An Automated Ear Identification System Using Gabor Filter Responses

Maarouf Korichi[1], Abdallah Meraoumia[1], Salim Chitroub[2], and Aiadi Kamal eddine[1]

[1]Univ Ouargla, Fac. des nouvelles technologies de l'information et de la communication,
Lab. de Génie Électrique, Ouargla 30 000, Algeria
[2]Signal and Image Processing Laboratory, Electronics and Computer Science Faculty, USTHB.
P.O. box 32, El Alia, Bab Ezzouar, 16111, Algiers, Algeria

Email: korichi.maarouf@gmail.com, Ameraoumia@gmail.com, S_chitroub@hotmail.com, Aiadi@yahoo.com

***Abstract*—About some years ago, several biometric technologies are considered mature enough to be a new tool for security and ear-based person identification is one of these technologies. This technology provides a reliable, low cost and user-friendly viable solution for a range of access control applications. In this paper, we propose an efficient online personal identification system based on ear images. In this purpose, the identification algorithm aims to extract, for each ear, a specific set of features. Based on Gabor filter response, three ear features have been used in order to extract different and complementary information: phase, module and a combination of the real and imaginary parts. Using these features, several combinations are tested in the fusion phase in order to achieve an optimal multi-representation system which leads to a better identification accuracy. The obtained experimental results show that the system yields the best performance for identifying a person and it is able to provide the highest degree of biometrics-based system security.**

***Index Terms*—Biometrics, Identification, Ear, Gabor filter, Data fusion.**

I. Introduction

THE capability of automatically establishing the identity of individuals, called as person verification or identification is essential to the reliability of several applications such as physical buildings, access control and information systems [1]. Biometric verification or identification is emerging as a powerful means for automatically recognizing a person's identity with a higher reliability [2]. Biometrics is the science of identifying people using physiological or behavioral traits, it has significant advantages over traditional authentication techniques namely passwords, ID card and key, due to biometric characteristics of an individual are not transferable and unique for every person and are not lost, stolen or broken [3].

In the past years researchers have exhaustively investigated the use of a number of biometric characteristics, including fingerprint, face, iris, retina, palmprint, hand geometry, voice, gait, signature, etc. Recently, a novel biometric trait, ear, has attracted an increasing amount of attention. Like any other biometric identifiers, ears are believed to have the desirable properties of universality, uniqueness, permanence and collectability for personal recognition [4]. In addition, there are several motivations for ear biometric [5]. Firstly, the ear data can be captured using conventional cameras. Secondly, the data collection is nonintrusive (i.e., requires no cooperation from the user). Thirdly, ear based access systems are very suitable for several usages. Finally, ear features are more stable over time and are not susceptible to major changes.

Biometric systems that use a single trait for recognition (i.e., unimodal biometric systems) are often affected by several practical problems [6] like noisy sensor data, non-universality and/or lack of distinctiveness of the biometric trait and unacceptable error rates. Multimodal biometric systems overcome some of these problems by consolidating the evidence obtained from different sources [7]. Among these sources, multiple representations (multiple algorithms for feature extraction) for the same biometric, is one combination that has been systematically used to improve the biometric based identification system accuracy. In this context, we propose in this paper a new scheme for improving the ear identification by combining the results of several representations extracted by Gabor filter technique. By using the real and imaginary responses of the Gabor filter, three feature vectors are then extracted: phase ($\mathcal{V}_{PH}$), module ($\mathcal{V}_{AM}$) and a combination of the real and imaginary parts ($\mathcal{V}_{RI}$). In this work, a series of experiments were carried out using the IIT Delhi touchless Ear Database. To evaluate the efficiency of our method, the experiments were designed as follow: the performances under different representations were compared to each other, in order to determine the best representation at which the ear identification system performs. However, because there are three representations, an ideal ear identification system should be based on the fusion of these representations at matching score level using different fusion rules.

The rest of the paper is organized as follows. The proposed scheme of the unimodal biometric system is presented in section II. Section III gives a brief description of the region of interest extraction. Feature extraction is discussed in section IV. This section including also an overview of the Gabor filter. In Section V we will discuss the similarity matching. The experimental results, prior to fusion and after fusion, are given and commented in section VI. Finally, the conclusion and further works are presented in sections VII.

II. Unimodal system description

Fig. 1 shows the block diagram of the proposed unimodal biometric identification system based on the ear images. In

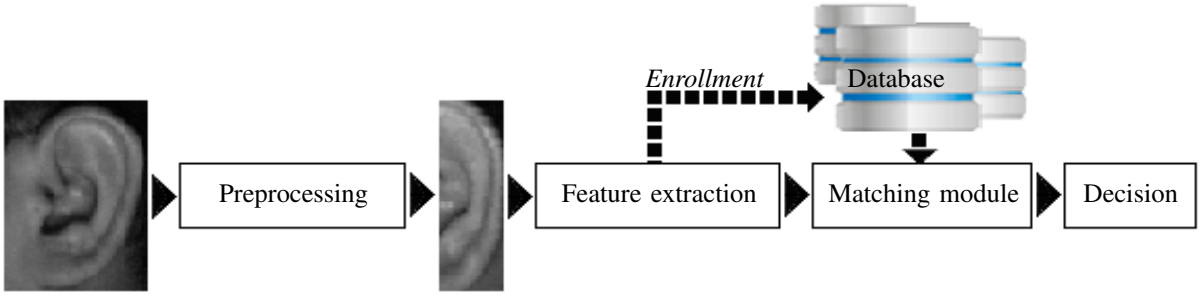


Fig. 1. Block diagram of the person identification system using ear images based on Gabor filter responses.

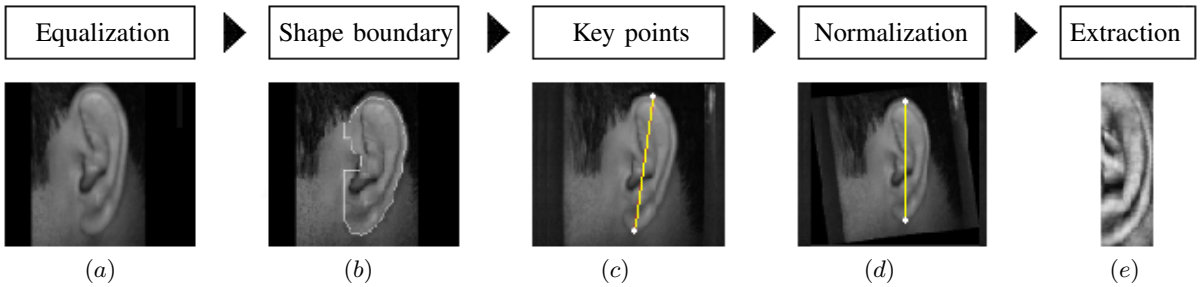


Fig. 2. Various steps in a typical region of interest extraction algorithm. *(a)* The image after histogram equalization, *(b)* The localization of ear shape boundary, *(c)* The extraction of stable key points, *(d)* The rotation of image (localization of the ROI), and *(e)* The preprocessed result (ROI).

the first phase the Region Of Interest (ROI) is located and extracted. In the second phase the texture of the ear image is extracted using Gabor filter responses (real, imaginary, module and phase parts). Finally, matching of the test ear image to the templates stored in the database. Based on this matching score, a decision about whether to accept or reject a user is made. The proposed multimodal system is composed of two sub-systems exchanging information in matching score level. Each sub-system exploits one of three feature extraction techniques.

III. Region of Interest Extraction

In order to localize the ROI area, the first step is to preprocess the ear images; we use the preprocessing technique described in [8] to align the ear images. In this technique, Gaussian filter which helps to suppress noise is used to smoothing the image, and then an histogram equalization is applied. After that, the boundary tracing of ear shape image is then employed to generate the ear shape boundary. The two key points on the located contour which achieve the maximum distance between them are selected as reference points on the reconstructed ear shape contour; these key points are used to align the ear ROI. Then, the image is rotated for normalized the ear ROI sub-image. Finally, the ROI part of the image, which is 180×50, is then cropped to represent the whole ear ROI sub-image. Fig. 2 shows the ear pre-processing steps.

IV. Feature Extraction

The feature extraction module processes the acquired biometric data and extracts only the salient information to form a new representation of the data. Ideally, this new representation should be unique for each person. In our system, feature representation is based on Gabor filter responses.

A. Gabor filter overview

The feature vectors are generated from the ROI sub-images by filtering the image with 2D Gabor filter [9]. Gabor filters (Gabor wavelets) can be used to extract components corresponding to different scales and orientations from images. In this work, we use the circular Gabor filter which has the following general form [10]:

$$h(x,y) = \frac{1}{2\pi\sigma^2} e^{-\{(x^2+y^2)/2\sigma^2\}} e^{2\pi i u(x\cos\theta + y\sin\theta)} \quad (1)$$

Where $i = \sqrt{-1}$, u is the frequency of the sinusoidal signal, θ controls the orientation of the function, and σ is the standard deviation of the Gaussian envelope. The parameters of Gabor filters were empirically determined for the acquired ear images. These were set as; orientation θ = 45^o, frequency u = 0.00568, and deviation σ = 1.000. The Gabor filters were implemented as $N \times N$, with N=16, spatial masks.

B. Feature vector generation

In our work, an efficiency method for personal identification using ear modality is presented. From each ear image, we design three sub-systems which employ different feature vectors. Thus, filtering the image $I(x,y)$ with the Gabor filter, $h(x,y)$, can be defined by:

$$\begin{aligned} I_G(x,y) &= h(x,y) * I(x,y) \\ &= \sum_{m=0}^{N-1}\sum_{n=0}^{N-1} h(m,n) I(x-m, y-n) \end{aligned} \quad (2)$$

Where $*$ denotes discrete convolution. Thus, the results $Re\{I_G\}$, and $Im\{I_G\}$ of a pair of a real and an imaginary filter are combined in order to produce three vectors {combined

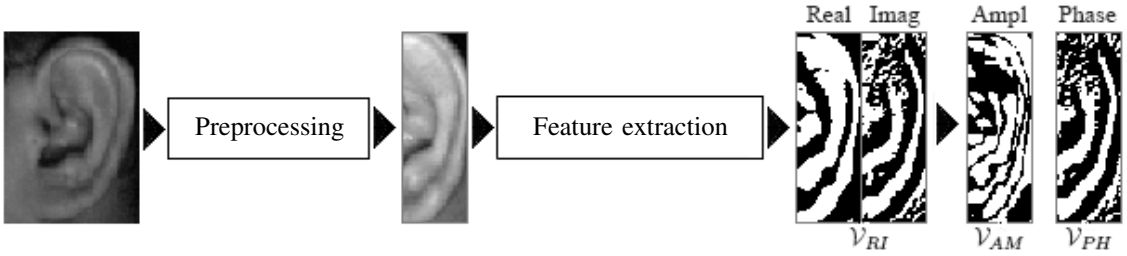


Fig. 3. Block diagram of the feature extraction process based on Gabor filter responses.

Real-Imaginary ($\mathcal{V}_{RI}$), PHase response ($\mathcal{V}_{PH}$) and AMplitude response ($\mathcal{V}_{AM}$)} :

$$\mathcal{F}_{RI} = [Re\{I_G\} \quad Im\{I_G\}] \tag{3}$$

$$\mathcal{F}_{PH} = \arctan\left[\frac{Im\{I_G\}}{Re\{I_G\}}\right] \tag{4}$$

$$\mathcal{F}_{AM} = \sqrt{Re\{I_G\}^2 + Im\{I_G\}^2} \tag{5}$$

These vectors should be binarized by a proper threshold value in order to generate the different feature vectors. It is important to find the proper threshold value in order to separate the lines from ear image. For our method, $\mathcal{F}_{RI}$ and $\mathcal{F}_{PH}$ are binarized by the threshold value equal to 0 and $\mathcal{F}_{AM}$ by the threshold equal to *mean*($\mathcal{F}_{AM}$). Finally, the feature vectors are obtained by:

$$\mathcal{V}_X(i,j) = \begin{cases} 1 & \text{if} \quad \mathcal{F}_X(i,j) \geq T_X \\ 0 & Otherwise \end{cases} \tag{6}$$

where $X = \{RI, PH, AM\}$. Fig. 3 shows the feature extraction method using in our work.

V. Matching and normalization method

A. Feature matching

Each ear image is represented by a unique binary feature (Template). The matching between an input and a stored template consists of computing matching scores between them. The matching task in our experimental schemes based on a normalized Hamming distance [11]. It is defined as the number of places where two templates differ. The Hamming distance (d_h) can be defined as:

$$d^h = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\mathcal{V}_X^t(i,j) \oplus \mathcal{V}_X^r(i,j) \tag{7}$$

where $\mathcal{V}_X^t$ and $\mathcal{V}_X^r$ are the test (input) and stored templates. The $\oplus$ is the Boolean operator (XOR) and N×N is the size of the templates. It is noted that d_h is between 1 and 0. For perfect matching, the matching score is zero. In order to further reduce the variation of the translation, all the ROI sub-images are translated by some pixels (-2, -1, 1, 2). For the ear based identification application, the score vector is given by:

$$\mathcal{D}_{\mathcal{H}} = [d_1^h,\ d_2^h,\ d_3^h,\ d_4^h \cdots d_w^h] \tag{8}$$

were w is the system database size.

B. Normalization method

The matching scores output by the various sub-systems are heterogeneous; score normalization is needed to transform these scores into a common domain, prior to combining them. Thus, a *Min-Max* normalization scheme was employed [12] to transform the scores computed into similarity scores in the same range. Thus,

$$\widetilde{\mathcal{D}}_H = \frac{\mathcal{D}_{\mathcal{H}} - min(\mathcal{D}_{\mathcal{H}})}{max(\mathcal{D}_{\mathcal{H}}) - min(\mathcal{D}_{\mathcal{H}})} \tag{9}$$

where $\widetilde{\mathcal{D}}_H$ represent the normalized vector. However, these scores are compared, and the lowest score is selected. Therefore, the best matching score is d_0^h and its equal to $min(\widetilde{\mathcal{D}}_H)$.

VI. Experimental results

A. Experimental database

The used touchless ear image database [13] is collected from the students and staff at IIT Delhi, India. All the images are acquired from a distance (touchless) using simple imaging setup and the imaging is performed in the indoor environment. The currently available database is acquired from the 221 different subjects and each subject has at least three ear images. All the subjects in the database are in the age group 14-58 years. The resolution of these images is 272×204 pixels and all these images are available in jpeg format.

B. Unimodal systems test results

In the all experiments, one image is randomly selected of the available images of each person was used in the enrolment stage to create the system database; the remaining images were used for testing. In the system performance evaluation, we setup a database with size of 221 persons, which are similar to the number of employees in small to medium sized companies. Thus, the error rates were performed by comparing the test images with the stored templates in the database. A total of 126412 comparisons were made.

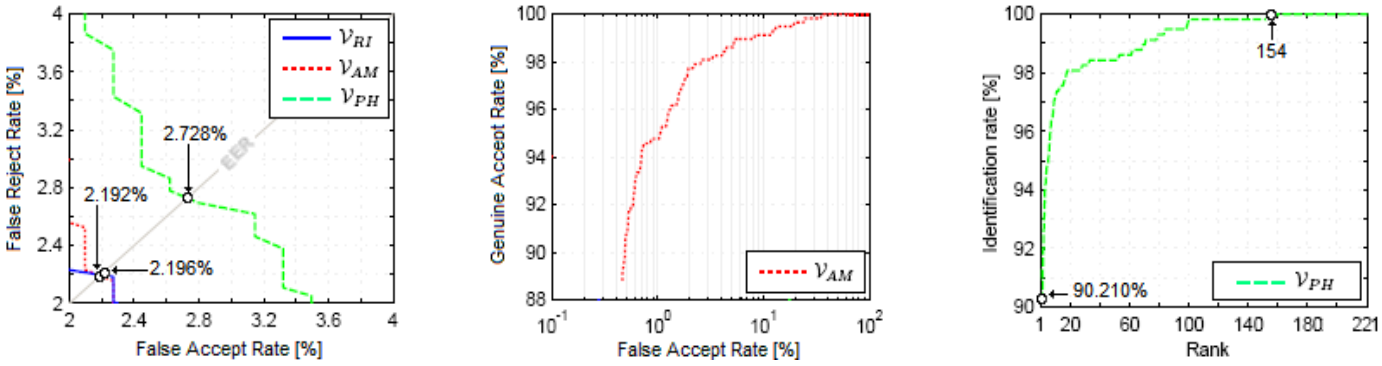
At the first stage, we conducted several experiments to selection the best feature vector of the three proposed vectors used ($\mathcal{V}_{PH}$, $\mathcal{V}_{AM}$ and $\mathcal{V}_{RI}$). This is carried out by comparing all these vectors and finding the vector that gives the best identification rate (choose the feature vectors such that the Genuine Acceptance Rate (GAR) is maximized). Thus, in the case of *open set* identification, the Receiver Operating Characteristic (ROC) curves for the three feature vectors are shown in Fig. 4.*(a)*. The experimental results indicate that

TABLE 1 : UNIMODAL OPEN SET IDENTIFICATION SYSTEM PERFORMANCE.

DB	$\mathcal{V}_{RI}$		$\mathcal{V}_{PH}$		$\mathcal{V}_{AM}$	
	T_o	EER	T_o	EER	T_o	EER
221 users	0.2352	2.728	0.2617	2.196	**0.2624**	**2.192**

TABLE 2 : UNIMODAL CLOSED SET IDENTIFICATION SYSTEM PERFORMANCE.

DB	$\mathcal{V}_{RI}$		$\mathcal{V}_{PH}$		$\mathcal{V}_{AM}$	
	ROR	RPR	ROR	RPR	ROR	RPR
221 users	87.762	191	**90.210**	**154**	88.462	133



(a) (b) (c)

Fig. 4. Results of unimodal *open/closed* set identification system. (*a*) The ROC curves with respect to the differents feature vectors (*b*) The ROC curve in the case of $\mathcal{V}_{AM}$ based *open set* identification system and (*c*) the CMC curve in the case of $\mathcal{V}_{AM}$ based *closed set* identification system.

the $\mathcal{V}_{AM}$ perform better than the $\mathcal{V}_{PH}$ and $\mathcal{V}_{RI}$ vectors in terms of Equal Error Rate (EER). Therefore, the system can achieve a higher accuracy at the $\mathcal{V}_{AM}$ vector compared with the other vectors with an EER equal to 2.192% at the threshold T_o = 0.2624. The ROC curve, which plot the GAR against False Accept Rate (FAR), for the best case ($\mathcal{V}_{AM}$ vector) is shown in Fig. 4.*(b)*. Finally, the performance of the *open set* identification system under all vectors is shown in Table 1.

In the case of a *closed set* identification, a series of experiments were carried out to select the best feature vector, this has been done by comparing all feature vectors and finding the vector that gives the best identification rate. Table 2 present the experiments results obtained for all vectors. From Table 2, the best results of Rank-One Recognition (ROR) produce 90.210 % with lowest Rank of Perfect Recognition (RPR) of 154 in the case of $\mathcal{V}_{PH}$ vector. The results expressed as a Cumulative Match Curves (CMC) obtained by the proposed scheme, in the case of $\mathcal{V}_{PH}$ vector, is plotted in Fig. 4.*(c)*.

C. Multimodal systems test results

The goal of the fusion process is to improve the unimodal system performance by fusing the information from different feature vectors. The system is then considered as a multimodal system where their inputs are different feature extraction methods (different vectors) for each ear image. Therefore, several multimodal systems are tested in order to choose the best one. In this work, the fusion is performed at the matching score level were several fusion rules are tested.

In the case of *open set* identification case, the individual scores using the three vectors are combined to generate a single scalar score, which is then used to make the final decision. Table 3 provides the performance of the identification system for several combination and fusion rules. From this Table, it is clear that our *open set* identification system achieves a best performance when using the fusion of $\mathcal{V}_{AM}$-$\mathcal{V}_{RI}$ and WHT fusion rule (EER = 1.193% and T_o = 0.2697). Compared with the previous results ($\mathcal{V}_{AM}$ based unimodal system), the proposed multimodal identification has achieves better results expressed in terms of the EER ($\simeq$ 54.500 % improvement). Fig. 5.*(a)* shows the comparison test. Finally, graphs showing the ROC curve, plot GAR against FAR, for the open set identification using unimodal and multimodal systems, were generated, see Fig. 5.*(b)*.

We also investigated the *closed set* identification system performance, thus, a series of experiments were carried out using the ear database to selection the best fusion rule that maximize the ROR rate. Thus, to determine the best fusion rule, Table 4 can be established. We can observe that the SUM rule based fusion and the fusion of all feature vectors has the best performance. Thus, the best result of ROR is given as 92.398 % with lowest RPR of 119. From this result, the performance of the *closed set* identification system is significantly improved by using the fusion. Finally, the comparison of the unimodal and multimodal *closed set* identification system is plotted in Fig. 5.*(c)*.

VII. CONCLUSION AND FURTHER WORK

This work describes the design and development of a multi-representations biometric personal identification system based on features extracted from ear image. Furthermore, the

TABLE 3 : MULTIMODAL OPEN SET IDENTIFICATION SYSTEM PERFORMANCE.

FUSION	SUM		WHT		MAX		MIN		MUL	
	T_o	EER	T_o	EER	T_o	EER	T_o	EER	T_o	EER
$\mathcal{V}_{RI}$-$\mathcal{V}_{PH}$	0.2542	2.098	0.2575	2.098	0.2820	2.098	0.2198	2.448	0.0650	2.098
$\mathcal{V}_{RI}$-$\mathcal{V}_{AM}$	0.2763	2.083	**0.2697**	**1.193**	0.3110	2.331	0.2120	2.098	0.0786	2.098
$\mathcal{V}_{PH}$-$\mathcal{V}_{AM}$	0.2732	1.985	0.2854	2.128	0.3256	2.149	0.1852	2.061	0.0322	2.310
$\mathcal{V}_{RI}$-$\mathcal{V}_{AM}$-$\mathcal{V}_{PH}$	0.2900	1.981	0.2952	2.028	0.3364	2.249	0.1858	1.761	0.0209	1.759

TABLE 4 : MULTIMODAL CLOSED SET IDENTIFICATION SYSTEM PERFORMANCE.

FUSION	SUM		WHT		MAX		MIN		MUL	
	ROR	RPR	ROR	RPR	ROR	RPR	ROR	RPR	ROR	RPR
$\mathcal{V}_{RI}$-$\mathcal{V}_{PH}$	91.674	128	91.855	122	91.674	149	90.317	137	90.317	125
$\mathcal{V}_{RI}$-$\mathcal{V}_{AM}$	91.493	156	91.493	147	91.312	123	88.869	134	88.869	131
$\mathcal{V}_{PH}$-$\mathcal{V}_{AM}$	92.217	102	92.217	102	91.674	116	89.864	102	89.864	102
$\mathcal{V}_{RI}$-$\mathcal{V}_{AM}$-$\mathcal{V}_{PH}$	**92.398**	**119**	92.036	114	91.674	110	88.597	118	88.597	153

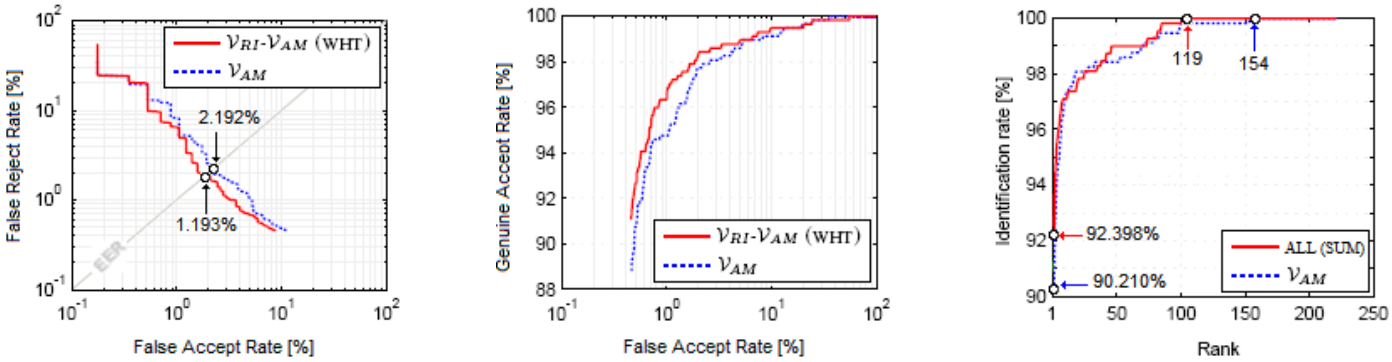


Fig. 5. Results of multimodal *open/closed set* identification system. *(a)* The ROC curves, FRR against FAR, with respect to the unimodal and multimodal systems, *(b)* The ROC curves, GAR against FAR, with respect to the best case and *(c)* The CMC curves with respect to the unimodal and multimodal systems.

unimodal systems suffer from various problems affecting the system performance. These problems are effectively handled by multimodal systems. In this paper, three different sub-systems derived from each modality were used. Fusion of the proposed sub-systems is performed at the matching score level to generate a fused matching score which is used for recognizing an image. The experimental results, obtained on a database of 221 users, shown accepted *open/closed* set identification accuracy. They also demonstrate that combining different system does significantly perform the accuracy of the system. For further improvement, our future work will project to combine this modality with other modalities like face, iris and retina as well as the use of other fusion level like feature and decision levels. Also we will focus on the performance evaluation in both phases (verification and identification) by using a large size database.

REFERENCES

[1] Jinfeng Yang, Yihua Shi, Jinli Yang, "Personal identification based on finger-vein features", Computers in Human Behavior, Vol. 27,pp. 1565-1570, 2011.

[2] Anil K. Jain, Arun Ross, and Sharath Pankanti, "Biometrics: A Tool for Information Security", IEEE transactions on information forensics and security, Vol. 1, No. 2, June 2006

[3] Jin-Rong Cui, "Multispectral palmprint recognition using Image Based Linear Discriminant Analysis", International Journal of Biometrics, Vol. 4, No. 2 pp. 106-115, 2012.

[4] K. Chang, K.W. Bowyer, S. Sarkar, and B. Victor, "Comparison and combination of ear and face images in appearance based biometrics", IEEE Trans. Pattern Anal. Mach. Intell., Vol. 25, No. 9, pp. 1160-1165, Sept. 2003.

[5] Mu Zhichun, Yuan Li, Xu Zhengguang, "Shape and Structural Feature Based Ear Recognition", Advances in Biometric Person Authentication, Guangzhou, China, pp. 663-670, 2004.

[6] A. Meraoumia, S. Chitroub and A Bouridane, "Fusion of Finger-Knuckle-Print and Palmprint for an Efficient Multi- Biometric System of Person Recognition", IEEE International Conference on Communications (ICC), Kyoto, Japan, pp. 1-5, June 2011.

[7] Lin Zhang, Lei Zhang, David Zhang and Hailong Zhu, "Ensemble of Local and Global Features for Finger-Knuckle-Print Recognition", Pattern Recognition, Vol 44, No 9, pp. 1990-1998, 2011.

[8] Ajay Kumar, Chenye Wu, "Automated human identification using ear imaging", Pattern Recognition, Vol xx, No x, pp. xxxx-xxxx, 2011.

[9] S. Senapati, G. Saha, "Speaker Identification by Joint Statistical Characterization in the Log-Gabor Wavelet Domain", International Journal of Intelligent Systems and Technologies, winter 2007

[10] D. Zhang, W. Kong, J. You, and M. Wong, "On-line Palmprint Identification", IEEE Trans. on PAMI, Vol. 25, No. 9, pp. 1041-1050, 2003.

[11] F. Wang, J. Han, "Iris recognition method using Log-Gabor filtering and feature fusion", Journal of Xian Jiaotong University, Vol.41, 2007.

[12] A. Meraoumia, S. Chitroub and A. Bouridane, "Multimodal Biometric Person Recognition System Based on Multi-Spectral Palmprint Features Using Fusion of Wavelet Representations", Advanced Biometric Technologies. Published by InTech, pp. 21-42, 2011.

[13] IIT Delhi Touchless Palmprint Database version 1.0, Available online at: http://webold.iitd.ac.in/∼biometrics/Database_Ear.htm