2006 9th Int. Conf. Control, Automation, Robotics and Vision
Singapore, 5-8th December 2006

Verifying a User in a Personal Face Space

Q. Tao and R.N.J. Veldhuis
Signals and Systems Group, University of Twente
Postbus 217, 7500AE Enschede, Email: q.tao@utwente.nl

Abstract **- For user verification on a personal digital assistant (PDA), a fast and simple system is developed. In the enrollment phase, face detection and registration are done by a Viola-Jones based method, taking advantage of its accuracy and speed. The face feature vectors obtained this way are then used to build up a face space specific to the user by principal component analysis (PCA). Furthermore, the face variations caused by small registration shifts are also modeled, in order to better capture the variation in the face space, and simplify the enrollment. Current experiments show that this system is fast, efficient, and accurate.**

Index Terms – user verification, Viola-Jones method, face space, PCA.

I. INTRODUCTION

The topic of face recognition has been actively researched in the past few decades. One of the most popular approaches is to explore the underlying face subspaces. In subspace methods the classifier is trained on face images from a large set of users, to obtain a universal representation such as overall eigenfaces [1] or a discriminative representation such as Fisherfaces [2]. By investigating the probability density of different classes, the unknown face can be classified. In our application of a user-verification system on a PDA, only one authentic user is involved. There are several points specific for this case. Firstly, the user is exposed extensively to the device, thus a large sample set can be obtained, which makes it easy to model the within-class variations of the specific user accurately. This is a big advantage that we can make use of. Secondly, the images acquired (when the user is looking at the PDA) are mostly frontal, and relatively constant, with a small range of rotations, shifts and scales. Finally, for the security of the personal device, real-time ongoing verification is needed, which requires very fast algorithms.

A potentially more effective representation than overall eigenfaces or Fisherfaces is proposed in this paper, by constructing a personal face space exclusively from the user's data. In such a way, the specific face variations of the user are better captured. This face space is more accurate and compact to describe this user, when compared with an overall eigenface space, or a general intra-personal space [6].

This paper is organized as follows: section II describes the feature extraction methods applicable to a personal device (of limited computational resources); section III described the construction of the personal face space; and section IV presents the experimental results for this proposed face verification system.

II. FEATURE EXTRACTION

In order to verify a user, feature vectors have to be extracted first from the image. The current application on a PDA device requires real-time verification of the user. Within this context, a very fast feature extraction scheme is implemented.

A. Face Detection

Traditional detection methods suffer from the fact that for the unknown target, an exhaustive search should be carried out across a large range of scales and at every position of an image. This calculation is therefore very expensive and slow. To solve this problem, Viola and Jones proposed a detection scheme based on combined Haar-like features and cascaded classifiers [3]. This method is very fast to search through different scales because of the simple rectangular features and the introduction of the integral image, and is robust to varying background and foreground because of the training algorithm AdaBoost. Furthermore, in most PDA scenarios the face occupies a large portion of the image, with small changes in scale, and this makes the detection even faster by setting scaling parameters. The Viola-Jones method, therefore, is very suitable for face detection in the PDA application.

B. Face Registration

Before recognition the faces should be aligned to a standard orientation and size. This is called face registration, which plays an essential role in accurate face recognition [5]. We consider a very fast way of face registration by facial landmarks. The 5 most prominent facial features, 2 eyes, 1 nose, and 2 mouth corners, are separately detected. Again Viola-Jones method is used, considering the real-time requirement of the PDA application. From a database, the facial landmark detectors can be trained beforehand and stored in the device. Although some of the mentioned facial landmarks do not have such a constant structure as the face, and thus give rise to more false detections, the previous face detection step acts to relieve this problem. By first constraining the corresponding region of interest (ROI) of the facial landmarks, the false detections can be largely reduced. The search also becomes much faster in a smaller region, making the real-time face registration possible.

After detection of the 5 facial landmarks, their coordinates are combined as a shape, which is then rigidly registered to the reference shape. Furthermore, the face region is masked so that only the important region of the face is retained. The face detection and registration process is illustrated in Fig 1.

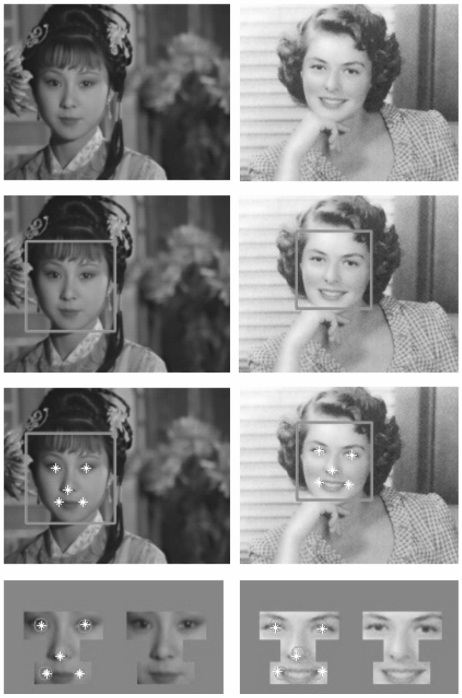

Fig 1. Face detection and registration procedure (from top down): original image, face detection, landmark detection, and registration.

III. Verification in the Personal Face Space

A. Construction of the Personal Face Space

In the enrollment phase, feature vectors from the user's face are obtained. The individual face space, therefore, is easily constructed by principal component analysis (PCA). Suppose we have matrix $\mathbf{X}$ containing N feature vectors $[\mathbf{x}_1, \ldots, \mathbf{x}_N]$, then the matrix $\mathbf{X}^0$ whose columns have zero mean can be calculated by subtracting the column mean $\overline{\mathbf{x}}$ from every vector of $\mathbf{X}$. Then the eigenfaces is computed by means of singular value decomposition (SVD).

$$\mathbf{X}^0 = \mathbf{USV}^T \tag{1}$$

Where the columns of $\mathbf{U}$ are the eigenvectors of PCA, and $\mathbf{S}$ is a diagonal matrix with the diagonal entries being the square root of the eigenvalues. Let the first k columns of $\mathbf{U}$ be $\mathbf{U}_k$, which spans the face space containing the most significant variances, the projection of $\boldsymbol{x}_i$ to the face space is

$$\mathbf{y}_i = \mathbf{U}_k{}^T(\mathbf{x}_i - \overline{\mathbf{x}}) \tag{2}$$

Fig 2 gives an example of the personal face space constructed from a specific user. It is illustrated in Fig 2 that the face space fits to this specific user. If the user wears glasses, most of the eigenfaces will give an indication of glasses.

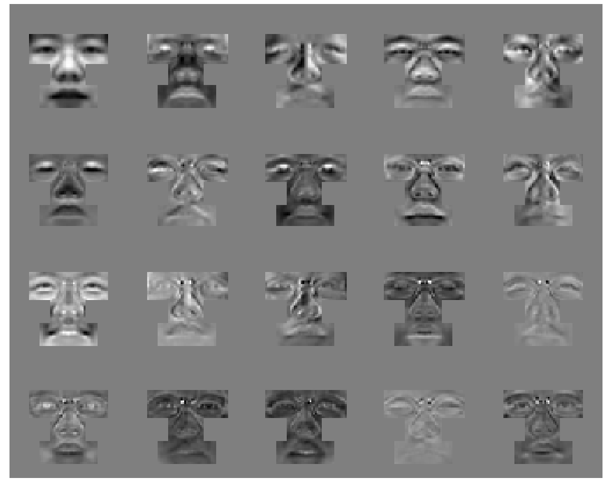

Fig 2. The average user face and the first 19 eigenfaces of the personal face space in lexicographical order

There are several sources accounting for the variations in face space. The first one comprises the user's different expressions and poses. The second source is environmental, such as varying illuminations and background. A third source arises from our feature extraction method. For real-time processing on a small PDA, only 5 facial landmarks are detected independently for rigid registration. Compared to ASM and AAM methods [4] which utilize many more landmarks as well as texture information, our registration depends on the accuracy of landmark detection. Although we found that the Viola-Jones detection yields very accurate results on landmark detection [5], minor shifts of landmarks still end up in small rotations and shifts of the registered image, which propagate to the feature vector.

To include the first type of variations in the personal face space, users can be required to take on some expressions. The second kind of variation can be reduced by conducting some high pass filtering of the original image. Obviously, more feature vectors are beneficial for capturing the feature variations, and ensure better verification performance. However, when thousands of samples are required, the enrollment may be too tedious a task. To better model the variances in the personal face space with less user's effort, we further generate feature vectors out of the already obtained feature vectors. Two types of generations are made. The first type is based on symmetry, by flipping the image horizontally. The second mode is based on the registration method, by adding some random minor shifts on the landmark positions. The first mode models the variations due to light sources on opposite sides, and the latter mode models the third type of variations introduced by our feature extraction methods.

B. Verification Metrics

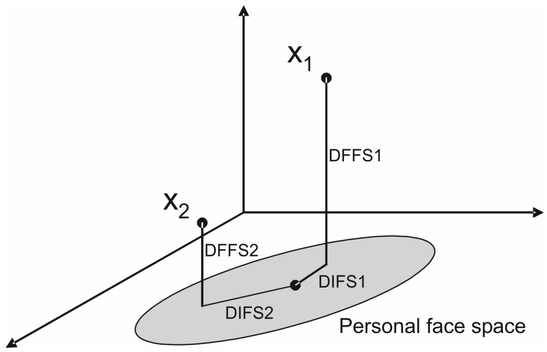

Fig 3. DFFS and DIFS metrics in a personal face space

Once the personal face space has been constructed, we can verify the unknown feature vector by computing and thresholding certain distance measure. Moghaddam and Pentland [6] proposed two metrics, DIFS (distance in feature space) and DFFS (distance from feature space). DIFS indicates the distance from the input feature vector to the center of user face space, and DFFS indicates the distance of the input feature vector from the user face space, as illustrated in Fig 3.

By constructing the personal face space, it is assumed that the user data will be distributed along this space within certain DFFS and DIFS range. The DFFS metric measures how likely a point $\mathbf{x}_i$ belongs to this space. As shown in Fig 3, although $\mathbf{x}_1$ has a smaller Mahalanobis distance to the center of the space, it is less likely to be the authentic user than $\mathbf{x}_2$**,** because in the first place it is farther away from the user space.

The DFFS measure here is the normalized residual error of the projection onto face space:

$$DFFS(\mathbf{x}_i) = \frac{\left\|(\mathbf{I} - \mathbf{U}_k \mathbf{U}_k{}^T)(\mathbf{x}_i - \overline{\mathbf{x}})\right\|}{\left\|\overline{\mathbf{x}}\right\|} \tag{3}$$

Where **I** is the identity matrix. The DIFS measure is the Mahalanobis distance between the projection of $\mathbf{x}_i$ and the centre of the space

$$DIFS(\mathbf{x}_i) = \sqrt{\mathbf{y}_i{}^T (\mathbf{S}_k{}^2)^{-1} \mathbf{y}_i} \tag{4}$$

Where $\mathbf{y}_i$ is computed from equation (2), and $\mathbf{S}_k$ is the upper left $k \times k$ submatrix of **S** in equation (1).

IV. Experiments and Results

In our application, the acquired images are webcam images of 320 pixels by 240 pixels. For training facial landmarks we use BioID database [7], because it has been hand-labeled so that the positive and negative samples for training are readily obtainable, and the detection result can be evaluated. The performance of face registration, together with the subsequent face recognition performance, has been reported in [5].

To construct the personal face space, we create 9 extra feature vectors out of every obtained feature vector, as described in III.A. In total 800×10 feature vectors are obtained, and 4,000 of them are used to calculate the dimensionality-reduced subspace which preserves most of the variances. 4,000 impostor data are also collected under the same condition, which are used to train the classifier.

Once the personal face space is built up, the DFFS and DIFS can be calculated for every feature vector, and these two metrics make up a further-reduced feature vector. We use a simple linear minimum-square-error (MSE) classifier [8], which minimizes the error in a least square sense. Fig 4 shows one example of scatter plot of the DFFS-DIFS distribution, as well as the MSE decision boundary.

We also investigate the influence of dimensionality to the verification error. We construct separate classifiers for one-dimensional data DFFS, DIFS, and two-dimensional data DFFS-DIFS. The error includes both falsely accepted and falsely rejected points. It is shown in Fig 5 that when the dimensionality is above 100, the error is already low for DFFS and DFFS-DIFS classifiers; and it is indicated that DFFS contributes much more in the two-dimensional classifier than DIFS.

The lowest verification error 0.86% (with false accept rate 0.95% and false reject rate 0.77%) occurs when the reduced dimensionality is 280. The MSE linear classifier, though simple and effective, is not optimal in this case, because the distribution of the user data and impostor data are unbalanced. In the further work we will explore the underlying distribution of the two metrics, and choose more suitable classifiers, such as support vector machine (SVM) classifier which relies on the most critical points along the boundary, rather than the overall distributions of the data.

It is stated in [6] that DFFS is a more discriminative metric than DIFS, and this is also observed in our experiment as in Fig 5. DFFS measures the distance from the input vector to its projection in the personal face space, i.e. the residual error, therefore, if this projection is not optimal (in the sense of minimum square error), this distance will be large. The purpose of constructing a personal face space is to minimize this DFFS distance for the specific user, while not take care of other people, thus possibly yielding large and unpredictable errors for the impostors.

The processing speed is a major concern in our application. The viola-Jones detection methods implemented by C++ can work in real time, and the subspace recognition method is also fast with only linear operations. Currently the algorithms are tested on a personal computer of Pentium IV, 3.2G Hz frequency, with 1G memory. It processes around 10 frames per second. (This frame rate can be set lower in real applications.) With optimization of the algorithm to PDA hardware, the real implementation of the face verification system on a PDA device will be possible.

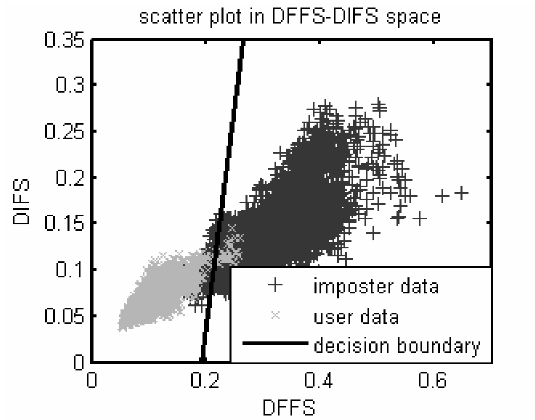

Fig 4. The DFFS and DIFS distribution of the user and imposters

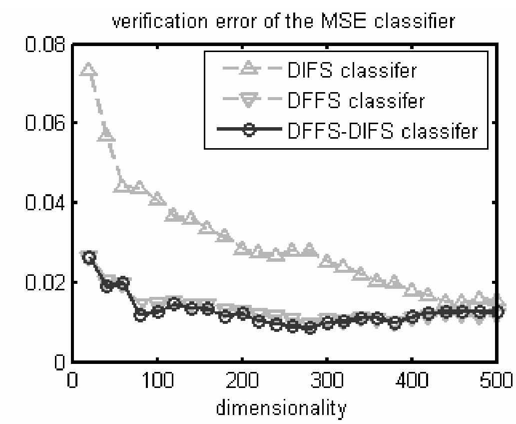

Fig 5. The error of the MSE classifier

V. Conclusions

For the verification system on a PDA, the fast face detection, registration, and verification algorithms are described. A personal face space unique to this user is built up. This is done by doing PCA of the face feature vectors, which is obtained during the enrolment phase. The face variations due to small registration shifts are additionally modelled, to improve the accuracy of recognition, and reduce the enrolment efforts. The experiments show that personal face space is an efficient representation in the verification case.

VI. Acknowledgement

The authors are grateful to the support of Freeband PNP2008 project.

References

[1] M.A. Turk and A.P. Pentland, “Face Recognition Using Eigenfaces”, Proc of IEEE Conf on Computer Vision and Pattern Recognition, June 1991, pp. 586-591.

[2] P.N. Belhumeur, J.P. Hespanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", European Conf. Computer Vis*ion,* 1996, pp. 45-58.

[3] P. Viola and M. J. Jones, “Robust Real-time Object Detection”, In Proc. of IEEE Workshop on Statistical and. Theories of Computer Vision, 2001.

[4] T. F. Cootes, G. J. Edwards, and C. J. Taylor, “Active Appearance Models”, Proc. 5th European Conference on Computer Vision, pp. 484 - 498, 1998.

[5] G.M. Beumer, Q. Tao, A.M. Bazen and R.N.J. Veldhuis. “Comparing landmarking methods for face recognition”, Proc. ProRISC 2005 16th Annual Workshop,Veldhoven, The Netherlands, Nov 2005, pp. 594-597.

[6] B. Moghaddam, A. Pentland. “Probabilistic Visual Learning for Object Representation”, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.19, No.7, July 1997, pp. 696-710.

[7] HumanScan, “BioID face database,” http://www. humanscan.de/.

- R.O. Duda, P.E. Hart, and D.G. Stork, Pattern classification, John Wiley