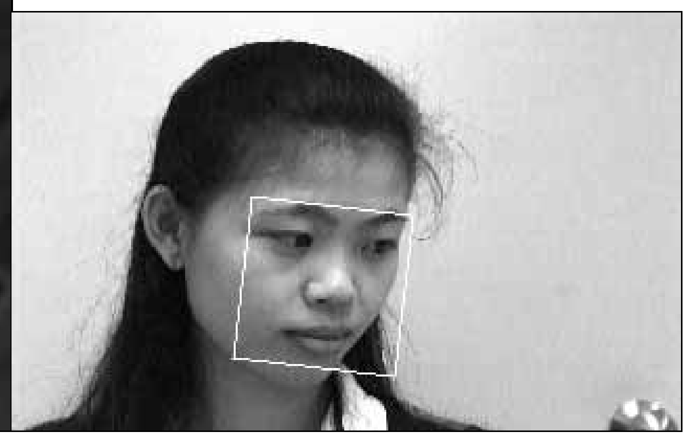
(c) (d)

(e) (f)

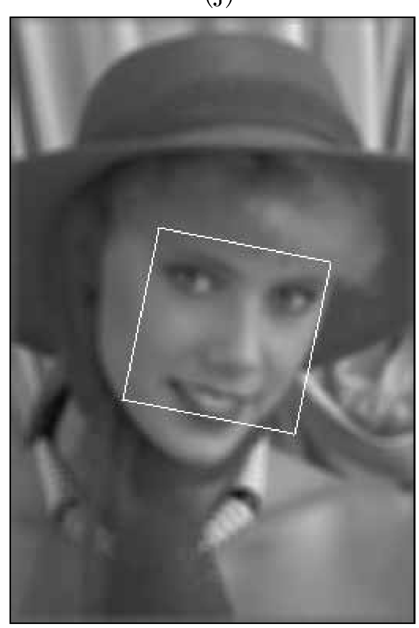
Figure 3: Testing examples.



(g) (h)

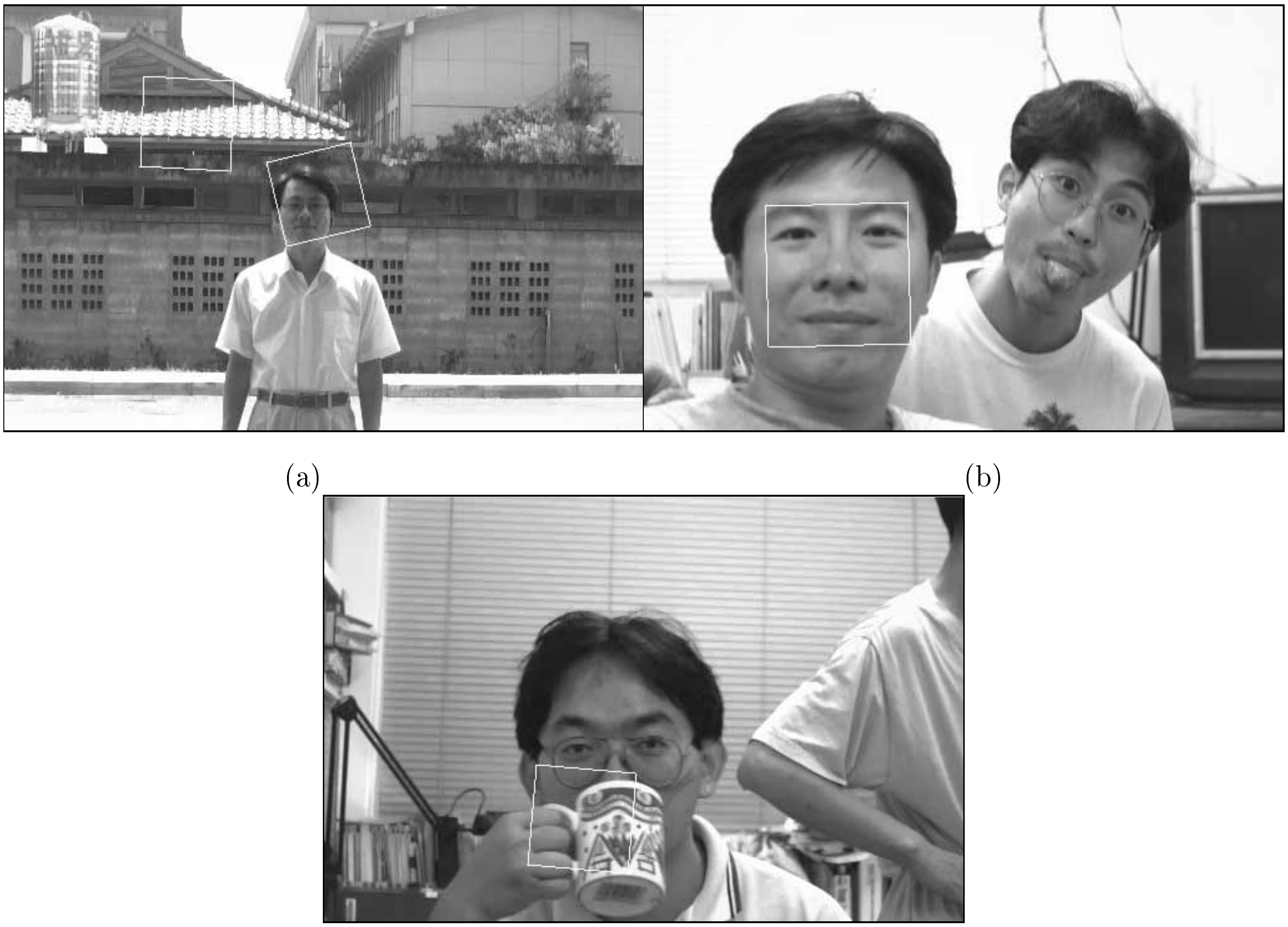


(i) (j)



(k) (l)

Figure 3: Testing examples(Continued).



(a) (b)

(c)

Figure 4: The mis-detected examples.