

LETTER

Discriminative Projection Selection Based Face Image Hashing

Cagatay KARABAT^{†a)}, Student Member and Hakan ERDOGAN^{††}, Nonmember

SUMMARY Face image hashing is an emerging method used in biometric verification systems. In this paper, we propose a novel face image hashing method based on a new technique called discriminative projection selection. We apply the Fisher criterion for selecting the rows of a random projection matrix in a user-dependent fashion. Moreover, another contribution of this paper is to employ a bimodal Gaussian mixture model at the quantization step. Our simulation results on three different databases demonstrate that the proposed method has superior performance in comparison to previously proposed random projection based methods.

key words: face image hashing, biometric security, privacy

1. Introduction

In recent years, biometrics has achieved wide-spread usage in various applications. In most of these applications, biometric templates are stored in databases and/or smart cards, thus raising questions such as data security and privacy. Various biometric hashing methods, which mostly depend on random projections, are proposed to protect the biometrics data [1]-[5] in the literature. See Figure 1 for an illustration of the basics steps of such biometric hashing methods.

Ngo *et al.*[1], [2] employ feature extraction methods (i.e. Principle Component Analysis (PCA), Fisher Linear Discriminant (FLD), Wavelet transformation, Wavelet Transform with PCA, Wavelet Transform with Fourier Mellin Transform etc.) to the face images and then make and use a random projection (RP) matrix for reducing the dimension of the feature vectors. Finally, they employ binary quantization to obtain face image hash vectors. We improve upon this method in our work.

In this work, we develop a new face image hashing method based on a proposed technique that we call "discriminative projection selection" to reduce verification errors. This technique selects the rows of an RP matrix, which is a user dependent dimension reduction matrix, by using the Fisher criterion [6]. Moreover, we employ Gaussian mixture model at the quantization step to obtain more distinct face image hash vectors for each user.

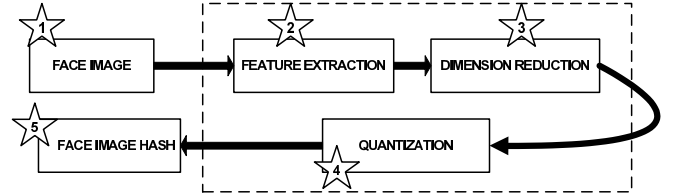


Fig. 1 Basic steps of the biometric hashing methods

2. The Proposed Biometric Verification Method

In this section, we introduce a biometric verification system which employs the proposed face image hashing method based on the discriminative projection selection technique.

2.1 Enrollment Stage

There are three main steps at the enrollment stage: (1) Feature extraction, (2) Dimension reduction, (3) Quantization.

2.1.1 Feature Extraction

At the feature extraction phase, we use two sets of data: training set and "others" set. The training set has training face images of registered users, $\mathbf{I}_{i,j} \in \mathcal{R}^{m \times n}$, $i = 1, \dots, K$ where K denotes number of users and $j = 1, \dots, L$ where L denotes number of training images per user. We lexicographically re-order them and obtain training face vectors, $x_{i,j} \in \mathcal{R}^{(mn) \times 1}$. The others set contains randomly selected face images which do not belong to any registered users $\tilde{\mathbf{I}}_s \in \mathcal{R}^{m \times n}$, $s = 1, \dots, M$ where M denotes the number of face images belonging to the others set. We again lexicographically re-order them and obtain face vectors, $\tilde{x}_s \in \mathcal{R}^{(mn) \times 1}$ of the others set. We apply PCA to the face images in the training set for feature extraction.

$$\mathbf{y}_{i,j} = \mathbf{A}(x_{i,j} - \boldsymbol{\mu}), \quad (1)$$

where $\mathbf{A} \in \mathcal{R}^{q \times (mn)}$ is the PCA matrix trained using the face images in the training set, $\mathbf{y}_{i,j} \in \mathcal{R}^{q \times 1}$ is the PCA coefficient vector belonging to the j^{th} training image of the i^{th} user and $\boldsymbol{\mu}$ is the mean face vector. We project the face images in the others set onto the PCA subspace as follows:

$$\tilde{\mathbf{y}}_s = \mathbf{A}(\tilde{x}_s - \boldsymbol{\mu}), \quad (2)$$

where $\tilde{\mathbf{y}}_s \in \mathcal{R}^{q \times 1}$ is the PCA coefficient vector belonging to

Manuscript received June 17, 2011.

[†]The author is with TUBITAK BILGEM (Center of Research for Advanced Technologies on Informatics and Information Security)

^{††}The author is with the Faculty of Engineering and Natural Sciences, Sabanci University

^{†††}The author is with the Faculty of Engineering and Natural Sciences, Sabanci University

a) E-mail: cagatay@uekae.tubitak.gov.tr

DOI: 10.1587/transinf.E0.D.1

the s^{th} image of the others set. We use these PCA coefficient vectors in the discriminative projection selection technique to find the most valuable features, which maximize the distance between the face images of a user in the training set and the face images in the others set, in the lower-dimensional subspace.

2.1.2 Dimension Reduction

At the dimension reduction phase, we generate an RP matrix, $\mathbf{T}_i \in \mathfrak{R}^{\ell \times q}$, for each user to reduce the dimension of her feature vector. The RP matrix elements are identically and independently distributed (*i.i.d*) and generated from a Gaussian distribution with zero mean and unit variance by using a random number generator (RNG) with a seed derived from the user's secret key. We apply the Gram-Schmidt (GS) procedure to obtain an orthonormal projection matrix $\mathbf{R}_i \in \mathfrak{R}^{\ell \times q}$ to have more distinct projections. Then, we project the PCA coefficient vectors of the i^{th} user onto a lower ℓ -dimensional subspace.

$$\mathbf{z}_{i,j} = \mathbf{R}_i \mathbf{y}_{i,j}, \quad (3)$$

where $\mathbf{z}_{i,j} \in \mathfrak{R}^{\ell \times 1}$ is the intermediate face image hash vector belonging to the j^{th} training image of the i^{th} user.

To determine the competing hash vectors, we also project the others set using the RP matrix as follows:

$$\tilde{\mathbf{z}}_s = \mathbf{R}_i \tilde{\mathbf{y}}_s, \quad (4)$$

where $\tilde{\mathbf{z}}_s \in \mathfrak{R}^{\ell \times 1}$ is the intermediate face image hash vector of the s^{th} face image of the others set and \mathbf{R}_i is the orthonormal random projection matrix of the i^{th} user.

The proposed discriminative projection selection technique selects the rows of the matrix \mathbf{R}_i using the Fisher criterion [6] and creates the discriminative random projections. Thus, we aim to increase discriminability due to mapping the PCA coefficient vectors into a more discriminant subspace. Fisher criterion is a feature selection method and in this case the features are the features obtained after the random projection of a PCA coefficient vector \mathbf{y} , namely:

$$z(k) = \mathbf{r}_{i,k}^T \mathbf{y}, \quad (5)$$

for $k = 1, \dots, \ell$, where $z(k)$ is a scalar value and $\mathbf{r}_{i,k}^T \in \mathfrak{R}^{1 \times q}$ denotes the k^{th} row of \mathbf{R}_i . The k^{th} feature for the i^{th} user's j^{th} image is $z_{i,j}(k)$ and the same feature for the s^{th} element of the others set is $\tilde{z}_s(k)$ which are the k^{th} elements of the corresponding vectors defined in Equations (3) and (4). The features are already uncorrelated due to the GS procedure which ensures $\mathbf{r}_{i,k}^T \mathbf{r}_{i,m} = 0$ for $k \neq m$.

We define

$$\mathbf{e}_i^k \triangleq [z_{i,1}(k), \dots, z_{i,L}(k)], \quad (6)$$

which is a collection of the k^{th} dimension (or bit position) coefficients of the intermediate hash vectors belonging to the i^{th} user for each $k = 1, \dots, \ell$.

First, we compute the sample mean value, $\hat{\mu}_i^{1,k}$, of each

\mathbf{e}_i^k vector for each bit position $k = 1, \dots, \ell$ as follows:

$$\hat{\mu}_i^{1,k} = \frac{1}{L} \sum_{j=1}^L e_i^k(j), \quad (7)$$

where $e_i^k(j)$ is the j^{th} element of the vector \mathbf{e}_i^k which is defined in Equation (6).

Then, we compute the sample standard deviation, $\hat{\sigma}_i^{1,k}$, of each \mathbf{e}_i^k vector for each bit position $k = 1, \dots, \ell$ as follows:

$$\hat{\sigma}_i^{1,k} = \left(\frac{1}{L} \sum_{j=1}^L (e_i^k(j) - \hat{\mu}_i^{1,k})^2 \right)^{\frac{1}{2}}. \quad (8)$$

Similarly, we collect together the k^{th} dimension values of the intermediate hash vectors of the "others" data set as

$$\mathbf{q}^k \triangleq [\tilde{z}_1(k), \dots, \tilde{z}_M(k)] \quad (9)$$

for all $k = 1, \dots, \ell$.

First, we compute the sample mean value, $\hat{\mu}_i^{2,k}$, of each \mathbf{q}^k for each bit position $k = 1, \dots, \ell$ as follows:

$$\hat{\mu}_i^{2,k} = \frac{1}{M} \sum_{s=1}^M q^k(s), \quad (10)$$

where $q^k(s)$ is the s^{th} element of the vector \mathbf{q}^k which is defined in Equation (9).

Next, we compute the sample standard deviation, $\hat{\sigma}_i^{2,k}$, of each \mathbf{q}^k for each bit position $k = 1, \dots, \ell$ as follows:

$$\hat{\sigma}_i^{2,k} = \left(\frac{1}{M} \sum_{s=1}^M (q^k(s) - \hat{\mu}_i^{2,k})^2 \right)^{\frac{1}{2}}. \quad (11)$$

By applying the Fisher criterion, we try to select the rows that have higher contrast between genuine user's data and the others set. In other words, we aim to reduce the distance between the genuine user's different face image hash vectors while at the same time we aim to maximize the distance between the i^{th} user's data and the others set. We compute the Fisher score for each row of \mathbf{R}_i as follows:

$$\eta_i(k) = \frac{|\hat{\mu}_i^{1,k} - \hat{\mu}_i^{2,k}|^2}{(\hat{\sigma}_i^{1,k})^2 + (\hat{\sigma}_i^{2,k})^2}. \quad (12)$$

We use these Fisher scores obtained for each dimension k to rank the rows of the random projection matrix \mathbf{R}_i . We define $\mathbf{r}_{i,k}^T$ to be the k^{th} row of \mathbf{R}_i . That is

$$\mathbf{R}_i = \begin{pmatrix} \mathbf{r}_{i,1}^T \\ \vdots \\ \mathbf{r}_{i,\ell}^T \end{pmatrix}. \quad (13)$$

We choose top ranking rows from the random projection matrix. Let \mathbf{c}_i be the index vector which ranks the rows of the matrix in a descending manner from 1 to w where w

is the number of desired rows. That is, $c_i(1)$ is the row index of \mathbf{R}_i that has the highest Fisher score, $c_i(2)$ is the index with the second highest score and so on. Thus, we obtain the discriminative random projection matrix $\hat{\mathbf{R}}_i \in \mathfrak{R}^{w \times q}$ and the index vector, \mathbf{c}_i , which contains the indices of top w rows for each user. We define

$$\hat{\mathbf{R}}_i = \begin{pmatrix} \hat{\mathbf{r}}_{i,1}^T \\ \vdots \\ \hat{\mathbf{r}}_{i,w}^T \end{pmatrix}, \quad (14)$$

where $\hat{\mathbf{r}}_{i,p} = \mathbf{r}_{i,c_i(p)}$ for $p = 1, \dots, w$. We only store the index vector, \mathbf{c}_i , for the i^{th} user in the database for verification at the test stage.

Next, we project the PCA coefficients, which belong to the training face images of the i^{th} user, onto a lower w -dimensional subspace by using the calculated $\hat{\mathbf{R}}_i$ as follows:

$$\mathbf{f}_{i,j} = \hat{\mathbf{R}}_i \mathbf{y}_{i,j}, \quad (15)$$

where $\mathbf{f}_{i,j} \in \mathfrak{R}^{w \times 1}$ is the raw face image hash vector belonging to the j^{th} training image of the i^{th} user.

Ngo *et al.* [1], [2] uses FLD as a feature extraction method which is applied before the random projection step in the algorithm. Their FLD transform is not user specific and aims to discriminate face images belonging to different users in the database. In our case, we employ the Fisher criterion for projection selection for biometric verification. Therefore, in our case, the projection selection is user-specific and aims to discriminate the claimed user's biometric hash vector from all other possible ones that may come from other face images. In our case, other face images may even be from outside the database which is more realistic in a real scenario. Our others set is chosen from another database for this purpose. The projection selection is done after the random projection step. FLD can reduce dimension at most to $K - 1$ dimensions, where K is the number of users in the database, due to the maximal rank of the between class covariance matrix. However, in our method, we do not have such a limitation since the selection is done by ranking the Fisher criteria obtained from each projection. In summary, there are fundamental differences between using FLD as a dimension reducing feature extraction method and using the Fisher criteria for selection of random projections that best separate the claimed identity from all others.

2.2 Quantization

In this subsection, we discuss the quantization methods used in this work. We employ two different quantization methods: (1) Binary quantization (BQ) [1], [2] and (2) The proposed Gaussian mixture model (GMM) based quantization method. In our simulations, we employ these quantization methods separately to show the performance of the system.

2.2.1 Binary Quantization Method with a Fixed Threshold

This technique is employed in Ngo *et al.*'s method [1], [2].

The raw face image hash vector $\mathbf{f}_{i,j}$ elements are binarized with respect to a pre-determined fixed threshold as follows:

$$\lambda_{i,j}(k) = \begin{cases} 1 & \text{if } \mathbf{f}_{i,j}(k) \geq \tilde{\mu}, \\ 0 & \text{Otherwise.} \end{cases} \quad (16)$$

where the threshold $\tilde{\mu}$ is chosen as the sample mean value of elements of the vector $\mathbf{f}_{i,j}$ and $k = 1, \dots, w$. The computed reference face image hash vectors $\lambda_{i,j} \in \mathfrak{R}^{w \times 1}$ are stored in the database.

2.2.2 The Proposed GMM Based Quantization Method

To the best of our knowledge, there is no face image hashing method employing GMM in the quantization step. GMM is one of the most widely used data clustering methods in the literature [7]. Let us assume that we have a set of numbers which are obtained by collecting the k^{th} elements of raw face image hash vectors and we want to binarize the element of this set. Since our aim is to make binarization, we fit two Gaussian distributions to the histogram of the k^{th} elements of the raw face image hash vectors by using the GMM. Then, we choose the average of the mean values of these two Gaussian distributions as an optimum threshold for partition of these two distributions. We repeat it for each bit location $k = 1, \dots, w$ separately. In other words, we employ bimodal GMM to find an optimum threshold for each bit position for binarization. Let $f_{i,j}(k)$ denote the k^{th} bit of $\mathbf{f}_{i,j} \in \mathfrak{R}^{w \times 1}$, we define the vector \mathbf{d}_k as the collection of all k^{th} dimension values of the raw image hashes in the database.

$$\mathbf{d}_k \triangleq [\mathbf{f}_{i,j}(k) : i = 1, \dots, K, j = 1, \dots, L] \in \mathfrak{R}^{r \times 1}, \quad (17)$$

where $k = 1, \dots, w$, $r = K \times L$, K is the number of users and L is the number of training images per user. Assume that the elements of the vector \mathbf{d}_k are observations of a single random variable d .

$$p(d | \Psi^k) = \sum_{s=1}^S \alpha_s^k p(d | \theta_s^k), \quad (18)$$

where α_s^k is a mixture weight, $\sum_{s=1}^S \alpha_s^k = 1$ where $S = 2$ due to binarization, $\Psi^k = \{\alpha_1^k, \alpha_2^k, \theta_1^k, \theta_2^k\}$ and $p(d | \Psi^k)$ is a one-dimensional Gaussian density with its own parameters $\theta_s^k = \{\mu_s^k, \sigma_s^k\}$ as follows:

$$p(d | \theta_s^k) = \frac{1}{\sigma_s^k \sqrt{2\pi}} e^{-\frac{(d - \mu_s^k)^2}{2(\sigma_s^k)^2}}, \quad (19)$$

where μ_s^k and σ_s^k denote mean variance for the s^{th} component of the GMM respectively for $s = \{1, 2\}$. We find an optimum threshold for each bit position as follows:

$$T_k = \frac{\mu_1^k + \mu_2^k}{2}, \quad (20)$$

where T_k denotes the optimum threshold for the k^{th} bit position of the raw face image hash vector $\mathbf{f}_{i,j}$, $\forall i, j$. Note that,

the GMM is trained using the whole training set for each bit position. Thus, the GMM parameters are not user dependent. Finally, the elements of $f_{i,j}$ are binarized with respect to the optimum system-level thresholds as follows:

$$\lambda_{i,j}(k) = \begin{cases} 1 & \text{if } f_{i,j}(k) \geq T_k, \\ 0 & \text{Otherwise.} \end{cases} \quad (21)$$

where $k = 1, \dots, w$. The computed reference face image hash vectors $\lambda_{i,j} \in \mathbb{R}^{w \times 1}$ are stored in the database.

2.3 Test Stage

At the test stage, a claimer claims that she is the i^{th} user and sends her face image and her secret key to the system. The system computes her test face image hash vector by using her face image, her secret key (to generate a RP matrix) and the index vector, c_i , which belongs to the i^{th} user. Recall that index vectors and the reference face image hash vectors of the registered users are stored in the database; however, the secret keys are not stored in the database. Then, the Hamming distance [8] is computed between the test face image hash vector and the reference face image hash vectors which belong to the i^{th} user and were generated at the enrollment stage. If it is below the pre-determined distance threshold, the claimer is accepted; otherwise, the claimer is rejected.

We simulate two scenarios in our experiments. These scenarios are described in detail below.

1. Key-Unknown Scenario: In this scenario, an unauthorized impostor has neither the secret key nor the face image template belonging to the genuine user. Note that the index vectors of the users are stored in the database. Therefore, whenever a claimer claims that she is the i^{th} user and sends her face image and a secret key to the system, the system computes a test face image hash vector by using the data sent by the claimer and the index vector, c_i , which belongs to the i^{th} user.

2. Key Stolen Scenario: In this scenario, an unauthorized impostor acquires the secret key of the i^{th} genuine user but does not have the claimed person's face image. When an impostor sends her face image and the secret key of the i^{th} user to the system, the system computes a test face image hash vector by using the data sent by the impostor and the index vector c_i that belongs to the i^{th} user which is stored in the database.

3. Simulation Results

In this section, we discuss our experimental results. We test the performance of the proposed method on AT&T [9], AR [10] and the Sheffield (previously UMIST) face databases [11]. AT&T database has 400 different face images corresponding to 40 distinct people. AR database has 3120 face images belonging to 120 different people's faces. The Sheffield database has 564 different face images belonging to 20 different people. Besides, we randomly select 104 face images from Carnegie Mellon University database [12] and create the others set.

We compare the performance of the proposed method to the Ngo *et al.*'s PCA+RP and FLD+RP methods that were introduced in [1], [2] as shown in Table 2. We automatically select face images for training and test sets and evaluate the performance of the proposed method and Ngo *et al.*'s methods [1], [2]. We use 1024-length PCA coefficient vectors for the face images belonging to the training, test and others sets in the simulations. In our experiments, pre-processing techniques such as eye alignment, head region masking, lighting adjustment are not applied to the face images. In our simulations for both scenarios; for impostor tests, each face image of each user in the test set is compared against each face image of all other users in the training set. The failed impostor test results in False Acceptance error. For the genuine tests, each face image of each user in the test set is compared against all face images of the same user in the training set. The failed genuine test results in False Rejection error. The detailed information on the data sets used in the experiments are given in Table 1. The proposed

Table 1 Data sets and Experimental Set-up

Database	Number of Face Images	Train set	Test set
AR	3120 images from 120 people	The first 7 images of each user	The last 2 images of each user
AT&T	400 images from 40 people	The first 5 images of each user	The rest 5 images of each user
Sheffield	564 images from 20 people	The first 8 images of each user	The following 8 images of each user

method has better performance in terms of equal error rate (EER) in comparison to the Ngo *et al.*'s methods [1], [2] whereas Ngo *et al.*'s PCA+RP and FLD+RP methods have comparable performances with each other as shown in Table 2. As the length of face image hash vector decreases, the proposed method shows better improvement since the proposed dimension reduction matrix better preserves the

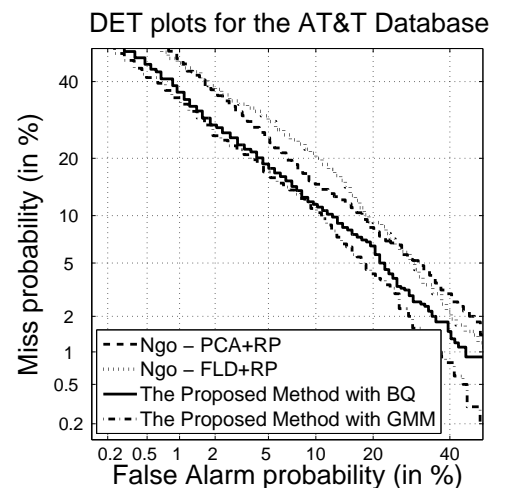


Fig. 2 DET plots for the methods with 256 bit face image hash vector length for key stolen scenario - AT&T database

Table 2 EER Performances of the Proposed Face Image Hashing Method and Ngo *et al.*'s Methods [1], [2]

Length of Face Hash Vector	EER (%) of Ngo <i>et al.</i> 's Method [1], [2] (PCA+RP)	EER (%) of Ngo <i>et al.</i> 's Method [1], [2] (FLD+RP)	EER (%) of The Proposed Method with Binary Quantization Method with Fixed Threshold	EER (%) of The Proposed Method with GMM Based Quantization Method	Scenario	Database
64 bit	% 12.19	% 10.46	% 3.83	% 2.73	Key Unknown	AT&T
128 bit	% 7.36	% 8.51	% 2.23	% 1.57	Key Unknown	AT&T
256 bit	% 5.81	% 5.50	% 1.80	% 1.15	Key Unknown	AT&T
512 bit	% 3.79	% 4.17	% 2.48	% 2.10	Key Unknown	AT&T
64 bit	% 16.93	% 18.13	% 14.40	% 13.58	Key Stolen	AT&T
128 bit	% 13.97	% 16.73	% 12.01	% 11.14	Key Stolen	AT&T
256 bit	% 12.76	% 14.50	% 10.80	% 10.23	Key Stolen	AT&T
512 bit	% 12.34	% 13.55	% 10.15	% 9.73	Key Stolen	AT&T
64 bit	% 23.63	% 23.34	% 9.08	% 8.96	Key Unknown	AR
128 bit	% 18.24	% 18.05	% 8.67	% 8.72	Key Unknown	AR
256 bit	% 15.82	% 13.93	% 7.81	% 8.12	Key Unknown	AR
512 bit	% 11.38	% 11.92	% 8.33	% 8.57	Key Unknown	AR
64 bit	% 28.27	% 28.51	% 18.07	% 18.46	Key Stolen	AR
128 bit	% 27.17	% 27.56	% 18.06	% 18.05	Key Stolen	AR
256 bit	% 25.50	% 26.44	% 19.10	% 18.83	Key Stolen	AR
512 bit	% 24.89	% 25.04	% 20.95	% 20.41	Key Stolen	AR
64 bit	% 17.09	% 22.00	% 15.75	% 16.23	Key Unknown	Sheffield
128 bit	% 16.38	% 19.10	% 13.33	% 14.03	Key Unknown	Sheffield
256 bit	% 15.05	% 14.93	% 11.45	% 11.05	Key Unknown	Sheffield
512 bit	% 14.97	% 14.12	% 10.44	% 12.20	Key Unknown	Sheffield
64 bit	% 21.40	% 24.50	% 19.38	% 20.68	Key Stolen	Sheffield
128 bit	% 21.92	% 24.30	% 17.51	% 19.71	Key Stolen	Sheffield
256 bit	% 22.53	% 22.02	% 16.96	% 17.80	Key Stolen	Sheffield
512 bit	% 23.47	% 22.55	% 19.27	% 18.22	Key Stolen	Sheffield

pair-wise distances between feature vectors in the reduced dimension subspace in comparison with the traditional random projection matrix. The best results are usually obtained with 128 or 256 bits. Besides, we plot the detection error trade-off (DET) curves [13] for key stolen scenario of the 256 bit face image hash length with the AT&T database in Figure 2.

Ngo *et al.* [1], [2] employs binary quantization method with a fixed threshold for all bits. This method may be sub-optimal in some cases. The proposed GMM-based quantization method reduces EER most of the time in comparison to the binary quantization since it finds approximately optimum threshold values for each bit position.

4. Conclusion

In this paper, we propose a novel face image hashing method for a biometric verification system. The proposed method is based on discriminative projection selection depending on Fisher criteria. Another novelty of the proposed method is to employ bimodal GMM in the quantization step. The simulations demonstrate that the proposed method has better performance in comparison with the random projection based face image hashing methods proposed in the literature.

References

[1] D.C.L.Ngo, A.B.J. Teoh, A.Goh, "Biometric Hash: High Confidence Face Recognition," Proc. IEEE Trans. on Circ. and Sys. for

Video Tech., vol.16, no.6, June 2006.
 [2] D.C.L.Ngo, A.B.J. Teoh, A. Goh, "Eigenspace-based face hashing," Proc. of the International Conference on Biometric Authentication (ICBA) in LNCS, vol. 3072, pp. 195-199, Hong Kong, July 2004.
 [3] C. Karabat, H.Erdogan, "A Cancelable Biometric Hashing for Secure Biometric Verification," Proc. IEEE IHH-MSP 2009, pp.1082-1085, Kyoto, Japan, 2009.
 [4] A. Lumini, L. Nanni, "An Improved BioHashing for Human Authentication," Proc. of the Pattern Recognition, vol.40, no.3, pp.1057-1065, March 2007.
 [5] A.Kong, K.H. Cheung, D. Zhang, M. Kamel, J. You, "An Analysis of BioHashing and Its Variants," Proc. of the Pattern Recognition, vol.39, no.7, pp.1359-1368, 2007.
 [6] G. J. McLachlan ed., Discriminant Analysis and Statistical Pattern Recognition, John Wiley&Sons, Inc., 1992.
 [7] R.O. Duda and P.E. Hart ed., Pattern Classification and Scene Analysis, John Wiley&Sons, Inc., 1973.
 [8] R.W. Hamming, "Error Detecting and Error Correcting Codes," Proc. Bell Sys. Techn. Journal, vol.29, no.2, pp.147-160, 1950.
 [9] Cambridge University AT&T face database, <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>.
 [10] A.M.Martinez and R.Benavente, "The AR Face Database," CVC Technical report #24, June 1998.
 [11] Daniel B. Graham and Nigler M. Allinson, "Face Recognition: From Theory to Applications," Proc. NATO ASI Series F, Computer and Systems Sciences, Vol. 163. H. Wechler, P.J.Philips, V.Bruce, F. Fogelman-Soulie and T.S. Huang (ed.), pp. 446-456, 1998.
 [12] Carnegie Mellon University Face Database, <http://amp.ece.cmu.edu>.
 [13] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M.A. Przybocki, "The det curve in assessment of detection task performance," in Proc. of Eurospeech Rhodes, Greece, Sept., 1997, pp. 1895-1898.