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

Hong-Yuan Mark Liao†[1],Chin-Chuan Han†, and Gwo-Jong Yu‡,
†Institute of Information Science,
Academia Sinica, Nankang, Taipei, Taiwan
‡Institute of Computer Science and Information Engineering,
National Central University, Chung-Li, Taiwan
E-mail : {liao,cchan,yugj}@iis.sinica.edu.tw

Abstract

Face recognition, by definition, should be a recognition process in which recognition is based on the content of a face. The problem is: what is a "face" ? Goudail *et al.*[1] and Swets and Weng[2] have recently proposed state-of-the-art face recognition systems. Through careful implementation, the results have shown that both methods are valid and efficient. However, we have a question about the "face" images they have adopted. They used "face" images that included hair, shoulders, face and background. Our intuition tells us that only a recognition process based on a "pure" face portion can be called face recognition. From a series of experiments, we have shown that the "real" face portion in their face recognition process did not play a key role at all. Instead, the combination of hair, shoulder and background dominated the whole recognition process. We suggest that future research on face recognition should use a face-only database, not a face combined with other irrelevant portions.

[1]To whom correspondence should be addressed.

1 Introduction

Face recognition has been a very hot research topic in recent years [3, 4, 5]. Some successful face recognition systems have been developed and reported in the literature[1, 2, 6, 7, 8]. In this paper, we shall report a very interesting and important finding. That is, the face images used by Goudail *et al.*[1] and Swets and Weng[2] were incorrect. In fact, the "face" images used in [1] and [2] were combinations of face, hair, shoulders and background. Through a series of experiments, we have found that this kind of "face" image will mislead the recognition process. Our experimental results have shown that the "real" face portions of their "face" images "didn't" play the role in the face recognition process they have suggested. On the contrary, the other portions of their "face" images, i.e., the hair, shoulders and background, played a key role in the entire recognition process. Our suggestion in response to this important finding is that future researches in this field should be based on a "correct" database, i.e., a face-only database. In the next section, we shall report in detail how the experiments were conducted. In Section 3, we shall report how a "correct" face database can be built. Concluding remarks will be given in Section 4.

2 Face + Hair + Shoulders + Background = Face ?

In this section, we shall describe a series of experiments and show that the databases used in [1] and [2] were incorrect. In order to prove that our above statement is correct, 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 include two frontal views, two 3/4 frontal 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 [1] and [2], respectively. Figure 2 shows a series of experimental

results. Figures 2(a)-(c) are the results obtained by applying the autocorrelation plus linear discriminant analysis(LDA) [9, 10] method proposed by Goudail *et al.*[1]. 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 closest Euclidean distances to the query image. The retrieved database images were ordered from left to right and top to bottom. Figures 2(d)-(f) show another set of results obtained by applying the principal component analysis (PCA)[7] plus LDA method proposed by Swets and Weng[2]. From the results shown in Figure 2, it is obvious that the results obtained by applying the method of [2] are superior to those of [1] in terms of accuracy. At this point, we have to raise a critical question, i.e., "Is this the correct manner to conduct face recognition ?" Our point is that, in both [1] and [2], the training images used to build the databases took hair, background, and shoulders into account. However, since both methods are statistics-based approaches, we wonder 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 portions of the query images in Figures 2(b) and (c) and attached them, respectively, to the face portion of the query image shown in Figure 2(a). Then, we used these two synthesized images as query images to retrieve database images. Figures 3(a) and (b) are the results obtained by autocorrelation plus LDA[1] and PCA plus LDA [2], respectively. From the results shown in Figure 3, it is not difficult to find that, no matter how the method proposed in[1] or [2] was applied, the none-face part (hair + shoulders + background) dominated the face recognition process. In other words, it was not "face recognition." Instead, it should be called "background + shoulders + hair" recognition. In order to make sure that our results were consistent at all times, we conducted 100 sets of similar experiments. We found that, still, the none-face portions dominated the whole recognition process. Based on these experiments, we conclude that face images used to train a face database should not include none-face portions, i.e., shoulders, hair and background. 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.

3 Building a New Face 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. At this stage, we apply a previously developed face detection algorithm[11] to perform the task. Since we already had a face database containing 768 face images from 128 persons, the algorithm in[11] 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[11], the detected face portions of the images in Figure 1 were those shown in Figure 5.

4 Concluding Remarks

In this paper, we have proposed a new database concept regarding face recognition. That is, face recognition should be based solely on a face portion, not on a face portion plus other irrelevant portions. We have shown through experimentation that Goudail *et al.*[1] and Swets and Weng[2] used incorrect face databases for face recognition. We suggest that future research in this field should be based on a face-only database for the face recognition task.

References

[1] 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.

[2] 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.

[3] 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.

[4] 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.

[5] 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.

[6] S. Lawrence, C. L. Giles, A. Tsoi, and A. Back, "Face recognition: A convolutional neural-network approach," *IEEE Trans. on Neural Networks*, vol. 8, pp. 98–113, January 1997.

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

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

[9] 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.

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

[11] S. H. Jeng, H.-Y. Mark 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.

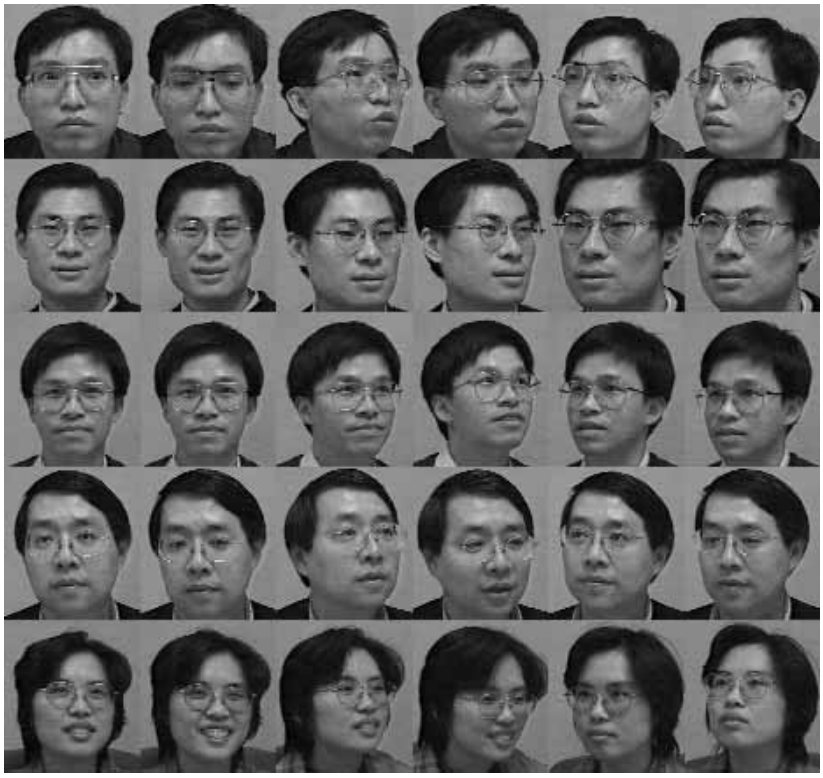
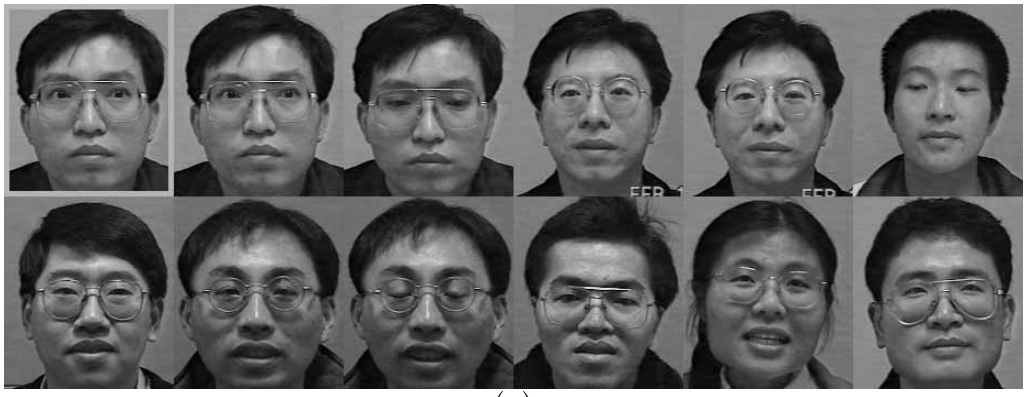
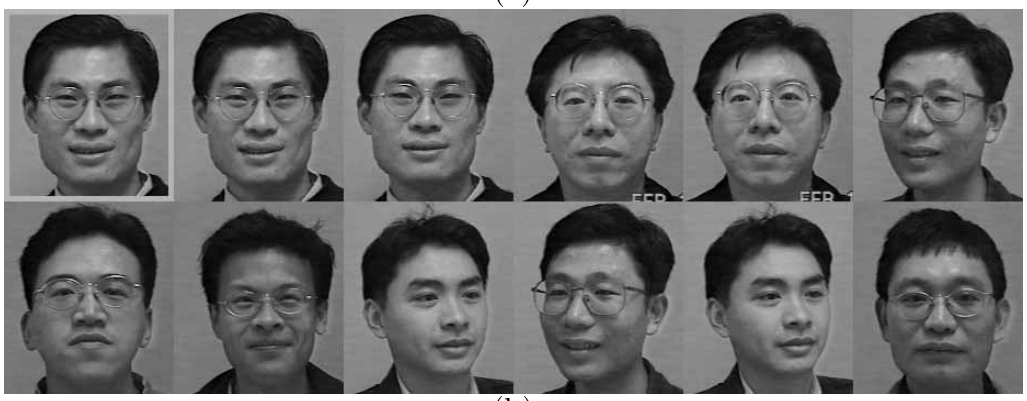


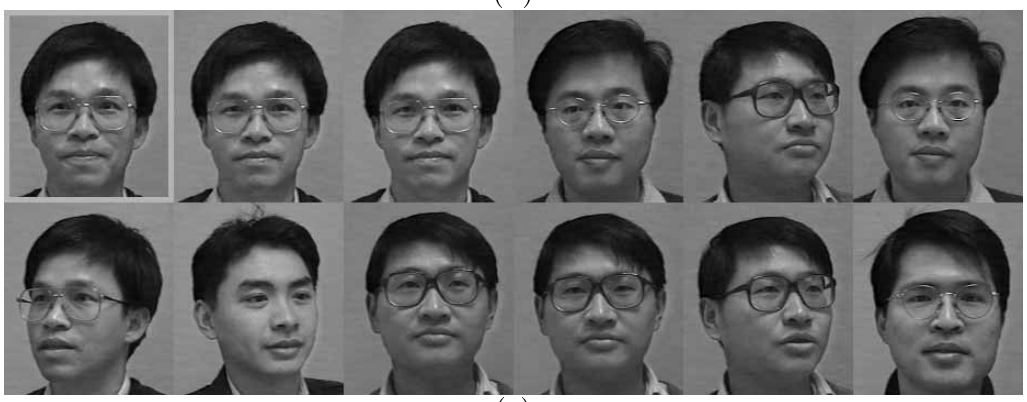
Figure 1: Part of the original face database.



(a)



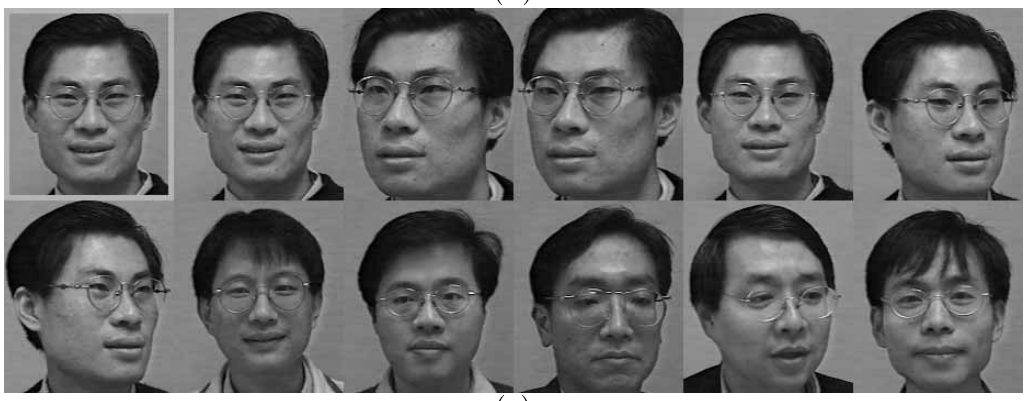
(b)



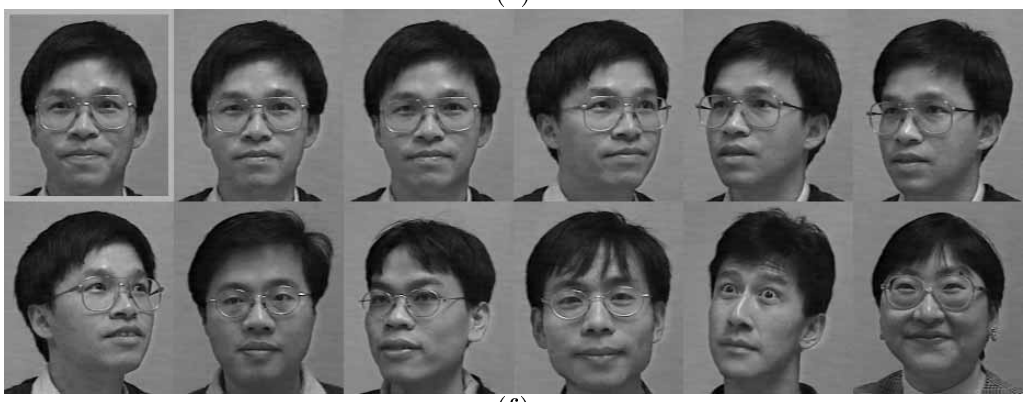
(c)

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

(d)

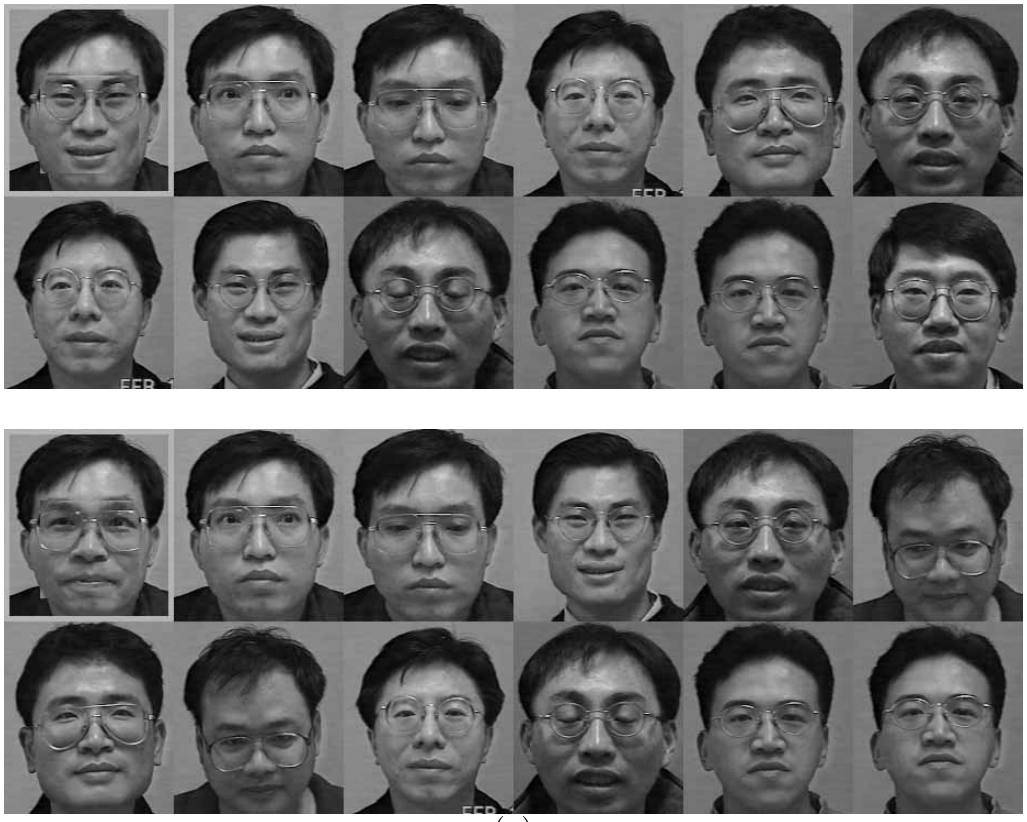


(e)



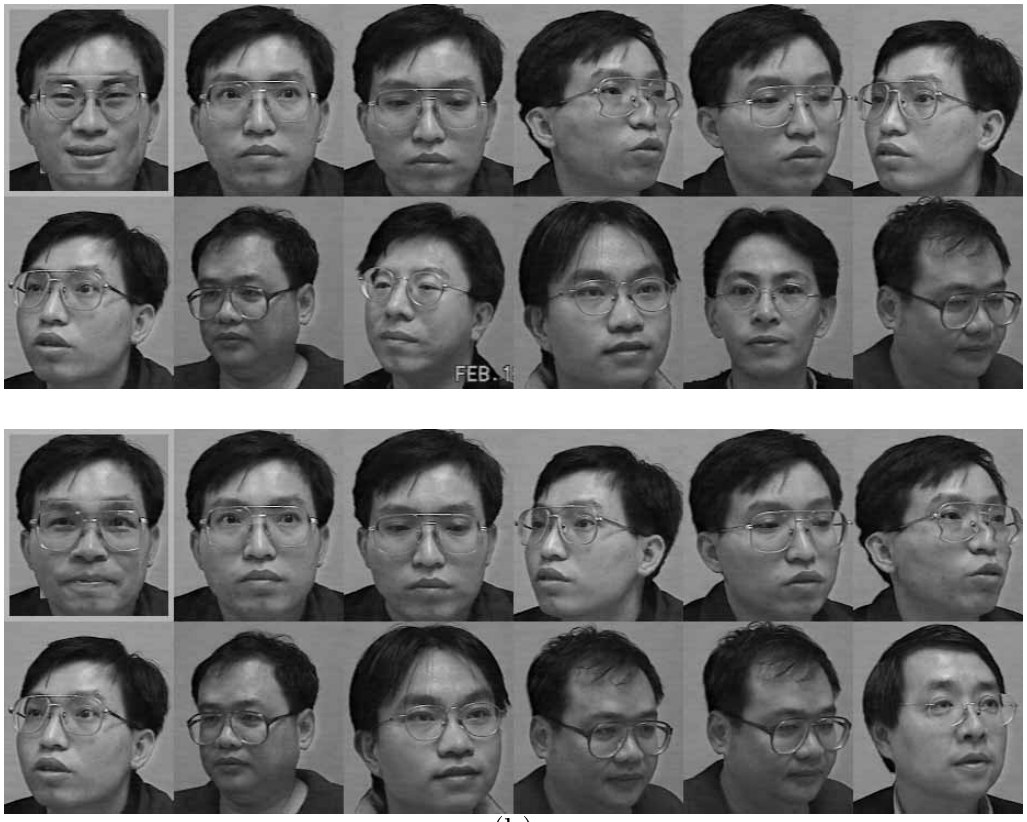
(f)

Figure 2: (d)-(e) Three sets of retrieved results obtained by applying the autocorrelation plus LDA method(continued.)



(a)

Figure 3: (a) Two sets of retrieved results obtained by applying the autocorrelation plus LDA method. The upper-left face images are synthesized query images.



(b)

Figure 3: (b) Two sets of retrieved results obtained by applying the PCA plus LDA method. The upper-left face images are synthesized query images(continued.)

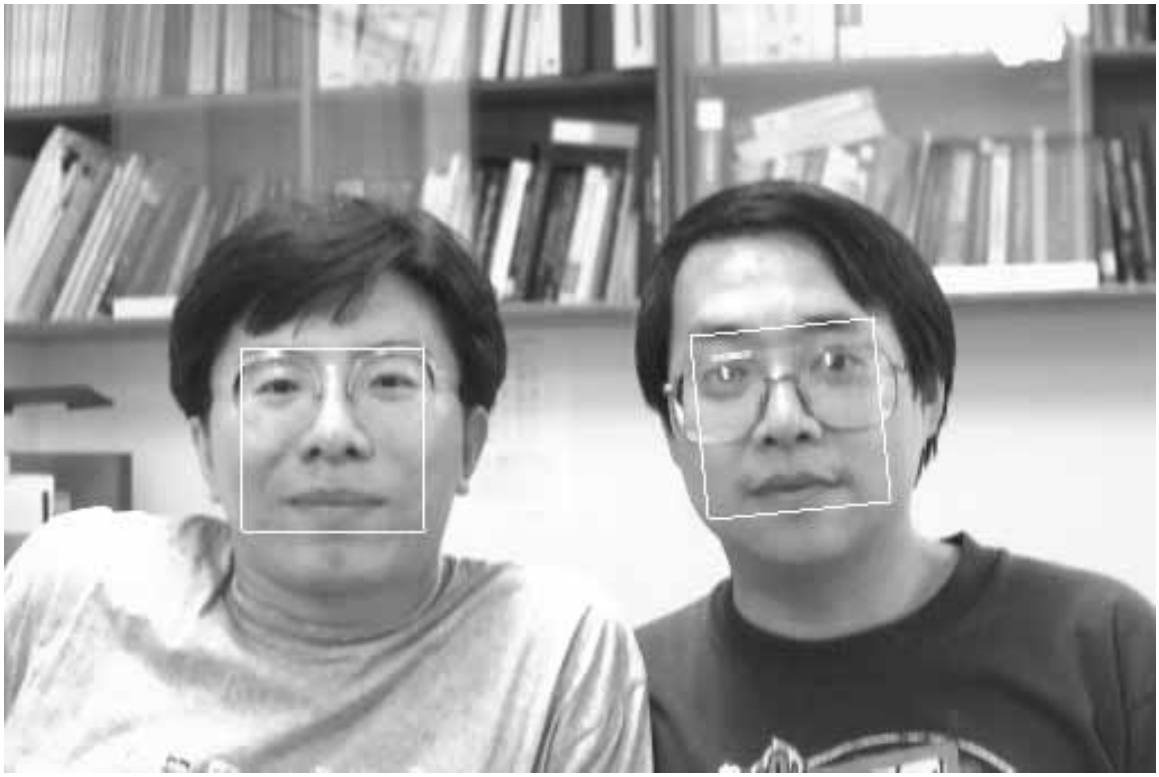


Figure 4: The detected face portions obtained by using the algorithm proposed in [11].

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