International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-1 Issue-4, April 2012

Multimodal Person Authentication using Qualitative SVM with Fingerprint, Face and Teeth Modalities

A. Jameer Basha, V. Palanisamy, T. Purusothaman

Abstract - Multimodal biometrics systems are becoming increasingly efficient over the unimodal system, especially for the securing handheld devices. However, the challenge with this authentication system is the relative degradation of the biometric modalities involved in the development and test data respectively. To overcome this problem, in this paper we propose a novel Qualitative Support Vector Machine (SVM) classifier with Face, teeth, and fingerprint as biometric traits. The test scores of individual modalities are adjusted according to their relative quality and then passed to binary SVM classifier. The experiments were conducted over a database collected from 20 individuals with three instances of all the three traits. The performance analysis of the fusion techniques revealed that the Equal Error Rates (EER) of 1.22%, 1.46%, and 1.88% for the qualitative SVM, raw score SVM and weighted summation rule classifiers respectively. On the other hand, the equal error rates for unimodal systems are 7.4%, 5.09% and 4.6% for teeth, face and fingerprint biometrics traits respectively. Hence, we confirmed that the proposed qualitative SVM method outperformed other raw score fusion techniques and unimodal classifiers.

Keywords: Multimodal biometrics, fingerprint verification, teeth recognition, face recognition, SVM classifier

I. INTRODUCTION

Biometrics, which refers to identify an individual based on his or her physiological or behavioral characteristics, can distinguish between an authorized person and an imposter as proposed by Jain, Ross and Pankanti, (2004)[1].

Multimodal biometric systems as proposed by Veeramachaneni, Osadciw and Varshney, (2005) [2] and Kumar, Kanhangad and Zhang (2010)[3] seek to alleviate some of these problems by providing multiple pieces of evidence of the same identity or different biometric traits. These systems help to achieve an increase in performance that may not be possible using a single biometric indicator.

Fusion techniques can be subdivided into adaptive and non-adaptive ones. Non-adaptive fusion techniques are those where all the fusion parameters are found using the development set, as seen in the previous chapter. With adaptive fusion techniques, all the parameters, or some of them, are found based on the test set. From the definition of the non-adaptive fusion technique, it can be seen that the drawback of this technique is the possible mismatch between the relative variation of the biometric modalities involved in the development and test data respectively. To tackle this problem, it would be logical to consider the relative levels of contamination in different biometric data not only in the development phase, but also at the test stage. Therefore we propose a multimodal biometrics technique using Qualitative SVM fusion with fingerprint, face and teeth modalities. A personal identification using teeth image, which is relatively new biometric traits considered for authentication was first proposed by Tae-Woo KIM and Tae-Kyung CHO (2006)[4] based on Linear Discriminant Analysis (LDA) as sequential steps. .and few other studies were researched by using various algorithms. Few other versions of teeth authentication were also proposed using modified principal component analysis (PCA) by Prajuabklang, Kumhom, Maneewarn, Chamnongthai (2004)[5] and Nadee, Kumhom, Chamnongthai (2005)[6] proposed an improved PCA based individual identification method using the invariance moment.

In our work, experiments were conducted over a database collected from 20 individuals with multiple instances of all the three traits and classified using various fusion techniques including qualitative SVM, raw score SVM and weighted summation rule techniques.

II. OVERVIEW OF RELATED WORKS

In this section, we present an abstract view of the biometric techniques, fingerprint recognition techniques and fusion methods K-NN classifier and weight-summation rule.

Yang, J., L. Liu and T. Jiang (2002)[7] in their work suggested that fingerprints are the ridge and furrow patterns on the tip of the finger and have been used extensively for personal identification of

Manuscript published on 30 April 2012.
* Correspondence Author (s)
A. Jameer Basha*, Department of Information Technology, Sri Krishna College of Technology, Coimbatore, India, +919159569755 (e-mail: a_jameerbasha@rediffmail.com).
V. Palanisamy, Principal, Info Institute of Technology, Coimbatore, India (e-mail: vpsamyin@gmail.com).
T. Purusothaman, Department of Computer Science and Engineering, Government College of Technology, Coimbatore, India (e-mail: purushgct@yahoo.com).

people. The quality of the ridge structures in a fingerprint image is an important characteristic, as the ridges carry the information of characteristic features required for minutiae extraction. Fingerprints have many conspicuous landmarks and any combination of them could be used for establishing a reference point. It is defined the reference point of a fingerprint as the point of maximum curvature of the concave ridges in the fingerprint image. Chan, Moon and Cheng (2004)[8] has proved that the minutiae can be extracted by scanning the local neighborhood of each ridge pixel in the image using a 3 x 3window. The CN value is then computed, which is defined as half the sum of the differences between pairs of adjacent pixels in the eight-neighborhood. Once the reference point is located, all minutiae extracted from a master fingerprint image can be aligned with the reference point to generate a circular sub region in the original image.

Duda and Hart (1973)[9] demonstrated that the KNN classifier assigns an object described by a set of relevant features to the class with the highest occurrence frequency among k nearest neighbors in the classifier's training dataset. Reference data points are required are for both classes representing the genuine and the imposter. An unknown test data point y is then attributed with the same class label as the label of the majority of its k nearest reference neighbors.

To find the k nearest neighbors, the Euclidean distance between the test point and all the reference points is calculated. The distances are then ranked in ascending order, and the reference points corresponding to the k smallest Euclidean distances are taken. The exhaustive distance calculation step during the test phase leads rapidly to large computing time, which is the major drawback of this otherwise a very simple algorithm.

The fusion approach using the weighted-summation rule does not requiring training phase in advance. Erzin, Yemez and Tekalp (2005)[10] and Wark and Sridharan (2001)[11] in their works reveal that the weighted-summation fusion is expressed as $S=\sum_{n=0}^{N} w_n S_n$. Where, S_n is the normalized score of n-th modality, and S denotes the fused score. Also, w_n denotes the weighted coefficient of the n-th modality, such that $\sum_n w_n = 1$. The weighting coefficients, w_n , can be usually set to some fixed values using a priori knowledge such as the performance of the n-th modality, or they can be adaptively estimated at each decision phase using various methods.

SVM binary classifiers has been studied in related works of Verlinde, Chollet and Acheroy (2000)[12]. It performs classification by constructing an N-dimensional hyper plane that optimally separates the data into two categories. SVM models are closely related to neural networks. In fact, a SVM model using a sigmoid kernel function is equivalent to a two-layer, perceptron neural network. Using a kernel function, SVMs are an alternative training method for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training.

III. Materials and Methods

A. Fingerprint Verification with Localized sub region

The authentication speed of the fingerprint verification system can be improved by extracting the features from a selected sub region rather the full fingerprint image. Once the reference point is located as shown in figure 1, all minutiae extracted from a master fingerprint image as shown in figure 1, can be aligned with the reference point to generate a circular sub region in the original image as suggested by Chan, Moon and Cheng (2004). This sub region contains a fixed number of minutiae to be matched with similar minutiae contained in a live template during an authentication process.

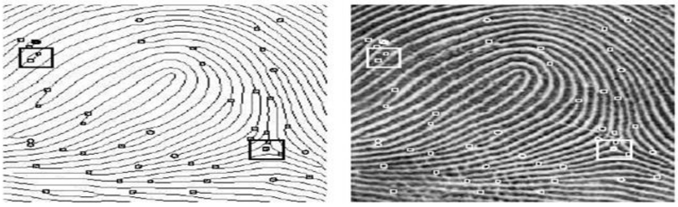


Fig. 1 Results of performing minutiae extraction on a fingerprint image (a) skeleton image (b) original image

First, the Cartesian coordinates of the extracted minutiae in a master fingerprint image are converted into Polar coordinates using the following equations:

$$r_i = \sqrt{(x_i - core_x)^2 + (y_i - core_y)^2}$$
$$\theta_i = \tan^{-1}(y_i - core_y, x_i - core_x) - core_{orient}$$
$$\vartheta_i = \phi_i - core_{orient}$$
(1)

Where:

(x_i,y_i)	=	Cartesian coordinates of minutia i
φ_i	=	Minutia orientation
(r_i,θ)	=	Polar coordinates of minutia i
ϑ_i	=	Normalized minutia orientation
$(core_x, core_y)$	=	Cartesian coordinates of the reference point
Core $_{orient}$	=	Reference point orientation

In polar coordinate representation, the minutiae are rotational and transitional invariant with respect to their reference point. After the coordinate's transformation, the minutiae are sorted in ascending order according to their distances from the reference point. To compute a minimum area that covers a predetermined number of minutiae points, we select the first minutiae from the list to form a master feature template.

Especially in the Arch fingerprints, that some reference points are located near the boundaries of the images. Such cases can lead to large bounding circle size as shown. As a remedy, we construct an average center (Xcenter, Ycenter) as shown in Figure 2 & 3;

$$x_{centre} = \sum_{i=0}^{N} \frac{x_i}{N}, \quad y_{centre} = \sum_{i=0}^{N} \frac{y_i}{N} \qquad (2)$$

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Fig. 2(a) Size of a bounding circle is large if the reference point is near boundary (b) Size of the bounding circle decreases when a "centralized" reference point is used.

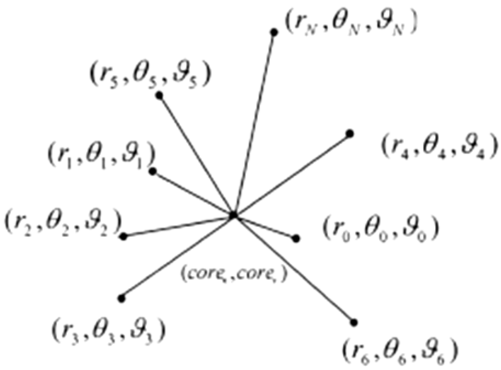
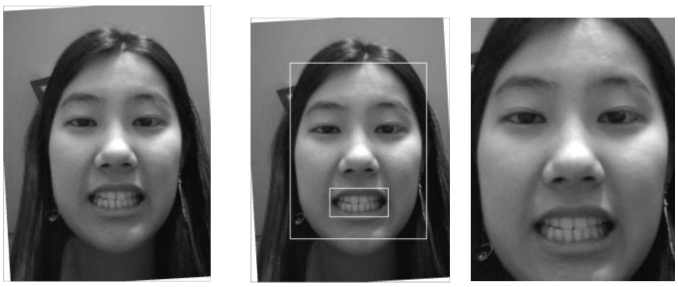


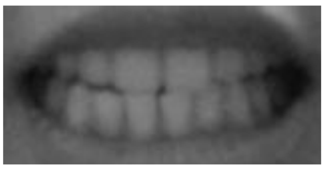
Fig. 3. The first N minutiae and their reference point formed

To obtain the adequate face subject, we utilize a pre-processing procedure, i.e., rotated-angle compensation, of previous system. Rotated-angle compensation can be successfully performed since both sides of the teeth image are usually dark, and we can calculate the rotated angle with a horizontal line connecting both centers of the image.

Then, the entire input image is rotated with the computed angle information as shown in Figure 4 (a). The face and teeth regions are redetected using the AdaBoost algorithms as shown in Figure 4 (b), and the face and teeth subjects used in authentication are acquired by cropping their corresponding regions. The resultant face and teeth subjects are shown in Figures 4 (c) and (d), respectively. After subject acquisition, the authentication phases are similarly performed with face and teeth subjects using EHMM algorithms, leading to the corresponding probabilities, respectively.



(a) (b) (c)



(d)

Fig. 4. Subject acquisition procedure: (a) input image rotation using computed angle information, (b) face and teeth region detection in rotated image, (c) face image as a biometric trait, (d) teeth image as a biometric trait.

B. Face and Teeth Authentication using EHMM Algorithm

The image-based authentication is comprised of the following sequential steps: image acquisition, region detection, and the authentication phase based on the EHMM algorithm as reported by Dong-Ju Kim, Jeong-HoonShin and Kwang-SeokHong (2010)[13]. These stages are equivalently applied to face and teeth authentication. However, the number of Embedded hidden Markov model (EHMM) states is somewhat different for each procedure. In order to model the face, we compose the state structure of EHMM using five super-states with three, five, five, five and three embedded-states, respectively. Each super-state represents the vertical face features such as forehead, eyes, nose, teeth and chin in the face image, and each embedded state in the super-state represents the horizontal local features. Also As illustrated by Dong-Ju Kim and Kwang-Seok Hong (2008)[14] the Nefian's EHMM the teeth image is modeled using three super-states and embedded-states of three, five and three in each super-state like that previous system.

Subsequently, the matching scores of three modalities are normalized to eliminate the unbalance and get good performance, we are motivated to normalize the feature before fusion using z-score model.

C. Qualitative SVM Classifier

The proposed approach is based on adjusting the weights for each of the two modalities according to their relative quality. This is performed by estimating the relative quality aspects of the test scores and then passing them on into the Support Vector Machine either as features or weights. In this technique the quality aspect of the test samples is quantified and then passed on into a SVM. This process involves estimating the quality of the development data by measuring some parameters for the development score data and then incorporating these parameters in the quality estimation of the test scores. This quantification is similar to that described in C. Sanderson and K. K. Paliwal (2000) [15] and is described as in the case of a two-class problem (Clients / Impostors), let *M*□□*f* / *s*□□be the development scores for fingerprint, face or teeth, (where □□*f* / *s*□□is used to denote that a measure is applied to either fingerprint, face or teeth modality) and let the client and impostor scores from each modality be given as

$$C_{M(f/s)} \cong \{ \mu^C_{M\left(\frac{f}{s}\right)}, \sigma^C_{M\left(\frac{f}{s}\right)} \} \tag{4}$$

$$I_{M(f/s)} \cong \{ \mu^I_{M\left(\frac{f}{s}\right)}, \sigma^I_{M\left(\frac{f}{s}\right)} \}$$

(5)

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Where, $\mu^{C}_{M\left(\frac{f}{s}\right)}$, $\sigma^{C}_{M\left(\frac{f}{s}\right)}$ are the mean and variance for the client scores from each modality. $\mu^{I}_{M\left(\frac{f}{s}\right)}$, $\sigma^{I}_{M\left(\frac{f}{s}\right)}$ are the mean and variance for the impostor scores from each modality. The quality of samples of a modality (face or speech in this chapter) is determined by the characteristics of the scores obtained with the development and test samples of that modality. The quality of the face scores (Q_f), fingerprint score (Q_p) and teeth scores (Q_t) are calculated as follows:

$$Q_{(f/s)} = D_{M(f/s)} \times T_{E(f/s)/E1(f/s)} \quad (6)$$

where $Q_{(f/s)}$ is the quality for face or speech, *D* is the quality of the development data, *T* is the quality of the test (sample) data, *E(f /s)* is the subset of scores from the test data which is used to determine the quality of the test samples and *E*1*(f /s)* is the rest of the scores from the test data which is used to investigate the performance for the proposed scheme.

Following the addition of quality aspects the SVM fusion is performed as shown in the work of P. Verlinde, G. Chollet and M. Acheroy (2000)[16]. Using a kernel function, SVMs are an alternative training method for polynomial, radial basis function and multi-layer perceptron classifiers in which the weights of the network are found by solving a quadratic programming problem with linear constraints, rather than by solving a non-convex, unconstrained minimization problem as in standard neural network training.

Vapnik also proved that the optimal hyper plane can be obtained solving the quadratic programming problem:

$$\text{Minimize} \frac{1}{2} ||w||^2 + C\sum_{i=1}^{l} \xi_i$$

(7)

$$Y_i(\langle w, |\phi(X)\rangle\rangle_F + b) >= 1 - \xi_I \quad i=1,..,l$$

$$\xi_{I >=0} \quad i=1,..,l$$

where constant C and slack variables, ξ_i re introduced to take into account the eventual non-separability of $\phi(X)$ into F. Applying the Karush-Kuhn-Tucker conditions to the problem in (7), the following sparse expression is obtained for the optimum:

$$W^* = \sum_{i \in SV} \alpha_i \, y_i \, \phi(X)$$

(8)

where SV = { i|α_i >0} is the set of support vectors. Taking into account that the decision function D(·) that classifies a test pattern $\mathbf{x}_T$ is:

$$D(x_T) = \text{sign}\{\langle w^*, \phi(xT)\rangle_F + b^*\}$$

(9)

Defining $K(x_i,x_j) = \langle \phi(x_i), \phi(x_j) \rangle_F$ as the kernel function and using (8) we obtain:

$$D(x_T) = \text{sign}\{\textstyle\sum_{i \in SV} \alpha_i \, y_i K(x_i, x_j) + b^*\}$$

(10)

The choice for the kernel has been in this case a Radial Basis Function (RBF):

$$K(x_i,x_j) = \exp(-||x_i-x_j||^2 / 2\sigma^2)$$

(11)

The fusion strategy relied on the computation of the decision function D(·). A modification in order to obtain not a final classifier decision, but a combined multimodal score based on the proximity of the test pattern to the separating surface, is proposed here. The combined score $S_T \in \mathbb{R}$ of the multimodal pattern will $X_T \in \mathbb{R}^R$ be calculated as:

$$S_T = \textstyle\sum_{i \in SV} \alpha_i \, y_i \, K(x_i, x_j) + b^*$$

(12)

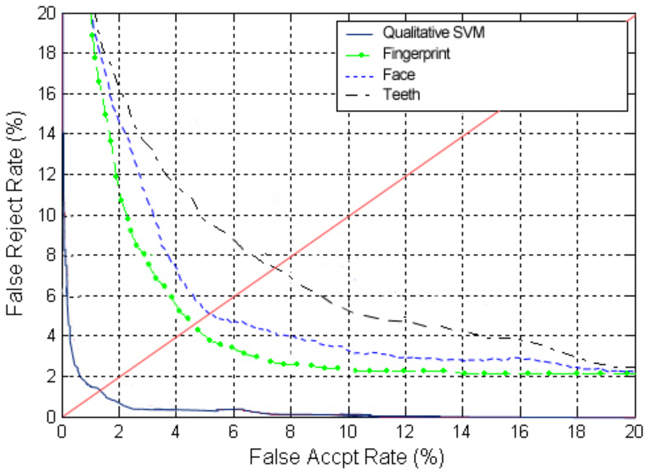
Following this approach, the decision threshold parameter can be adjusted to reach different working points.

IV. Results and Discussions

A. Database Collection

The database for our Multimodal system was collected using fingerprint sensor, camera and speaker attached to a mobile device. Three instances of fingerprint print image were scanned from each of the 20 individuals with biometric scanner (500 dpi 640 X 480). It should be noted that the several instances were obtained to accommodate orientation and pressure of the impressions during the training phase. Three images of teeth were captured at a resolution of 840 X 1120 pixels using 8MP mobile camera and compressed to 80 X 40 pixels from individuals with indoor and outdoor environments to accommodate lighting variations.

Initially, the fingerprint biometric system was tested with



6, 10 and 18 local minutiae points and found 10 points to be optimal for identification. The results showed detection rates of 99.2% and 98.8% for face and teeth, respectively.

B. Performance Evaluation

Initially the performance analyses of all individual modalities were calculated. The performances of fingerprint, speech and teeth modalities are measured using False Acceptance Rate (FAR), False Rejection Rate (FRR) and EER. EER is the error rate where the FAR and FRR assume the same value.

Figure 5 represents the FAR and FRR curves with different number of matching minutiae in fingerprint authentication with sub regions.

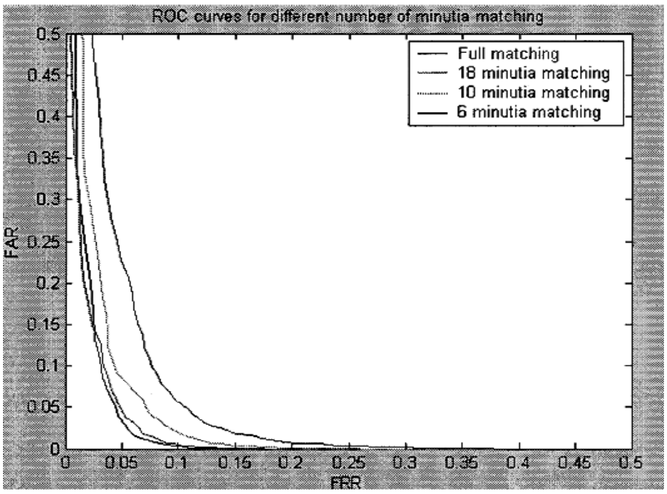


Fig. 5 FAR and FRR of fingerprints with different minutia counts

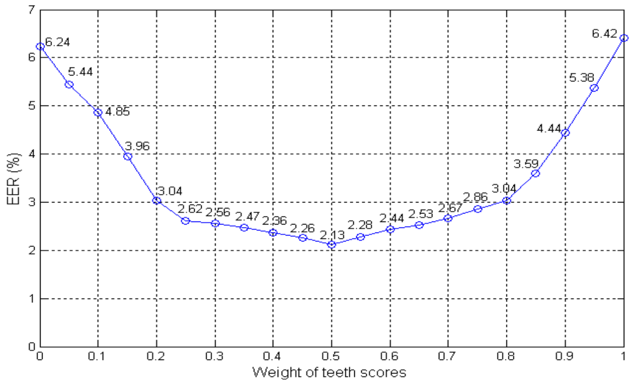


Fig. 6 Resultant EER against weight of teeth scores.

Fig. 7 Error rates of Fingerprint, face and teeth traits along with qualitative SVM fusion results

Figure 6 plots the EER obtained by combining the normalized scores using the weighted-summation method along with changing *p* of weight of teeth scores from 0 to 1.

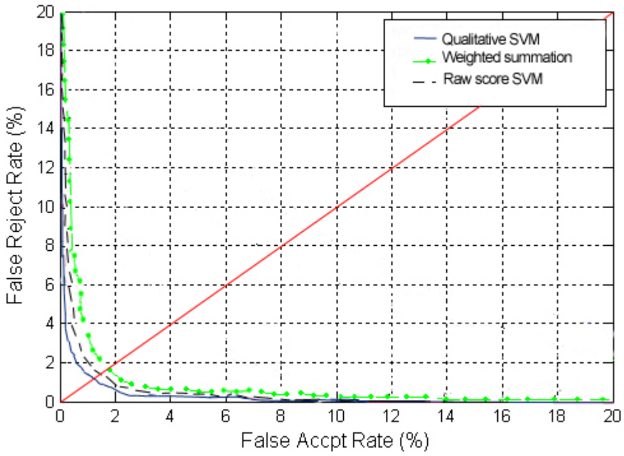


Fig. 8 Error rate comparison of various fusion techniques

From our experiments it is observed the equal error rates of individual biometric systems as shown in Figure 7 are 7.4%, 5.09% and 4.6% for teeth, voice and fingerprint biometrics traits respectively.

The results of the various fusion techniques showed that the qualitative SVM classifier has the best EER rate with 1.22%, whereas the raw score SVM and weighted summation rule revealed error rates of 1.46% and 1.88% respectively as shown in Figure 8.

CONCLUSIONS

Usage of handheld devices like smart phones, PDA and, etc., are increasingly become the target for theft for not only its physical value but also for the invaluable data like banking passwords, email accounts and, etc. Hence, we present a new biometric framework to secure the mobile devices using fingerprint, teeth and voice fused using qualitative SVM binary classifier. Our work uses the sub regions of the fingerprint image to improve the speed of authentication and reduce the amount of data stored. Face and Teeth regions are detected using Ada Boost algorithm and verified using EHMM technique. Finally, the normalized scores of all the unimodal systems are fed into qualitative SVM classifier for reject/accept the claim. The performance analyses of the proposed system were conducted over a database collected from 20 individuals with multiple instances of all the three traits. The performance analysis of the fusion techniques revealed that the equal error rates of 1.22%, 1.46%, and 1.88% for the qualitative SVM, raw score SVM and weighted summation rule classifier respectively. On the other hand, the equal error rates for unimodal systems are 7.4%, 5.09% and 4.6% for teeth, voice and fingerprint biometrics traits respectively. Hence, we confirmed that the proposed qualitative SVM fusion method outperformed other fusion techniques and unimodal classifiers.

REFERENCES

[1] Jain, A.K., A. Ross and S. Pankanti, 2004. " An introduction to biometric recognition" . IEEE Trans. Circuits Syst. ideo Technol., 14: 4-20. ISSN: 1051-8215

[2] Veeramachaneni, K., L.A. Osadciw and P.K. Varshney, 2005. " An adaptive multimodal biometric management algorithm" . IEEE Trans. Syst. Man Cybern. C Appl. Rev., 35: 344-356. ISSN: 1094-6977

[3] Kumar, A., V. Kanhangad and D. Zhang, 2010. "A new framework for adaptive multimodal biometrics management". IEEE Trans. Inform. Forensics Security, 5: 92-102. SSN: 1556-6013

[4] Tae-Woo KIM, Tae-Kyung CHO, "Teeth Image Recognition for Biometrics", IEICE TRANSACTIONS on Information and Systems Vol.E89-D No.3 pp.1309-1313, 2006.

[5] K. Prajuabklang, P. Kumhom, T. Maneewarn, K. Chamnongthai, "Realtime Personal Identification from Teeth-image using Modified PCA", Proceeding, the 4-th information and computer Engineering Postgraduate Workshop, Vol. 4, No. 1, pp.172-175, 2004.

[6] C. Nadee, P. Kumhom, K. Chamnongthai, "Improved PCA-Based Personal Identification Method Using Invariance Moment", The thirdInternational Conference on Intelligent Sensing and Information Processing, December 14-17, 2005.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

[7] Yang, J., L. Liu and T. Jiang, 2002. “An improved method for extraction of fingerprint features.” Proceeding of the 2nd International Conference Image and Graphics, Anhui, China, Aug.
[8] Chan, K.C., Y.S. Moon and Cheng, 2004. “Fast fingerprint verification using subregions of fingerprint images.” IEEE Trans. Circuits Syst. Video Technol., 14.
[9] R. O. Duda and P. E. Hart, “Pattern Classification and Scene Analysis”. John Wiley & Sons, 1973.
[10] E. Erzin, Y. Yemez, A. Tekalp, "Multimodal speaker identification using an adaptive classifier cascade based on modality reliability", IEEE Transaction on Multimedia, vol. 7, no. 5, pp. 840-852, 2005.
[11] T. Wark, S. Sridharan, "Adaptive fusion of speech and lip information for robust speaker identification", Digital Signal Process, vol. 11, no. 3, pp. 169-186, 2001.
[12] P. Verlinde, G. Chollet and M. Acheroy, "Multi-Modal Identity Verification using Expert Fusion", Information Fusion, no. 1, pp. 17-33, Elsevier, 2000.
[13] Dong-Ju Kim, Jeong-HoonShin and Kwang-SeokHong, “Teeth recognition based on multiple attempts in mobile device”, Journal of Network and Computer Applications 33 (2010) 283–292, Elsevier 2010.
[14] Dong-Ju Kim and Kwang-Seok Hong, “Multimodal Biometric Authentication using Teeth Image and Voice in Mobile Environment”, IEEE Transactions on Consumer Electronics, Vol. 54, No. 4, NOVEMBER 2008.
[15] C. Sanderson and K. K. Paliwal, "Adaptive Multi- Modal Person Verification System," Proceedings of the First IEEE Pacific-Rim Conference on Multimedia, 2000.
[16] P. Verlinde, G. Chollet and M. Acheroy, "Multi-Modal Identity Verification using Expert Fusion", Information Fusion, no. 1, pp. 17-33, Elsevier, 2000.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)