Biometric Personal Identification System using the ECG Signal

Emna Rabhi[1], Zied Lachiri[1,2]

[1]Ecole Nationale d'ingénieurs de Tunis, Tunis, Tunisia
Université Tunis El Manar, Tunis, Tunisia
[2]Département Génie physique et instrumentation,
Institut National des sciences Appliqués de Tunis, Tunisia

Abstract

The electrocardiogram has unique cardiac features to each individual, which motivated us to use it as a biometric, hence, its robust nature against falsification makes it rather reliable for security systems, as it offers ultimate security in all situations. This paper presents a new approach applying this ECG particularity. A robust ECG Biometrics based on the features extraction with fiducial detection in the time domain is proposed.

After preprocessing, ten morphological descriptors are extracted from each heartbeat and which were divided into homogenous groups (amplitude, surface, interval and slope). Later, sixty Hermite Polynomials Expansion (HPE) coefficients are extracted from the ECG signal. Finally, classification is based on the Hidden Markov Model (HMM) with the (HTK) recognition toolkit using a Bakis model with one Gaussian. A particular strategy is adopted for personal identification: the groups of morphological parameters were used separately in the classification then were made associations between these groups them in one input vector. On the other hand, the Hermite Polynomials Expansion coefficients were classified apart.

In order to improve performance, a combination between 10 morphological parameters and 60 HPE coefficients was performed in a one system. Our algorithm is tested on 18 healthy signals of the MIT_BIH database. Each recording lasts about 30 minutes, 20 min of the data were used for training and the last 10 min were used for testing. The analysis of different groups separately showed that the best recognition performance is 96.7% for the Hermite Polynomials Expansion coefficients and the results of experiments showed that the proposed hybrid approach has led to an overall maximum of 99%.

1. Introduction

The traditional strategies of human identity such as passwords, secret codes and personal identification numbers (PIN) can be easily shared, observed, stolen or forgotten.

Recognition of the person biometric refers to the process of automatically recognizing a person using behaviors (approach, signature, keystroke, lip movement, handle) or physiological traits (face, voice, iris, fingerprint, hand geometry, electroencephalogram EEG, ECG electrocardiogram, ear shape, body odour) [5]. The idea of using medical attributes of the human body for identification system is relatively recent [4,6].

Biel et al [1] conducted the biometric experiment on 20 persons using Siemens ECG equipment to record the characteristics of the medical diagnosis. A multivariate analysis was performed by the SIMCA (Soft Independent Modeling of Class Analogy) model for the classification using principal components analysis (PCA).

Ming Li [8,9] used a method based on both temporal and cepstral features. He used only the healthy individuals. The temporal features were classified by a Support Vector Machines and the cepstral features were classified by a GMM. To improve the performance of recognition, he made a weighted sum fusion and he reaches a rate equal 98.3%.

The main problems of the previous works are the heart physiological changes and the choice of the features extraction [10].

In this paper, we used temporal parameters which are the morphological descriptors and the Hermite Polynomials coefficients. In order to improve the identification rate, we combined all of the parameters in a one input vector. For the classification, we used the Markov model (HMM) under HTK platform.

The remainder of this paper is organised as follow: Section2 describes the proposed identification system. The extracted features are given in section 3. Section 4 introduces the adopted classifier identification .Section 5 presents the experimental results showing the effectiveness of the proposed system. A conclusion is presented in section 6.

2. Description of the proposed identification system

The proposed system contains three steps as shown in Figure 1 and which are preprocessing, features extraction and classification.

Improve the overall performance of recognition; we tested our system with an input vector formed by the combination of all the morphological parameters and the coefficients of Hermite polynomials. The parameter vector contains the 10 morphological parameters and the 60 coefficients of each beat. We adopted this association in the topology of Hidden Markov model toolkit HTK.
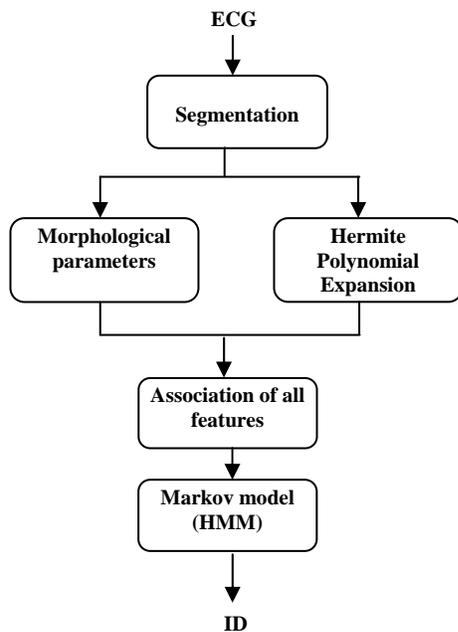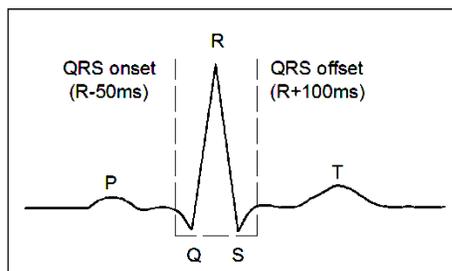
Preprocessing of the ECG signal is made by sampling the QRS complex and determining the extremum of the complex which is the wave (R) [3]. For this, we used the list of reference signals given by the MIT_BIH database. Each wave R determines the beginning of the QRS complex (Q = R - 50 ms) and end (S + R = 100 ms) as shown in Figure. 2.

The morphological descriptors such as amplitude, surface interval illustrated in Figure 3 and others have been widely used by several researchers for the classification of heart beats and which gave good results [2]. So we thought to use these parameters for eventual classification with HMM. Only ten descriptors extracted from channel 1 of ECG recording were used and which are described in table1.
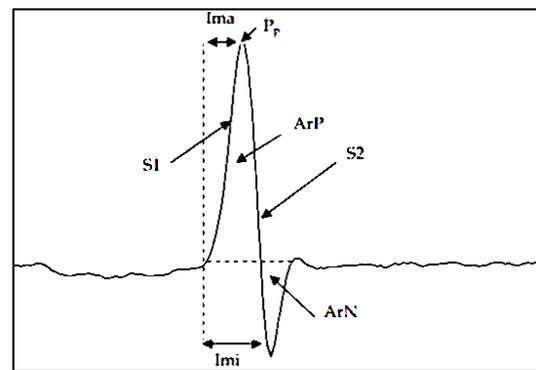

Figure 1. Block diagram of the algorithm adopted.

3. Recognition parameters

3.1. Morphological descriptors

Figure 2. Description of QS interval.

Figure 3. The morphological descriptors.

Table 1. Description of morphological descriptors.

Type	Description
P_p	Maximal amplitude of the positive peak
P_n	Maximal amplitude of the negative peak
ArP	Area of the positive samples
ArN	Area of the negative samples
Ar	Number of samples with 70 higher amplitude
No	Number of samples with 70 higher amplitude
Ima	Time-interval from the QRS complex onset to the maximal positive peak
Imi	Time-interval from the QRS complex onset to the maximal negative peak
S1	QRS slope velocity calculated for the time-interval between QRS complex onset and the first peak
S2	QRS slope velocity calculated for the time-interval between the first peak and the second peak

3.2. Hermite polynomial expansion (HPE)

The application of Hermite model is very suitable for the ECG signal modelisation due to the similarity existing between their forms [7]. Each complex QRS was normalized and segmented using standard methods of ECG toolbox [4].

The Hermite polynomials are represented by:

$$a(t) = \sum_{n=0}^{L-1} c_n \phi_n(t, \delta) \qquad t \in [-M, M] \qquad (1)$$

with a(t) the vector curve modelling each heart beat, L the polynomial order, $c_n (n = 0, ..., L-1)$ the coefficients of Hermite polynomials and $\phi_n(t, \delta)$ Hermite functions which are defined as follows:

$$\phi_n(t, \delta) = \frac{1}{\sqrt{\delta 2^n n! \sqrt{\pi}}} e^{\frac{-t^2}{2\delta^2}} H(t/\delta) \qquad (2)$$

The physicists' Hermite polynomials represent the Hermite functions which are defined recursively by:

$$H_n(t) = 2t H_{n-1}(t) - 2(n-1) H_{n-2}(t) \qquad (3)$$

$$H_n(t) = 2t H_{n-1}(t) - 2(n-1) H_{n-2}(t) \qquad (4)$$

The Hermite Polynomials Expansion coefficients are determined from ECG signal extract from a window of width equal 200 ms centered on the position of the R-wave.

4. The classifier identification

4.1 Hidden Markov Models (HMM)

An HMM is a probabilistic automaton of finite state. It consists of states interconnected by transitions. A transition between a state s_i and a state s_j makes possible the transition between these two states [12].

4.2 Model topology

We adopted the model of order 1 called Bakis. It is a model left to right of 5 states; the initial and final vectors do not emit observations. The Markov process is continuous, each state has a Gaussian. The adopted model is defined by an average vector initialized to zero and a variance vector initialized to one. The transition matrix is of diagonal form. HTK reads to topology of HMM model and builds a new model of averages and variances [13].

Recognition of a set of data requires the training of parameters associated with their respective HMM from a training corpus. The estimation of model parameters is made by applying the Baum-Welch optimal algorithm until convergence and to estimate probabilities of emission and transition.

Identification is achieved by the Viterbi algorithm; it is to find the sequence of states most likely corresponding to the parameters observed in a composite model. The recognition result is compared with the reference labels by dynamic alignment to count the labels identified, omitted, substituted or inserted by another label and finally calculates the recognition rate to evaluate the quality of recognition and quantify the differences with the ideal result expected on the labeling of the reference test database.

5. Experimental results

The assessment of performance is achieved using the Physiological database MIT / BIH [11]. Only 18 ECG signals from different healthy persons were used for evaluations.

We used for human identification with Electrocardiograms a HMM. Hence, we applied the approach two-thirds one-third for both classifiers. The training stage lasts 20m that is about 2/3 of the data and the test stage lasts 10m that is about the last 1/3.

First for the classification of morphological parameters we changed the topology of the HMM that is to say, we varied each time the number of Gaussian(1 to 5) and number of states(5 to 7), the best rates are obtained for HMM model to 5 states and 1 Gaussian, within this framework, the parameters are divided according to their morphological characteristics into four groups namely: Amplitude (Pp and Pn), Surface (Ar, ArP, ArN), interval (No, Ima, Imi) and slope (S1, S2). to try to better the contribution of these descriptors, we vary each time the groups of parameters used in the classification and we also make associations between groups of parameters.

We started our classification by surface group (ArN + Ar + ArP) which gave the lowest identification rate 80.3%, then we see that there have been improvements every time when changing group parameters and especially for the group of amplitude($P_p + P_n$) which reaches a rate of 95.08%. Then we adopted the approach of association between different groups of morphological parameters and the worst rate of association for the group Surface-Interval (ArN+Ar+ArP) and (No+Ima+Imi) is 80.39% and the highest rate is for the group Amplitude-Slope ($P_p + P_n$) and (S1+S2) which reaches 95.14%, so we can conclude that the more we associate parameters more we will obtain a best rates of recognition, why we have put the 10 morphological parameters in a same system and we got a very satisfactory rate equal to 96.03%.

Implementation of Hermite model is very suitable for

estimation and ECG signal modeling. For that purpose, we chose calculated coefficients of these polynomials, we chose a polynomial order equal to 60, M = 100et δ = 10 and we achieved a classification rate equal to 96.79% using the same model topology of HMM used for classification of morphological parameters. So we see that the coefficients of the Hermite polynomials are given for the moment the best rate because due to the similarity existing between the forms and functions of Hermite waveforms of the ECG.

Finally, to improve overall performance, we combined morphological parameters with the coefficients of the Hermite polynomials which gave a rate equal to 99.02%.

Table 2. Global rate with different features.

Parameters group	Global rate (%)
$P_p + P_n$	95.08
ArN+Ar+ArP	80.38
No+Ima+Imi	93.14
S1+S2	88.41
($P_p + P_n$) and (ArN+Ar+ArP)	87.49
(ArN+Ar+ArP)and (No+Ima+Imi)	80.39
(ArN+Ar+ArP)and (S1+S2)	92.48
($P_p + P_n$)and(No+Ima+Imi)	90.25
(S1+S2) and (No+Ima+Imi)	90.29
($P_p + P_n$) and(S1+S2)	95.14
($P_p + P_n$),(ArN+Ar+ArP),(No+Ima+Imi) and (S1+S2)	96.03
$C_n (n = 1...60)$	96.79
($P_p + P_n$),(ArN+Ar+ArP)(No+Ima+Imi), (S1+S2) and C_n	99.02

6.   Conclusion

The paper presented a new method to assign the ECG signal for identifying individuals. The classification of persons was carried out for 18 healthy signals of the MIT_BIH database, using two features extraction techniques such as morphological descriptors and HPE coefficients and a HMM with Bakis model for classification. We have tested different combination of parameters groups. The best result is achieved by associating the set of all groups. For further work, we can include more number of ECG signals having various pathologies and improve the method by using multimodality technique while merging ECG signals with other biometric parameters to improve the rate of individual identification.

References

[1] Biel L, Pettersson O, Philipson L, Wide P. ECG analysis: a new approach in human identification. IEEE Transactions on Instrumentation and Measurement 2001: 808–812.

[2] Christov I, Herrero GG, Krasteva V, Jekova I, Gotchev A, Egiazarian K. Comparative study of morphological and time-frequency ECG descriptors for heartbeat classification. Medical engineering & physics 2006; 28:876-887.

[3] De Chazal P, O'Dwyer M, Reilly RB. Automatic classification of heartbeats using ECG morphology and heartbeat interval features. IEEE Transactions on Biomedical Engineering 2004; 51: 1196-1206.

[4] Irvine JM, Israel SA, Scruggs WT, Worek WJ. EigenPulse: robust human identification from cardiovascular function. Pattern Recognition 2008:3427-343.

[5] Jain A, Ross A, Prabhakar S. An introduction to biometric systems. IEEE Transactions on Circuit and Systems for Video Technology 2004; 14:4–20.

[6] Kyoso M. A technique for avoiding false acceptance in ECG identification. Proc IEEE EMBS Asian-Pacific Conf on Biomedical Engineering 2003:190-191.

[7] Lagerholm M, Eterson C, Braccini G, Edenbradt, Sornmo I. Clusturing ECG complexes using Hermite function and self-organizing maps. IEEE Trans. Biomed. Eng 2000; 47: 838-847.

[8] Li M, Narayanan S. Robust ECG biometrics by fusing temporal and cepstral information pattern recognition (ICPR) 2010: 1326-1329.

[9] Li M, Rozgic V, Thatte G, Lee S, Emken A, Annavaram M, Mitra U, Spruijt-Metz D, Narayanan S. Multimodal physical activity recognition by fusing temporal and cepstral information. IEEE Transactions on Neural Systems & Rehabilitation Engineering 2010;18: 369 - 380.

[10] Martinez JP, Almeida R, Olmos S, Rocha AP, Laguna P. A wavelet-based ECG delineator: evaluation on standard databases. IEEE Transactions on Biomedical Engineering 2004; 51: 570–581.

[11] MIT-BIH Normal Sinus Rhythm Database. http://www.physionet.org/physiobank/database.

[12] Rabiner LR. A tutorial on hidden markov models and selected applications in speech recognition. Proceedings of the IEEE 1989; 7: 257-286.

[13] Young S, Evermann G, Kershaw D, Moore G, Odell J, Ollason D, Povey D, Valtchev V, Woodland P. The HTK Book 3.2. Cambridge University Engineering Department - Speech Group, 2001-2006.

Emna Rabhi,
21 rue 18 Janvier 2034, Ez-zahra, Tunis
emna.rabhi@gmail.com