Incorporating Cohort Information for Reliable Palmprint Authentication

Ajay Kumar
Biometrics Research Laboratory, Department of Electrical Engineering
Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India
Email: ajaykr@ieee.org

Abstract

This paper presents a new approach to achieve the performance improvement for the traditional palmprint authentication approaches. The cohort information is used in the matching stage but only when the matching scores are inadequate to generate reliable decisions. The cohort information can also be utilized to achieve the significant performance improvement for the combination of modalities and this is demonstrated from the experimental results in this paper. The rigorous palmprint authentication results presented in this paper are the best in the literature and confirm the utility of significant information that can be extracted from the imposter scores. The statistical estimation of confidence level for the palmprint matching requires an excellent match between the theoretical distribution and the real score distribution. The performance analysis presented in this paper, from over 29.96 million imposter matching scores, suggests that Beta-Binomial function can more accurately model the distribution of real palmprint matching scores.

1. Introduction

The hand-based biometric authentication has higher user-acceptance and receiving increasing interest in recent years. The palmprint images have larger area and thus more abundant features which are quite unique even among identical twins. The usage of these systems for large scale personal authentication requires further efforts to achieve significant performance improvement. In this work my efforts are focused to achieve further performance improvement on the promising approaches presented in the literature. Daugman [8] has presented an optimal method for iris representation based on the distribution of a Hamming distances from the irisCode. This distribution is estimated from more than 9 million different iris comparisons and shows a Binomial trend which peaks at a normalized hamming distance of 0.5. Despite the popularity of palmprint biometric, there has not been any effort to establish the theoretical distribution of real palmprint matching scores to ascertain the confidence in decisions.

1.1 Prior Work

Several approaches for palmprint authentication using line features, appearance-based features, and texture-based features have been presented in the literature [1]-[6]. Recent comparative study of palmprint authentication approaches in [5] has suggested that the ordinal representation delivers best performance on peg-free palmprint database. The authors in [5] have also compared the performance of the ordinal features with those based on Gabor phase encoding presented earlier in [4], [6] and illustrated promising results on PolyU palmprint database [12]. The hand geometry features can be simultaneously extracted from the hand images and utilized to enhance the palmprint identification performance [10]. The usage of such approach for hand identification using contactless imaging has been illustrated in [3]. Jain and Demirkus [2] have recently presented a new approach for automated palmprint matching using latent palmprints. Authors have suggested to utilized the friction ridges, flexion creases and minutiae points that can be simultaneously extracted from 500 dpi palmprint images.

1.2 Proposed Approach

In this paper a new approach is investigated to improve the performance of traditional palmprint identification systems that have been presented in the literature. The performance improvement is investigated by integrating cohort information in the decision making. The experimental results are presented on the publicly available palmprint database [12] from 386 different palms and also from the publicly available *touchless palmprint database* [15] from 235 subjects. This work also details the performance improvement results using the cohort information for the score level combination of bimodal dataset from which hand geometry and palmprint features are simultaneously extracted. The experimental results suggest significant improvement in the performance as compared to the previously presented approaches in the literature [4]-[6]. Another aspect of this work is focused on the accurate estimation of theoretical model for the palmprint matching scores. The rigorous experimental results from over 29.96 million comparisons on the publicly available database suggest that the Beta-Binomial distribution is the most appropriate model for the palmprint score characterization.

2. Integrating Cohort Information

The block diagram for the palmprint authentication system using cohort information is shown in figure 1. The acquired palmprint images from the every user are used to extract the (ordinal) features and matched with the corresponding user template (ordinal features). If this matching score S_m is less than the decision threshold (T) than the user is authenticated as genuine. However, we do

978-0-7695-3476-3/08 $25.00 © 2008 IEEE
DOI 10.1109/ICVGIP.2008.73

not reject this user if his score S_m is more than threshold but employ cohort information to ascertain if he/she is genuine or imposter. This requires additional computations of matching scores S_i ($\forall\ i = N - 1$) for the training data from all the cohort templates. The matching scores S_i between the two palmprint samples f_i^1 and f_i^2 from each of the $i = 1,\ldots, N$ users is defined as;

$$S_i = \Theta(f_i^1, f_j^2) \quad \text{for } i \neq j \tag{1}$$

where Θ denotes the matching scores between sample f_i^1 and f_i^2. If the score S_m is less than all the cohort matching scores Si, then the user is authenticated as genuine user. However, even if any of the cohort scores, *i.e.* S_i, is less than S_m then the user is authenticated as imposter user.

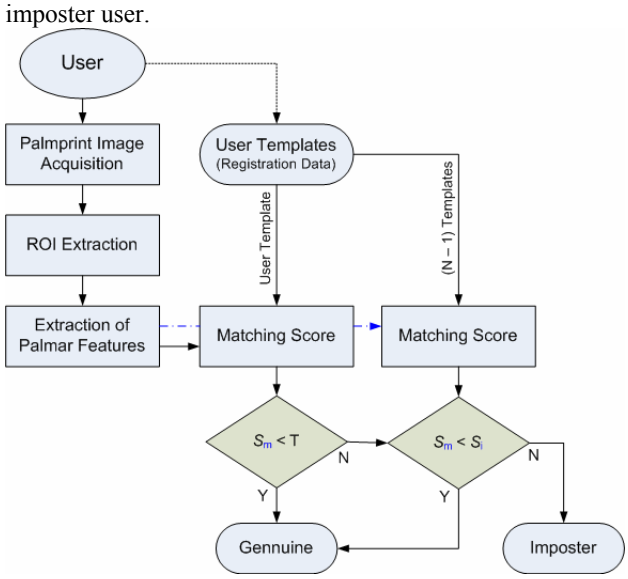


Figure 1: Palmprint authentication using cohort information.

3. Palmprint Score Distribution Model

The performance evaluation of biometric system requires accurate methodologies for assessing the distribution of real matching scores. The beta-binomial distribution is a generalization of binomial distribution and a more appropriate model for performance evaluation since it also incorporates the intra-user correlation between in matching attempts. The preference of Beta-Binomial distribution, over binomial distribution, has been illustrated in several publications [11], [13]-[14] and cited as most appropriate. However, the binary nature of outcomes from the matching decisions suggest that the Binomial distribution can be more appropriate for analyzing the biometric matching scores [8]. In this work, the accuracy of best fit model is ascertained from the empirical estimation of the mean square error from the best fit generated from (i) Beta (ii) Binomial, (iii) Beta-Binomial and (iv) Gaussian distribution.

Let us assume that the probability associated with each of ith individual $i = 1, 2, \ldots m$ is p_i. Each of these p_i's come from the conditionally independent draws that are characterized by beta distribution. The Beta distribution is characterized by two parameters (α and β) and its probability distribution is as follows.

$$f(p_i|\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p_i^{\alpha-1}(1-p_i)^{\beta-1} \tag{2}$$

The mean and variance for the above beta function $B(\alpha, \beta)$ is as follows:

$$E[p_i|\alpha,\beta] = \alpha/(\alpha+\beta) = \pi \tag{3}$$

$$Var[p_i|\alpha,\beta] = \pi(1-\pi)/(\alpha+\beta+1) \tag{4}$$

The Binomial distribution $Bin(n_i, p_i)$ for the m individuals, with each of the individuals tested n_i times with X_i number of successes, has the following probability functional form:

$$f(x_i|n_i) = \binom{n_i}{x_i} p_i^{x_i}(1-p_i)^{n_i-x_i} \tag{5}$$

The unconditional distribution of X_i having Beta-Binomial distribution has the following probability function:

$$f(x_i|n_i,\alpha,\beta) = \binom{n_i}{x_i} \frac{B(\alpha+x_i, \beta+n_i-x_i)}{B(\alpha,\beta)} \tag{6}$$

If we assume that the X_i are conditionally independent, then the mean and variance of beta-binomial distribution $Betabin(\alpha,\beta,n_i)$ can be estimated as follows:

$$E[X_i] = n_i\pi, \quad Var[X_i] = n_i\pi(1-\pi)\eta \tag{7}$$

where $\eta = (\alpha+\beta+n_i)/(\alpha+\beta+1)$ The detailed derivation of (7) and various fundamental properties of beta-binomial distribution have been described in [11], [13].

4. Experiments and Results

The performance improvement using cohort information for the palmprint biometric is regirously investigated in two separate set of experiments. Firstly, the peg-free hand image database that can also simultaneously extract the hand-geometry features is employed. The extraction of palmprint and hand-geometry images, feature extraction and the matching criteria employed is same as detailed in [3]. Each of the 300 × 300 pixels palmprint image is divided into 24 × 24 pixels overlapping blocks. The extent of this overlapping has been empirically selected as 6 pixels. Thus we obtain 144 separate blocks from each palmprint image. The standard deviation of discrete cosine transform (DCT) coefficients, obtained from each of the overlapping blocks, is used to characterize the palmprint region. The 17 features that can characterize every hand-shape images; *i.e.* perimeter, 4 finger length, 8 finger width, palm width, palm length, hand area, and hand length, are extracted. The first five images were

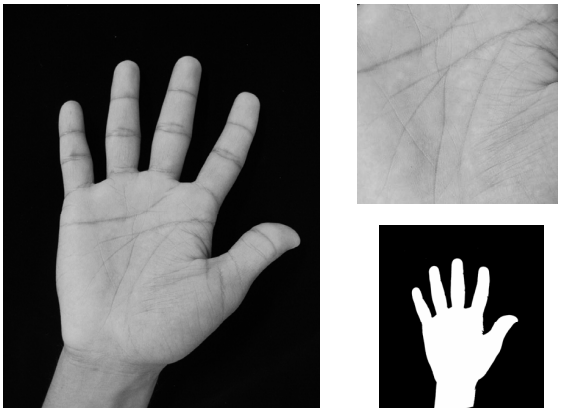


Figure 2: Acquired hand image (left), extracted palmprint (right-top) and corresponding hand geometry image (right-bottom)

employed for the training and rest images were employed for the testing. As detailed in section 2 (figure 1), the cohort information, from each of the imposter matches, was integrated in the decision making separately for the palmprint and hand geometry authentication. The receiver operating characteristics (ROC) from the palmprint matching for two cases, *i.e.* without and with the usage of cohort information, is shown in figure 3. Similarly, the ROC for the hand geometry authentication, with and without usage of the cohort information is shown in figure 4 (upper). In addition, the utility of the cohort information for the performance improvement using cohort information was also ascertained.

The palmprint and hand geometry matching scores were combined using hyperbolic product combination to ascertain the performance from the combination of two modalities. The hyperbolic product combination generates the combined score from $\tanh(s_p * s_h)$, where the s_p is the normalized palmprint scores and s_{hg} is the normalized hand geometry scores. This combination was empirically evaluated against the sum, product, weighted sum rule

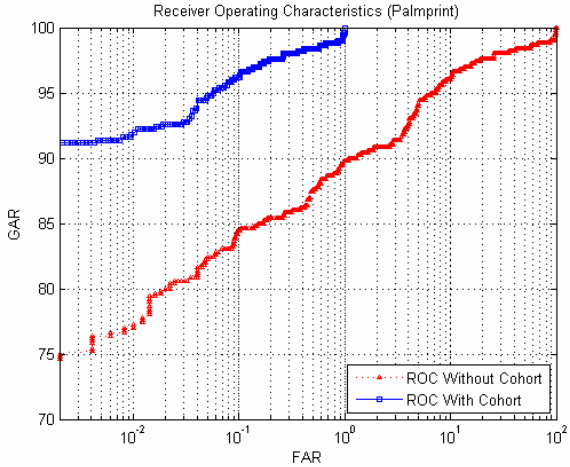


Figure 3: The performance from the palmprint features with and without the usage of cohort information.

Table 1: Improvement in Equal Error Rate using Cohort Information

	Palmprint	Hand Geometry	Palmprint + Hand Geometry
Without Cohort	5.6 %	6.4 %	2.6 %
With Cohort	0.96 %	0.85 %	0.40 %

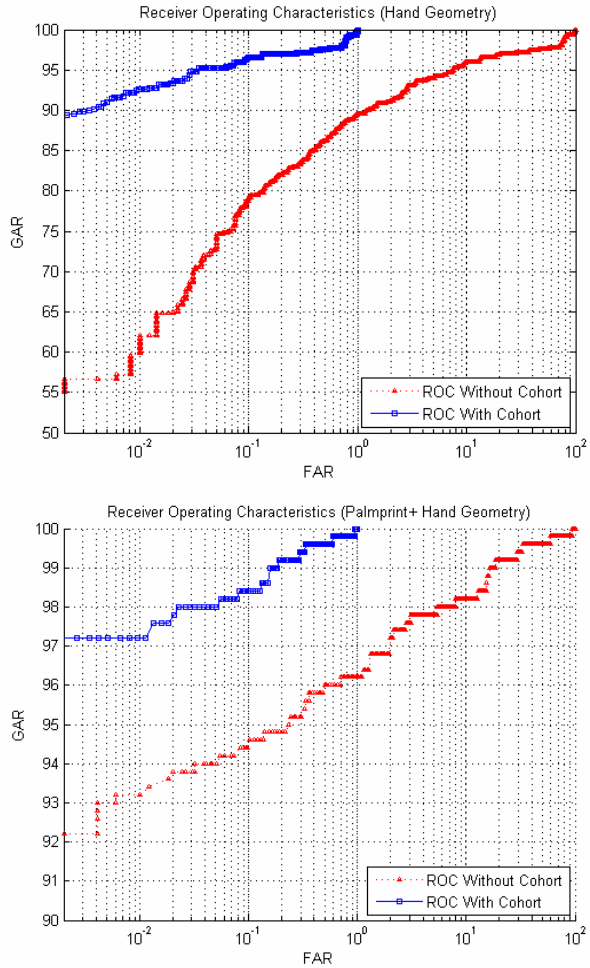


Figure 4: The performance from the hand geometry features (top), score level combination of palmprint and hand geometry features (bottom) using cohort information.

and found to generate better performance. The ROC from this score level combination, with and without usage of the cohort information is also shown in figure 4. The equal error rate from this first set of experiments is summarized in table 1. These experimental results consistently suggest significant improvement in the performance with the usage of cohort information.

The second set of experiments was exclusively focused to ascertain the performance improvement from the publicly available PolyU palmprint database [12] using three approaches. This database contains palmprint samples from the 386 palms (193 individuals) and the number of images employed from each of the users were kept uniform (18 samples) in these set of experiments.

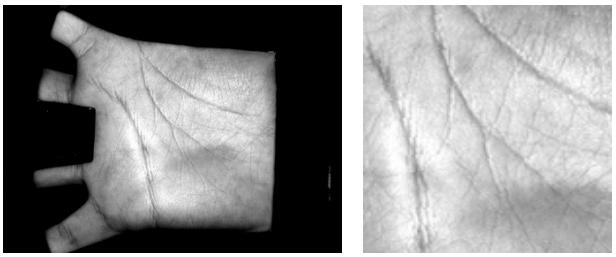


Figure 5: Image sample (left) from palmprint database in [12] and the extracted corresponding palmprint image (right)

The method of automatically extracting the region of interest from the available images was same as detailed and employed in [4]. The binarized feature extraction, from the phase-encoding of region of interest, using the OrdinalCode [5] PalmCode [4], and the CompCode [6] was investigated. The parameters of the Gabor filter and the method for extracting features was the same as detailed in [4] and [6]. However, the parameters for the Ordinal feature extraction were empirically selected (filter size 35 × 35, δ_x = 3, δ_y = 10) as the corresponding parameters available in [5] are selected of smaller version of this database (version 1.0).

The Hamming distances from these three approaches, using the parameters as detailed in the corresponding references, for all the possible genuine and imposter matches were generated while ensuring that no genuine and imposter pair matches are repeated [8]. Thus we employed 59058 (153×386) gennuines and 22737330 (18×385×386) imposters to investigate the performance improvement. The ROC from the method in [5], [6] and [4], along with the usage of cohort information, is illustrated in figure 6, 7, and 8 respectively. The equal error rate from each of these three cases is summarized in table 2. These experimental results are consistent and suggest that significant performance improvement can be achieved with the usage of cohort information.

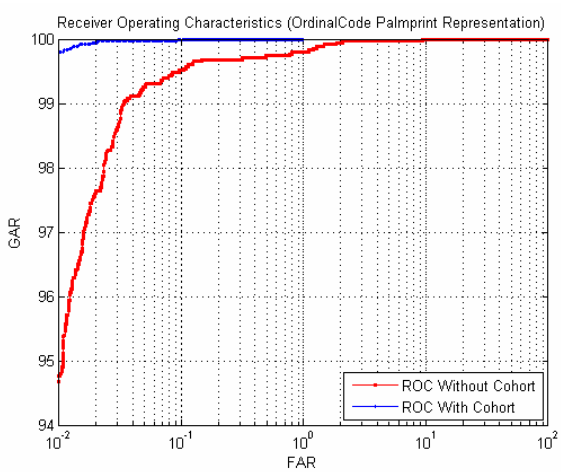


Figure 6: The performance from the *OrdinalCode* features with and without the usage of cohort information.

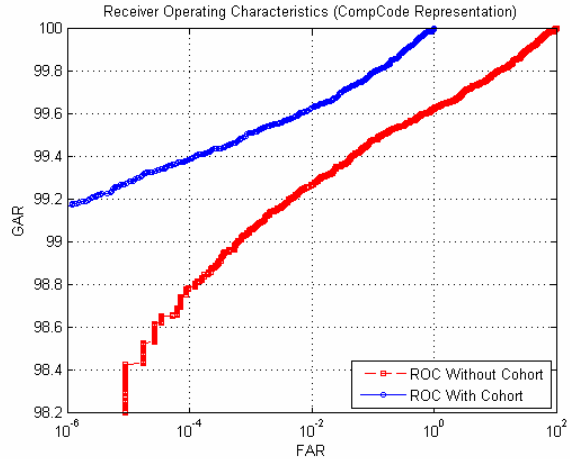


Figure 7: The performance from the *CompCode* features with and without the usage of cohort information.

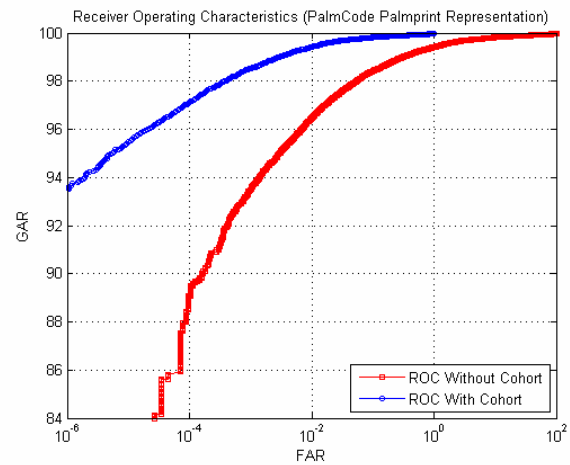


Figure 8: The performance from the *PalmCode* features with and without the usage of cohort information.

Table 2: Improvement in Equal Error Rate using Cohort Information

	PalmCode [4]	CompCode [6]	OrdinalCode [5]
Without Cohort	0.70	0.43	0.89
With Cohort	0.15	0.17	0.13

Another set of experiments were focused to investigate the performance improvement from peg-free and *touchless* palmprint image database acquired at IIT Delhi. This database has been acquired from 235 subjects using simple imaging setup detailed in [15]. The images in IITD palmprint database have high pose, translation, and scale variations resulting from unconstrained and touchless imaging. The inter-finger valleys (first and third) points are used as reference points for the extraction of palmprint region from the acquired images. The varying size palmprint region/images are automatically normalized to fixed size of 150 × 150 pixels. Figure 9 shows sample images from this database and the corresponding segmented images after the normalization.

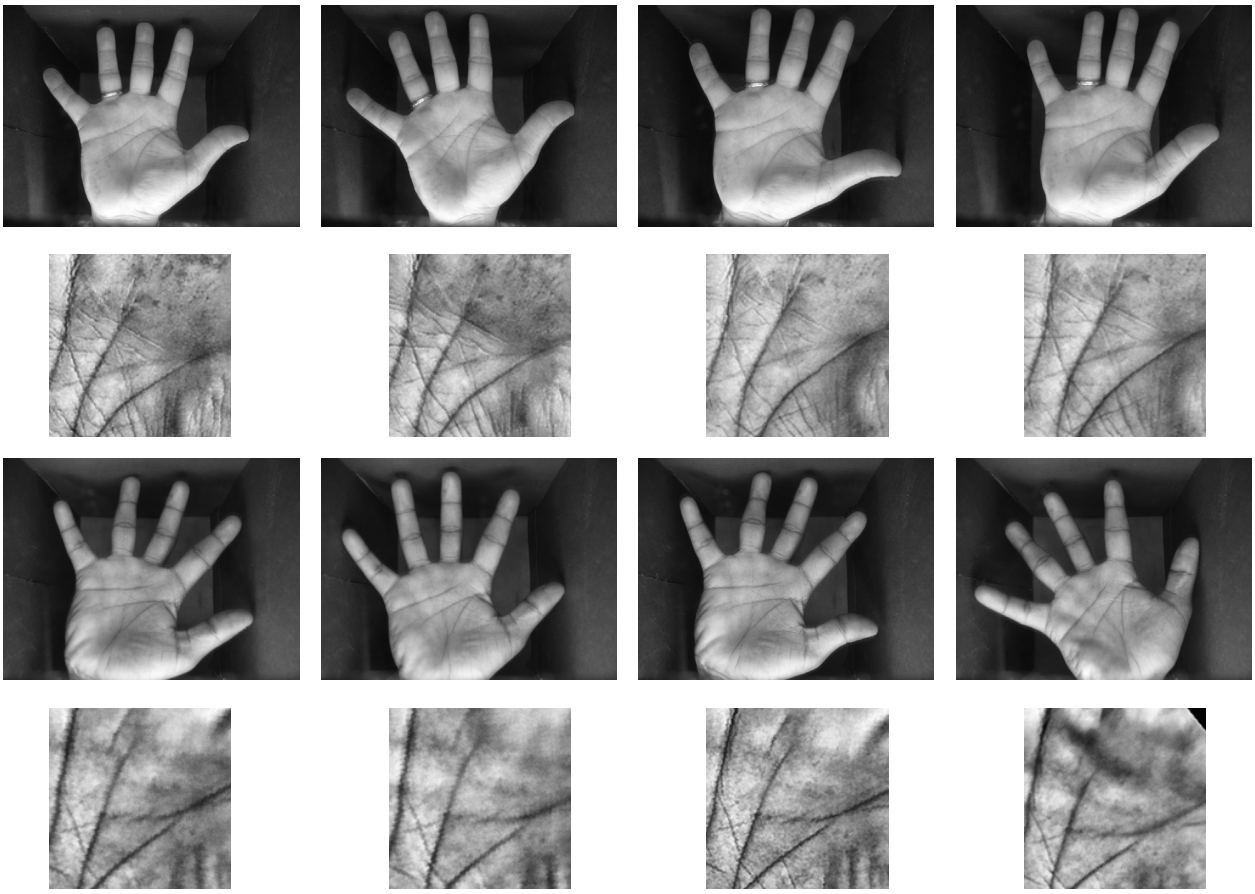


Figure 9: Acquired image samples from two subjects in touchless IITD palmprint database and the corresponding segmented images of 150 × 150 pixels after normalization.

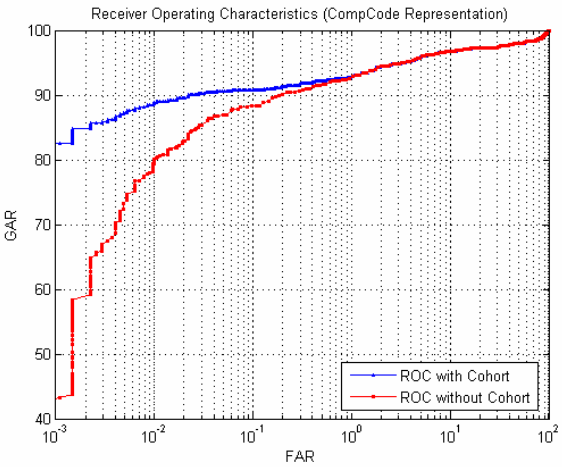


Figure 10: The performance from the *CompCode* features with and without the usage of cohort information.

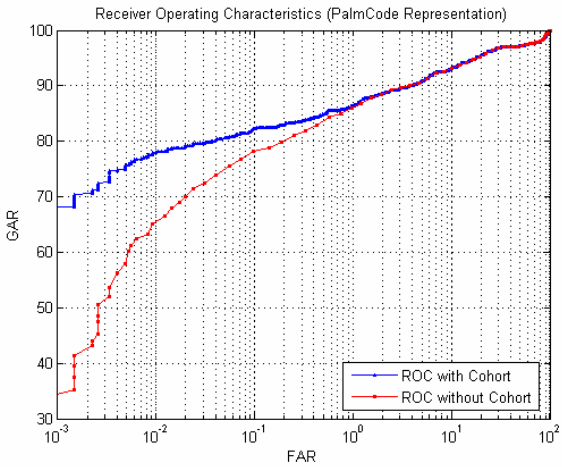


Figure 11: The performance from the *PalmCode* features with and without the usage of cohort information.

It can be observed from this set of sample images that the touchless imaging results in uneven image scale changes and the variation of pixel density in the palm region. Therefore varying details of palmprint texture are observed from the normalized images of same subject, unlike PolyU palmprint database which has negligible or

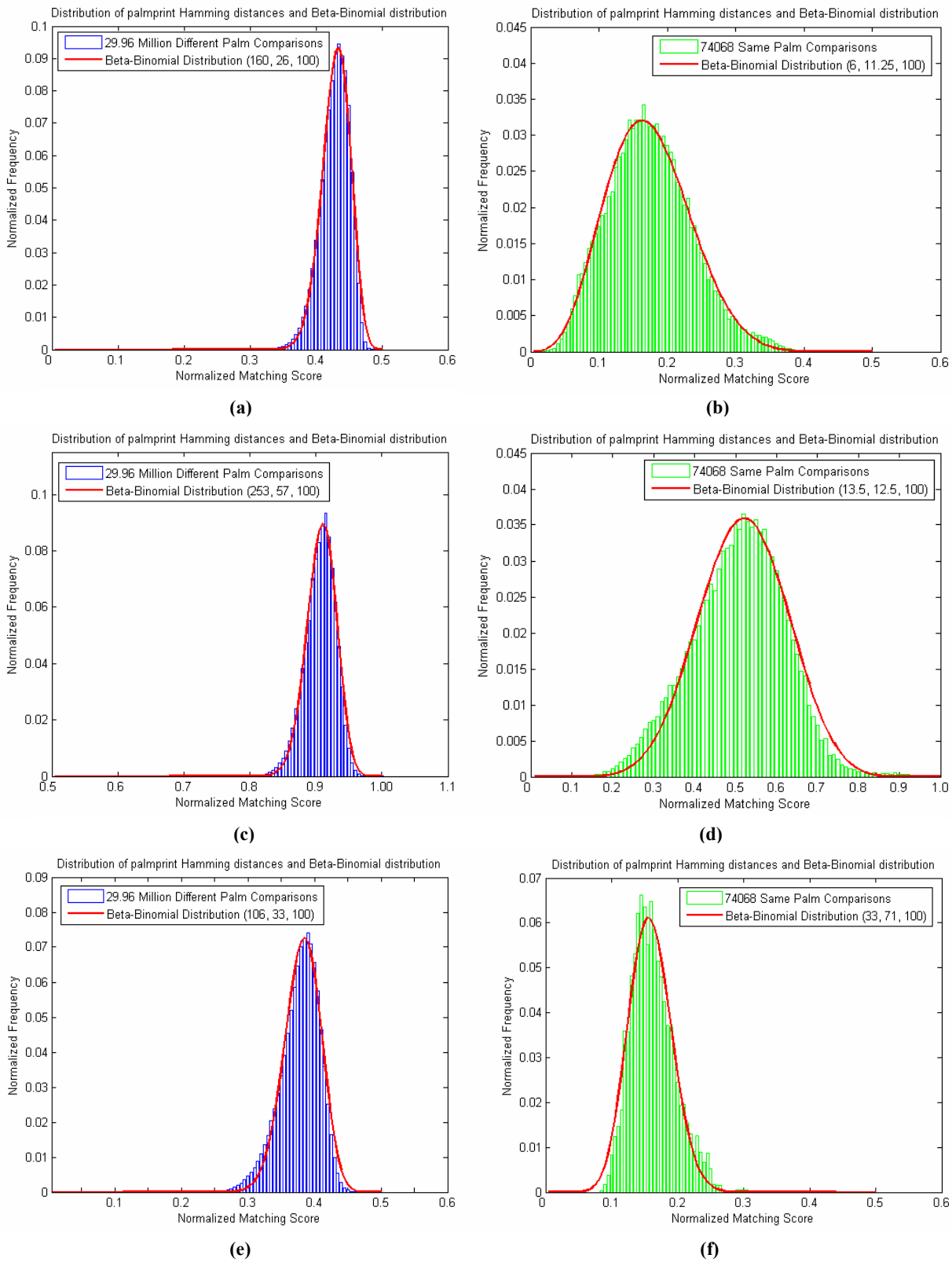
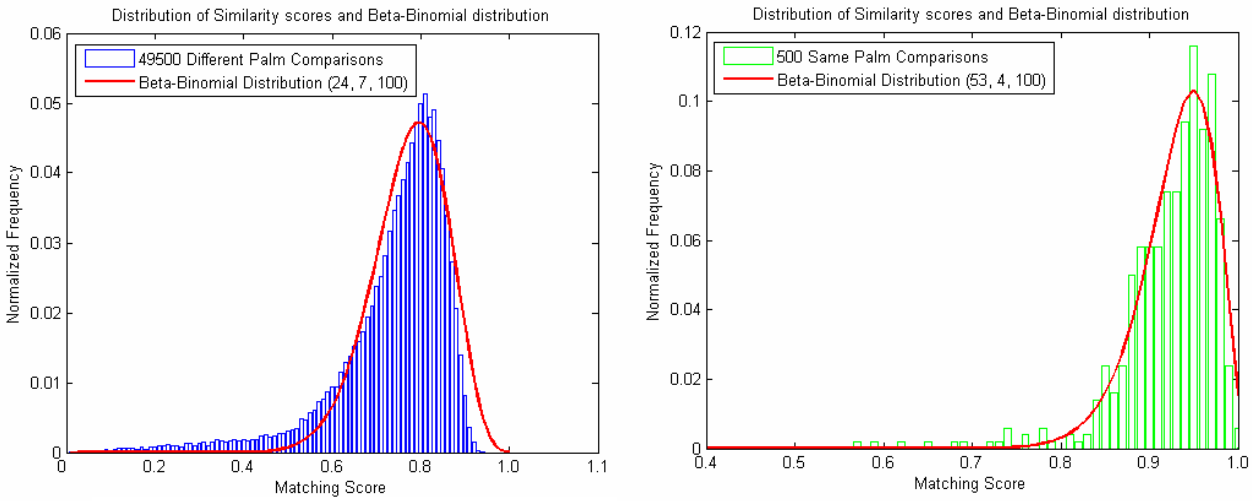


Figure 12: Estimation of matching score distribution for imposter and genuine matches from Ordinal representation in (a) and (b), from PalmCode in (c) and (d), from CompCode in (e) and (f) respectively.



(a) (b)

Figure 13: Estimation of matching score distribution for imposter and genuine matches from the DCT representation in (a) and (b) respectively.

Table 3: Norm of the error between the theoretical and actual score distributions

	Beta		Beta-Binomial		Binomial		Gaussian	
	G	I	G	I	G	I	G	I
Ordinal Codes [5]	0.0157	0.0567	**0.0088**	**0.0154**	0.1367	0.0674	0.0187	0.0409
PalmCode [4]	0.0172	**0.0099**	**0.0142**	0.0146	0.0999	0.0347	0.0161	0.0338
CompCode [6]	0.0265	0.0445	**0.0260**	**0.0197**	0.0457	0.0571	0.0402	0.0411
DCT Features [3]	0.0488	0.0626	**0.0463**	**0.0310**	0.1446	0.1276	0.1022	0.0780

G: Genuine, **I** - Imposter

very little scale variations. In this work, four images were employed for the training and the remaining image was used for testing. The average of test results, when each of the five images are used for testing, are reported in this work. The Gabor filters of size 65 × 65, centered at frequency 2√2, were employed to extract CompCode and PalmCode features. The performance from these two set of features is shown in shown in figure 10 and figures 11. It can be observed from this figure that the usage of cohort information significantly improves the performance, especially at lower FAR, which is most likely to be the operating point of the system usage. The experimental results from the OrdinalCode also illustrate significant performance improvement but are not reproduced due to space constraints.

The next set of experiments were focused to ascertain the best statistical model for the real distribution of genuine and imposter scores obtained from the palmprint biometric. In this set, all the available images from every subject, from the PolyU database [12], were employed to generate the maximum possible number of genuine and imposter matches. Thus all 74068 genuine and 29968808 imposter matching scores were generated fromall the 7752 palmprint images using the PalmCode, OrdinalCode and CompCode representations and were employed to generate the best fit plot from Beta, Binomial, Beta-Binomial and Gaussian models. The mean square error from the best set of parameters was computed to ascertain the best match. Thus the parameters for the each of these four models were different and empirically selected such that the mean square error between resulting distribution and the real palmprint matching score distribution is minimum. Figure 12 and 13 illustrate the best Beta-Binomial distribution plots, corresponding to minimum error, from the various palmprint matching scores. The parameters of Beta-Binomial function in these figures are in the following order: (α, β, n_i). The table 3 summarizes the error for the best fit parameters using the four models. It can be observed that for the genuine matching score distribution; the Beta-Binomial distribution generates the minimum error and hence offers the best model for palmprint feature representations

obtained from the same palm matching. The best fit error obtained from the imposter matching scores is minimum when the Beta-Binomial distribution is employed. However, the error for the imposter scores obtained from the PalmCode features is an exception in this as Beta distribution achieves marginally better error distribution error for the best fit imposter. The best set of experimental results (figure 7, table 2) are obtained from the CompCode features, for which the Beta-Binomial distribution achieved the best fit. In summary, the experimental results suggest that the Beta-Binomial distribution achieved the minimum error in most palmprint feature distributions, both for genuine and imposter matches, and therefore more appropriate model for the palmprint score distributions.

5. Conclusions

This paper has investigated a new approach to achieve the performance improvement for traditional palmprint authentication approaches presented in the literature. The performance improvement is achieved by integrating cohort information in the decision making. *The cohort information is used in the matching stage only when the matching scores are inadequate to generate reliable decisions*. The rigorous experimental results presented in this paper from the peg-free hand database and PolyU palmprint database illustrate significant improvement in the performance, *i.e.*, decrease in equal error rate by 78.57% for approach in [4], 60.47% for the approach in [6], and 85.39% for the approach detailed in [5]. Although the performance improvement is consistently significant, it comes with some added computational complexity but only for false rejects. The classifier has to perform additional matching operations, with all the imposters, every time the resulting matching score from the genuine user is more than the fixed decision threshold. An important aspect of this work is on the estimation of the accuracy of various theoretical models for the palmprint matching scores. The rigorous experimental results from over 29.96 million comparisons on the publicly available database suggest that the Beta-Binomial distribution is the most appropriate model for the palmprint score characterization.

This paper, for the first time in the biometrics literature, has presented experimental results from the touchless and peg-free palmprint database. The new touchless database [15] has significantly higher intra-class variations (translation, scale, and orientation) and is now freely/publicly made available for further palmprint research and development. The experimental results illustrated in this paper from the IITD palmprint database are from right hand images while results from the left hand image also achieved significant performance improvement, but could not be included (also the matching score distributions) due space constraints in this paper.

6. Acknowledgement

This work was partially supported by the research grant from Department of Science and Technology, Government of India (grant no. 100/IFD/1275/2006-2007). Author thankfully acknowledges the support of Mr. Anshu Vaid and all the volunteers in acquiring the touchless palmprint images for IITD palmprint database.

7. References

[1] P. H. Hennings-Yeomans, B. V. K. Kumar, and M. Savvides, , "Palmprint classification using multiple advanced correlation filters and palm-specific segmentation," *IEEE Trans. Info Forensics & Security*, vol. 2, no. 3, pp. 613-622, Sep. 2007.

[2] A. K. Jain and M. Demirkus, "On latent palmprint matching," MSU Technical Report, May 2008.

[3] A. Kumar and D. Zhang, "Personal recognition using shape and texture," *IEEE Trans. Image Process*., vol. 15, no 8, pp. 2454-2461, Aug. 2006.

[4] D. Zhang, W. K. Kong, J. You, and M. Wong, "On-line palmprint identification," *IEEE Trans. Patt. Anal. Machine Intell*., vol. 25, pp. 1041-1050, Sep. 2003.

[5] Z. Sun, T. Tan, Y. Yang, and S. Z. Li, "Ordinal palmprint representation for personal identification," *Proc. CVPR 2005*, pp. 279-284, 2005.

[6] W. K. Kong and D. Zhang, "Competitive coding scheme for palmprint verification," *Proc. ICPR 2004*, pp. 520-523, 2004,

[7] A. K. Jain, A. Ross, and S. Pankanti, "A Prototype hand geometry-based verification system", *Proc. of 2nd Internatinal Conference on Audio and Video-Based Biometric Person Authentication*, Washington DC, pp.166-171, Mar 1999.

[8] J. Daugman, "The importance of being random: Statstical principles of iris recognition," *Pattern Recognition*, vol. 36, no. 2, pp. 279-291, 2003.

[9] S. Pankanti, S. Prabhakar, and A. K. Jain "On the Individuality of Fingerprints," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 24, no. 8, pp. 1010-1025, 2002.

[10] A. Kumar and D. Zhang, "Hand geometry recognition using entropy-based discretization," *IEEE Trans. Info. Security Forensics*, vol. 2, pp. 181-187, Jun. 2007.

[11] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions*, 3rd edition, New York, Wiley, 2005.

[12] The PolyU Palmprint Database (version 2.0); http://www.comp.polyu.edu.hk/~biometrics

[13] M. E. Schuckers, "Using the beta-binomial distribution to access performance of a biometric identification device," *Int. J. Image & Graphics*, vol. 3, no. 3, pp. 523-529, 2003.

[14] E. T. Bradlow, P. J. Everson, "Bayesian inference for the Beta-Binomial distribution via polynomial expansions," *J. Comput. & Graphical Statistics,* vol. 11, no. 1, pp. 200-207, Mar. 2002.

[15] IITD Palmprint Database, http://web.iitd.ac.in/~ajaykr/Database_Palm.htm