

# Random Permutation-Based Linear Discriminant Analysis for Cancelable Biometric Recognition



P. Punithavathi and S. Geetha

**Abstract** The increased use of biometrics in the present scenario has led to the concerns over security and privacy of the enrolled users. This is because the biometric traits like face, iris, ear, etc., are not cancelable or revocable. In case if the templates are compromised, the imposters may gain illegitimate access. To resolve such issues, a simple yet powerful technique called “random permutation-based linear discriminant analysis” for cancelable biometric recognition has been proposed in this paper. The proposed technique is established on the notion of a cancelable biometric system through which the biometric templates can be revoked and renewed. The proposed technique accepts the cancelable biometric template and a key (called PIN) issued to the user. The user’s identity is recognized only when both cancelable biometric template and PIN are valid, else the user is prohibited. The performance of the proposed technique is demonstrated on the freely available face (ORL), iris (UBIRIS), and ear (IITD) datasets against state-of-the-art methods. The key benefits of the proposed technique are (i) classification accuracy remains unaffected by using random permutation and (ii) robustness across different biometric traits.

**Keywords** Biometric template security · Cancelable biometric recognition · Linear discriminant analysis · Random permutation · Template revocation

## 1 Introduction

Biometrics is a unique attribute possessed by every individual. It can be either physiological (e.g., face, iris, fingerprint, palmprint, etc.) or behavioral (e.g., gait, keystroke

---

P. Punithavathi · S. Geetha (✉)

School of Computer Science and Engineering, Vellore Institute of Technology, Chennai Campus, Chennai, Tamil Nadu, India

e-mail: [geethabaalan@gmail.com](mailto:geethabaalan@gmail.com)

P. Punithavathi

e-mail: [p.punithavathi2015@vit.ac.in](mailto:p.punithavathi2015@vit.ac.in)

dynamics, mouse dynamics, etc.) attribute to identify the user. With rapid technological advancements, biometrics have replaced passwords or personal identification number (PIN) in several access control applications like financial, healthcare, immigration, surveillance, etc. Principal component analysis (PCA) [1] and linear discriminant analysis (LDA) [2, 3] are two popular approaches used in the biometric recognition process. PCA is an unsupervised learning approach which learns eigenfaces from covariance matrix generated by unlabeled training data. PCA preserves the distributions in training data but has no information about the classification. Hence, it fails miserably in pattern recognition (PR) applications. On the other hand, LDA is a supervised learning approach. It generates optimal projection by maximizing the between-class distance and minimizing the within-class distance, thereby LDA is more influential than PCA in PR applications.

During biometric recognition, it is obvious that the dimension of the modality like face, iris, etc., is higher than the number of sample images in the central database. This leads to “Small Sample Size” (3S) problem [4]. The 3S problem can be alleviated by increasing the number of sample images per person or by using a dimension reduction technique. The former one becomes infeasible due to the storage cost, effort, and time to be spent in collecting several sample images, hence selecting a dimension reduction technique is better than increasing the number of sample images per person. Among several linear dimension reduction techniques, random projection (RP) [5, 6] is a feasible approach as it maps a set of points in a high-dimensional space to a new set of points in lower-dimensional space while approximately preserving the pairwise distances between them.

Cancelable biometric system (CBS) [7] is a template securing mechanism which generates cancelable biometric templates out of original biometric attributes for a user-specific key. The cancelable biometric template is generated out of original biometric attributes using a secure and discriminability-preserving transformation function and user-specific key. In case of a database breach, the cancelable biometric templates are alone compromised. Hence, the privacy of the biometric attributes is preserved during attacks. The CBS also provides an added advantage of generating numerous and diverse cancelable biometric templates for various applications just by changing the transformation function and/or user-specific key, thereby preventing privacy threats. In this way, CBS provides high level of security, privacy, and revocability to biometrics that may help to increase public confidence for acceptance of biometric-based systems.

## 2 Literature Survey

CBS is a biometric template securing technique in which enrollment and authentication are performed in the transformed domain. The CBS transformation techniques have been broadly classified into biometric salting and non-invertible transforms as in [8]. Apart from these two techniques, several other categories of CBS have emerged in recent times.

Biometric salting can be defined as a transformation technique which generates cancelable biometric templates by mixing in an artificial pattern. The mixing patterns can be a random/synthetic pattern or a pure random noise. Non-invertible transformation technique can be defined as a cancelable biometric template generation technique which uses a secret key as a constraint for the transformation function. Several other techniques have risen in recent times. For instance, hashing-based transformation [9–11] uses the index positions of the biometric feature template derived from several basic hashing techniques to generate cancelable templates. The Bloom filter-based transformations were introduced in [12]. The Bloom filters were initially introduced by Rathgeb et al. in [13–15] which map multiple code words to identical position to generate cancelable biometric templates.

The following research gaps have been identified from the literature survey.

- The transformation is mostly applied at feature level which becomes time. There exists a research gap for generating cancelable templates by applying the transforms at the signal level
- The 3S problem persists if the dataset contains less number of sample images
- The security, privacy, revocability, and diversity of the biometric information of user must be preserved simultaneously while achieving good recognition rate

This work proposes a novel cancelable biometric recognition technique called *Random Permutation-based Linear Discriminant Analysis (RPLDA) method* that addresses the above issues. The approach aims to generate secure, revocable, non-invertible, privacy-preserving, and performance preserving templates even in the stolen token scenario.

The research contributions of the proposed system have been listed as follows:

- The proposed cancelable biometric template generation system—RPLDA is capable of generating cancelable biometric templates which have been transformed at the signal level
- The 3S problem has been alleviated by employing LDA
- The recognition performance of RPLDA has been proved to be better in the transformed domain
- The RPLDA satisfies the basic properties of CBS such as non-invertibility, diversity, unlinkability, and revocability which are evident from the research outcomes

The rest of the paper is organized as follows. Section 3 explains the proposed RPLDA technique, analysis of the relationship between LDA and RPLDA and its applicability as a cancelable biometric system. Section 4 mentions the experimental setup. Section 5 describes the results. Section 6 gives a brief conclusion.

### 3 Proposed Technique

Intermediate templates are generated by employing a random permutation matrix on the given training images of the biometric traits. The random permutation matrix is chosen to be the PIN. The random permutation matrix is selected to be a matrix whose entries are “0” and “1”, distributed randomly. Through the proposed technique, the cancelable features in the intermediate templates are recognized using LDA [2, 3]. The finally extracted cancelable features are the discriminant vectors which comprise the cancelable template.

#### 3.1 Preliminaries of LDA

LDA is a popular feature extraction technique. The main objective of the LDA is to derive the direction along which the variance in the data is high.

Assume that  $x \in \mathbb{R}^d$  is a column vector and it represents each image in “d” dimensional space. If there are “c” users, i.e.,  $\{1, 2, 3, \dots, c\}$ , such that each user has  $N_i$  images. Thus, the total number of training images is given by Eq. (1)

$$N = \sum_{i=1}^N N_i \quad (1)$$

Let  $\hat{x}$  represent the mean image vector of the training data as shown in Eq. (2)

$$\hat{x} = \frac{1}{N} \sum_{i=1}^N N_i \hat{x}_i \quad (2)$$

where the value of  $\hat{x}_i$  is given by Eq. (3)

$$\hat{x}_i = \frac{1}{N_i} \sum_{i=1}^N x \quad (3)$$

The total scatter matrix ( $M$ ) for the training data is given by Eq. (4) as below:

$$M = M_w + M_b \quad (4)$$

where  $M_w$  and  $M_b$  are within-class scatter matrix and between-class scatter matrix, respectively as given by Eqs. (5) and (6).

$$M_w = \sum_{i=1}^c (x - \hat{x}_i)(x - \hat{x}_i)^t \quad (5)$$

$$M_b = \sum_{i=1}^c N_i (\hat{x}_i - \hat{x})(\hat{x}_i - \hat{x})^t \quad (6)$$

The criterion function ( $J$ ) for a projection matrix  $\psi = \{w_1|w_2|\dots|w_{c-1}\}$  is given by Eq. (7)

$$J(\psi) = \frac{|\psi^t M_b \psi|}{|\psi^t M_w \psi|} \quad (7)$$

For an optimal projection matrix, say  $\psi^*$  the eigenvectors corresponding to the largest eigenvalues and can be modeled as

$$\psi^* = \{w_1^*|w_2^*|\dots|w_{c-1}^*\}.$$

Thus, the transformed data points ( $T$ ) are given by Eq. (8)

$$T = \psi^t x \quad (8)$$

### 3.2 Proposed RPLDA Scheme

A random permutation matrix is projected on the sample image of the biometric traits of the users thereby generating an intermediate template which is a column vector. The random permutation matrix is chosen to be an involutory matrix to convert the given sample image into a column vector. The cancelable templates are generated by extracting the LDA features out of the intermediate template. Thus, the biometric patterns can be renewed or revoked whenever required just by changing the random permutation matrix.

If  $x$  is an input sample, then a random permuted image ( $x'$ ) or an intermediate template can be generated from the input sample  $x$  using a random permutation matrix  $R$ , as represented in Eq. (10).

$$x' = Rx \quad (10)$$

The LDA features of the intermediate template are determined to construct the cancelable template.

The various entities involved in the generation of the cancelable templates using the proposed RPLDA scheme have been illustrated in Fig. 1. The users are enrolled in an application during the enrollment phase. The users are verified during the verification phase. During the enrollment phase, the biometric image sensor captures the image samples of the biometric trait of the user. The random permutation matrix projection unit generates a random permutation matrix which is an involutory matrix.

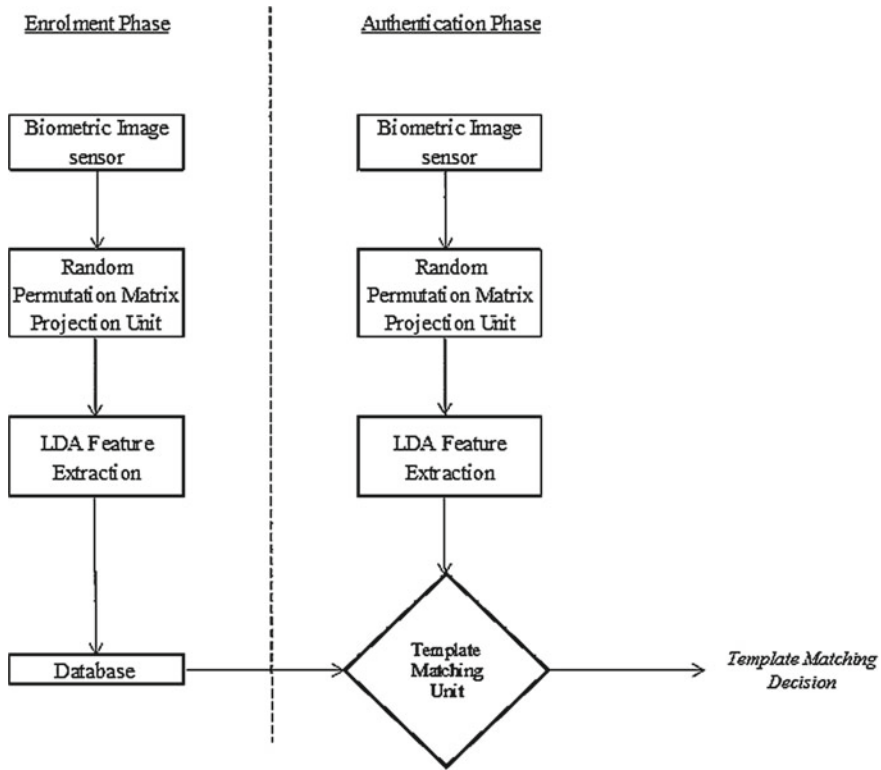
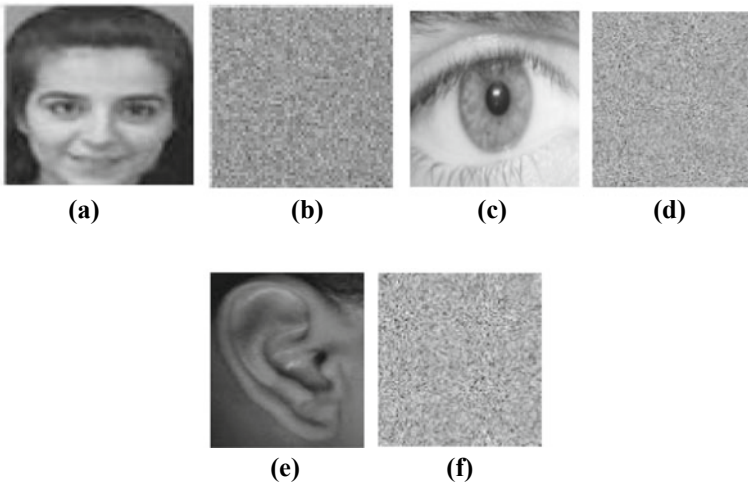


Fig. 1 Enrollment phase and authentication phase in proposed system

This random permutation matrix is projected on the image sample of the user to generate an intermediate template. The LDA feature extraction unit extracts the features from the intermediate template to generate a cancelable template which is stored in the database.

During the authentication phase, the user again produces the biometric trait to the biometric imaging sensor. The random permutation matrix projection uses the same matrix which was generated in the enrollment phase from the user to generate a query intermediate template. The LDA feature extraction unit extracts LDA features from the query intermediate template and generates a query cancelable template. This query template is matched with the cancelable template stored in the database by the matching unit to grant access to the application, in case if the templates match.

A sample face image and its corresponding cancelable template are shown in Fig. 2. The major contribution of the research work is the relationship between eigenvalues and eigenvectors of original biometric images and cancelable templates. The eigenvalues possessed by the original biometric patterns and the cancelable templates are the same. But the eigenvectors of the cancelable template are randomly



**Fig. 2** **a** Sample face image from ORL dataset; **b** cancelable template generated from a sample face image in **(a)** using proposed RPLDA; **c** sample iris image from UBIRIS dataset; **d** cancelable template generated from sample iris image in **(c)** using proposed RPLDA; **e** sample ear image from IITD dataset; **f** cancelable template generated from sample ear image in **(e)** using proposed RPLDA

permuted form of the eigenvectors of the original biometric images. The randomly permuted intermediate templates correspond to the random permutation matrix.

### 4 Experimental Setup

The cancelable template has been generated using proposed RPLDA technique using the iris, face and ear biometrics are chosen from publicly available datasets like UBIRIS [16], ORL [17], and IITD [18] databases, respectively. Table 1 displays the details of the databases.

The performance of the proposed techniques is evaluated in terms of the security provided by the proposed technique, average classification accuracy, and average training time of the algorithm. The training image from each identity is selected randomly and the remaining images are used as testing set. This process is repeated 40 times to achieve a stable classification accuracy and average training time. All

**Table 1** Summary of datasets used in experiments

Dataset	No. of subjects	Image size	Total
ORL	40	112 × 92	400
UBIRIS	241	150 × 200	1877
IITD	125	180 × 50	493

the experiments are performed on Intel Xeon E3 CPU 2.4 GHz with Windows 7 and 8 GB memory.

## 5 Results and Discussions

In this section, the performance of the proposed technique has been analysed both qualitatively and quantitatively with other state-of-the-art methods, viz. Gray-Salt (GS) PCA, Block-Remapping (BR) PCA, and RPPCA [19]. The cancelable templates were generated independently using each technique. The classification accuracy for each technique has been determined using nearest neighborhood technique. Then the classification accuracies have been compared with the classification accuracy of the proposed RPLDA technique.

The classification accuracy measures the percentage of identities correctly classified by some technique or algorithm. It depends on factors such as the number of training used and the number of dimensions in the transformed representation. The classification accuracy of the proposed RPLDA technique has been reported in Table 2. The classification accuracy of the proposed technique is higher when compared to the state-of-art techniques like GSPCA, BRPCA and RPPCA as shown in Table 2.

The equal error ratio (EER) is the metric which is used to measure the matching performance of a CBS. The EER value must be as low as possible to indicate that CBS has a good matching performance. The EER value of the proposed system is calculated and compared with the EER values of the RPPCA, GSPCA, and BRPCA. The EER comparative results have been listed in Table 3. It can be inferred from Table 3 that the EER of the proposed system is lower than the other state-of-the-art techniques,

**Table 2** Classification accuracy achieved by proposed RPLDA technique

Techniques	Classification accuracy (%)		
	ORL	UBIRIS	IITD
RPPCA	93.84	92.98	90.24
GSPCA	87.68	85.35	83.66
BRPCA	86.99	84.96	83.13
Proposed RPLDA	95.33	94.89	92.67

**Table 3** EER (%) achieved by the proposed RPLDA technique

Techniques	EER (%)		
	ORL	UBIRIS	IITD
RPPCA	6.72	6.28	8.33
GSPCA	11.43	14.13	15.39
BRPCA	12.38	17.38	14.36
Proposed RPLDA	4.21	5.43	6.52



**Table 4** Training time (in seconds) required by the proposed RPLDA technique

Dataset	No. of training images			
	2	3	4	5
	Training time (in seconds)			
ORL	0.165	0.269	0.612	1.478
UBIRIS	5.111	6.547	9.124	12.96
IITD	0.412	0.613	0.948	1.631

The average training time required by the proposed techniques is shown in Table 4. The RPLDA has been trained using only few training images of each user to alleviate 3S problem [20]. The average training time over 40 runs of the proposed algorithm was measured on all the datasets as shown in Table 4. It is observed that the training time of the proposed technique increases with an increase in the number of training images. It can be readily observed from Table 4 that the training time required by RPLDA is significantly less and feasible.

The cancelable templates generated using the proposed system are non-invertible. If a brute-force attack has been simulated against the cancelable templates generated using the proposed system, then the number of iterations required to reverse-engineer the cancelable templates is completely dependent on the input image size. It will take  $(112 \times 92)!$ ,  $(150 \times 200)!$ , and  $(180 \times 50)!$  for achieving a preimage of the original biometric image. It is very difficult to reverse-engineer the cancelable templates generated using the proposed system. Hence, the cancelable templates generated using the proposed system are non-invertible.

The cancelable templates generated using the proposed system are diverse. A user can register to different applications using different cancelable templates which correspond to a single biometric pattern. This is achieved just by changing the random permutation matrix. With the change in the pattern of the entries in the random permutation matrix, the cancelable templates also change. Assume if there are “ $n$ ” entries in a random permutation matrix, then “ $n$ ” different cancelable templates can be generated from the input biometric pattern.

## 6 Conclusion

In this research work, a simple yet powerful technique called as random permutation-based linear discriminant analysis (RPLDA) has been proposed. The proposed technique is applicable in cancelable biometric recognition. The users are required to provide their biometric trait and the PIN which is the random permutation matrix. The random permutation matrix is projected on the biometric image and an intermediate template is generated. The LDA extracts cancelable features from the intermediate template and generates the cancelable template. The effectiveness of the proposed technique has been illustrated by the results of the experiments conducted

on different datasets. The classification accuracy of the proposed technique is found to be better than the state-of-the-art techniques. The training time is also found to be very less. The proposed technique is also proved to be effective against 3S problem. The future study can be directed toward application of diagonal LDA for cancelable biometric recognition.

**Acknowledgements** The authors would like to thank the Management and Staff of Vellore Institute of Technology, Chennai Campus. The first author is supported by Visvesvaraya Ph.D. Scheme, sponsored by Digital India Corporation, held by the Ministry of Electronics and Information Technology (MeitY), Government of India.

## References

1. S. Wold, K. Esbensen, P. Geladi, Principal component analysis. *Chemometrics Intell. Lab. Syst.* **2**(1–3), 37–52 (1987)
2. K. Fukunaga, *Introduction to Statistical Pattern Recognition* (Academic, New York, 2013)
3. R.O. Duda, P.E. Hart, D.G. Stork, *Pattern Classification*, vol. 2 (Wiley, New York, 1973)
4. W.J. Krzanowski, P. Jonathan, W.V. McCarthy, M.R. Thomas, Discriminant analysis with singular covariance matrices: methods and applications to spectroscopic data. *Appl. Statistics* 101–115 (1995)
5. E. Bingham, H. Mannila, Random projection in dimensionality reduction: applications to image and text data, in *Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2001)
6. P. Punithavathi, S. Geetha, Dynamic sectored random projection for cancelable Iris template, in *2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (IEEE, 2016), pp. 711–715
7. N.K.C.J.H. Ratha, R.M. Bolle, Enhancing security and privacy in biometrics-based authentication systems. *IBM Syst. J.* **40**(3), 614–634 (2001)
8. H. Kaur, P. Khanna, Biometric template protection using cancelable biometrics and visual cryptography. *Multimedia Tools Appl.* **75**(23), 16333–16361 (2016)
9. A.B.J. Teoh, D.N. C. Ling, A. Goh, Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recogn.* **37**, 2245–2255 (2004)
10. A.B.J.C.S. Teoh, J. Kim, Random permutation Maxout transform for cancellable facial template protection. *Multimedia Tools Appl.* **77**, 1–27 (2018)
11. H. Li, J. Qiu, A.B.J. Teoh, Palmprint template protection scheme based on randomized cuckoo hashing and MinHash. *Multimedia Tools Appl.* **1**(1), 1–25 (2020)
12. J. Bringer, C. Morel, C. Rathgeb, Security analysis and improvement of some biometric protected templates based on Bloom filters. *Image Vis. Comput.* **58**, 239–253 (2017)
13. C. Busch, C. Rathgeb, F. Breiting, H. Baier, in *On the application of Bloom filters to Iris biometrics*. *IET Biometrics* (2014)
14. S. Ajish, K.S. AnilKumar, Iris template protection using double bloom filter based feature transformation. *Comput. Secur.* **97**, 101985 (2020)
15. M. Gomez-Barrero, C. Rathgeb, J. Galbally, J. Fierrez, C. Busch, Protected facial biometric templates based on local gabor patterns and adaptive bloom filters, in *2014 22nd International Conference on Pattern Recognition (ICPR)* (2014)
16. H. Proenca, L. Alexandre, UBIRIS: a noisy iris image database, in *13th International Conference on Image Analysis and Processing (ICIAP 2005)* (2005)
17. F. Samaria, A. Harter, Parameterisation of a stochastic model for human face identification, in *Proceedings of 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, FL (1994)

18. A. Kumar, C. Wu, Automated human identification using ear imaging. *Pattern Recogn.* **41**(5), 956–968 (2012)
19. N. Kumar, S. Singh, A. Kumar, Random permutation principal component analysis for cancelable biometric recognition. *Appl. Intell.* **48**(9), 2824–2836 (2018)
20. R. Bellman, *Adaptive Control Processes* (Princeton, Princeton University Press, 1961).