Multimodal Biometrics System's Resistance to noise

Fingerprint and Voice

Youssef ELMIR
Dept. of computer science
University of Adrar
Adrar, Algeria
y.elmir@univ-adrar.dz

Omelkheir GHAZAOUI, Farida BOUKENNI
Dept. of computer science
University of Adrar
Adrar, Algeria

***Abstract*— This paper should give an idea about the resistance of multi biometrics to different level of noise which can affect the data obtained from different sensors. In comparing with the ability of monomodal biometrics to resist against noise using fingerprints and voice recognition and the fusion of both of them, we could find as results that multimodal biometrics systems resist much better than separated monomodal biometrics systems.**

***Keywords*— Multibiometrics, SVM, Gabor filters, MFCC, Fingerprints, voice.**

I. INTRODUCTION

Most biometric systems use a single biometric to find identity of persons. Some of the challenges commonly encountered by biometric systems are listed below.

1. Noise in sensed data,
2. Non-universality,
3. Upper bound on identification accuracy,
4. Spoof attacks,

In this work, we tried to propose a solution to the problem of noise in sensed data by using fusion of Multi-biometrics in one system. In our case, we fused fingerprint and voice data and we used SVM to train the system and to evaluate its performance.

In the following sections, we present in detail our proposed algorithm of recognition using SVM. The principle of multi-biometric identification is presented in Section II. Sections III and IV present the problem statement and our proposed design for it. The algorithm used to codify fingerprint images and voice records using bank of Gabor filters and MFCCs are given respectively in section V and VI, the experimental results obtained on the FingerCell and EPA databases, are presented in Section VII. Finally, conclusion is given in Section VIII.

II. PRINCIPLE OF MULTIBIOMETRICS

A multi-biometric system relies on the evidence presented by multiple sources of biometric information.

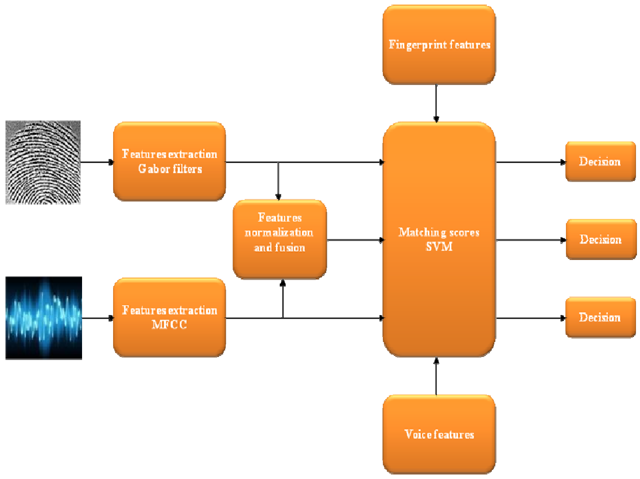

Figure 1. Scheme of our multi-modal biometric system.

Based on the nature of these sources, a multi-biometric system can be classified into one of the following six categories:

- multi-sensor,
- multi-algorithm,
- multi-instance,
- multi-sample,
- Multimodal and hybrid.

III. PROPOSED DESIGN

The problem of multi-biometric recognition is one of much broader topics in scientific and engineering. The goal of multi-biometric recognition is to improve person recognition of using a number of their biometrics. This biometrics are in our case are images of fingerprints matrix called vectors codes or fingercodes and records of voices that are extracted from an input data using the techniques described in the later section. The classes here refer to individual person. Since the classification process in our case is applied on extracted features, it can be also referred to as feature matching.

This article demonstrates how multi-biometric identification can resist against noise, but before the original data has to be converted into a vector code, by using Gabor filter bank for fingerprints and MFCCs for voices.

IV. PROBLEM STATEMENT

An SVM is to be designed and trained to recognize the codes of the databases that are actually used. An imaging system that converts each minutiae image obtained from a fingerprint image in fingercode or minutiae matrix code by using a bank of Gabor filters. An MFCC algorithm is to be used to convert data obtained from records into vector codes. The result is that each fingerprint image is represented as a vector of 256 real values and each voice is represented as a vector of 160 real values, and to get the final code we need concatenate both of them into one code which give a code of 416 elements. Finally, results obtained by the fusion of fingerprint and voice features will be compared to those obtained by fingerprint and voice recognition systems.

V. CREATING FINGERCODE

The following steps are observed to create the fingercode:

1. Preprocessing of the image (to remove noise) by window wise normalization, Histogram Equalization, low pass and median filtering [1].
2. Core point location using max concavity estimation [2].
3. Tessellation of circular region around the reference point.
4. Sector wise normalization followed by application of bank of Gabor filters which has following general form in the spatial domain [3].

$$G(x, y; f, \theta) = \exp\left\{ \frac{-1}{2} \left[\frac{x'^2}{\delta_{x'}^2} + \frac{y'^2}{\delta_{y'}^2} \right] \right\} \cos(2\pi f x'). \tag{1}$$

$$x' = x \sin\theta + y \cos\theta. \tag{2}$$

$$y' = x \sin\theta - y \cos\theta. \tag{3}$$

Where f is the frequency of the sine plane wave along the direction θ (0, 45, 90 and 135 degrees) from the x-axis, $\delta_{x'}$ and $\delta_{y'}$ are the space constants of the Gaussian envelope along X' and Y' axes, respectively.

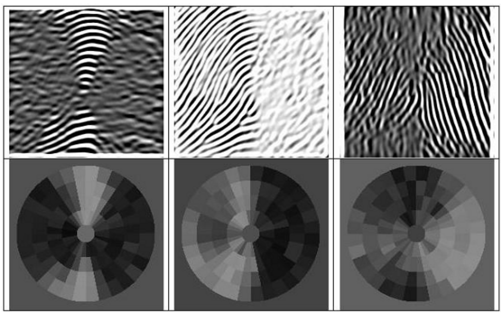

Figure 2. Filtered images and their corresponding feature vectors for the orientations 0°, 5°. 22.5° and 45° are shown [11].

Finally feature code generation by obtaining standard deviation values of all the sectors [2].

VI. CREATING VOICEPRINT

The signal of a voice is first processed by software that converts the speech waveform to some type of a parametric representation (at a considerably lower information rate) for further analysis and processing. The speech signal is a slowly timed varying signal (it is called quasi-stationary). An example of speech signal is shown in Figure 2. When examined over a sufficiently short period of time (between 5 and 100 ms), its characteristics are fairly stationary. However, over long periods of time (on the order of 1/5 seconds or more) the signal characteristic change to reflect the different speech sounds being spoken. Therefore, short-time spectral analysis is the most common way to characterize the speech signal. A wide range of possibilities exist for parametrically representing the speech signal for the speaker recognition task, such as Linear Prediction Coding (LPC), Gaussian mixture models (GMM) [4], Mel-Frequency Cepstrum Coefficients (MFCC), and others.

MFCC is perhaps the best known and most popular, and these will be used in this paper. MFCC's are based on the known variation of the human ear's critical bandwidths with frequency, filters spaced linearly at low frequencies and logarithmically at high frequencies have been used to capture the phonetically important characteristics of speech(to obtain voiceprint or voice signal matrix code). This is expressed in the mel-frequency scale, which is linear frequency spacing below 1000 Hz and a logarithmic spacing above 1000 Hz. The process of computing MFCCs is described in more detail in [5, 6]. A block diagram of the structure of an MFCC processor is given in Figure 1. The speech input is typically recorded at a sampling rate above 10000 Hz. This sampling frequency was chosen to minimize the effects of aliasing in the analog-to-digital conversion. These sampled signals can capture all frequencies up to 5 kHz, which cover most energy of sounds that are generated by humans. As been discussed previously, the main purpose of the MFCC processor is to mimic the behavior of the human ears. In addition, rather than the speech waveforms themselves, MFCC's are shown to be less susceptible to mentioned variations. The voice signal matrix is immediately encrypted to eliminate the possibility of identity theft and to maximize security. For example, here is the voice signal from database [6] and the voiceprint matrix of this voice signal (see Fig. 3 and Fig. 4):

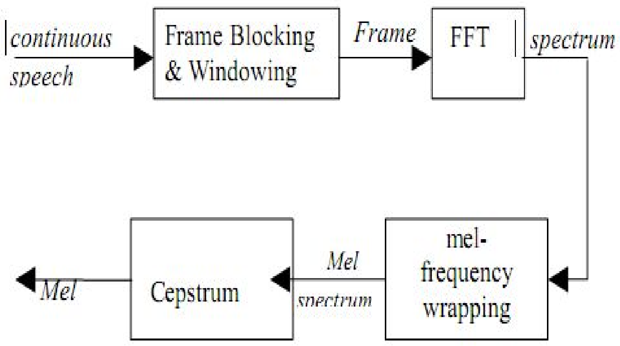

Figure 3. Block diagram of the MFCC processor

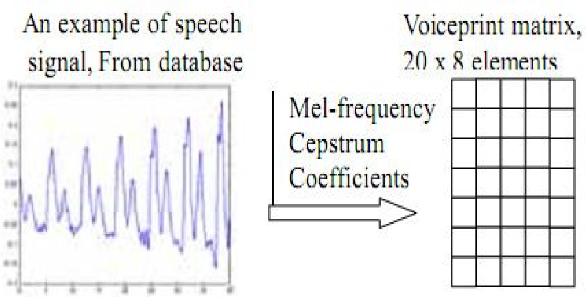

Figure 4. Voiceprint

Perfect classification of N ideal input voiceprint matrix of voice signal is required and reasonably accurate classification of speech waveform (N is equivalent to a number of distinguishing class of speaker in each database). The N 160-element input voiceprint matrix of voice signals are defined as a matrix of input matrixes (voiceprint matrix size ~ 20 x 8). The target vectors are also defined with a variable called target. Each target vector is a N-element vector with a 1 in the position of the voiceprint it represents, and 0's everywhere else. For example, the voiceprint number one is to be represented by a 1 in the first element (as this example is the first voiceprint of the database), and 0's in elements two through N.

VII. RESULTS

As it is shown in Fig. 5, the multi-biometric system produces the lowest error rate (30%) in recognition in comparing with monomodal biometric systems. In addition, it proofs more efficiency in resistance against applied noise till noise level equal to 1.5 and higher when the monomodal systems start producing a considered error rates (50%, 60%) from the firsts levels of noise.

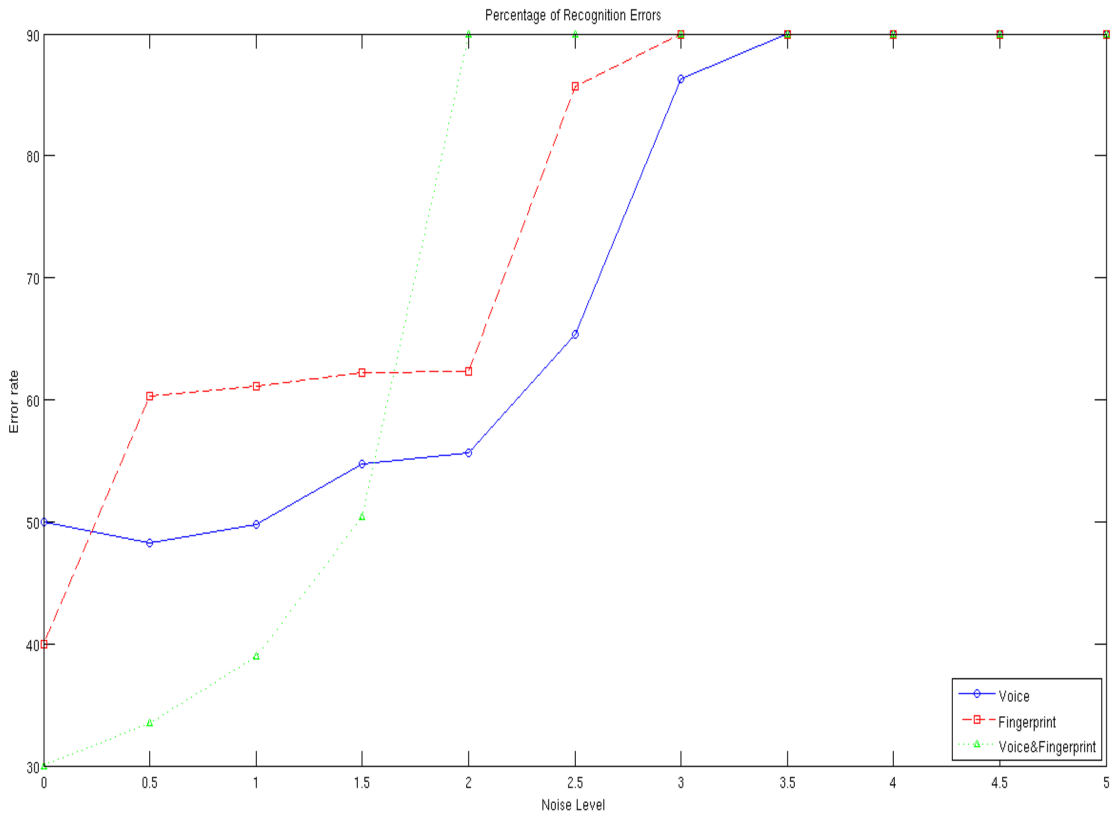

Figure 5. Resistance of the studied systems against different levels of applied noise.

VIII. CONCLUSION

This work presents a biometric person identification system based on fusing two common biometric traits: fingerprint and voice. The fusion is carried out by a simple scheme that combines the independent features from both fingerprint and voice traits. The performance of the recognition system is assessed under deferent levels of noise. These noise levels might affect the samples used to build the classifiers, and/or the test samples the system must identify. It is shown that the proposed fusion framework provides a viable identification system under deferent contamination conditions, even when the independent classifiers have low single performance. As future work, we will investigate more robust feature extraction tools that provide better results under this scheme.

REFERENCES

[1] S. Greenberg, M. Aladjem, D. Kogan and I. Dimitrov, "Fingerprint Image Enhancement using Filtering Techniques", Proc of the 15th International Conference on Pattern Recognition (ICPR'00), Barcelona, Spain, September 2000.

[2] A. K. Jain, D Maltoni, D Maio and S Prabhakar, "Handbook of Fingerprint Recognition", Springer Professional Computing, 2003.

[3] A K. Jain, S Prabhakar, and L. Hong, "A Multichannel Approach to Fingerprint Classification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 21, no. 4, pp. 348–359, April 1999.

[4] Y. Mami, "Reconnaissance de locuteurs par localisation dans espace de locuteurs de références", PhD thesis at « Ecole nationale supérieure des télécommunications », Paris, France, October 2003.

[5] M. N. Do, "An Automatic Speaker Recognition System", Audio Visual Communications Laboratory, Federal Institute of Technology, Lausanne, Switzerland, February, 2003.

[6] C. Cornaz, U. Hunkeler and V. Velisavljevic, "An automatic speaker recognition system", Digital Signal Processing Laboratory, Federal Institute of Technology, Lausanne, Switzerland, February, 2003.