DICTA2002: Digital Image Computing Techniques and Applications, 21--22 January 2002, Melbourne, Australia

Robust Face Recognition Using Multiple Eye Positions

Jiaming Li, Rong-Yu Qiao, Jason Lobb, Geoff Poulton
Image & Signal Processing Discipline
CSIRO Telecommunications & Industrial Physics
Australia
Tel: 612 9372 4104, Fax: 612 9372 4411, Email: jiaming.li@csiro.au

Abstract

This paper describes a robust face recognition algorithm using multiple candidate eye positions to improve recognition. Face recognition systems consist of four major stages. They are face detection, eye detection, face normalisation and face recognition. Most recognition schemes (eg. PCA) assume accurate knowledge of eye positions. By using multiple candidate eye positions, inaccuracies in eye detection can be overcome. An application of this method is given for a scheme with orthogonal complement PCA (OCPCA) features, and training and test image sets chosen from different databases. Experiments on face recognition have shown that about 5.4% performance improvement has been achieved by using multiple eye positions.

1. Introduction

The face is one of several features that can be used to uniquely identify a person. It is the characteristic that we most commonly use to recognise people and it plays a vital role in our social interactions. Since no two human faces are identical, faces are well suited for use in identification schemes, in the same way that fingerprints or DNA samples are used.

Automatic face identification is a challenging task. Its potential applications include access control and surveillance. Compared with competing methods, the obvious advantage of a face recognition system is its low level of intrusion. It does not require more than looking into a camera.

The recognition of faces is done by finding the closest match of a newly presented face to all faces known to the system. Popular face recognition methods include Principal Component Analysis (PCA) [1,2], Independent Component Analysis (ICA) [3], Neural Network [4], etc. [5]. For most methods, face recognition performance in real-time systems largely depends on the accuracy of face detection. However, a detected face is rarely close to its database equivalent, in a pixel-by-pixel sense. This difference comes from lateral and vertical shifts due to eye detection errors, as well as other factors such as pose, lighting conditions and facial expression. In a real-time face recognition system, face detection [6,7] and facial feature (such as eyes) detection [8,9] are the most important steps. A lot of effort has been put into accurate detection of human eyes in an arbitrary scene, and robust eye detection in real time is still under intensive investigation.

We present a method for improving recognition with an imperfect eye-finder, by using multiple candidate eye positions. This method is applicable to many face recognition methods. Here it is applied to a global method using a feature set derived by OCPCA. Experimental results on the FERET database [10] are presented. These results show that recognition performance is improved by up to 5.4% compared to a system, which does not employ a multiple eye position approach.

The remainder of this paper is organised as follows. Section 2 introduces the orthogonal complement PCA face recognition method. Section 3 gives the modified OCPCA face recognition algorithm using multiple eye positions. Section 4 presents and discusses the experimental methodology and results, and conclusions are given in Section 5.

2. Orthogonal Complement PCA (OCPCA) Face Recognition

In face recognition based on conventional PCA, there is no differentiation of variations in images caused by several different factors. The most important of these factors are:

Type (A) Fundamental variations between images of different individuals.

Type (B1) Variations in images of a single individual due to change of expression, hairstyle, facial hair, aging etc.; and

Type (B2) Variations in images due to differences in the mode of image capture.

For effective verification it is necessary to be able to discriminate images on the basis of Type (A) variations whilst ignoring as far as possible variations of Type (B). To fulfill this, the orthogonal complement PCA (OCPCA) method has been investigated [11]. The OCPCA method concentrates on difference between images instead of the images themselves.

Suppose there are two sets of images, P_B and P_{AB}. P_B consists of pairs of images of the same people. P_{AB} consists of pairs of images of different individuals.

Consider the following sets of image differences:

D_B: $\{d_{11}, d_{12} ...\}$ - differences between pairs in P_B,

and

D_{AB}: $\{d_{21}, d_{22}, ...\}$ - differences between pairs in P_{AB}.

D_B contains information about image differences of Type (B), whilst D_{AB} has information about both Type (A) AND Type (B) differences. This happens because all the factors causing differences from image to image of a single person can also operate to cause part of the difference between images of two people.

Two orthonormal bases S_B and S_{AB} are then generated, using PCA or a similar method, for both sets of difference images D_B and D_{AB}. What is required is a basis spanning only Type (A) variations, and this basis may readily be obtained by finding the *orthogonal complement* S_{OC} of S_B in S_{AB}. This process of deriving an orthogonal complement (OC) basis is illustrated schematically in Figure 1.

This OC basis will account only for differences between individuals, and should be independent of variations between images of a single person and the imaging modality. The method retains the simplicity and computational speed of PCA or similar global methods.

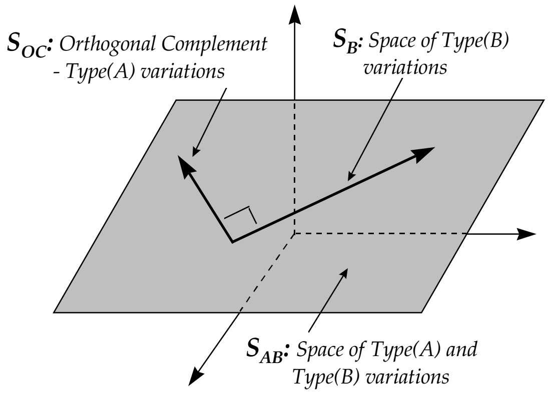

Figure 1: Schematic Diagram Illustrating the Generation of an Orthogonal Complement (OC) Basis

3. Face Recognition Using Multiple Eye Positions

Figure 2 describes the block diagram of a general face recognition system. The video sequence is input from devices such as a camera, VCR or DVD. Usually, a real-time system may use a number of attributes, such as motion, skin colour and face features, to detect faces. Once a face is detected, more accurate eye positions can be obtained using a second eye detection algorithm. Then the face image is normalised in size according to this final eye position. The normalised face image is passed to the face recognition stage, whose output indicates whether the subject is recognised or not.

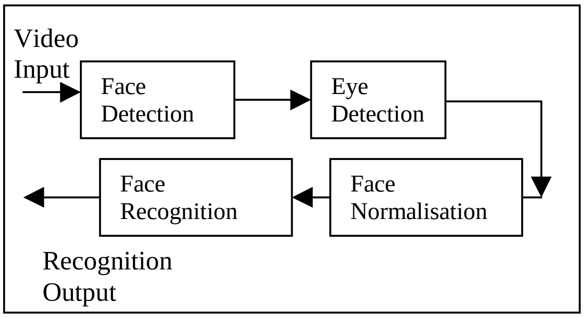

Figure 2: Block Diagram of a General Face Recognition System

In systems like that above, although the face detector can efficiently detect the face, the eye locations for each face are often not very accurate. This accuracy can greatly affect the recognition performance. To overcome this problem and make the system more robust to eye detection, the eye detector may be asked to output several possible eye positions for each image, in order of likelihood. For each of these candidate eye positions, orthogonal complement PCA is employed in the face recognition stage to generate for each a recognition distance. A recognition decision is then made on the basis of the minimum of these distances. To summarise, the process involves three steps as shown below.

Step 1: Get multiple eye positions for face image.
Step 2: For each eye position calculate its recognition distance. The recognition distance is defined as the Euclidean, Mahalanobis or other distance between the test image and database image.
Step 3: Define the minimum recognition distance as the final recognition distance, and use this to make a decision on recognition.

4. Experiments and Results

(1) Training Image Set

In the example given below, the system is first trained on a set of images of 28 individuals with strictly controlled lighting and pose. In total, 196 images are used, comprising 7 instances of each of the 28 individuals with different lighting conditions and expressions.

(2) Test Image Set

Once training is complete, recognition performance is tested on part of the Face Recognition Technology (FERET) database sponsored by the DoD Counterdrug Technology Development Program Office [10]. The particular database used is the first release of FERET images. It consists of 3737 greyscale images of human heads with views ranging from frontal to left and right profiles. In this database there are male and female subjects. A few of the subjects wear glasses.

To analyse the recognition performance under different conditions, we have chosen three test sets as below.

- Set1 (Facial expression test set): frontal images with different expression captured on the same day.
- Set2 (Illumination test set): frontal images captured on different days.
- Set3 (Facial pose test set): frontal images differing by $\pm 22.5^0$ in horizontal pose.

(3) Recognition experiments

To evaluate the performance of the multiple eye approach three different experiments were conducted. The first experiment used manually located eyes to establish an optimal recognition result. The second experiment used the automatically selected “best” candidate eye position. The third experiment used the best five candidates and the procedure described in Section 3. Each experiment was carried out on each of the three test sets.

(4) Results and discussions

The experiment results are shown in Figure 3. In the figures, the horizontal axes are the recognition distance between test image and a database image, the vertical axes give the cumulative fraction of image distances which differ by less than a given value. The solid curve represents the false-positive recognition rate, while the dotted curve represent the false-negative recognition rate. The recognition performance is often measured as the value of the crossover point of the two curves.

By comparing Experiment 2 and Experiment 3, it can be seen that the multiple eye position method can improve the crossover point by 2.6% to 5.4% compared to the method using only a single pair of eye positions. The improvement is test-set related. It is important to note that the recognition performance of the multiple eye position approach is as good as that of the real eye position approach when testing on pose variation.

5. Conclusion

We have introduced a robust face recognition method using multiple eye positions. The experiments have shown that accurately detecting eye position is very important for recognition performance. However by using multiple pairs of candidate eye positions we can compensate for eye detector inaccuracies. The recognition performance can be improved by up to 5.4% compared to using one pair of eye positions. Especially for the facial pose test, the recognition performance was substantially the same as that obtained using manually located eye positions.

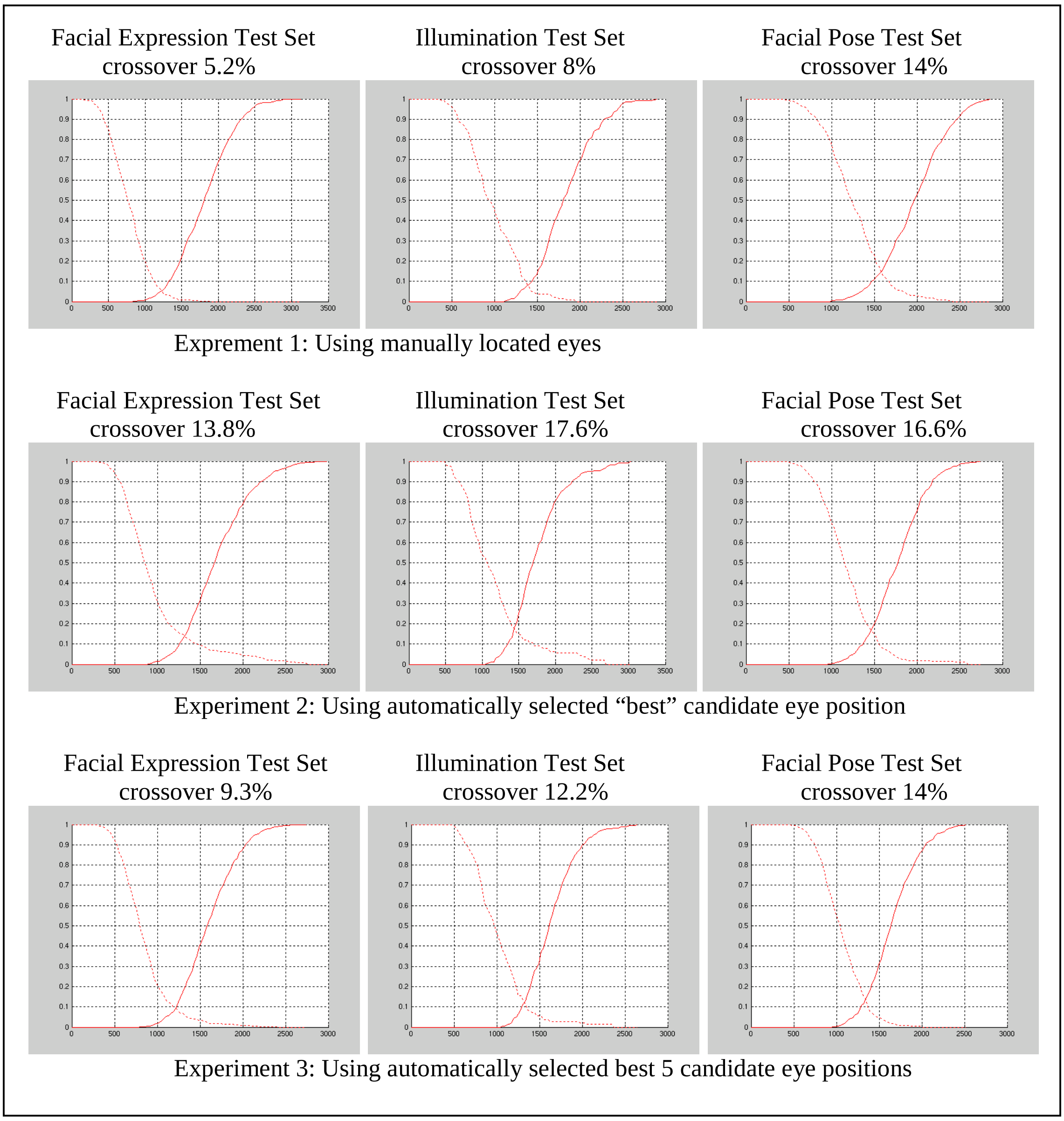

Figure 3: Recognition Results

References

[1] A. Pentland, B. Moghaddan, T. Starner, “View-Based and Modular Eigenspaces for Face Recognition”, Computer Vision and Pattern Recognition, 1994, pp.84-91.

[2] M. Turk, A. Pentland, “Eigenfaces for Recognition”, Journal of Cognitive Neuroscience 3(1), 1991, pp.71-86.

[3] M. S. Bartlett, T. J. Sejnowski, “Independent Components of Face Images: A Representation for Face Recognition”, Proceedings of the 4th Annual Jount Symposium on Neural Computation, Pasadena, CA, May 17, 1997.

[4] S. Lawrence, C. L. Giles, A. C. Tsoi, A. D. Back, “ Face Recognition: A Hybrid Neural Network Approach”, U. of Maryland Technical Report CS-TR-3608.

[5] F. Samaria and F. Fallside, “Face Identification and Feature Extraction Using Hidden Markov Models”, Image Processing:

Theory and Applications, edited by G. Vernazza, Elsevier.

[6] J. Cai, A. Goshtasby, "Detecting Human Faces in Colour Images", Image and Vision Computing 18, 1999, pp.63-75.

[7] K.K. Sung, T. Poggio, "Example-Based Learning for View-Based Human Face Detection", IEEE Transactions Pattern Analysis and Machine Intelligence 20(1), 1998, pp.39-51.

[8] L. Bala, K Talmi, J. Liu, "Automatic Detection and Tracking of Faces and Facial Features in Video Sequences", Picture Coding Symposium 1997, 10-12 September 1997.

[9] X. Xie, R. Sudhakar, H. Zhuang, "On Improving Eye Feature Extraction Using Deformable Templates", Pattern Recognition, 1994-27(6), pp. 791-799.

[10] "Face Recognition Technology (FERET) database" in web site http://www.itl.nist.gov/iad/humanid/feret

[11] G. Poulton, "Optimal Feature Sets for Face Recognition", Proceedings, IASTED Int. Conf. On Signal Processing & Communications, Feb. 11-14, 1998, pp. 269-272.