Gaussian Mixture Model Coupled with Independent Component Analysis for Palmprint Verification

Raghavendra.R, Bernadette Dorizzi, Ashok Rao, Hemantha Kumar G

Abstract In this paper we present a new scheme for Palmprint verification. The proposed method can be viewed as a combination of Gaussian Mixture Model (GMM) followed by Independent Component Analysis (ICA I and ICA II) applied directly on the pixels. This approach follows the path opened by previous works making use of GMM followed by Principal Component Analysis (PCA) and Linear Discriminate Analysis (LDA) projection methods for face recognition and is expected to be efficient to tackle major variations in the data. Extensive experiments have been carried out on PolyU palmprint database. We show that ICA I performs better than PCA, LDA and ICA II in the non Mixture Model case, while in the Mixture Model case, ICA II MM outperforms the three other Mixture Models (PCA MM, LDA MM, ICA I MM). Moreover, using artificially corrupted data (noisy palmprints), we show the robustness of Mixture Model approaches to noise, a property which can be very interesting in realistic situations.

1 Introductiong

Gaussian Mixture Model (GMM) has continued to receive a lot of attention over years [2]. GMM provides a mathematical approach to statistical modeling of a wide variety of random phenomena and also provides semi parametric framework

Raghavendra.R
Institut TELECOM, TELECOM and Management SudParis, Evry, France and University of Mysore,India.

Bernadette Dorizzi
Institut TELECOM, TELECOM and Management SudParis, Evry, France.

Ashok Rao
Mentor, C.I.T, Gubbi, India.

Hemantha Kumar G
University of Mysore, Mysore, India

to model an unknown probability density distribution shapes. The combination of GMM with linear projection schemes is becoming popular as it provides more than one transformation matrix. Here the idea is to capture the variability in data while keeping a linear projection method for feature extraction. Two powerful approaches making use of GMM followed by linear projection methods such as Principal Component Analysis (PCA) and Linear Discriminate Analysis (LDA) are currently in vogue [3][4] for face verification. In this paper, we want to investigate two other projection methods ICA I and ICA II on the different transformation matrices obtained using GMM. Indeed ICA will allow one to represent data in terms of statistically independent variables, it also captures higher order statistics (phase spectrum) with a set of non orthogonal vector basis. We chose to experiment this approach on palmprint because this biometric has several advantages compared with other hand biometrics such as fingerprints [1]. It also presents some variability in illumination by example which is of interest to be tackled using GMM models.

Indeed palmprint contains more information than fingerprint and this way it is expected to be more distinctive. Palmprint can be captured using low resolution devices (as low as 200 dpi), it is invariant with time and is widely accepted by end users. Palmprint contains many features like geometry features, wrinkles, ridges, principal lines etc. In the literature[1], there have been many approaches proposed for palmprint Verification/Identification. Kong et al.[1], used a texture based approach where 2D Gabor filter is used, which have shown better performance than Principal line extraction method. As palmprint contains more distinctive features such as principal lines and wrinkles, much of the work is reported on extraction of the principal lines. Zhang et al.[5], extracted palm lines using twelve templates. Han et al.[6], proposed principal line extraction using Sobel and morphological operation. Hu et al.[7], uses Principal Component Analysis (PCA), PCA and Linear Discriminate Analysis (LDA), PCA and Locality Preserving Projection (LPP) and 2DLPP for palmprint recognition. They show that 2DLPP outperform all other methods. Zang et al.[1], use Fourier transform to extract frequency domain features of palmprint and obtain improved result. Shang et al.[8], proposed the use of Fast ICA on palmprint images. Both ICA I and ICA II architectures are explored and finally the classification is done using RBF neural networks.

In this paper, we propose two novel schemes called ICA I Mixture Model (ICA I MM) and ICA II Mixture Model (ICA II MM) for building a palmprint verification system. As our concern is mainly to show the improvement that can be expected from the joint use of GMMs and ICAs, we work directly on the pixels and do not perform any prior feature extraction on the images. This way, our system will not be optimal compared to state of the art, but we can nevertheless perform some experiments showing the superiority of our approach compared to non mixture models. We also show improvement in performance using ICA I or ICA II instead of PCA or LDA associated to MM. Extensive experiments are carried out on PolyU palmprint database to prove the efficacy of proposed scheme.

The rest of the paper is organized as follows: Section 2 describes the palmprint verification system designed using the proposed methods, Section 3 describes Experimental set up, Section 4 describes results and discussion and Section 5 draws the conclusion.

2 Proposed Method

Figure 1 shows the block diagram of our palmprint verification system. The proposed system contains three important steps: Preprocessing, Feature extraction and Classification.

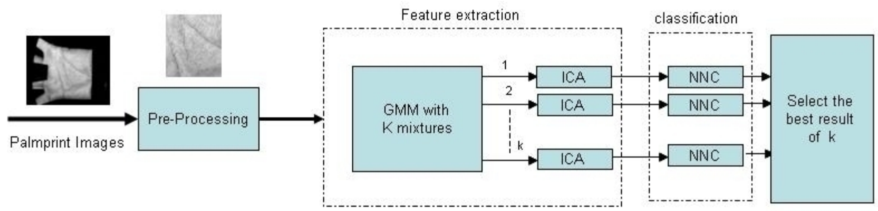

Fig. 1 Block Diagram of Proposed Palmprint Veification System

2.1 Preprocessing

Before feature extraction, it is necessary to obtain a sub-image from the captured palmprint image, and to eliminate the variations caused by rotation and translation. The sub image is also called ROI. This region contains prominent information about palmprint such as: Principal lines, Ridges, Wrinkles etc. Hence, quality of ROI has a great influence on the accuracy of the palmprint verification system. In this paper, we use the method given in [1] to extract the ROI.

2.2 Feature Extraction

This section describes the proposed methods namely ICA I Mixture model(ICA I MM) and ICA II Mixture Model(ICA II MM) for palmprint verification. In our feature extraction model, we first learn the different Gaussian mixtures as done classically in the literature[2]: in practice after determining the right number of mixtures, we use the EM algorithm in order to learn the GMM parameters using the learning(training) database as defined in Section 3. Usually, GMM is used for likelihood estimation of data, but in our approach we use GMM to obtain different transformation (covariance) matrices. Thus, a GMM with k mixtures gives k different transformation matrices. For each transformation matrix, we determine the associated ICA transformation(ICA I & ICA II separately).

2.2.1 Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a statistical signal processing technique and belongs to a class of Blind Source Separation (BSS) methods for separating data into underlying information components. According to literature [9], there are two ways in which ICA architecture can be implemented in image recognition task. In Architecture I (ICA I) input images in X are considered as a linear mixture of image of a statistically independent basis S combined with an unknown mixing matrix M. The ICA I algorithm learns the weight matrix w that corresponds to the coefficients of the linear mixture [9]. Architecture II (ICA II) finds the statistically independent coefficients for input data. The source separation is performed on the pixels, and each row of learned weight matrix w is an image of M, the inverse matrix of w contains the basis images in the column [9]. In practice, ICA II separates the data taking into account higher statistics while ICA I addresses the variation up to second order statistics. Details of ICA I & ICA II are discussed in [9].

2.3 Classification

After projecting the k transformation matrices using ICA (ICA I & ICA II separately), we do the classification using Nearest Neighbor Classifier (NNC). That is, for each test and training images, we calculate k distances (k is the number of mixtures) using NNC and, at the end, we select the transformation matrix that gives the minimal distance (see Figure 1).

2.4 Model Order

Before using the Mixture model, one has to determine the number of mixture components i.e number of mixtures. Choosing few components may not accurately model the distinguishing features present in the considered modality (i.e palmprint). Also, choosing too many components may over fit the data and reduce the performance and also result in excessive computational complexity both in training and classification. In our experiments, we find the model order by cross validation. Given a training dataset, we evaluate the performance over different numbers of mixture components. We then select the number of mixture components which give the best performance.

3 Experimental set up

The proposed algorithms are validated on polyU palmprint database[1]. The polyU palmprint database contains 7752 gray scale images corresponding to 386 different palm images. Each palm contains twenty samples, out of which ten samples are taken in first session and another ten are taken in second session. The average interval between the first and second session is two months. From this database, we have selected palmprint images corresponding to 150 different users. For training, we selected a total of six samples such that three samples are taken from first session and remaining three are taken from second sessions. Thus, we have 900 palm images for training. For testing, we have chosen 2 samples from each session resulting to a total of 600 palmprint images. We conducted two different experiments:

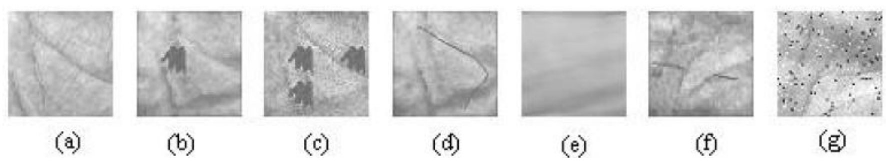

Fig. 2 (a) Clean palmprint (b)-(g). Noisy Palmprints

1. Experiment on clean data (Database I): We use good quality samples to evaluate the performance of proposed algorithms.
2. Experiment on Noisy data (Database II): We use low quality(noisy) samples to evaluate the performance of proposed algorithms. For obtaining low quality samples we artificially introduce five different types of noise such as smudging, lines across & along principal lines, salt & pepper and blurring effect. These noises simulate two different real time conditions. First, blurring of the palmprint image and salt & pepper noise could indicate improper maintenance of the equipment. Second, smudged palmprints and lines indicate improper biometric traits presented by the user. These noisy palmprints along with clean palmprint are shown in Figure 2.

4 Results and Discussion

This section describes the results obtained using proposed algorithms on two different palmprint databases. On each of these databases the results of proposed algorithms are compared with other algorithms such as LDA Mixture model[4], PCA Mixture Model[3] and also with non mixture model approaches (PCA, LDA, ICA I & ICA II).

4.1 Results on Database I

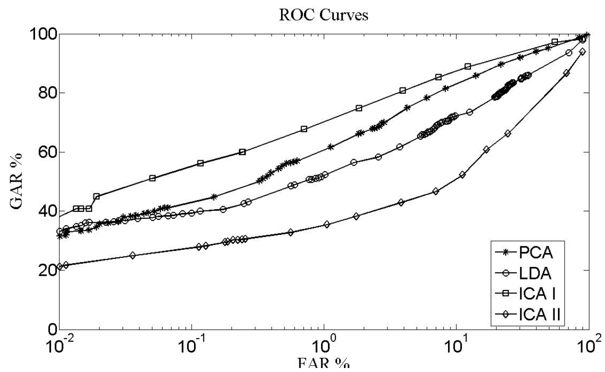

Fig. 3 Performance of Non Mixture Model Approaches on Database I

Figure 3 shows the performance of Non-Mixture Model approaches on clean palmprint samples(Database I). It can be observed that ICA I outperforms all other methods with $GAR = 39.86\%$ at $FAR = 0.01\%$. Figure 4 shows the performance of proposed ICA I MM and ICA II MM along with LDA MM and PCA MM. Here, the best performance is observed for ICA II MM with $GAR = 62.14\%$ at $FAR = 0.01\%$. In order to explain why ICA II is performing better with GMM while it is the worst method in the Non Mixture Model case, we suggest the following interpretation. We think that the GMM will effectively model the higher order statistics while ICA II is able to properly address these higher order statistics presented by GMM. Thus, the combination of GMM followed by ICA II will capture higher order stastistical information that is richer than information used in PCA MM, LDA MM & ICA I MM. For this reason, the proposed ICA II MM shows best performance over the other methods. It is also observed from Figures 3 & 4 that performance of mixture model based approaches are better than Non Mixture Model approaches(average increase of about 22.7% in GAR at $FAR = 0.01\%$).

4.2 Results on Database II

The robustness to noise of the proposed algorithm is validated on Database II. Figure 5 and 6 show the performance of Non Mixture Model along with Mixture Model approaches on the noisy database(Database II). Similar to what occurred on Database I in the Non Mixture Model case, here also ICA I gives the best result with $GAR = 24.2\%$ at $FAR = 0.01\%$ (see Figure 5). Figure 6 shows the performance of proposed ICA I MM and ICA II MM along with LDA MM and PCA MM. Here, it is observed that, ICA II MM outperforms other methods with $GAR = 39.3\%$ at

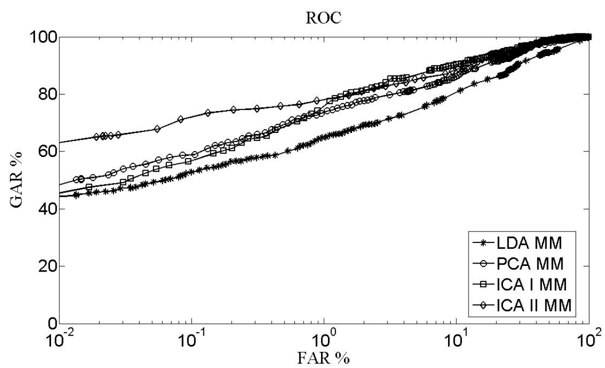

Fig. 4 Performance of Mixture Model Approaches on Database I

$FAR = 0.01\%$ as was already the case for non noisy data (Database I). However the gap in performance (in terms of GAR) between all the algorithms (in both Mixture Model & Non Mixture Model approaches) at low FAR (at $FAR = 0.01\%$) is less improved than in Database I. Let us note, however, a global degradation of the noisy

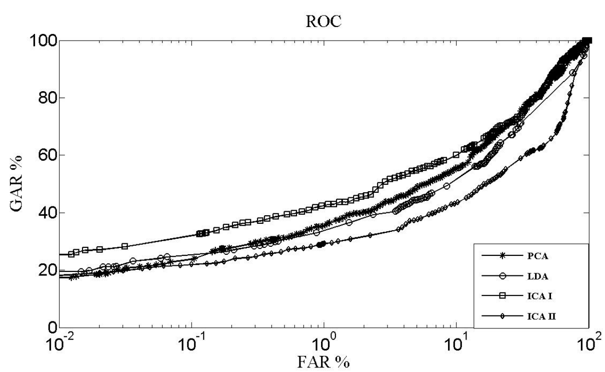

Fig. 5 Performance of Non Mixture Model Approaches on Database II

results compared to those of the clean case for both Mixture Model & Non Mixture Model approaches. Note that the use of Mixture Model based Approaches on noisy data gives a high level of performance (decrease of only roughly 20% as compared to clean case). This is very important in practical application.

5 Conclusion

In this paper we propose two methods for feature extraction namely ICA I MM and ICA II MM. We have conducted extensive experiments on both clean and noisy

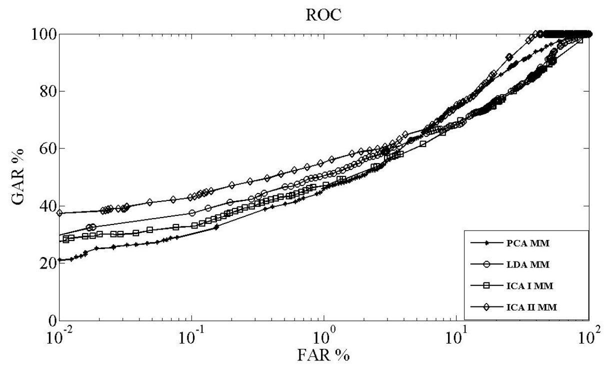

Fig. 6 Performance of Mixture Model Approaches on Database II

databases to prove the efficacy of the proposed methods. Experimental results on both type of data show that ICA II MM out performs the three other mixture model approaches(PCA MM, LDA MM & ICA I MM) which is due to the appropriate use of higher order statistics present in the data. Our experiments also show that Mixture Model approaches out performs Non Mixture Model approaches as they use more than one transformation matrices. This is specially true in the noisy case which is important in practice.

References

1. David D, Zhang.: Palmprint Authentication. Springer-Verlag, 2004.
2. Geoffrey J. McLachlan, David Peel.: Finite Mixture Models. Wiley Series in Probability and Statistics, 2000
3. Hyun-Chul, Kim., Daijin, Kim., Sung Yang, Bang.: Face recognition using the mixture-of-eigenfaces method. Pattern Recognition Letters, 23, 1549-1558 (2002)
4. Hyun-Chul, Kim., Daijin, Kim., Sung Yang Bang.: Face recognition using LDA mixture model. Pattern Recognition Letters , 24, 2815-2812 (2003)
5. Zhang,D., Kong,W., You,J., Wong,M.: Online palmprint identification. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25, 9, 1041-1050 (2003)
6. Han,C C., Cheng, H L., Lin, C L., Fan, C K.: Personal authentication using palmprint features. Pattern Recognition, 37, 10, 371-381 (2003)
7. Dewen, Hu., Guiyu, Feng., Zongtan, Zhou.: Two-dimensional locality preserving projections (2DLPP) with its application to palmprint recognition. Pattern Recognition, 40, 339-342, (2007)
8. Li, Shang., De-Shuang, Huang.,Ji-Xiang, Du., Zhi-Kai, Huang.: Palmprint Recognition Using Fast ICA and Radial Basis Probabilistic Neural Network. Nerocomputing, 69, 1782-1786 (2006).
9. Marian Stewart, Bartlett.,Javier, R Movrllan.,Terrence, J Sejnowski.: Face Recognition by Independent Component Analysis. IEEE Transaction on neural networks, 13 06, 1450-1464 (2002)