Conférence Internationale en Automatique & Traitement de Signal (ATS-2017)
Proceedings of Engineering and Technology – PET
Vol.22 pp.78-83

Efficient Person Identification by Finger-Knuckle-Print Based on Discrete Cosine Transform Network and SVM classifier

Khaled Ben Sid[1], Djamel Samai[1], Fatima Zohra Laallam[1]and Amina Tidjani[1]

[1]Université Kasdi Merbah Ouargla, Laboratoire de Génie Électrique.
Faculté des Nouvelles Technologies de l'Information et de la Communication, Ouargla, 30000, Algérie
Email: khaled.sid@gmail.com, samai.djamel@univ-ouargla.dz, Flaalem@gmail.com, tidjani.amina1@gmail.com

***Abstract*—The recognition of individuals is more importance in daily human life. Because it ensures the transactions of people in differing domains in order to ensure a relevant security. Biometric technology systems constituted the basic rule for the secure identification of the identity. These systems have been known in different areas because it is convenient and easy to use. To identify persons, they rely on different distinguishing features of the human body such as fingerprints, face, the palm print and retina, etc. Finger knuckle print occupies a privileged position among the various biometric technologies because its benefits, and gives a best result, it was adopted in the proposed biometric system. In this paper, we design a multimodal finger knuckle print identification system based on new deep learning algorithm called Discrete Cosine Transform Network DCTNet and support vector machine (SVM) classifier.**

***Index Terms*—Biometric, finger knuckle print, multimodal, DCTNet, SVM.**

I. Introduction

BIOMETRIC technologies are crucial components of secure personal identification and verification systems, which control access to valuable information, to economic assets, and to parts of the national infrastructure. This technologies are automated methods for identify in a person or verifying a persons identity based on the pensons physiological or behavioral characteristics ; physiological characteristics include fingerprints, hand geometry, and facial, voice, iris, and retina features ; behavioral characteristics include the dynamics of signatures and keystrokes, biometric technologies capture and process a persons unique characteristics , and then verify that persons identity based on comparison of the record of captured characteristics with a biometric sample presented by the person to be verified [1], [2], [3].

Instead of the broad categories (physiological, behavioral and biological attributes), for convenience the physiological modalities can be further sub-divided into different subcategories according to their respective position in human body such as : hand region attributes, facial region attributes, ocular and periocular region attributes, behavioral attributes, and medico-chemical attributes [4].

After many years of research and development, biometric technologies have become reliable and cost-effective, and acceptable to users, however, new applications of biometrics are being somewhat successfully implemented in more secure travel documents, visas, and personal identity verification cards.

The hand is the body part at the end of your arm that includes your fingers and thumb, it composed by five parts contains rich texture information that provided the foundations for early recognition systems [5].

The multimodal biometrics is a promising area of information processing research which is directed towards understanding of traits and methods for more accurate and reliable personal information representation for subsequent decision making and matching. In the recent years, there is a significant increase in research activity directed at understanding all aspects of biometric information system representation and utilization for decision making support, for use by public and security services, medical diagnostics, and for understanding the complex processes behind biometric matching and recognition [6].

The feature extraction, or extraction of the observation vector, is an essential step in a biometric system. A feature vector (observation vector) is used to represent the discriminant characteristic of the biometric modality of a person, a reduced overall dimension with respect to the image. This vector can be modeled using the different techniques in order to have a better identification [7].

In this study, we chose a person's identification system using finger knuckle print features. This system uses The shape of the outer part of the finger knuckle as a biometric features. These characteristics are permanent and stable throughout life, as unique to each and another. This work aims at achieving the unimodal and multimodal biometric systems based on multi-sample finger knuckle print images and from feature extraction we use a deep learning DCTNet.

The goal of multi-sample images is to increase the performance of biometric system and also increases the value of safe and trust to security systems based on biometric technologies and for DCTNet is to view different representation of several levels to give together upper-level characteristics can efficiently represent the discriminating characteristics of the FKP images.

The remainder of the paper is organized as follows.The system design for multimodal finger knuckle print identification is exposed in section II. Section III includes also an overview of DCTNet Architecture. The proposed feature matching method based on Support Vector Machines(SVM) algorithm and the

Copyright IPCO-2017
ISSN 2356-5608

fusion rules used to merge the information provided by the feature extraction are detailed in section IV. The experimental evaluation, prior to fusion and after fusion, are given and commented in section VI. Finally, section VII is devoted to the conclusion and future work.

II. System Design

In this paper, we proposed to fuse different samples of finger knuckle print features. this system is composed of two biometric based subsystems. Each subsystem exploits different biometric techniques which are left index finger (*LIF*), left middle finger (*LMF*), right index finger (*RIF*) and right middle finger (*RMF*) modalities. Any biometric system compose of two phases: an enrollment phase and an identification/verification phase. It consists of pre-processing process, matching process, normalization and decision process. Figure 1 shows the schematic diagram of the proposed system using finger knuckle print images.

III. Feature extraction

A. DCTNet Architecture

After the success of PCANet [8] in image classification, the DCTNet [9] is the new variant of a deep learning method. Its a very similar structure to PCANet except there is an extra layer at the histogram output for histogram normalization. Thus, the block diagram of DCTNet algorithm presented in Fig. 2 can be described as follows:

1) Convolution Layer : As shown in Fig. 2 in convolutional layers, the output of input image I_d of size nm with D channels given by

$$O_d^p = \{I_d * W_l^p\}_{p=1}^{p_l} \tag{1}$$

where $*$ denotes the discrete convolution and the size of output O_d^p is same as I_d and $W_p^l \in R^{k \times k}, p = 1, 2, ..., P_l$ is 2D-DCT bases from P_l filters at layer l.
Of each layer, we have d outputs. The cascaded position of this layer can form a deeper network. Since there is no non-linear operation between the previous convolution layer and the next layer, DCT bases of each layer can be combined to form a flat single layer network. The number of bases formed is $\prod_{i=1}^{L} P_i$ where L represents the number of convolution layers.

2) Binary hashing: The outputs of convolution layer of DCTNet give real values. In this step, the output obtained in the last layer is turned into binary format with the comparison of responses with threshold to zero (value 1 for positive response, zero otherwise)denoted by $BIN(.)$.

$$BIN(O_d^p) = \begin{cases} 1 & \text{if } O_d^p \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $BIN(O_d^p)$ is a binary image. In addition, around each bit is encoded as a single integer and forms an image for each output defined $\sum_p^{P_L} 2^{p-1} BIN(O_d^p)$ where each pixel has an integer range of $[0, 2^{P_L - 1}]$.

3) Block-wise Histograms: Then, after the binary step, each of these binary images is divided into B non-overlapping blocks. The characteristics of these images are obtained by concatenating all the histograms of each block B such as:

$$H = \{H_b^d\}_{b=1,d=1}^{B,D} \tag{3}$$

where $b = 1, 2, ..., B; d = 1, 2, ..., D$. The combination of binary hashing and block-wise histograms should be able to extract discriminating characteristics.

4) Histogram Tied Rank Normalization (TR Normalization): Each H_b^d is ranked with the linked ranking without considering the locker with a zero occurrence noted by $\bar{H}_b^d$. This is because bin with zero occurrences is not a sample in the histogram, it must be ignored in the ranking process. To make $\bar{H}_b^d$ more equitably distributed, we calculate the square root of $v_b^d = \sqrt{\bar{H}_b^d}$, and we use the idea of normalization of the norm $L2$ [10] to obtain $\hat{v}_b^d$. The final TR normalized histogram feature vector of the input image is obtained by concatenating all the $\hat{v}_b^d$

$$v = [\hat{v}_1^1, \hat{v}_2^1, ..., \hat{v}_B^1, \hat{v}_1^2, ..., \hat{v}_B^D] \in R^{(2^{P_L})BD} \tag{4}$$

Finally, the TR normalization step is not necessary because we can use the final histogram of Block-wise Histograms step as the feature vector.

5) Features histogram reduction using (WPCA) : A dimension reduction technique based on the traditional PCA [11] transforms the X dataset to a new set of Y dimension, keeping up the essential information of the starting set. Generally, neither the geometry of the variety or the dimension d are known. Size reduction techniques can be classified into several groups. The main classification criterion is the linear appearance or not the methods. Linear methods assume that the data is based on a linear variety of the large space.
In PCA, feature components in different dimensions are unequally treated due to different eigenvalues. Actually, it is hard to determine which component is more important than others without prior knowledge in the recognition process. So, we need to normalize the feature components according to the eigenvalues. a WPCA feature space with unit variance by weighting the features and then palm recognition is performed in the new subspace [12].

IV. Feature matching, Normalization and fusion method

Matching process determines the similarity/dissimilarity between two given templates. This step involves modeling the parameters extracted from one modality of an individual based on their common characteristics.The top match can be determined by examining the match scores pertaining to all the comparisons and reporting the identity of the template corresponding to the largest similarity score. A model is a set of useful information, discriminatory and non-redundant characteristic of one or more individuals with similarities, they will be grouped in the same class, and these classes vary according to the type of decision.

Copyright IPCO-2017
ISSN 2356-5608

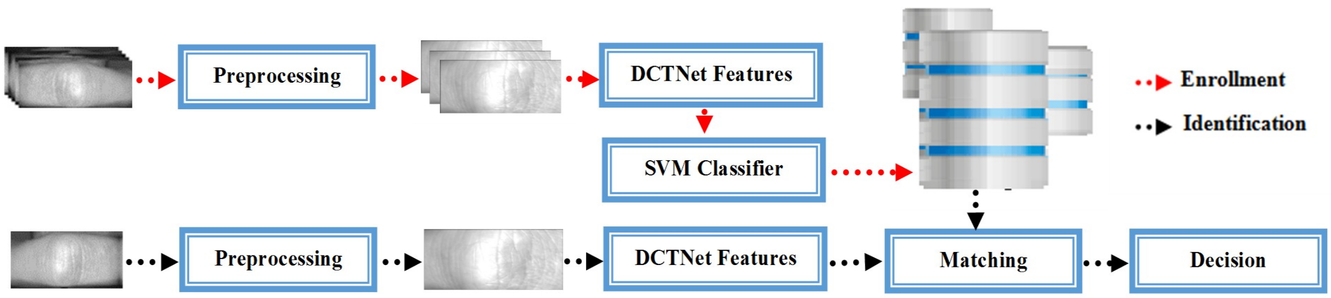

Fig. 1. Multimodal finger knuckle print identification system .

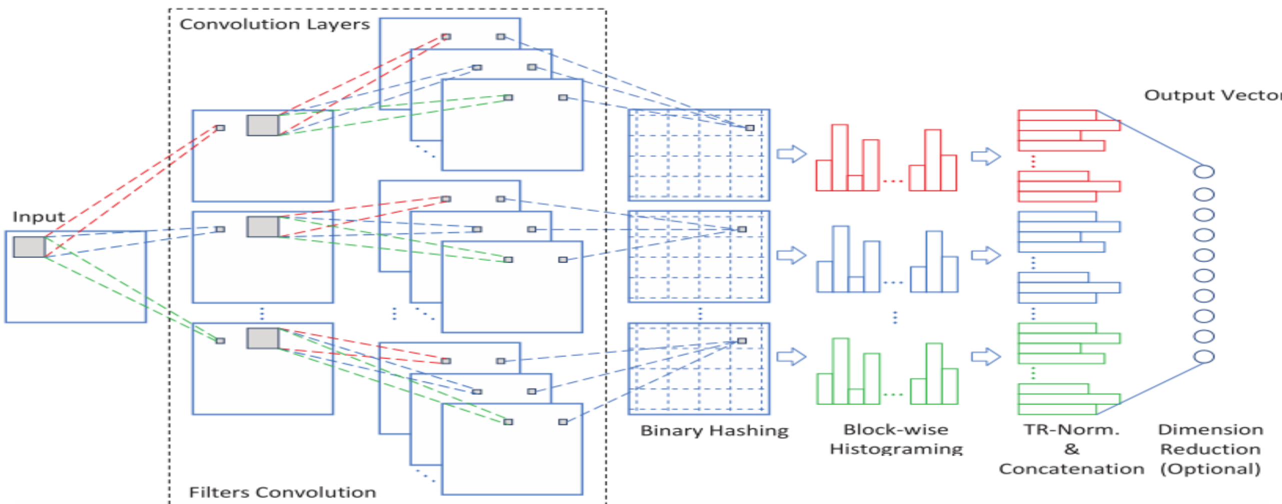

Fig. 2. The block diagram of the proposed DCTNet [9].

A. Support Vector Machine SVM

A support vector machine is a technique of discrimination, it is a supervised learning method for classification and regression. It consists in separating two or more sets of points by a hyperplane. Depending on circumstances and configuration points. The original idea of SVM is based on using kernel core functions that allow optimal separation of the points of the plan in different categories. The method uses a set of training data. which enables a hyperplane separating the best points. In this paper we use the multi class SVM [13].

B. Normalization method

Score normalization is needed to transform these scores into a common domain, prior to combining them. Thus, a Min-Max normalization scheme was employed. To transform the scores computed into similarity scores in the same range. Thus,

$$\tilde{D} = \frac{D - \min(D)}{\max(D) - \min(D)} \tag{5}$$

Where $\tilde{D}$ represent the normalized vector. However, these scores are compared, and the lowest score is selected. For perfect matching, the matching score is zero.

V. Fusion process

Fusion at the matching score level is the most popular and frequently used method because of its good performance and simplicity. The outputs of the two or more matching modules (LIF, LMF, RIF, RMF) are combined using fusion at the matching-score level.
The object is to combine these scores to generate a single score which is then used to make the final decision. The fusion is expressed by different rules on the matching scores obtained from these matching modules.

A. Score fusion rule

There are several matching score fusion rules integrate normalized matching scores of a user to produce the final matching score. In this work we used Score fusion rule [14].

- Simple Sum rule : The Simple Sum rule takes the sum of the R matching scores of the (k)th user as the final matching score S_k of this user. S_k is calculated as follows:

$$S = \frac{1}{N}\sum_{i=1}^{N} S_i \tag{6}$$

- The product rule : This rule defines the new scores for

Copyright IPCO-2017
ISSN 2356-5608

each matcher, is calculated as follows:

$$S = \frac{1}{N} \prod_{i=1}^{N} S_i \tag{7}$$

- The minimum rule : This rule simply sets a new scores as the minimum score of each matcher's scores, is calculated as follows:

$$S = \min(S_i) \tag{8}$$

- The maximum rule : This rule simply sets a new scores as the maximum score of each matcher's scores, is calculated as follows:

$$S = \max(S_i) \tag{9}$$

The final result of the fusion is a new matching score, which is the basis for the classification decision of the entire system.

VI. Experimental results and discussion

A. Experimental database

We experimented our approach on hong kong Polytechnic University (PolyU) Finger-Knuckle-Print Database [15]. The database has a total of 7920 images obtained from 165 persons. This database including 125 males and 40 females. Among them, 143 subjects are $20 - 30$ years old and the others are $30 - 50$ years old. These images are collected in two separate sessions. In each session, the subject was asked to provide 6 images for each of LIF, LMF, RIF and RMF. Therefore, 48 images from 4 fingers were collected from each subject.
If to develop a finger knuckle print recognition system, it is necessary to have two databases: a database to perform learning and other techniques to test and determine their performance, but there are no rules to determine what share of quantitative manner. It often results from a compromise based on the number of available data and time to perform learning. In the series of tests we have done the base was divided as follows:
Images of learning: the first, fifth and ninth image of each person to serve learning phase.
Tests Images : The remaining 9 images of each individual have helped us achieving different tests.

In order to properly analyze our identification system, and in order to achieve satisfactory results, we divided our work into three parts:
In the first we studied the influence of the parameter of DCTNet in the identification rate of biometric system, and we choose the best parameter of DCTNet feature extraction algorithm. After that we used the DCTNet algorithm for extracting the features of finger knuckle print. These algorithms are among the best current texture descriptors. We conducted several experiments to see what is the best finger that give powerful results.
In the end we have merged the different finger knuckle print samples for increased the performance of the identification system. in this case we use several rules.

B. DCTNet parameter adaptation

The GAR (Genuine Accept Rate) shows the performance of biometric system because it given the value of identification rate, another hand the FRR (False Reject Rate) and FAR (False Attempts Accepted) gives the biometric system error rate, on the one hand it rejects the users as well as the acceptance of the impostors. In the DCTNet structure shows in Fig. 2 five important parameter the number of layer and number of filter in each layer and filter size and block wise histogram size. In this stage, we select the optimal parameter for DCTNet this in order to obtain the best value for the identification rate.
From previous studies we concluded that two layer give the best result compared to one and three layer but if the number of layer larger of three it creates a calculation problems, which requires special equipment calculator intensive.
Fig. 3 illustrates the GAR against the number of filter in first

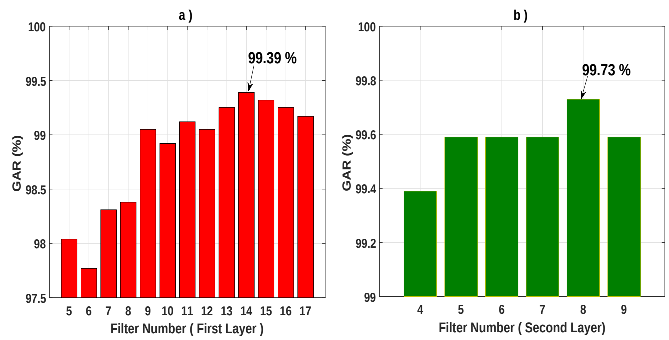

Fig. 3. DCTNet parameter adaptation a) GAR against Filter number in first layer, b) GAR against Filter number in second layer.

layer and the GAR against the filter number in second layer. In Fig. 3 a), we note that: the accuracy of the identification becomes very high at a number of filters 9 - 17, it gives the best result GAR = 99.39% when the number of filter equal 14. After that, we prove the number of filter in first layer in 14 and we change the number of filter in second layer. In In

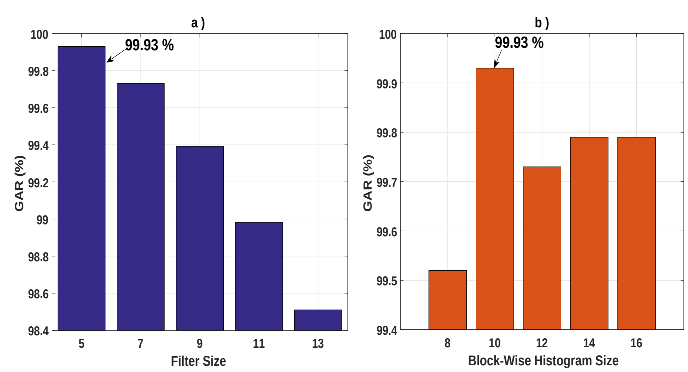

Fig. 4. DCTNet parameter adaptation a) GAR against Filter size, b) GAR against block wise histogram size.

Fig. 3 b), we note that : All filters give better results when they give the identification rate higher than 99.20%, but the

Copyright IPCO-2017
ISSN 2356-5608

perfect given by number of filter equal 8.
From the above we conclude that: The influence of filter number is great in the biometric system performance because the filter number is the essential of DCTNet algorithm.
Now, we will see the impact of the filter size and the block wise histogram size in the identification rate of identification system. Fig. 4 shows the GAR against the filter size and the GAR against the block wise histogram size.
In Fig. 4 a), we note that: as a general remark the filter size must be an odd number, the best results given when the filter size less than $[9 \times 9]$ but from the filter size larger to $[9 \times 9]$ we see a decline in the identification rate less than 98.00%. In this case, the perfect result given when the filter size equal $[5 \times 5]$ it gives GAR =99.93%.
In Fig. 4 b) we note that : as a general remark the block wise histogram size must be pair number, all the value gives a best result $> 99.00\%$ and the perfect is 99.93% from the block wise histogram size equal $[10 \times 10]$.
At the end, and after the comparison of the previous results, we conclude the best parameters of DCTNet Algorithm are:
The number of layers = 2
The number of filters in the layers = $[14 \quad 8]$
The filter size = $[5 \times 5]$
The block-Wise histogram size = $[10 \times 10]$

C. The application on a unimodal system

There are two types in the identification mode, open set identification as the system allows anyone outside the user and the identification as a closed set as the system allows users only.
The goal of this experiment was to evaluate the system performance when we using information from each modalities (each finger). For this, in Open Set identification we found the performance under different modalities (LIF, LMF, RIF, RMF). Table I and Fig 5 compares the performance of the unimodal system using DCTNet feature extraction for varying fingers. The experimental results indicate that the RIF perform better than the LIF, LMF and RMF in terms of EER. It gives EER $= 0.016\%$ and $T_0 = 0.8017$. But in the general case, the rest finger gives best result when gives EER equal 0.06%.
Fig 5 b) shows the CMC curve (the identification rate against the rank) of closed set identification system in this figure, all fingers give the best result when they give the identification rate higher than 99.70%, but the better is 99.86% from LMF, RIF, RMF fingers and the perfect is RIF because it is the first of up to 100% RPR equal 04.

From a practical standpoint, it is preferable to reduce the size of the samples after the DCTNet method with less effect on the results. For this, insert the WPCA method that has a very high reputation in terms of recognition. Thus, for effective classification of palm prints used, we used the famous SVM classifier.
✓The use of the reduction in the size of the feature vectors and the SVM does not affect the results obtained, on the contrary, they are a little improved. Indeed, this improvement is due the WPCA keeps the relevant information.

TABLE I
THE PERFORMANCE OF THE UNIMODAL SYSTEM USING DCTNET FEATURE EXTRACTION

DCTNet feature extraction					
Data Base		OPEN SET		CLOSED SET	
	Type	EER	T_0	ROR	RPR
165 person	LIF	0.060%	0.6657	99.73%	26
	LMF	0.067%	0.5770	99.86%	23
	RIF	0.016%	0.8017	99.86%	04
	RMF	0.0673%	0.6209	99.86%	67

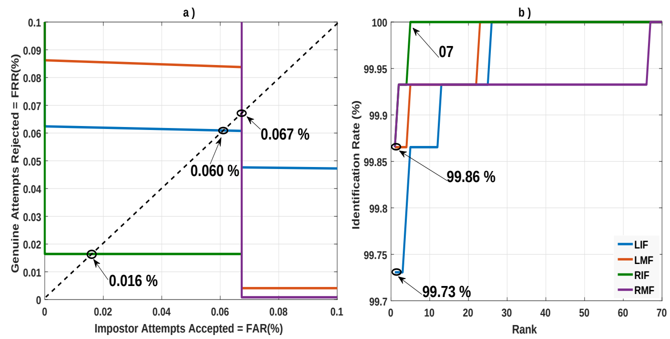

Fig. 5. Unimodal system results a) ROC curves b) CMC curves.

D. The application on a multimodal system:

he goal of the fusion process is to improve the performance by fusing the information from different modalities. To improve more our results, we will try to merge the different scores for different finger to obtain a multimodal system.
In this case, we merge the different samples of some fingers (LIFE and LMF, RIF and RMF) and at the end realize a system based on the fusion between the two fingers. Table II and Fig 6 shows the performance of the multimodal identification system the result indicate that : in open set identification all the fusion method give a better result EER equal 0.00%, and the same in closed set identification all fusion method give a better identification rate 100% in first rank.

TABLE II
THE PERFORMANCE OF THE MULTIMODAL SYSTEM

DCTNet feature extraction					
Data Base		OPEN SET		CLOSED SET	
	Fusion Type	EER	T_0	ROR	RPR
165 person	LIF+LMF	0.00%	0.719	100%	01
	RIF+RMF	0.00%	0.711	100%	01
	LF+RF	0.00%	0.387	100%	01

Table III and Fig. 7 show the performance of the multimodal identification system using different fusion rules, from the results we note that the sum rule gives the best result EEE equal 0.00% and the identification rate equals 100% the other rules gives the best result except the Max rule gives a best result in open set identification as a result less in closed set identification 99.79%.

Copyright IPCO-2017
ISSN 2356-5608

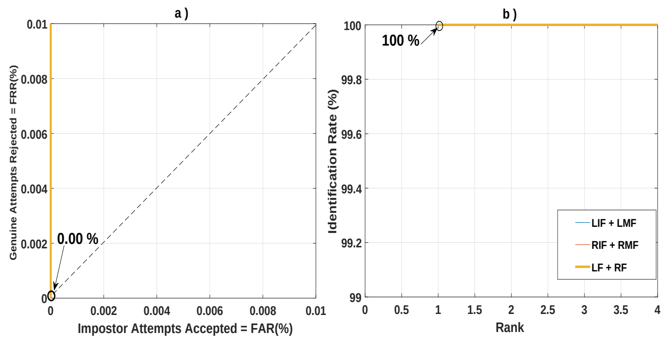

Fig. 6. Multimodal system results a) ROC curves b) CMC curves.

TABLE III
THE PERFORMANCE OF THE MULTIMODAL SYSTEM

Data Base	Fusion Rules	OPEN SET		CLOSED SET	
		EER	T_0	ROR	RPR
		DCTNet features			
165 persons	*Sum*	0.00%	0.7199	100%	01
	Mul	0.00%	0.417	100%	01
	Min	0.00%	0.768	100%	01
	Max	0.00%	0.995	99.97%	02

✓The accuracy of the multimodal system is better than the uni-modal system.

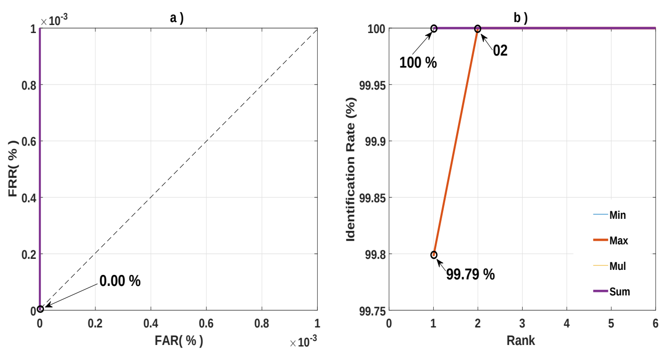

Fig. 7. Multimodal system results a) ROC curves b) CMC curves.

✓The multimodal system have a ($EER = 0.00\%$) and an ($ROR = 100\%$) and an ($RPR = 01$), thereby obtaining a perfect result. This ideal precision can be reduced to a large database.
✓The sum rule is the best compared to another rules because it give a perfect result and it sample to use.

VII. CONCLUSION AND FURTHER WORK

The objectives pursued in this paper suggest a path that is to improve the performance of the biometric identification via the finger knuckle print by a simple deep learning feature extraction methods called DCTNet. For this, we made the comparison in order to select the better parameter of the DCTNet. To make our system more convenient, we have reduced the size of the feature space by weighted principal component analysis algorithm key followed by support vector machine in matching module. In the end, the results are very interesting. In fact we got a great recognition rate of 100%, this rate is very interesting what makes our reliable system where it meets the objective that we set at the start, namely the implementation of a system for the recognition of individuals. For further improvement, our future work will project to use other finger knuckle print databases (CASIA). Or different biometric modalities (Face and Finger vein) as well as the use of other fusion level like sensor and feature and decision levels.

REFERENCES

[1] David D. Zhang, "Automated Biometrics Technologies and Systems", Originally published by Kluwer Academic Publishers, New York in 2000.
[2] Anil K. Jain, Ruud Bolle, Sharath Pankanti., "BIOMETRICS Personal Identification in Networked Society", Kluwer Academic Publishers New York, Boston, Dordrecht, London, Moscow, 2002.
[3] Peter Gregory And Michael A. Simon, "Biometrics For Dummies", Cisa, Cissp, 2008.
[4] Anil K. Jain, Arun A. Ross, Karthik Nandakumar, "Introduction to Biometrics", Springer Science+Business Media, LLC 2011.
[5] D.Maltoni, D.Maio, A.K. Jain, S. Prabhakar. Handbook of Fingerprint Recognition. In: Springer, 2003.
[6] Amine Nait-Ali, Rgis Fournier. Signal and image processing for biometrics. In: ISTE, 2012.
[7] David. Zhang, Zhenhua. Guo, Yazhou. Gong. Multispectral Biometrics Systems and Applications. In: Springer Science+Business Media, 2014.
[8] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma. PCANet: A Simple Deep Learning Baseline for Image Classification?. In: IEEE Trans. Image Process. Publ. IEEE Signal Process.Soc, 2015.
[9] Cong Jie Ng and Andrew Beng Jin Teoh. DCTNet : A Simple Learning-free Approach for Face Recognition. In: Proceedings of APSIPA Annual Summit and Conference, pp. 26212631, 2015.
[10] R. Arandjelovic and A. Zisserman. All About VLAD. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1578-1585, 2013.
[11] Jolliffe I. Principal Component Analysis. In: Springer, New York, 2nd ed, 2002.
[12] Qiao Yu, Huang Xiyue, Chai Yi. Face recognition based on Weighted PCA. In: Journal of Chongqing University, Vol. 27, No. 3, 2004.
[13] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. In: ACM Transactions on Intelligent Systems and Technology, Vol. 2, pp. 1-27, 2011.
[14] T. Joshi, S. Dey, D. Samanta. Multimodal biometrics: state of the art in fusion techniques. In: International Journal of Biometrics, Inderscience publisher, Vol. 1, No.4, pp. 393-417, 2009.
[15] PolyU Finger KnucklePrint Database. Available at: http://www.comp.polyu.edu.hk/biometrics/FKP.htm

Copyright IPCO-2017
ISSN 2356-5608