Face Recognition Using A Face-Only Database: A New Approach

Hong-Yuan Mark Liao†, **Chin-Chuan Han**†, **Gwo-Jong Yu**‡,
Hsiao-Rong Tyan§, **Meng Chang Chen**†, and **Liang-Hua Chen**¶

†Institute of Information Science,
Academia Sinica, Taipei, Taiwan
‡Institute of Computer Science and Information Engineering,
National Central University, Chung-Li, Taiwan
§Department of Information and Computer Engineering,
Chung Yuan Christian University, Taiwan
¶Department of Computer Science and Information Engineering,
Fu Jen University, Taiwan

Corresponding author's address –
Hong-Yuan Mark Liao
Associate Research Fellow
Institute of Information Science,
Academia Sinica, Nankang, Taipei, Taiwan
E-mail : liao@iis.sinica.edu.tw
TEL : 886-2-788-3799 ext. 1811
FAX : 886-2-782-4814

Face Recognition Using A Face-Only Database: A New Approach[1]

Hong-Yuan Mark Liao†[2], **Chin-Chuan Han**†, **Gwo-Jong Yu**‡,
Hsiao-Rong Tyan§, **Meng Chang Chen**†, and **Liang-Hua Chen**¶
†Institute of Information Science,
Academia Sinica, Taipei, Taiwan
‡Institute of Computer Science and Information Engineering,
National Central University, Chung-Li, Taiwan
§Department of Information and Computer Engineering,
Chung Yuan Christian University, Taiwan
¶Department of Computer Science and Information Engineering,
Fu Jen University, Taiwan
E-mail : {liao,cchan,yugj}@iis.sinica.edu.tw

Abstract

In this paper, a coarse-to-fine, LDA-based face recognition system is proposed. Through careful implementation, we found that the databases adopted by two state-of-the-art face recognition systems [4, 5] were incorrect because they mistakenly use some none-face portions for face recognition. Hence, a face-only database is used in the proposed system. Since the facial organs on a human face only differ slightly from person to person, the decision-boundary determination process is tougher in this system than it is in conventional approaches. Therefore, in order to avoid the above mentioned ambiguity problem, we propose to retrieve a closest subset of database samples instead of retrieving a single sample. The proposed face recognition system has several advantages. First, the system is able to deal with a very large database and can thus provide a basis for efficient search. Second, due to its design nature, the system can handle the defocus and

[1]This work was partially supported by National Science Council of Taiwan under grants NSC86-2745-E-001-004 and NSC86-2213-E-001-023.

[2]To whom correspondence should be addressed.

noise problems. Third, the system is faster than the autocorrelation plus LDA approach [4] and the PCA plus LDA approach [5], which are believed to be two statistics-based, state-of-the-art face recognition systems. Experimental results prove that the proposed method is better than traditional methods in terms of efficiency and accuracy.

1 Introduction

Face recognition has been a very hot research topic in recent years[1, 2, 3]. Some successful face recognition systems have been developed and reported in the literature[4, 5, 6, 7, 8]. Among these works, the systems developed by Goudail *et al.*[4] and Swets and Weng[5] represent two state-of-the-art face recognition systems. In [4], Goudail *et al.* investigated the performance of a technique for face recognition based on the computation of 25 local autocorrelation coefficients. They used a large database of 11,600 frontal facial images of 116 persons, organized in training and test sets, for evaluation. For database construction, they asked all the persons to wear dark company jackets and to sit down in front of a uniform, black background. They recorded the head part of the persons over two 30-second periods on videotape using a CCD camera. In the autocorrelation coefficient computation stage, they used 25 3x3 filters to convolve with the whole image and came up with 25 numerical coefficients. This process was used to reduce the dimensionality of the raw images. They used the set of transformed 25-dimensional database samples to determine the set of most discriminating projection axes and then calculated each sample's projective feature vector. When an unknown image appeared, its corresponding projective feature vector was calculated and compared with those of the database samples. Without rejection of unknown faces, they obtained a peak recognition rate of 99.9%. The advantage of this approach is its speed, especially when the size of the database is small. There are several drawbacks associated with this approach, however. First, they used an unorganized database. When the size of the database is huge, the sequential search nature of this approach leads to degradation of search efficiency. Second, the dimensionality reduction process is global in nature. This makes the approach very sensitive to blurred and noisy images. Third, this approach is worse than the principal component analysis(PCA)[9] plus linear discriminant analysis(LDA)[6, 7, 8, 10] based approach[9] in terms of accuracy. Further, although they keep the color of the background and cloth dark, their "face" image is actually a combination of face, hair, shoulders and background. It was proven in [11] that this kind of database is incorrect in terms of "face" recognition.

Another state-of-the-art system was proposed by Swets and Weng[5]. In this work, they applied the PCA technique to reduce the dimensionality of the original images. They selected the top 15 principal axes and used them to derive a 15-dimensional feature vector for every sample. These transformed samples are then used as bases to execute LDA, and they reported a peak recognition rate of more than 90%. Again, we found that their face image contained face, hair, shoulders and background, not solely face. The major advantage of this approach is its accuracy. However, it also suffers from several drawbacks. First, the PCA-based dimensionality reduction process requires intensive computation, which is mostly consumed in the principal axes determination process. Second, when new samples come in, the original set of principal axes has to be re-computed.

In this paper, we propose a coarse-to-fine, LDA-based approach to solve the above mentioned problems. In order to make the recognition process base its recognition on the "real" face portion, we build a face database with every sample containing only facial organs. In the feature extraction process, each sample image is equally partitioned into 2^{2i} regions at resolution $i (i = 1, 2, ...)$. The value of a partitioned region is obtained by averaging the gray values of all pixels in that region. Hence, the coarse-to-fine feature extraction process transforms all the database samples into 2^{2i}-dimensional feature vectors at resolution i (i=1,2,...). Based on the extracted feature vectors, LDA can be executed at different resolutions, and the 2^{2i} most discriminating projection axes can be determined at resolutions $i(i = 1, 2, ...)$. For resolution i, the transformed 2^{2i}-dimensional database samples are projected onto their corresponding most discriminating projection axes, and a set of 2^{2i}-dimensional projective feature vectors of database samples is obtained. When an unknown image comes in, the same coarse-to-fine feature extraction process is performed. Then, the set of different dimensional feature vectors is projected onto the most discriminating projection axes corresponding to their own resolutions, and a set of different dimensional projective feature vectors for the unknown image is obtained. Based on these projective feature vectors, a coarse-to-fine comparison can be executed, and the final results will be obtained at the finest resolution. There are several advantages associated with the proposed approach. First, the coarse-to-fine search structure can help deal with a very large database and provide a good basis for efficient search. Second, using an average gray level to represent a region can suppress the effect caused by defocus or noise. Third, the design nature makes the database extensible. Further, the search speed of this approach is faster than that of the approaches developed by Goudail *et al.*[4] and Swets and Weng[5]. As to accuracy, this approach is better than that of Goudail *et al.*. Finally and most importantly, our approach uses a face-only database to execute the face recognition task. Since the facial organs on a human face differ only slightly from person to person, the decision boundary set is hard to decide upon. However, experimental results have shown that our approach is still robust under these circumstances. The rest of this paper is organized as follows. In the next section, we shall discuss why a face-only database is adopted as our face database. A series of experiments will be described which support our point. Section 3 will cover the questions of how to construct a face-only database and how to perform coarse-to-fine feature extraction. Then, classification based on a coarse-to-fine LDA-based strategy will be described in Section 4. Section 5 will report a series of experimental results. Finally, conclusions and future work will be addressed in Section 6.

2 Face + Hair + Shoulders + Background $\neq$ Face

In this section, we shall describe a series of experiments and show that the databases used in [4] and [5] were incorrect. In order to prove that our above statement is true, we have built a 128-person face database. In the database, 6 face images were taken of each person as shown in Figure 1. The 6 face images of each person included two frontal views, two 3/4 fontal views with the right side, and two 3/4 frontal views with the left side. Therefore, the total number of training samples was 768 face images. We then implemented the methods proposed in [4] and [5], respectively. Figure 2 shows a series of experimental results. Figures 2(a)-(b) are the results obtained by applying the autocorrelation plus LDA method proposed by Goudail *et al.*[4]. The upper-left face image of each experiment was a query (test) image which had to be different from all the training images. The remaining images were the retrieved database images that had the 11 Euclidean distances closest to the query image. The retrieved database images are ordered from left to right and top to bottom. Figures 2(c)-(d) show another set of results obtained by applying the PCA plus LDA method proposed by Swets and Weng[5]. We found that, in both [4] and [5], the training images used to build the databases took hair, shoulders, and background into account. However, since both methods are statistics-based approaches, we wondered whether those none-face portions played a role in the face recognition process. In order to follow up on this suspicion, we cut out the face portion of the query image in Figure 2(b) and attached it to the face portion of the query image shown in Figure 2(a). Then, we used this synthesized image as a query image to retrieve database images. Figures 3(a) and (b) are the results obtained by autocorrelation plus LDA [4] and PCA plus LDA [5], respectively. From the results shown in Figure 3, it is not difficult to find that, no matter how the method proposed in [4] or [5] was applied, the none-face part dominated the face recognition process. Based on these experiments, we conclude that face images used to train a face database should not include none-face portions. This is a very important finding because training samples without hair, shoulders and background will generate a totally different decision boundary set in the feature space. One thing to be noted is that the facial organs on a human face only differ slightly from person to person. In other words, a set of decision boundaries based on features extracted from facial organs may not be as discriminating as a set based on features extracted from facial organs, hair, shoulders and background. In what follows, we shall explain how a face-only database can be built and then propose a coarse-to-fine feature extraction technique to reduce the dimensionality of the original image.

3 Building a Face-only Database and Feature Extraction

In the previous section, we have experimentally shown that the face databases used in [4, 5] were incorrect. In the next subsection we shall describe how a "correct" face database can be built. Then, in Section 3.2, a coarse-to-fine feature extraction mechanism is proposed.

3.1 Building a Face-only Database

A "correct" face here is defined as a face portion which includes the eyes, mouth, and nose. In the image acquisition process, it is very difficult to control the camera so as to photograph only the "face" portion. In this stage, we apply a previously developed face detection algorithm [12] to perform the task. Since we already had a face database containing 768 face images from 128 persons, the algorithm in [12] was used to "cut" out the face portion of each image. Figure 4 is an example showing this step. After performing the face detection algorithm proposed in [12], the detected face portions of the images in Figure 1 were those shown in Figure 5. In what follows, we shall discuss how to perform coarse-to-fine feature extraction based on the new database.

3.2 Coarse-to-fine Feature Extraction

The purpose of feature extraction is to reduce the dimensionality of the original images so that the small sample size problem[6, 7, 8, 10] will not happen. In [4], Goudail *et al.* calculated 25 local autocorrelation coefficients and used them to substitute the original image. Swets and Weng [5] applied PCA to reduce the dimensionality of the original images. In this paper, we propose a coarse-to-fine feature extraction mechanism. At the coarsest resolution, a face image is equally partitioned into 4 (2x2) regions as shown in Figure 6. For each region, the gray values of all the pixels are averaged, and the value is then used as the region's representative value. Therefore, at this resolution, every sample image is converted from its original size to a 4-dimensional feature vector. For the second resolution, every database image is partitioned into 16 (4x4) regions. Again, each region of this partition is represented by its corresponding average gray value(Figure 6). At this resolution, every sample image is represented as a 16-dimensional feature vector. Using a similar partition, every sample image at the third resolution should be represented by a 64-dimensional feature vector. Basically, the number of resolutions adopted in the coarse-to-fine feature extraction process should be provided in advance. Furthermore, this number is dependent on the scale of the face database. Figure 6 shows the details of the coarse-to-fine feature extraction mechanism. The proposed coarse-to-fine feature extraction strategy can guarantee a fast search process, which is a basic requirement in most nationwide face-based ID checking systems. In the next section, we shall report how, based on the coarse-to-fine features, LDA-based classification can be performed.

4 Classification Using Coarse-to-fine LDA

In Section 3, we proposed a coarse-to-fine feature extraction mechanism. In this section, we shall employ a multi-resolution linear discriminant analysis (LDA)[5, 13, 14] technique to classify the feature vectors extracted from images at different resolutions. We shall introduce the LDA approach in Section 4.1. Then, we shall explain how LDA is used for classification in our coarse-to-fine mechanism in Section 4.2.

4.1 Linear Discriminant Analysis (LDA)

In the coarse-to-fine feature extraction stage, different dimensional feature vectors are extracted from resolution to resolution. The LDA technique discussed here is a general description which fits feature vectors extracted at different resolutions.

Let the training sample set be comprised of M classes, and let M_k be the number of samples contained in class k. X_j^k denotes a feature vector representing the j^{th} sample of the k^{th} class. $\bar{X}^k$ denotes the mean feature vector of the k^{th} class, and $\bar{X}$ denotes the mean feature vector of the whole population. The physical meaning of LDA is to determine the mapping

$$Y_j^k = A' * X_j^k$$

that simultaneously maximizes the between-class scatter while minimizing the within-class scatter of all Y_j^k (where $k = 1, \ldots, M$, $j = 1, \ldots, M_k$) in the projective feature vector space.

The within-class scatter in the projective feature space can be calculated as follows:

$$WS = \sum_{k=1}^{M} \sum_{j=1}^{M_k} (Y_j^k - \bar{Y}^k) * (Y_j^k - \bar{Y}^k)'.$$

The between-class scatter can be calculated as follows:

$$BS = \sum_{k=1}^{M} (\bar{Y}^k - \bar{Y}) * (\bar{Y}^k - \bar{Y})'.$$

The way to find the required mapping A is to maximize the following quantity:

$$\mathbf{tr}(WS^{-1} \times BS).$$

An algorithm which can solve the mapping matrix A can be found in [13].

4.2 LDA-based Coarse-to-fine Classification

In this section, we shall discuss how LDA can be applied to form coarse-to-fine clustering at different resolutions. Assume that the current resolution is i. In Section 3, we described how to reduce the dimensionality of all training samples to 2^{2i} dimensions. These 2^{2i} dimensional samples are then used as bases to calculate the set of most discriminating projection axes thru LDA. In order to simplify the process, at each resolution i, we calculate the 2^{2i} projection axes, i.e., the upper bound of allowable projection axes. After finding the 2^{2i} most discriminating projection axes at resolution i, the next step is to project all the training samples onto these axes and to obtain their corresponding 2^{2i}-dimensional projective feature vectors. When an unknown query image comes in, the coarse-to-fine feature extraction process will first be executed. This process will reduce the dimensionality of the unknown image to a 2^{2i}-dimensional feature vector at resolution i ($i = 1, 2, \ldots$). Then, for each resolution i, the transformed 2^{2i}-dimensional unknown data are projected onto the corresponding 2^{2i} most discriminating projection axes, and a 2^{2i}-dimensional projective feature vector for the unknown image at resolution i is obtained. The unknown projective feature vector obtained at resolution i can be compared with all the qualified database projective feature vectors at resolution i. If the Euclidean distance between the unknown projective feature vector and that of a database sample is larger than a statistically predetermined threshold, then at the next resolution (a finer resolution), this sample will not be considered. Figure 7 shows the coarse-to-fine search mechanism. The above mentioned statistics-based threshold determination process is an automatic threshold determination process. It can reduce human intervention to the largest extent. In what follows, we shall describe the details of this process.

4.2.1 Determining Distance Threshold for Every Class at Different Resolutions

First, we have to use some extra training samples to achieve this goal. That is, in addition to the database training samples used to derive the most discriminating projection axes at different resolutions, a similar set (with similar size) of verification samples is needed. These verification samples will be used together with the database samples to determine the distance thresholds for all the classes at different resolutions. Assume that we have used the database samples to calculate the 2^{2i} most discriminating projection axes at different i resolutions($i = 1, 2, \ldots$). Since both the database samples and verification samples are given, the class of every sample is known in advance. Let $Y_i^k(j), j = 1, \ldots, J$ be the set of projective feature vectors of database samples belonging to class k, at resolution i, and let $Y_i^{*k}(l), l = 1, \ldots, L$ be the set of projective feature vectors of verification samples belonging to class k, at resolution i. From the previous section, we know that the projective feature vectors of these samples are of dimension 2^{2i} at resolution i. Define $\|Y_i^k(j) - Y_i^{*k}(l)\|$,

$j = 1, \ldots, J, l = 1, \ldots, L$ as the Euclidean distance between $Y_i^k(j)$ and $Y_i^{*k}(l)$. There are a total of $J \cdot L$ distances between the database samples and the verification samples for class k, at resolution i. Assume that these distances are normally distributed; then, the mean and variance of class k at resolution i can be calculated as follows:

$$M_i^k = \frac{1}{J \cdot L} \sum_{j=1}^{J} \sum_{l=1}^{L} \|Y_i^k(j) - Y_i^{*k}(l)\|,$$

$$V_i^k = \sqrt{\frac{1}{J \cdot L} \sum_{j=1}^{J} \sum_{l=1}^{L} (\|Y_i^k(j) - Y_i^{*k}(l)\| - M_i^k)^2}.$$

After M_i^k and V_i^k are determined, the distance threshold, T_i^k, for class k at resolution i can be defined as

$$T_i^k = M_i^k + \alpha V_i^k.$$

In statistics, if a pattern belongs to class k at resolution i, then the probability that the distance between this pattern and k's center is smaller than T_i^k is roughly about 98% if $\alpha = 2$. In the next subsection, we shall describe in detail the coarse-to-fine search algorithm.

4.2.2 Coarse-to-fine Search Algorithm

The LDA-based coarse-to-fine search algorithm is described as follows.

Coarse-to-fine Search Algorithm

Input : An unknown gray image $\mathcal{A}$.

Output : The m closest database face images.

Step 1: Convert the input image $\mathcal{A}$ into feature vectors, $U_1, U_2, \ldots, U_n$, at various resolutions, $i = 1, 2, \ldots, n$.

Step 2: Project $U_1, U_2, \ldots, U_n$ onto the most discriminant axes, and derive the projective feature vectors, $V_1, V_2, \ldots, V_n$, via the mapping matrices, $A_1, A_2, \ldots, A_n$, respectively.

Step 3: Initialize $\mathcal{R}_1 = \mathcal{R}_2 = \ldots = \mathcal{R}_n = \emptyset$.

Step 4: Set $\mathcal{R}_1 = \{(k, j) | k = 1, 2, \ldots, M, \text{ and } j = 1, 2, \ldots, M_k\}$.

Step 5: Initialize the search stage at the coarsest resolution, $i = 1$.

Step 6: Compute all the distances $||Y_i^k(j) - V_i||$ for all $(k, j) \in \mathcal{R}_i$.

Step 7: If $||Y_i^k(j) - V_i||$ is smaller than the predetermined threshold T_i^k(for all $(k, j) \in \mathcal{R}_i$), then add a new item (k, j) to set $\mathcal{R}_{i+1}$. Otherwise, $\mathcal{R}_{i+1}$ remains.

Step 8: If the finest resolution is reached, i.e., $i = n$, then output the m closest prototypes, $X_n^k(j)$'s, where (k, j)'s should be members of $\mathcal{R}_n$, and then terminate the search procedure.

Step 9: Get to the next finer resolution, i.e., $i = i + 1$, and then go to Step 6.

In the above search algorithm, m was set to be 11 throughout all experiments. The 11 closest retrieved database samples were arranged in order from left to right and top to bottom.

5 Experimental Results

In order to demonstrate the efficiency and accuracy of our method, a series of experiments was conducted. First of all, we used the morphology-based face detection algorithm proposed in [12] to detect the face portions of 768 database images. The detected face portions of each old database sample were put together to form a new face-only database, which contained

768 new database images. The original database was retained, and an index table was established to relate the old database to the new one. In order to avoid the small sample size problem, the number of resolutions selected was four (i.e., 2^1 to 2^4). In the recognition process, a set of test images which was different from the training set was used. In every experiment, a query image (test image) was presented to the system. The recognition system first performed face detection on the query image and then executed a normalization process on the detected face portion. The normalized face portion was then used as an input to retrieve the 11 closest samples (out of 768 samples) in the new database. The 11 retrieved closest samples were arranged in left to right, top to bottom order. Figure 8 shows one set of experimental results. Figure 8(a) shows the detected face portions. After normalization, the detected portion was used as a query to retrieve the new database. The 11 closest retrieved face-only samples together with the query image are shown in Figure 8(b). Figure 8(c) shows 11 original database images corresponding to the 11 retrieved new database samples. In this experiment, it is obvious that, although only the face portion was used to perform recognition, the top 6 of the 11 closest retrieved samples were correct identifications. Figures 9(a)-(b) showed the results of another two sets of experiments. The results shown in Figure 9(a) are as good as those shown in Figure 8. From the results shown at the bottom of Figure 9(b), it is obvious that the top 5 and the ninth retrieved samples were correct identifications. We ran 1000 sets of experiments; among these experiments about 95.02% of the first ranked retrieved results were correct idenfications. If the scope was extended to the retrieved subset (11 closest samples), then the success rate could be further increased to 99.15%.

In order to show that our approach can also adapt to blurred and noisy images, another set of experiments was conducted. Figure 10(a) shows a set of results obtained by using a clear test image. After blurring the test image, the corresponding retrieval results were those shown in Figure 10(b). Figure 10(c) shows another set of experiment results obtained when noises were introduced.

6 Conclusion and Future Work

In this paper, we have proposed a new face recognition system which bases its recognition process on a face-only database. In the feature extraction process, the average gray level technique is employed to reduce the dimensionality of the original image at different resolutions. Using this process, the effect caused by noises or defocus can be significantly suppressed. Unlike the PCA-based approach, our approach does not require recalculation of the projection axes, which makes the database extensible. Furthermore, based on a coarse-to-fine database architecture, our approach can handle an extremely large face database. As to the efficiency problem, our method is faster than the approaches proposed by Swets and Weng[5] and Goudail *et al.*[4]. As to the recognition rate, our approach is superior to Goudail *et al.*'s approach.

In the future, we shall integrate the face detection subsystem[12] and the LDA-based coarse-to-fine face recognition subsystem. Figure 11(a) shows the face portions detected using the morphology-based approach[12]. Through normalization, these detected faces were fed into the face recognition subsystem, and the expected retrieval results are shown in Figure 11(b). Since it is possible that the acquired images were greatly influenced by different factors in the environment, a number of difficulties are associated with the integration process. If too many constraints are forced onto the whole face recognition system, the robustness of the system will be degraded. Therefore, we shall focus our efforts on suppressing the effects caused by the environment.

References

[1] A. Samal and P. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern Recognition*, vol. 25, no. 1, pp. 65–77, 1992.

[2] R. Chellappa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of the IEEE*, vol. 83, no. 5, pp. 705–740, 1995.

[3] D. Valentin, H. Abdi, A. O'toole, and G. Cottrell, "Connectionist models of face processing: A survey," *Pattern Recognition*, vol. 27, no. 9, pp. 1209–1230, 1994.

[4] F. Goudail, E. Lange, T. Iwamoto, K. Kyuma, and N. Otsu, "Face recognition system using local autocorrelations and multiscale integration," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 1024–1028, October 1996.

[5] D. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, pp. 831–836, August 1996.

[6] Y.-Q. Cheng, Y.-M. Zhuang, and J.-Y. Yang, "Optimal fisher discriminant analysis using the rank decomposition," *Pattern Recognition*, vol. 25, no. 1, pp. 101–111, 1992.

[7] Z.-Q. Hong and J.-Y. Yang, "Optimal discriminant plane for a small number of samples and design method of classifier on the plane," *Pattern Recognition*, vol. 24, no. 4, pp. 317–324, 1991.

[8] Q. Tian, M. Barbero, Z.-H. Gu, and S. H. Lee, "Image classification by the foley-sammon transform," *Optical Engineering*, vol. 25, no. 7, pp. 834–840, 1986.

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

[10] D. H. Foley and J. J. W. Sammon, “An optimal set of discriminant vectors,” *IEEE Trans. Comput.*, vol. 24, pp. 281–289, 1975.

[11] H.-Y. M. Liao, C.-C. Han, and G.-J. Yu, “Face + hair + shoulders + background $\neq$ face,” Tech. Rep. TR-IIS-97-007, Institute of Information Science, Academia Sinica, Taipei, Taiwan, 1997.

[12] C. Han, H.-Y. M. Liao, G. Yu, and L.-H. Chen, “Fast face detection via morphology-based pre-processing,” Tech. Rep. TR-IIS-97-001, Academia Sinica, Taipei, Taiwan, 1997.

[13] K. Liu, Y. Cheng, and J. Yang, “Algebraic feature extraction for image recognition based on an optimal discriminant criterion,” *Pattern Recognition*, vol. 26, no. 6, pp. 903–911, 1993.

[14] K. Liu, Y. Cheng, and J. Yang, “A generalized optimal set of discriminant vectors,” *Pattern Recognition*, vol. 25, no. 7, pp. 731–739, 1992.

[15] S. H. Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, “Facial feature detection using geometrical face model:an efficient approach,” *to appear in Pattern Recognition*, 1997.

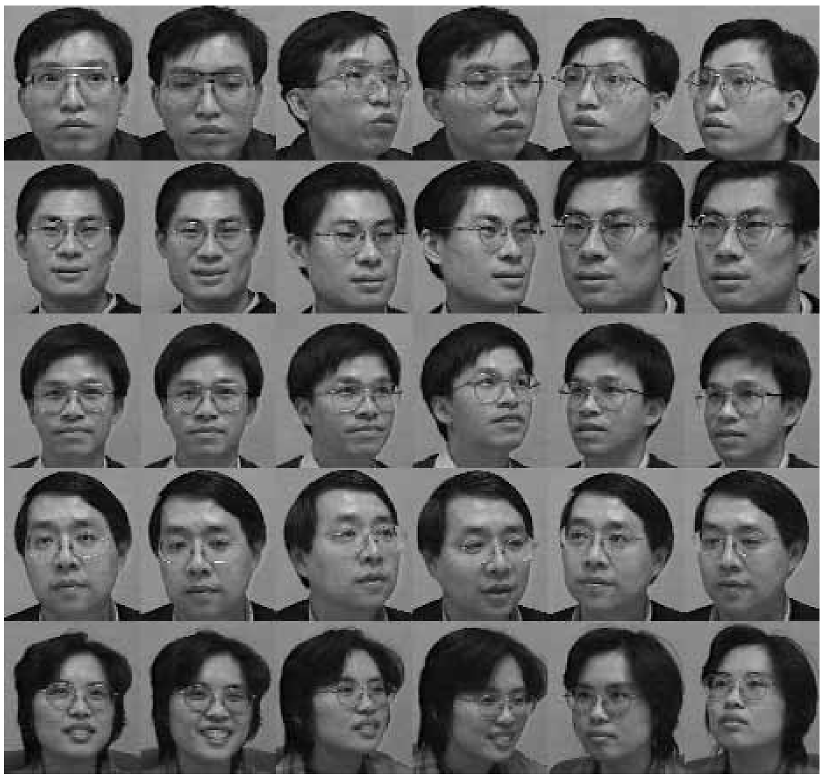
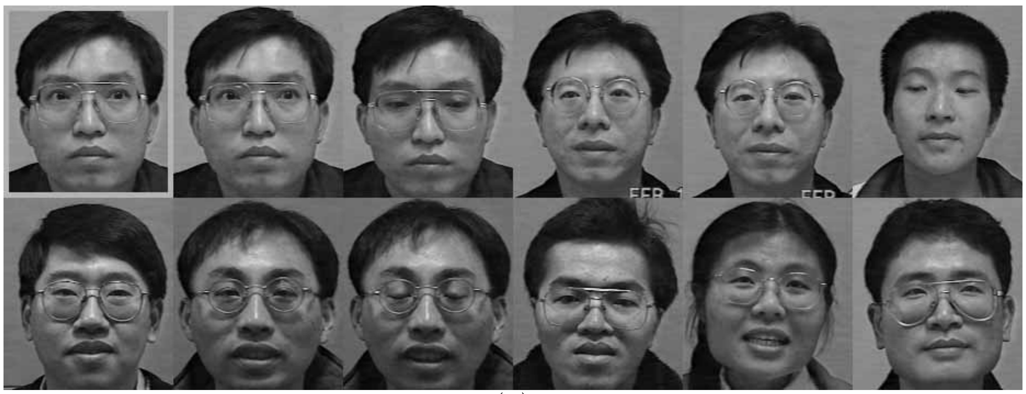
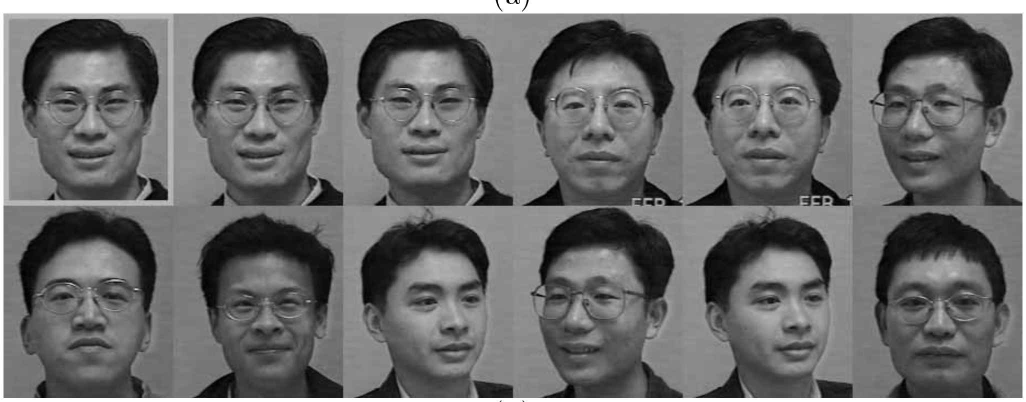


Figure 1: Part of the original face database.



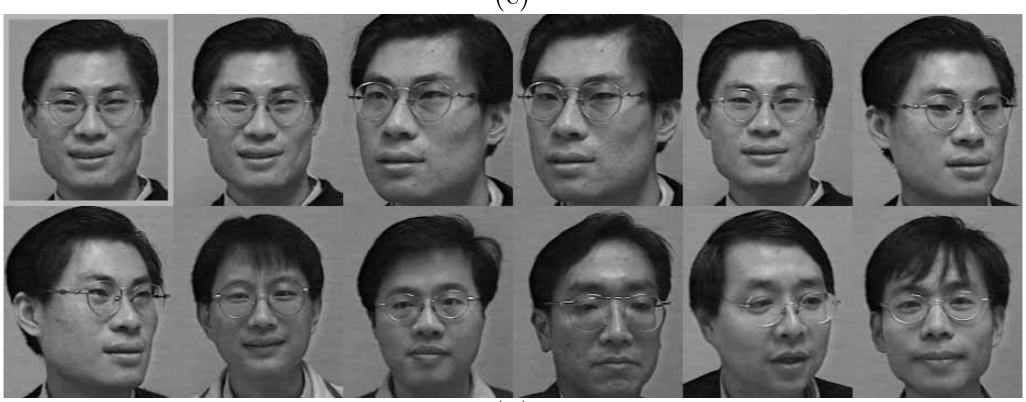
(a)



(b)

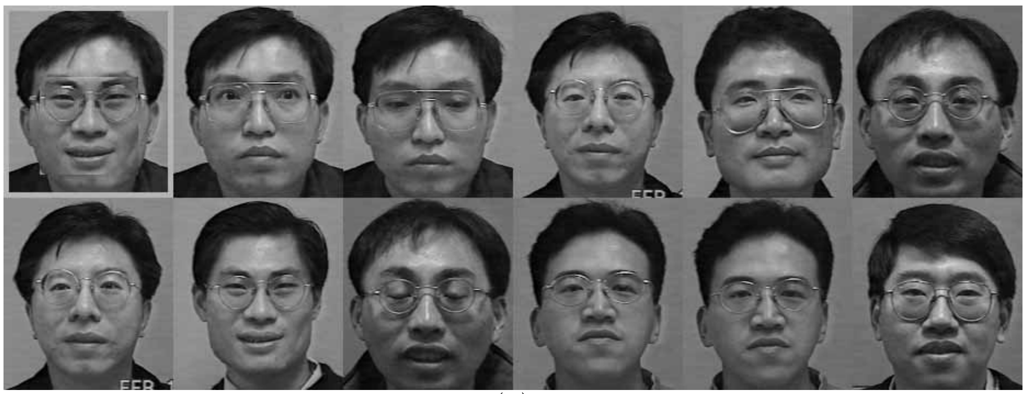
Figure 2: (a)-(b) Two sets of retrieved results obtained by applying the autocorrelation plus LDA method. The upper-left face image is a query image.

(c)



(d)

Figure 2: (c)-(d) Two sets of retrieved results obtained by applying the PCA plus LDA method(continued.)



(a)

(b)

Figure 3: (a) A set of retrieved results obtained by applying the autocorrelation plus LDA method. (b) A set of retrieved results obtained by applying the PCA plus LDA method. The upper-left face images of both cases are synthesized query images.

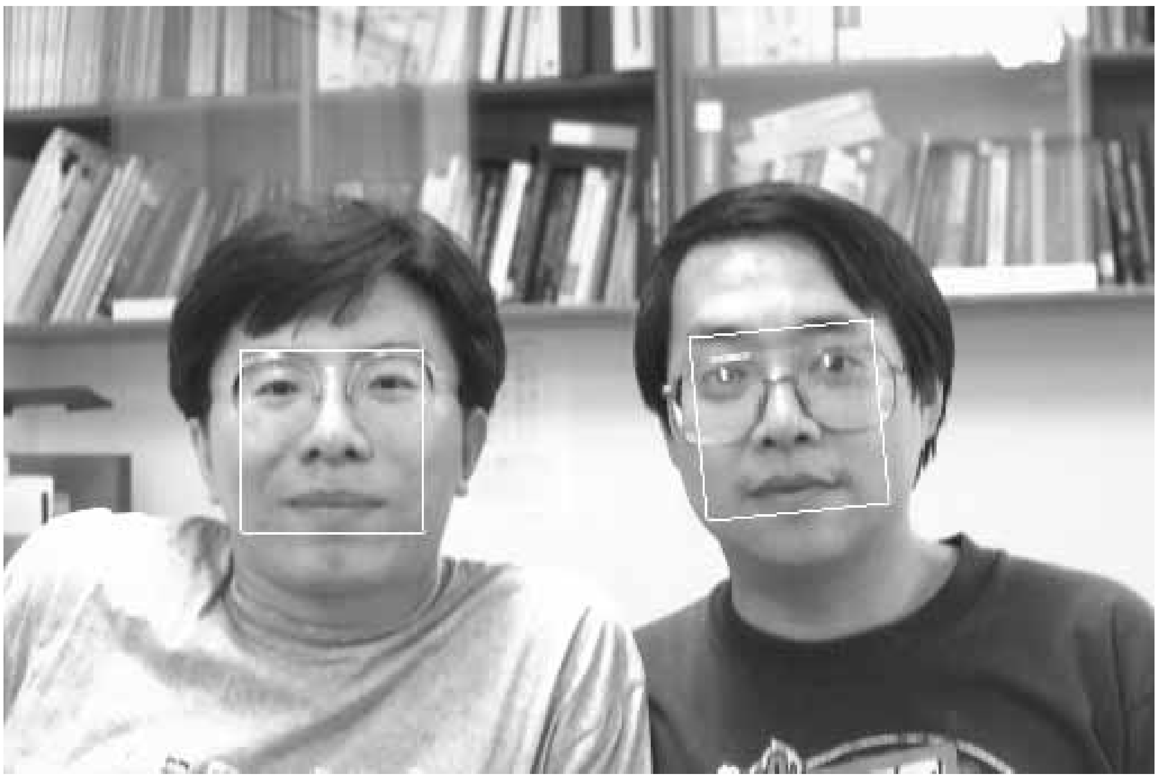


Figure 4: The detected face portions obtained by using the algorithm proposed in [14].

Figure 5: The detected face portions of the face images in Fig. 1.

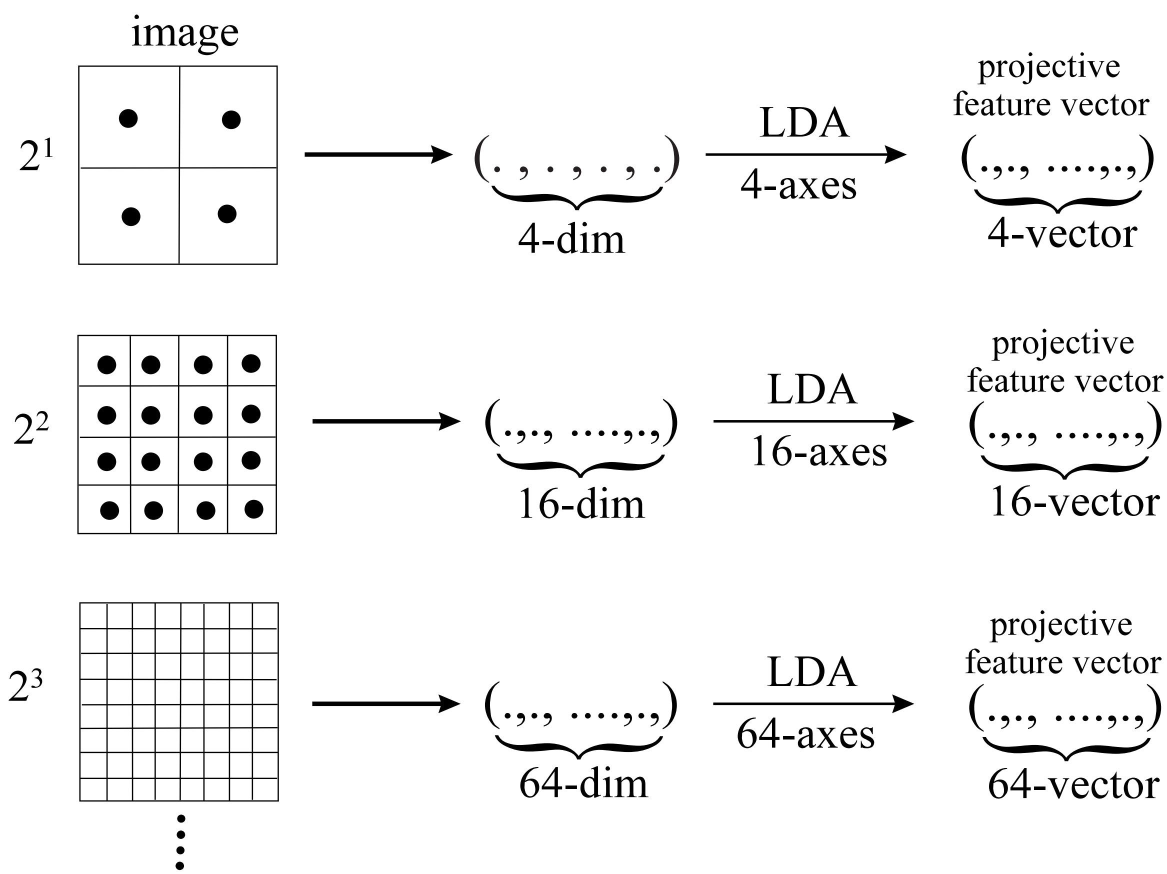


Figure 6: Coarse-to-fine feature extraction and multi-scale LDA.

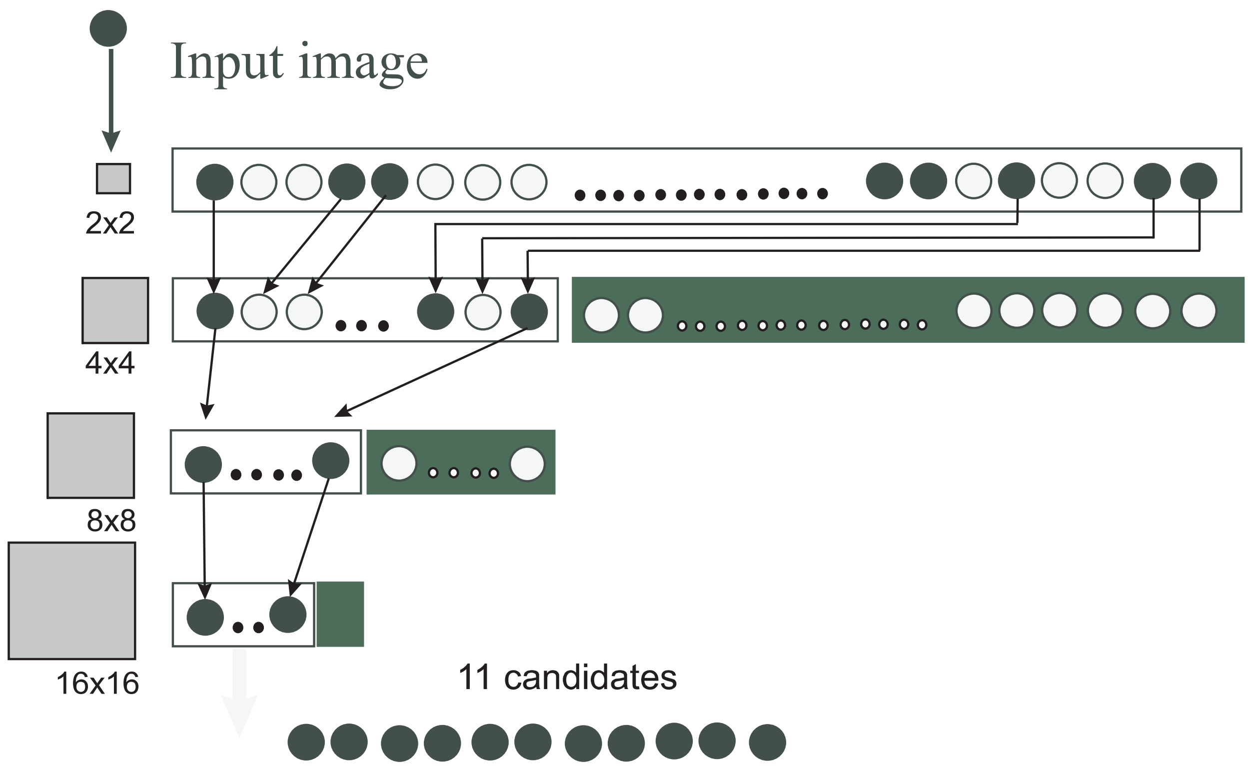
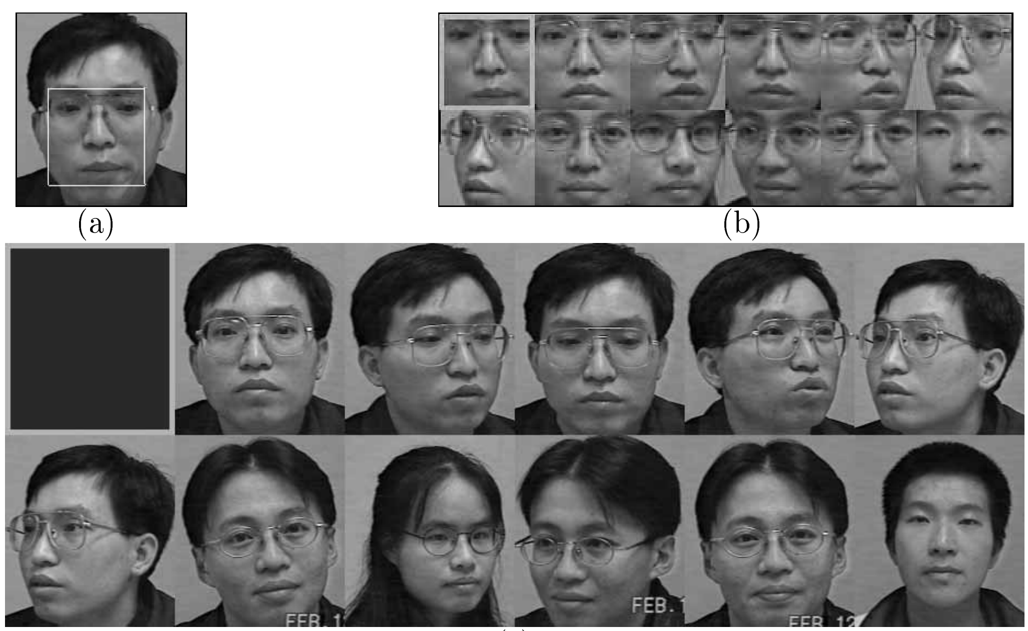


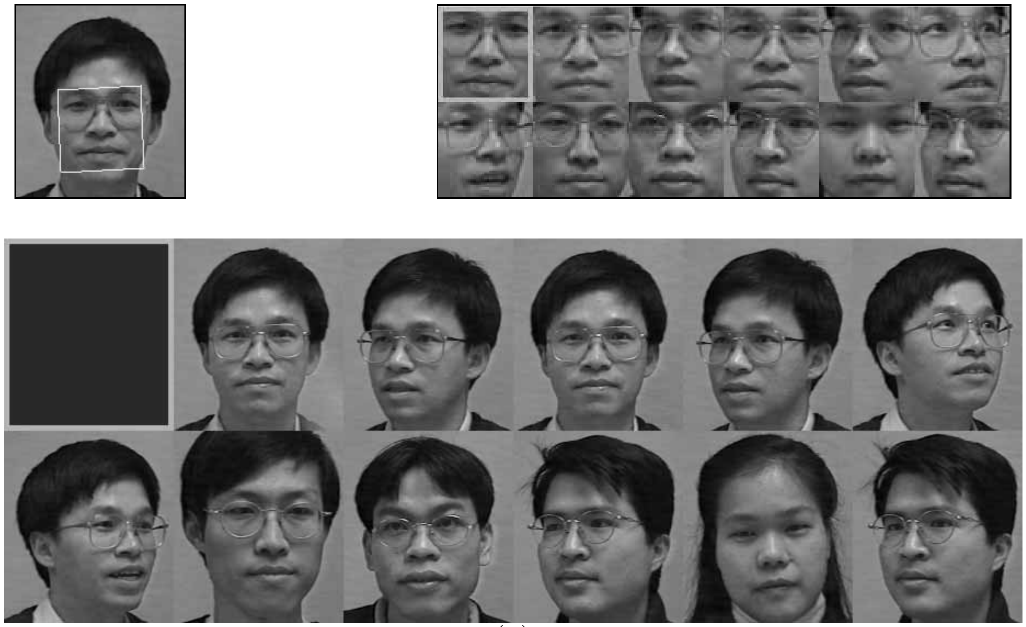
Figure 7: The coarse-to-fine search mechanism.



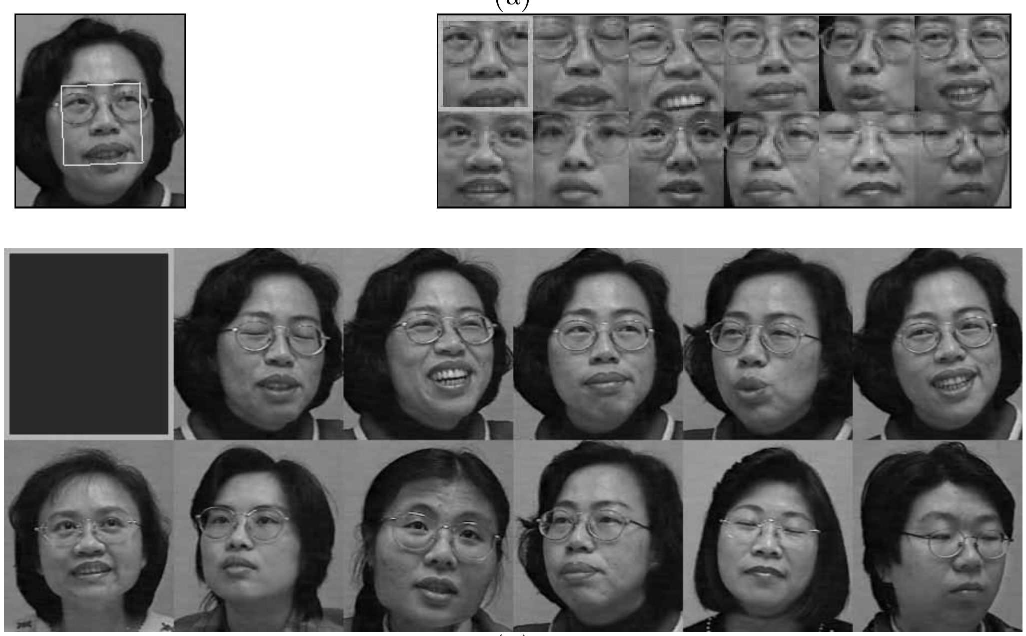
(a) (b)

(c)

Figure 8: (a) The detected face portion; (b) the 11 retrieved face-only samples and the query image; (c) the 11 original database samples which correspond to the 11 retrieved face-only database samples.

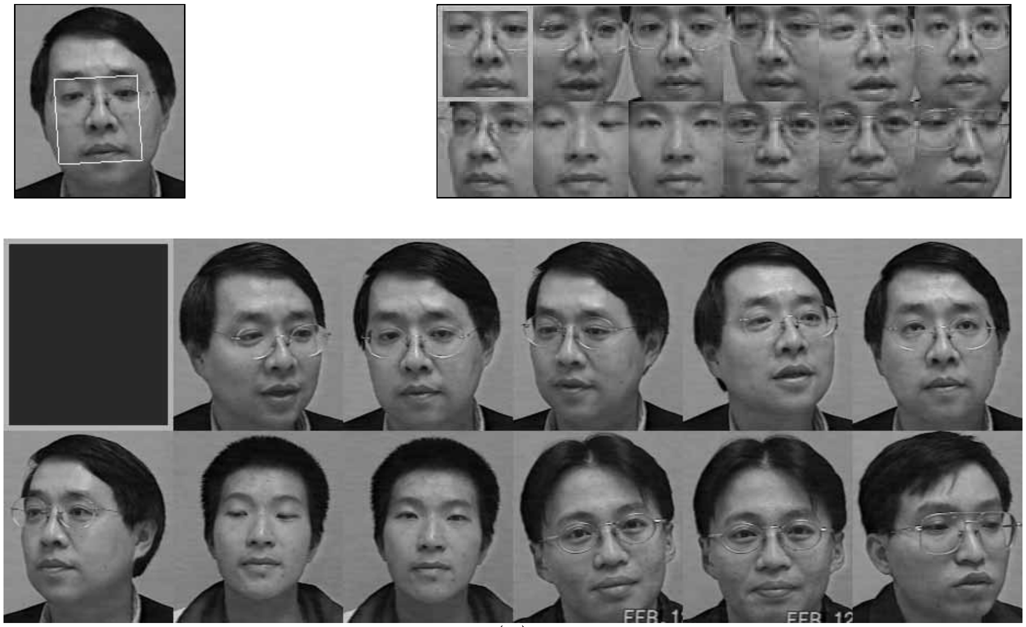


(a)

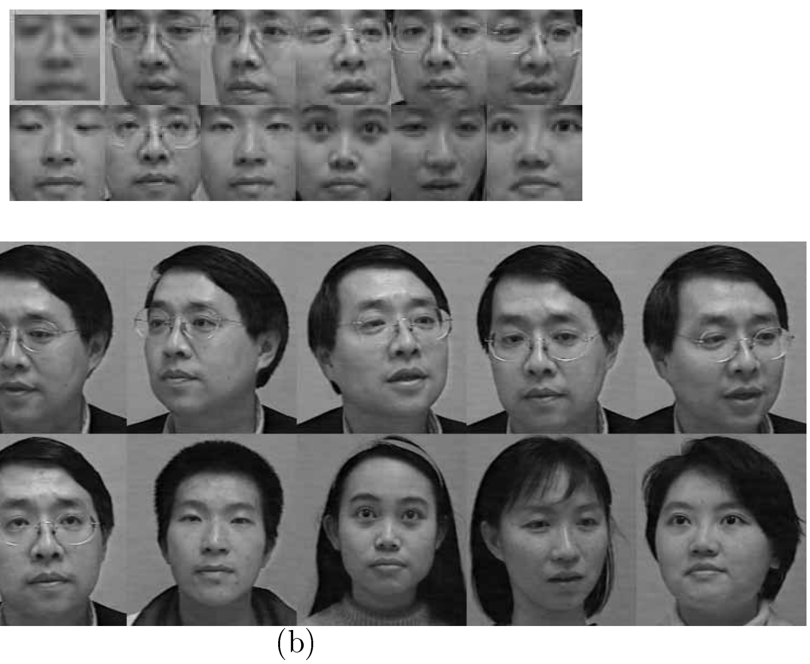


(b)

Figure 9: (a)-(b) The retrieval results of another two sets of experiments.

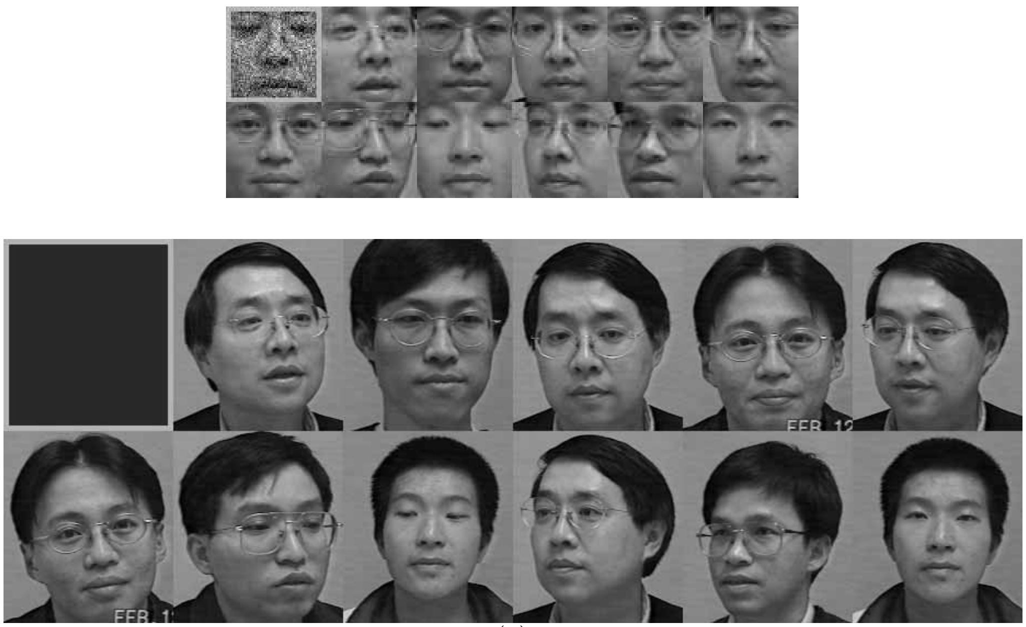


(a)



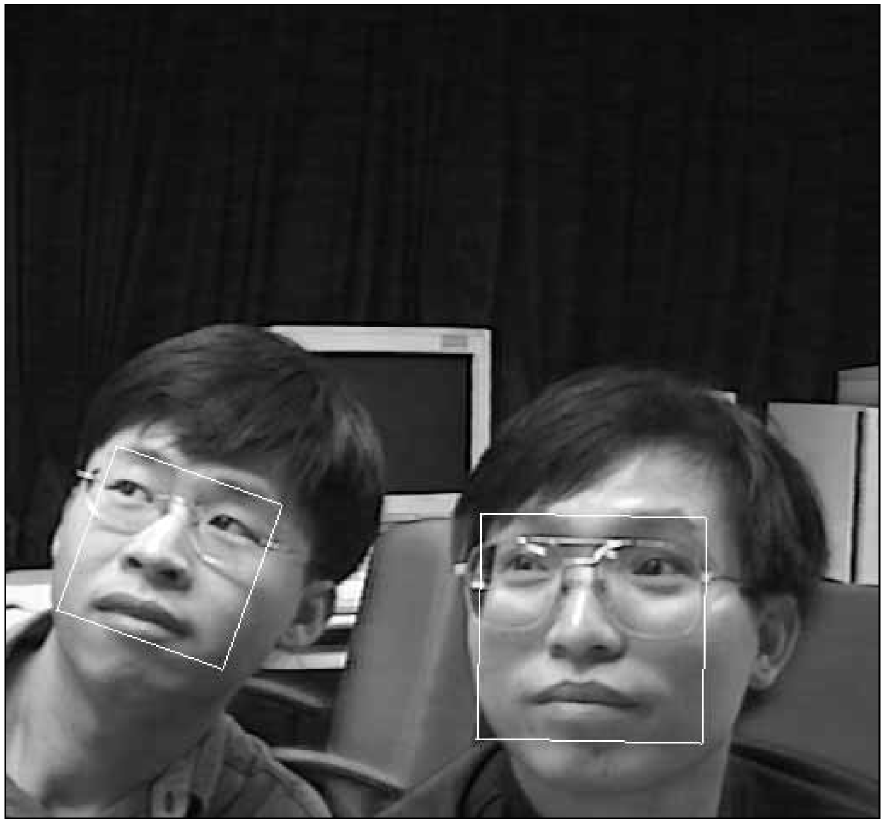
(b)

Figure 10: (a) A set of retrieval results using a clear test image; (b) retrieval results obtained using a blurred test image; (c) retrieval results when noises were introduced into the test image.

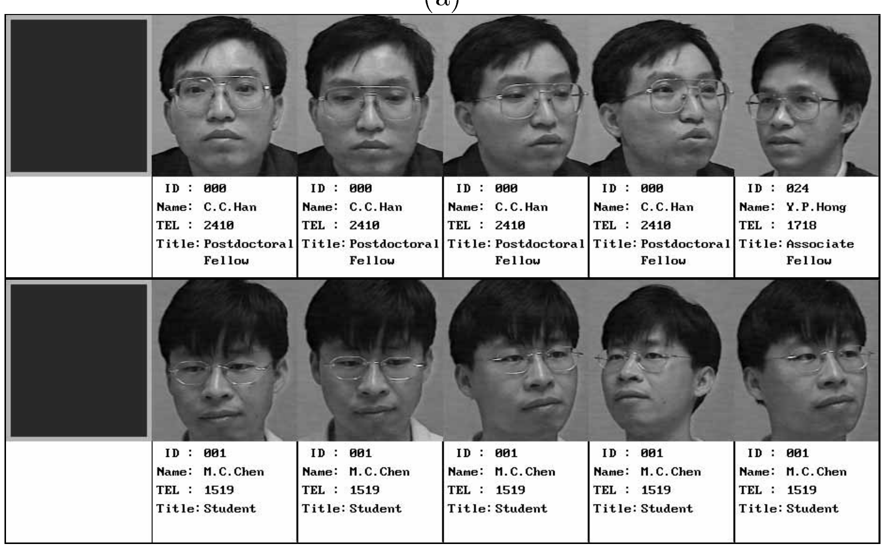


(c)

Figure 10: (Continued.)



(a)



(b)

Figure 11: Integration of the face detection subsystem and the face recognition subsystem.