Advanced Multimodal Fusion for Biometric Recognition System based on Performance Comparison of SVM and ANN Techniques

Mofdi Dhouib
ATMS research unit,
National school of engineering,
Sfax, 3038, Tunisia
Higher Institute of
Technological Studies
Department of
Information Technologies.
El Bousten BP 88A, Sfax, 3099,
Tunisia

Sabeur Masmoudi
ATMS research unit,
National school of engineering,
Sfax, 3038, Tunisia
Ahmed Ben Hamida
ATMS research unit,
National school of engineering,
Sfax, 3038,Tunisia

ABSTRACT
Multimodal fusion for biometrics recognition system had gained specific attention nowadays thanks to its remarkable valuable results. For this approach, **c**lassification methods have been the basis of important recognition accuracy improvements. The artificial neural networks (ANN) and support vector machines (SVM) belong to this class of methods. This paper presents comparison concerning the performances of the some methods that have been successfully applied to the fusion of scores for multimodal biometric recognition. After recognizing each single modality which was the fingerprint, the face as well as the voice, we recovered three similarity scores. These scores are then introduced into the classification system based on neural networks and on support vector machine techniques. Experimental results demonstrate that the identity established by such an integrated system is more reliable than the established identity by fingerprint recognition system, facial verification system and a voice verification system. Fusion phases are performed at score level. An average rate (= 57,69 %) is obtained by fusion with ANN. While fusion with the SVM gives an average rate equal to (= 63,31 %). A brief introduction is provided regarding the commonly used biometrics, including face, fingerprint and voice. Comparing Merger methods is made according to criteria of optimization of error rate.

General Terms
Security, Human Factors, Recognition, Verification, Identification, Image Processing

Keywords
Multimodal biometric system, Voice, Fingerprint, Face, Recognition, Score-level, Fusion, ANN, SVM..

1. INTRODUCTION
Nowadays, there is a strong demand for automatic and secure identity verification systems. Biometric identification is a new technology to solve this problem. The fingerprint, face and speech are among the most commonly used biometric features.

In the current environment, the computer security has become a research area of great importance. In particular devise a reliable, efficient and robust identification system is a priority task. The individual's identification has become essential to ensure the safety of systems and organizations. Faced with this increasing load, several methods for Biometric recognition have been proposed like: speaker recognition, facial recognition, fingerprint, iris recognition, retina, the hand shape.

Several works on multimodal biometric systems has already been done in the literature. Dieckmann et al. [1] proposed a summary level fusion scheme: "2-of-3 approach" that integrates the movement of the lips, face, and voice based on the principle that man uses, parallel, several indices identify a person, and Brunelli Falavian [2] proposed a system level measurement to combine the outputs of the sub-graders, Kitter et al. [3] demonstrated the effectiveness of an integration strategy that merges multiple snapshots of a biometric property initials using a Bayesian framework. Bigun et al. [4] proposed a Bayesian integration scheme of combining different evidence. Maes et al. [5] proposed to combine biometric data (e.g. fingerprint) with non biometric data (e.g. passwords). Hong and Jain [6], have developed a multimodal identification system that incorporates two different biometrics (fingerprints and face) that complement each other.

The unimodal biometric systems have been around for a few years but they are rather suitable for a medium security level. In fact, the higher the security level, the higher they tend towards the use of multimodal systems, more efficient and safer. In addition, systems based on a single biometric modality are vulnerable to attack. For the moment, no biometric indicator is 100% reliable according to [7]. This gave birth to the fusion of multimodal biometrics city before all the arguments over the results of various studies [8] [9] showed the performance of Multimodal Biometric systems over single-mode systems is a strong reason that we led to work on this topic (Add a term to a biometric system is to add a new source of information [8]).

The goal of this work is to provide a multimodal biometric system respecting several constraints comfort [10] and reliabilities (Increase rate recognition calculation inexpensive, robustness). In this context fusion allows address the lack of information resulting from the use of a single modality. This paper proposes an adaptive system of recognition of individuals by the merger of three biometric modalities: fingerprints, face and voice. Fusion was made using a hand machines support vector (SVM), and artificial neural networks (ANN) on the other hand. These classification methods have greatly enriched the biometric recognition methods. Finally the results are compared.

The remaining of this article is organized as follows: Section II describes the unimodal biometric systems. Section III presents the proposed multimodal system using respectively ANN and SVM. Section IV discusses the experimental results of these approaches. The performance of the proposed multimodal approach using ANN is analyzed and compared with respect to that of the proposed multimodal approach using SVM. The final section, section V presents the conclusion and discusses perspectives of this work.

2. UNIODAL RECOGNITION

2.1. Fingerprint Recognition

This method relies on the principle of extracting the minutiae; settings relevant characteristic footprint such as Bifurcation: the point where the ridge is divided into two (Figure 1-a) and Ridge ending: the point where the ridge is stopped (Figure 1-b).

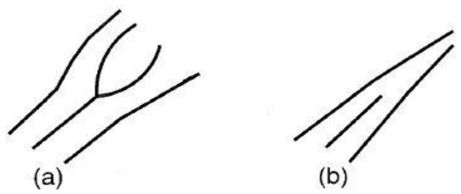

Figure 1: Fingerprint minutiae

The preprocessing phase is essential in a system for recognizing forms. To improve the quality of the information extracted from the images, one can specify regions of interest or enhance the contrast of images [5]. To avoid the extraction of false minutiae, several pretreatment steps have been performed like: Binarization, Skeletonization, (Thinning), Region of Interest, Minutiae extraction. The overall architecture of a fingerprint recognition system is described onfigure 2.

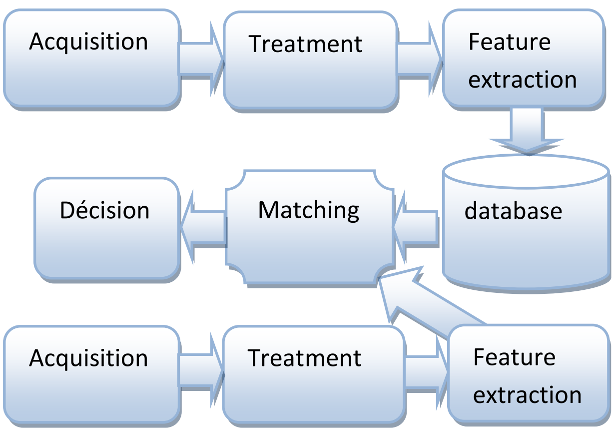

Figure 2: Principle of a fingerprint recognition system

2.2. Face recognition

Facial recognition is a task that humans naturally and effortlessly perform in their daily lives. It is one of the basic biometric technologies, took a share of more and more important in the field of research, this being due to rapid advances in technologies such as digital cameras, Internet and mobile devices, all associated with security needs constantly increasing.

Facial recognition has several advantages over other biometric technologies. It is an inexpensive used technique, very well accepted by the public and requires no action by the user (Non-intrusive and no contact).

The basic principle of operation of a facial recognition system is illustrated by (Figure 3). It can be summarized in four stages: detection [3] and standardization [4] of the face and the last two represent the recognition made by a subsequent extraction phase a comparison of the characteristics.

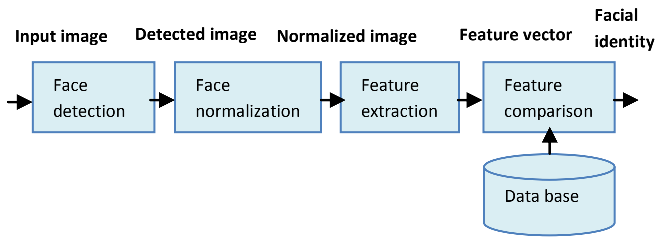

Figure 3: Principle of a facial recognition system

2.3.Voice recognition

This is a transformation of a speech signal into a sequence of symbols representative of the signal content. The most commonly used extracting algorithms are the Mel frequency cepstral coefficients (MFCC) that showed on the following figure 4.

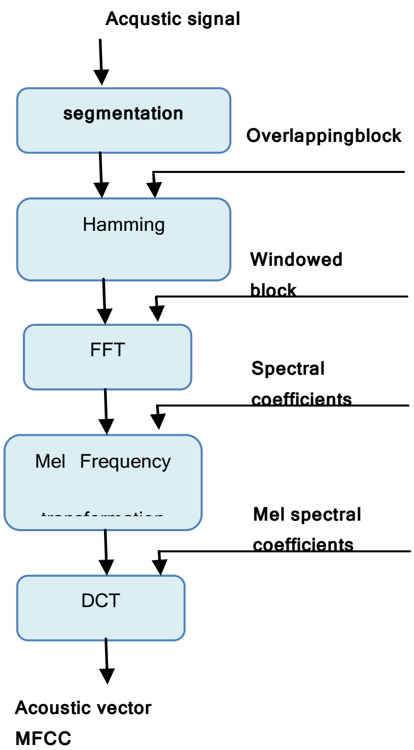

Figure 4: Principle of the extraction of MFCC coefficients

Although biometric recognition techniques can be very efficient, we cannot currently guarantee an excellent recognition rate with unimodal biometric systems based on a unique biometric signature.

Thus the error rate associated with unimodal biometric systems are relatively high, which makes them unacceptable for deployment of safety critical applications. To overcome these drawbacks, a solution proposed is the use of multiple biometric modalities in one system..

3. THE PROPOSED MULTIMODAL ARCHITECTURE

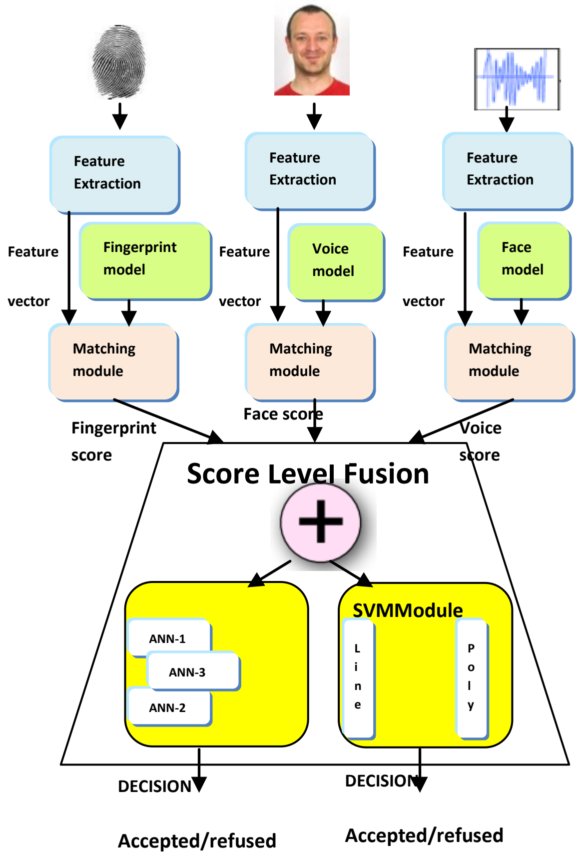

Figure 5: Fusion Score level for the multimodal biometric system

This proposed approach is to merge the output score of three different unimodal recognition systems use two types of classifiers. Then a performance comparison of the ANN merger with the SVM merger has been made.

3.1 Fusion with ANN

To achieve performance close to those observed in humans, the classifier based on artificial neural networks (ANN) have been used, associated with the fusion of the three modalities already cited. Indeed, using ANN for three separated biometrics, three different scores are recovered. These are supplied to a neural network composed of three classifiers to find the final score. Figure 5 shows the entire structure of the proposed system.

This proposed approach consists to fuse the output score of three different unimodal RNA. The fingerprint is combined with the face of a hand and the voice with the face of the other. Then the obtained output scores are combined to represent the input of a third network.

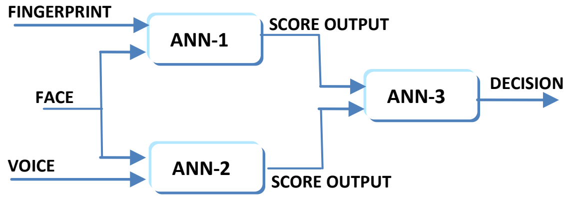

Figure 6: The proposed multimodal architecture based on ANN

3.2 Fusion with SVM

There are two approaches for combining the separately and individual matching score. The first approach is to formulate it as a classification access, while the other approach is to treat it as a combination access.

The idea of the classification approach is to construct a feature vector using the matching scores output by the separate matchers. After that, this feature vector is classified into one of two classes: “Accept” (genuine user) or “Reject” (impostor). In general, the classifier utilized for this aim is able of acquiring knowledge of the decision frontier without regard for how the feature vector is constructed.

3.2.1 Overview of Support Vector Machine (SVM)

In 1992, Boser, Guyon, and Vapnik introduced Support Vector Machine (SVM) which became rather popular since SVM are a set of related supervised learning methods used for classification and regression [23]. They appertain to a family of generalized linear classifiers.

Vapnik have developed the foundations of Support Vector Machines (SVM) [24] which have been gained popularity due to many promising features such as better empirical performance. The formulation utilizes the Structural Risk Minimization (SRM) principle, which has shown to be upper, to traditional Empirical Risk Minimization (ERM) principle, utilized by conventional neural networks. SRM minimizes a superior bound on the expected risk, where as ERM keep down the error on the training data.

In biometrics, Support Vector Machine has been utilized for different learning based operations such as face recognition and multimodal fusion.

SVM is therefore a classifier that executes classification by building hyper planes in a multidimensional space and separating the data points into different classes.

3.2.2 . Linearly separable data

Let {xi, yi} be a set of N data vectors with xi ∈ Rn, yi ∈ {+1, −1}, and i = 1, . . .,N. xi is the ith data vector that belongs to a binary class yi.

A binary classifier should find a function f that maps the points from their data space to their label space

f :Rn→ {+1,-1}

xi→yi

For the benefit of simplicity, we suppose that the data space is R2 and that a hyperplane separates the data. There are in fact an infinite number of hyperplanes that could divide the data into two classes. In accordance with the SRM principle, SVM utilizes an iterative training algorithm which maximizes the margin between two classes to construct just one optimal hyperplane.

Assuming that we have a hyper plane separating the positive data and negative data, xi belongs to the hyperplane which satisfies the relationship:

w .xi+b= 0 (1)

In this equation w is the normal to the hyperplane and it is also a vector, b is the parameter of the hyperplane.

For mathematical calculations we have,

w × xi + b = +1,yi= +1 (2)

w × xi + b = - 1, yi= - 1 (3)

These equations can be combined in the following inequality:

yi (w × xi + b) >= 1 (4)

The following figure shows the linearly separable case we have treated above:

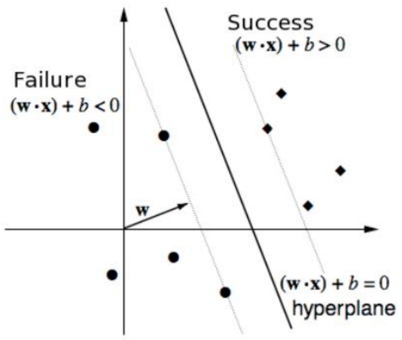

Figure7: Linear separation hyperplane for linearly separable data.

The points satisfying equality (2) belong to a hyper plane H1:

w × x + b = +1 (5)

Similarly, the checking point equal (3), belong to the hyper plane H2:

w × x + b = -1 (6)

The distance d (w, b; x) of a point x from the hyperplane (w, b) is,

$$d(w,b;x) = \frac{|(wXxi+b)|}{||w||} \quad (7)$$

Optimal hyper plane was constructed which the distance to the nearest points (margin) is Max. Maximize margin amounts to minimizing 2 / ||w||. For this, the problem is reformulated as Lagrangian. There are two reasons for this; the first is that the constraint (4) will be replaced by a constraint on the Lagrangian multipliers which will be easier to treat. In addition, in this reformulation of the problem, only data learning appear as a dot product. Thus, it introduces positive multipliers αi " i = 1... l in (4). Constraints in equation (4) are multiplied by αi and the equation becomes:

$$L\,(w,b,\alpha) = \frac{1}{2} \parallel w \parallel 2 - \sum_{i=1}^{l} \gamma i\ [yi\ (w \times xi + b) - 1], \alpha i >= 0 \quad (8)$$

L is called the Lagrangian primal.

It must minimize the Lagrangian with respect to w and simultaneously require its derivatives with compared to all of the Lagrangian multipliers αi disappears. By imposing that gradients of L with respect to w and b disappear and it obtained:

$$L' = \sum_{i=1}^{1} \alpha i - \frac{1}{2}\sum_{j=1}^{1} yiyj\sigma i\sigma jxixj \quad (9)$$

L' is called Lagrangian dual.

The points, which αi are strictly greater than 0, are called support vectors and they belong to one of the hyperplanes H1 or H2. These points are closest to the border decision and they form the separator plan.

B3. No linearly separable data

If no hyperplane can be found to separate the data, a nonlinear mapping function is then needed. To overcome the disadvantages of non-linearly separable case, the idea of SVM is to change the data space. The data will be mapped nonlinearly in a high-dimensional space and the optimal hyper plane is computed in the high-dimensional space. The nonlinear transformation of data can allow linear separation examples in a new space. So we will have a change in dimension. This new dimension is called "re-description of space." Indeed, intuitively, the more the size of the re-description space, the greater the probability to find a separating hyper plane between examples is high. This is illustrated by the following scheme:

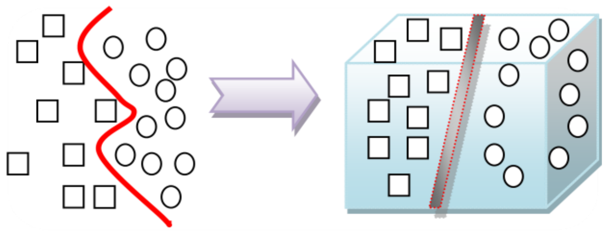

Figure 8 : Non linearly separable data.

Where examples are not linearly separable, the constraints (2) and (3) are released by introducing slack variables xi ≠0 " i = 1... l which become:

w × xi + b = +1 - xi, yi= +1 (10)

w × xi + b = - 1 + xi, yi= - 1 (11)

Therefore there is a transformation of a nonlinear problem of separation in the space of representation to a linear separation problem in an area of re-description of largest dimension. This non-linear transformation is performed using a specific kernel function.

4. EXPERIMENTAL RESULTS

The experimental results presented in this paper are divided into two parts. First the results obtained for each unimodal recognition system (fingerprint, face, and voice) are summarized. Secondly, the results of the proposed biometric multimodal fusion system used with three ANN classifiers are presented.

4.1. Experimental results for the proposed architecture with ANN

Table 2 summarizes the performance of the ANN fusion of the used biometric modalities.

Table 2: Performance of the modalities fusion

		Number of epochs		
Fusion of modalities	HN	1000	5000	10000
Voice/Face	5	18.39 %	20.11 %	22.14 %
	10	21.55 %	28.7 %	31.55 %
	50	34.60 %	40.37 %	43.75 %
	100	44.9 %	48.03 %	56.40 %
Fingerprint /Face	5	15.03 %	21.69 %	27.15 %
	10	23.90 %	28.00%	34.29 %
	50	35.85 %	37.12 %	43.65 %
	100	42.01 %	45.2 %	54 %
Fingerprint /Face / Voice	5	11.00 %	18.50 %	27.87 %
	10	22.30 %	28.32%	35.61 %
	50	35.85 %	41.64 %	48.30 %
	100	43.97 %	48.2 %	57.69 %

From the Table 2, we can notice that the recognition rate is improved by the third classifier. The recognition rate is not enough to evaluate the performance of a biometric system. So the following table summarizes values of other performance criteria.

Table 3: Performance evaluation of fusion system using three ANN

	FRR	FAR	EER
Fusion by three ANN	1,54 %	4,589 %	4,149

With: FRR is False Rejection Rate; FAR is Acceptance Refuse Rate and EER is Equal Error Rate.

4.2. Experimental results for the proposed architecture with SVM

Table 4: Performance evaluation of fusion system using SVM

	FRR	FAR	EER
Fusion by linear kernel	1,63 %	4,71 %	2,8158%
Fusion by polynomial kernel	1,48 %	4,52 %	2,3467%

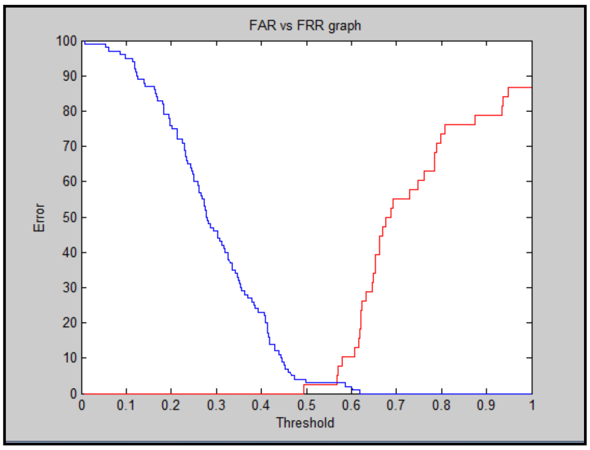

Figure 9: FAR and FRR based on threshold (linear method).

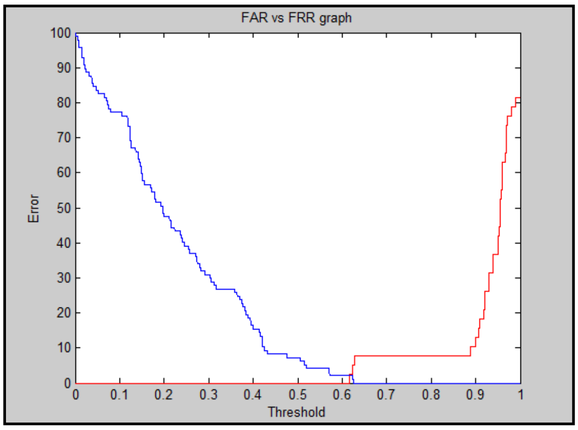

Figure 10: FAR and FRR based on threshold (polynomial method).

Figure 11 shows the ROC curves and EER of the following biometric system: only face verification, only fingerprint verification, only voice verification and the proposed multimodal verification.

Multimodal biometric authentication based on score level fusion using SVM 5

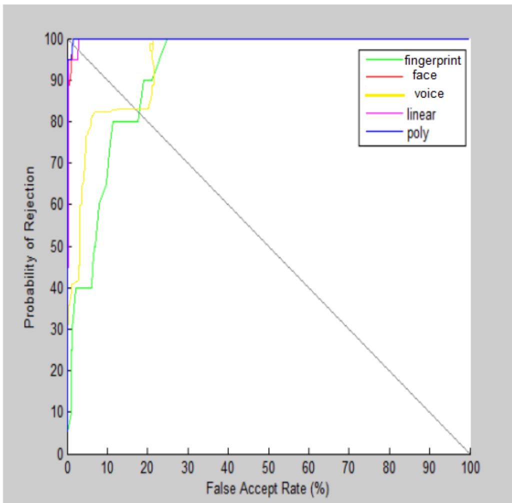

Figure 11: ROC curves of unimodal method and the proposed method

Place Tables/Figures/Images in text as close to the reference as possible (see Figure 1). It may extend across both columns to a maximum width of 17.78 cm (7").

Captions should be Times New Roman 9-point bold. They should be numbered (e.g., "Table 1" or "Figure 2"), please note that the word for Table and Figure are spelled out. Figure's captions should be centered beneath the image or picture, and Table captions should be centered above the table body.

5. CONCLUSION

In this paper, the concepts of recognition unimodal and multimodal biometrics are introduced. The principle is to design unimodal recognition systems and combine their scores from different biometric modalities to increase the power of identification

The errors come from the imperfection of one biometric have been remedied by the fusion process by ensuring better recognition rate.

In addition, the concept of classification by neural network and support vector machines for multimodal fusion are detailed. Among the various levels of existing fusion, the score level is chosen because it offers the best compromise between the wealth of information and the ease of implementation.

This work provides new contribution to the field of biometrics multimodal. In fact, it shows the authentication of individuals by multimodal fusion based on ANN and SVM using the fingerprint, face and speech recognition.

The experimental results showed a significant improvement of SVMs compared to ANNs. This is due to what they can suffer multiple local minima, the solution to an SVM is global and unique. Two other advantages of SVMs are that it has a simple geometric interpretation and give a sparse solution.

From experiment results we obtain the following conclusions:

- The accuracy of verification is more improved than simply unimodal biometrics using the fusion of three biometric modalities.
- By comparing the results of SVM using a linear kernel with those using a non-linear kernel, we note an advantage of non-linear kernels. This is due that convexity is an interesting and important property of nonlinear SVM classifiers
- Unlike SVMs computational complexity ANNs is proportional to the dimensionality of the input space. ANNs empirical use of risk minimization, while SVMs using structural risk minimization. Why SVMs outperform ANNs often in practice is that they deal with the biggest problem with ANNs, SVMs are less prone to overfitting.
- This method has the superiority over the previous methods due to the application of the new recognition algorithms and the SVM-based fusion rule.

Future work will investigate on better alternative recognition technique suitable for fusion of fingerprint, speech and face. The performance of multi-biometric systems can be improved if a suitable fusion strategy is used in particular for the system running an uncontrolled environment. Therefore, it would be interesting to apply other approach of fusion and to compare its results with those obtained by the ANN and SVM to maximize the performance of multi-biometric system.

6. REFERENCES

[1]. U. Dieckmann, P. Plankensteiner, and T. Wagner. Sesam: A biometric person identification system using sensor fusion. Pattern Recognition Letters, 18(9):827-833, 1997.

[2]. R. Brunelli and T. Poggio. Face recognition: Features versus templates. IEEE Trans. Pattern Analysis and Machine Intelligence, 15(10): 1042-1052,1993.

[3]. J. Kittler, Y. Li, J. Matas, and M. U. Sanchez. Combining evidence in multimodal personal identity recognition systems. In Proc. 1st Int. Conf. on Audio Video-based Personal Authentification, pages 327-334, Crans-Montana, Switzerland, March 1997.

[4]. E.S. Bigun, J. Bigun, Duc, and S. Fischer. Expert conciliation for multi modal person authentication systems by Baysian statistics. In Prooc. 1st Int. Conf. on audio Video-Based Personal Authentication, pages 327-334, Crans-Montana,Switzerland, March 1997

[5]. S. Maes and Beigi. Open sesame! Speech password or key to secure your door? In Proc. 3rd AsionConference on Computer Vision, pages 531-541, Hong Kong, China, 1998.

[6]. L. Hong and A. Jain. Integrating faces and fingerprints for personal identification. In Proc. 3rd AsionConference on Computer Vision, pages 16-23, Hong Kong, China, 1998.

[7]. Arif, fusion de données applications à l'Identification et à l'Authentification thèse de doctorat, Université François Rabelais Tours 2005.

[8]. Poinsot, Yang, fusion de biométries sans contact paume et visage Article, Université de Bourgogne, 2008.

[9]. Hammoud, Abidi,face biometrics for personal identification multimodal systems.

[10]. Doublet, Revenu, Olivier, reconnaissance biométrique sans contact de la main intégrant des informations de forme et de texture, France Telecom.2003

[11]. A. Ross and A.K. Jain, "Information fusion in biometrics", Pattern Recogn. Lett. 24, 2115–2125 (2003)

[12]. Y. Wang, T. Tan, Y. Wang, and D. Zhang, "Combining face and iris biometric for identity verification", Proc. 4th Int. Conf. on Audio− and Video−Based Biometric Person Authentication (AVBPA) 1, 805–813 (2003)

[13]. Chaohong, advanced feature algorithms for automatic fingerprint recognition system , university of new yorkatbuffalo.2007

[14]. Thibaud, Réseaux de neurones en cascade pour la localisation précise de points caractéristiques du visage Article Université Pierre et Marie Curie. 2008

[15]. Lorène,la biométrie multimodale: stratégies de fusion de scores et mesures de dépendance virtuelle. thèse de doctorat de l'institut des Télécom paris .2009

[16]. Morizet: reconnaissance biométrique par fusion multimodale du visage et de l'iris thèse de doctorat Télécom 2009.

[17]. Chaudhary, multimodal recognition system based on

fusion of Palm print, fingerprint and face, InternationalConference IEEE .2009

[18]. Chollet, vérification biométrique multimodal: Le projet incitatif GET- BIOMET Telecom paris.2010

[19]. D. Kisku, P. Gupta and J. Sing, “Multibiometrics Feature Level Fusion by Graph Clustering”, International Journal of Security and Its Applications Vol. 5 No. 2, April, 2011.

[20]. M. Kazi and Y. Rode, “multimodal biometric system using face and signature: a score level fusion approach”, Advances in Computational Research, Vol. 4, No. 1, 2012.

[21]. TejalChauhan, HemantSoni, SameenaZafar, “A Review of Automatic Speaker Recognition System”, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-4 September 2013.

[22]. SheetalChaudhary, RajenderNath , A New Multimodal Biometric Recognition System Integrating Iris, Face and Voice, International Journal of Advanced Research in Computer Science and Software Engineering, 2015.

[23]. C. Chen and C. Chu, “Fusion of face and iris features for multimodal biometrics”, Lect. Notes Comput. Sc. 3832, 571–580 (2006)

[24]. K. Nandakumar. MultibiometricSystems : Fusion Strategies And Template Security. PhD thesis, 2008

[25]. A.K. Jain, A. Ross, and S. Prabhakar, “An introduction to biometric recognition”, IEEE T. Circ. Syst. Vid. 14, 4–20 (2004)

[26]. Youssef Elmir, ZakariaElberrichi and RédaAdjoudj, A Hierarchical Fusion Strategy based Multimodal Biometric System, The International Arab Conference on Information Technology (ACIT’2013)

[27]. E.A. Zanaty, Support Vector Machines (SVMs) versusMultilayer Perception (MLP) in data classification, Egyptian Informatics Journal (2012) 13, 177–183

[28]. Yi-Li Tseng,Keng-Sheng Lin and Fu-Shan Jaw2, Comparison of Support-Vector Machine and Sparse Representation Using a Modified Rule-Based Method for Automated Myocardial Ischemia Detection, Hindawi Publishing Corporation Computational and Mathematical Methods in Medicine Volume 2016, Article ID 9460375, 8 page

[29]. RupinderSaini , Narinder Rana, COMPARISON OF VARIOUS BIOMETRIC METHODS, International Journal of Advances in Science and Technology (IJAST) Vol 2 Issue I (March 2014)

[30]. SheetalChaudhary, RajenderNath, A New Multimodal Biometric Recognition System Integrating Iris, Face and Voice, International Journal of Advanced Research in Computer Science and Software Engineering, Volume 5, Issue 4, April 2015

[31]. Asim Ali Khan1, Smriti Kumar, Mister Khan, Iris Pattern Recognition using Support Vector Machines and Artificial Neural Networks, International Journalof Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering Vol. 2, Issue 12, December 2014