Intelligent Person Recognition System Based on ECG Signal

Haryati Jaafar and Nurul Syazana Ismail
Faculty of Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus,
Sg. Chuchuh, 02100, Padang Besar, Perlis, Malaysia
haryati@unimap.edu.my

***Abstract*—Automated recognition based on ECG signal is now preferable to identify the person for security monitoring work. This approach is gradually replacing manual techniques that claimed to be outdated. However, it is challenging task to execute the automated system since it is in the infant stage. In this paper, an intelligent person recognition system based on ECG signal is proposed. Here, 79 recorded signals from 79 subjects are used. Three processes are i.e. pre-processing, feature extraction and classification is discussed. A combination of enhanced start and end point detection namely short time energy (STE) and short time average zero crossing rate (STAZCR) is employed in the pre-processing. Subsequently, an autocorrelation method is applied in feature extraction. Finally, support vector machine (SVM) is implemented to evaluate the performance of the system. The experimental demonstrate that 0.93% to 0.97% equal error rate (EER) is achieved when the training data is set to 35.**

***Index Terms*—Autocorrelation method; ECG signal; EER; STE and STAZCR; SVM.**

I. INTRODUCTION

Automated security is one of the major concerns of modern time where secure and reliable authentication is in great demand. However, traditional authentication methods such as password and smart card are now outdated due to they can be lost, stolen and shared [1]. Therefore, automatic biometric authentication has emerged as the backbone of the new-age solution to our society’s demand for the improved security requirements. There are many types of modalities that have been used in a biometric system which can be categorized into two categories i.e. physiological and behavioral [2]. Physiological characteristic are related to the shape or structure of the body for examples fingerprint, face, DNA, iris, retina, hand geometry and palm print [3]. Whereas, the behavioral characteristics are typing rhythm, gait and voice.

In recent times, some studies show that the use internal feature i.e. heartbeat signal which is known as electrocardiogram (ECG) signal has been documented to be suitable for biometric human recognition [4]. It is reported that the heartbeat signal is highly confidential and secure in comparison to the external features that are not robust against falsifications [5]. Compared to the external features like voice, it can be simply imitated, iris which can be duplicated by using contact lenses with copied iris features printed on it, a face which is exposed to artificial disguise, fingerprint and palmprint that can be faked by latex or gummy [4, 6]. The validity of using ECG for biometric recognition is supported by the fact that the physiological and pathological of the heart in different individuals display certain uniqueness in their ECG signals [7].

An ECG signal describes the electrical activity of the heart. The electrical activity is related to the impulses that travel through the heart. It provides information about the heart rate, rhythm, and morphology. Normally, ECG is recorded by attaching a set of electrodes on the body surface such as chest, neck, arms, and legs [7]. The typical ECG waveform comprises a P wave, QRS complex and T wave (Figure 1). The right and left atria or upper chambers make the first wave called a P wave, following a flat line when the electrical impulse goes to the bottom chambers. The delay between P wave and QRS complex is known as PR interval. The ventricles have relatively long ionic potential duration is called QT interval. The plateau part after QRS is known as ST interval. The T wave represents electrical recovery or returns to a resting state for the ventricles interval [5]. The right and left bottom chambers or ventricles make the next wave which is QRS complex. It is generally the strongest wave and most prominent within the electrocardiographic (ECG) signal, with normal duration from 0.06s to 0.1sBecause of this reason, the QRS complex serves as an entry point for almost all automated ECG analysis algorithms and detection of the QRS complex is the most important task in automatic ECG signal analysis [3].

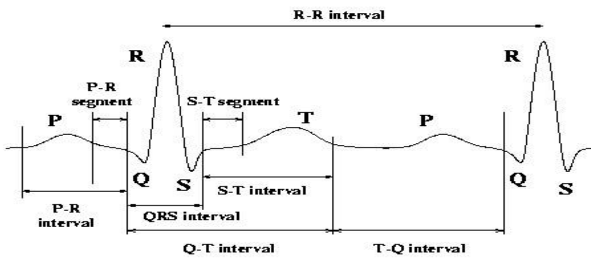


Figure 1: Typical electrocardiogram (ECG) waveform

Although the research of ECG is wide-spread use has been in the spotlight for clinical study, the research for ECG-based biometric recognition is still in its infant stage [5,8-11]. These studies were implemented to the healthy subjects that considered has a regular heartbeat. The main result shared by these studies is that the detailed electrical activity of the heart, as captured by the ECG, is of sufficient quality to be used in high-performance identity recognition systems.

This paper determines the suitable methods for intelligent person recognition system based on ECG signal. The standard design heartbeat biometric system is shown in Figure 2 [12]. The entire system can generally be subdivided into a number of separate processing modules such as preprocessing, feature extraction/selection, and pattern matching. Compares to the previous study, different methods are used in signal segmentation and feature extraction process. In the signal segmentation, the peak finding algorithm is employed to determine R-peak. In addition, the Short Time Energy (STE) and Short Time Average Zero Crossing Rate (STAZCR) are introduced in this process.

Consequently, in the feature extraction process, an autocorrelation method is introduced to find out similarity or relationship features. Finally, the support vector machine (SVM) is used as a classifier in the pattern matching process.

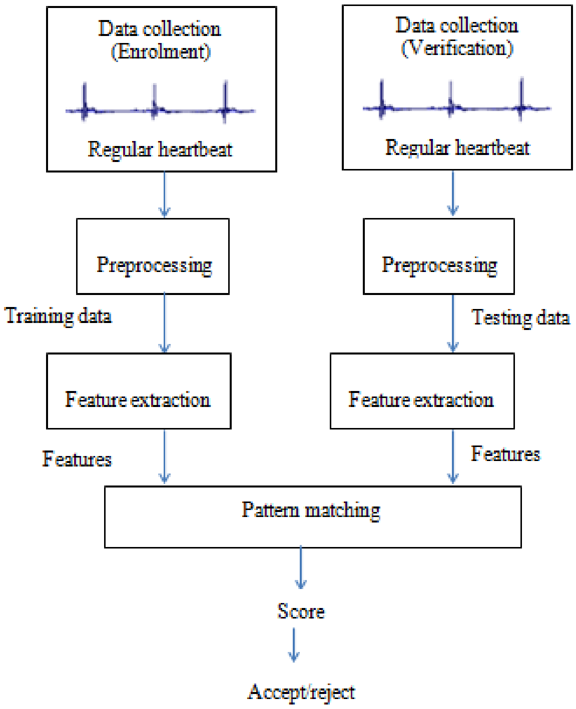


Figure 2: Architecture of heartbeat biometric system

II. Methodology

A. *Data Acquisition*

The ECG database is obtained freely from public heart sound database, assembled for an international competition, the PhysioNet/Computing in Cardiology (CinC) Challenge 2016. The archive comprises nine different heart sound databases sourced from multiple research groups around the world. In this study, the Shiraz University adult heart sounds database (SUAHSDB) was used where this database was constructed using recordings made from 79 healthy subjects and 33 patients (total 69 female and 43 male, aged from 16 to 88 years). During the recording, the subjects were asked to relax and breathe normally during the recording session. The database consists of 114 recordings (81 normal recordings and 33 pathological recordings). The recording length varied from approximately 30s–60s. The sampling rate was 8000 Hz with 16-bit quantization except for three recordings at 44 100 Hz and one at 384000 Hz. The data were recorded in the wideband mode of the digital stethoscope, with a frequency response of 20 Hz–1 kHz [13].

B. *Pre-processing*

The raw ECG signal is often contaminated by a lot of noise which has to be eliminated. The most common types of noise in ECGs are the baseline wander. Power line interference is generated by poor grounding or conflicts with nearby devices. The effect of this noise will be reduced by using a 7th order Butterworth high-pass filter. The cutoff frequencies of the filter are 1–40Hz based on empirical results [14]. Meanwhile, the baseline wander is caused by low-frequency components that force the signal to extend away from the isoelectric line. This occurs in chest-lead ECG signals by coughing or breathing with large movement of the chest. To remove the baseline noise, it is extracted by using a 7th order high-pass Butterworth digital filter with cut-off frequency 0.5Hz. Figure 3 shows the comparison of original signal and filtered signal.

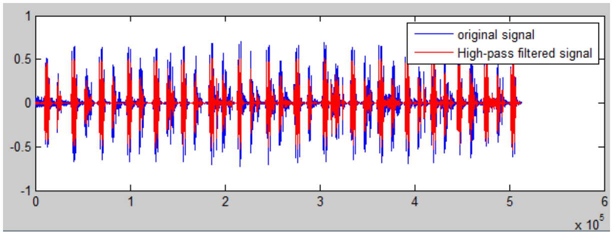


Figure 3: Comparison of original signal and filtered signal

C. *Signal Segmentation*

After filtering, each individual heartbeat waveform will be segmented from the full recording. Instead of using the Pan-Tompkins algorithm that had been employed in previous studies [14], this study applied two other techniques based on time-domain which are the Short Time Energy (STE) and Short Time Average Zero Crossing Rate (STAZCR)[15].

The STE is the energy of a short desired signal segment. It is used to estimate the initial signal in the detection of desired and undesired signal segments. In the meantime, the STAZCR indicates the presence or absence of sound in the input signal [15]. If the value of STAZCR is high, the frame is considered to be undesired signal and if it is low, the frame is considered to be desired signal frame.

These techniques are firstly implemented by the framing process where the signals are converted into frames. The windowing process is then applied to minimize the signal discontinuities at the beginning and end of each frame by zeroing out the signal outside the region of interest.

In this study, the Hamming window was used as the window function due to the side lobes of this window being lower than the other windows. Moreover, the Hamming window results in more attenuation at the outside of the bandpass than the other comparable windows [16]. The window is defined by the expression below:

$$w(k) = \begin{cases} 0.54 - 0.46\cos\left(2\pi k / N - 1\right) & k = 0, \ldots, N-1 \\ 0 & otherwise \end{cases} \tag{1}$$

where: $w(k)$ = Window function
N = Length of each frame

By introducing the framing and windowing processes, the STE function is defined by the following expression:

$$E_m = \frac{1}{N}\sum_{k=1}^{N}\left[x(k)w(m-k)\right]^2 \tag{2}$$

where: E_m = Function which measures the change of voice signal amplitude
$x(m)$ = Input signal in one frame
m = Temporal length of each frame and the operator
$w(m\text{-}k)$ = Frequency shifted window sequence

On the other hand, the STAZCR is defined as:

$$Z_m = \frac{1}{2N}\sum_{k=1}^{N} |\operatorname{sgn} x(k) - \operatorname{sgn}[x(k-1)]|\, w(m-k) \tag{3}$$

where: Z_m = Function which defines the zero crossing count

If the STE of the incoming frame is high, the frame is classified as a desired signal frame and if the STE of the incoming frame is low, it is classified as an undesired signal frame. In contrast, if the STAZCR is high, the frame is considered to be an undesired signal and if it is low, the frame is considered to be the desired signal frame.

Since the strongest wave is defined at QRS complex, it conveys useful information about the features that should be segmented. Due to R-peak is the highest amplitude, it is an obvious point to be captured and used as separation. Therefore, the peak finding algorithm is applied to determine the local maxima and minima of the signal. The local maxima and minima are considered as R-peak. In order to determine the local minima and maxima, the potential points are firstly determined. This is done by calculating the first derivative of E_m and Z_m.The potential points can be detected by considering the sign of the difference. A change from negative to positive number corresponds to local maximum and a change from positive to negative figure corresponds to a local minimum [14].

Subsequently, a selective point is used to ensure the local maxima is selected at least ¼ of the range of the data. For the STE, this point is given as:

$$E_{ms} = \frac{E_{m\max} - E_{m\min}}{4} \tag{4}$$

where: E_{ms} = Selective point for the STE
$E_{m\max}$ = Maximum value of the STE
$E_{m\min}$ = Minimum value of the STE

Meanwhile, the selective point for the STAZCR is given as:

$$Z_{ms} = \frac{Z_{m\max} - Z_{m\min}}{4} \tag{5}$$

where: Z_{ms} = Selective point for the STAZCR
$Z_{m\max}$ = Maximum value of the STAZCR
$Z_{m\min}$ = Minimum value of the STAZCR

The next step is to decide whether the potential points can be selected as local maxima or otherwise. The potential point is considered as local maxima if the point is satisfied with the condition written in Equations (8) and (9).

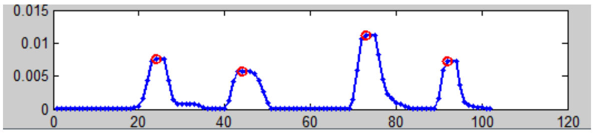
$$E_{mp} > E_{ms} + E_{mr} \tag{6}$$

where: E_{mp} = Value of the STE at current test point
E_{mr} = Value at the reference point

$$Z_{mp} > Z_{ms} + Z_{mr} \tag{7}$$

where: Z_{mp} =Value of the STAZCR at current test point
Z_{mr} = Value at the reference point

Figure 4 shows an example of the final result local maxima after obtaining the peak finding algorithm[15].



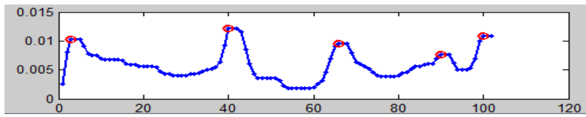


Figure 4: Final result local maxima after obtaining the peak finding algorithm

After determining the local maxima, the threshold level is given as:

After determining the local maxima, the threshold level is given as:

$$T_h = \begin{cases} 1, E_m \geq T_E \text{ and } \quad Z_m \geq T_z \\ 0, \text{otherwise} \end{cases} \tag{8}$$

where T_E and T_Z are the thresholds for STE and STAZCR, respectively and they are defined as:

$$T_E = \frac{W(E_{m\max,1}) + E_{m\max,2}}{W+1} \tag{9}$$

$$T_Z = \frac{W(Z_{m\max,1}) + Z_{m\max,2}}{W+1} \tag{10}$$

Figure 5 illustrates the outputs of the segmentation by using the STE and STAZCR techniques.

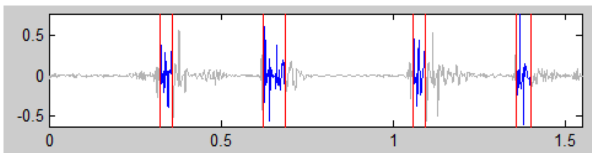


Figure 5: The outputs of the segmentation

D. Feature Extraction

Once the ECG signals have been segmented, the signals which consist of regular and irregular heartbeat segments will be divided into training and testing sets. A training set is used for parameter estimation and it is implemented to build up a model, while a test (or validation) set is to validate the model built [16]. Since heartbeat signals contain the useful feature, redundant features and leftover noises. It is important to pick only features that are unique, significant and least corrupted noise. For this propose, an autocorrelation method is used to find out similarity or relationship features among records of the same subject.

$$R_f(x) = \int_{-\infty}^{\infty} f(t) f^*(t-x) \quad (11)$$

With regard to the series in which successive observations are correlated, the first-order autocorrelation, the lag is one time unit. It is merely the correlation coefficient of the first N-1 observations, X_t , t = 1,2...N-1 and the next N-1observations X_{t+1} , t = 1,2...N-1. The correlation between X_t and X_{t+1} is given by:

$$r_1 = \frac{\sum_{t=1}^{N-1}\left(x_t - \overline{x_{(1)}}\right)\left(x_{t+1} - \overline{x_{(2)}}\right)}{\left[\sum_{t=1}^{N-1}\left(x_t - \overline{x_{(1)}}\right)^2\right]^{1/2}\left[\sum_{t=1}^{N-1}\left(x_t - \overline{x_{(2)}}\right)^2\right]^{1/2}} \quad (12)$$

where: $\overline{x_{(1)}}$ = Mean of the first N-1 observations
$\overline{x_{(2)}}$ = Mean of the last N-1 observations

Since the autocorrelation space is a high dimensional space, an algorithm such as principle component analysis (PCA) is applied to the autocorrelation coefficients for dimensionality reduction[16].

E. Classification

During the recognition, an unknown ECG or testing data is presented to the system. Subsequently, the process of computing similarity between testing data and reference ECG signal template (training data) is employed. In this study, support vector machine (SVM) classifier is employed where this classifier is based on the principle of structural risk minimization and become popular for solving problems in pattern recognition. The SVM is formulated in such a way that it is only capable of discriminating between two classes whereas most classification tasks typically involved more than two classes [16]. The solution of the linearly separable case is started by considering a problem of separating the set of training vectors belongs to two separate classes;

$$D = \{(x^1, y^1), ..., (x^L, y^L)\} \quad x \in \Re^{\eta} \quad y \in \{-1,1\} \quad (13)$$

with a hyperplane;

$$(w, x) + b = 0 \quad (14)$$

where: *w* = Direction in space
b = Position in space
w = Normal to the plane

The hyperplane has the same distance from the nearest points from each class and the margin is twice the distance for each direction, *w*. The support vectors which is a linear combination of a small subset of data, $x_s, s \in \{1,...,N\}$, is the solution for the optimal hyperplane. Equation (14) is minimized by the hyperplane that optimally separates the data which is equivalent to minimizing an upper bound on VC dimension.

$$\Phi(w) = \frac{1}{2}\|w\|^2 \quad (15)$$

VC dimension is a scalar value that measures the capacity of the learning function. The saddle point of the Lagrange functional (Lagrangian) is used to solve the optimization problem and given as;

$$\Phi(w,b,a) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{L} a_i (y^i[\langle w, x^i \rangle + b] - 1) \quad (16)$$

where: a_i = Lagrange multiplier

The Lagrangian has to be maximized with respect to $a \geq 0$ and minimized with respect to *w* and *b*. The solution of the linearly separable case is given by;

$$a^* = \arg\min_{\alpha} \frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i\alpha_j y_i y_j \langle x_i, x_j \rangle - \sum_{k=1}^{L}\alpha_k \quad (17)$$

with constraints;

$$0 \leq a_i \leq c \quad i = 1,...,L \quad and \quad \sum_{j=1}^{L} a_j y_j = 0 \quad (18)$$

The nonlinear mapping is used in the case of the linear boundary is inappropriate which the SVM can map the input vector, *x* into a manifold embedded in a high dimensional feature space *z.* The SVM construct an optimal separating hyperplane in the higher dimensional space [1,14]. The non-linear mappings are polynomial functions, radial basis function and certain sigmoid functions. In this study, the polynomial kernel is employed. Hence, the optimization problem becomes;

$$a^* = \arg\min_{\alpha} \frac{1}{2}\sum_{i=1}^{L}\sum_{j=1}^{L}\alpha_i\alpha_j y_i y_j K(x_i, x_j) - \sum_{k=1}^{L}\alpha_k \quad (19)$$

with constraints;

$$0 \leq a_i \leq c \quad i = 1,...,L \quad and \quad \sum_{j=1}^{L} a_j y_j = 0 \quad x(t) = s(t) - 0.95)x(t-1) \quad (20)$$

where: $K(x_i, x_j)$ = Kernel function that performs the non-linear mapping into feature space.

For the polynomial kernel, it is defined as;

$$K(x_i, x_j) = \Phi(x_i)^T \Phi(x_j) = (\gamma x_i^T x_j - r)^a \quad (21)$$

where: $\gamma > 0$
γ , r and *d* = Kernel parameters

III. Experimental Results

The experiments were implemented using Matlab R2010 (b) and have been tested in Intel Core i5, 2.1GHz CPU, 6G RAM and Windows 7 operating system. A selected signal from 79 healthy subjects with 79 recordings was employed. In this experiment, there were four different numbers of training samples, *j*=20, 25, 30 and 35 were randomly selected for the training and the remaining were used as the testing samples.

The recognition system performance is evaluated based on equal error rate (EER). The lower value of EER gives the better performance of the system in term of accuracy. The EER is obtained by applying two charts which are receiver operating characteristic (ROC) and detection error trade-off (DET). These plots indicate the biometric system false rejection rate (FRR), false acceptance rate (FAR) and genuine acceptance rate (GAR) as its characteristics [17].

Figure 6 illustrates the performance of four classifiers in terms of different numbers of training samples. As shown in this figure, the GAR rates of each training data ascended quickly until the FAR number was about 0.01% and then they either were kept stable or slowly increased as the number of training samples increased.

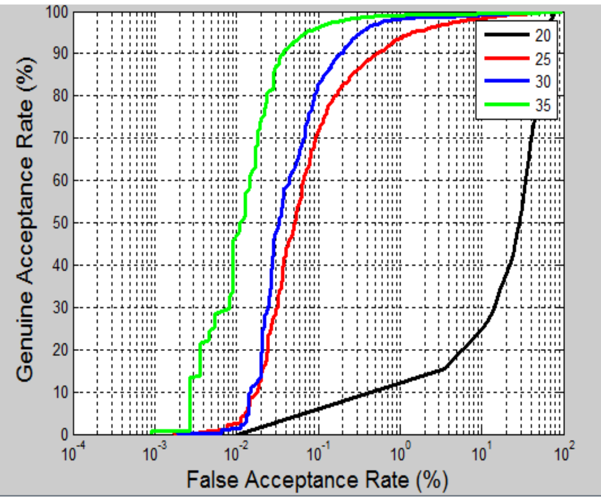


Figure 6: The comparison of ROC curves of different training data

Table 1 indicates the EER performances for a matching process for the different number of training data for ROC curve. It was found that 35 training data gives the best EER percentage compared to the others.

Table 1
Comparison of EER Based on Different Number of Training Data for ROC Curve

Training	EER (%)
20	36.67
25	3.26
30	1.52
35	0.97

The empirical comparisons based on different numbers of training samples for DET curve are presented in Figure 7. As shown in the figures, the SVM has a good performance in a different number of training data when FAR is less than 0.5%. However, when the values of FAR was increased, the performance of the SVM for training data 35 outperformed other number.

Table 2 shows the EER performances for a matching process for the different number of training data for DET curve. It shows that the larger the size of training data, the better the system performances.

Based on EER performances, 0.97% and 0.93% were achieved while from the ROC and DET curves, respectively. Based on these findings, improvement in segmentation and feature extraction shows that differential computing methods give a good result for human recognition system based on ECG signal.

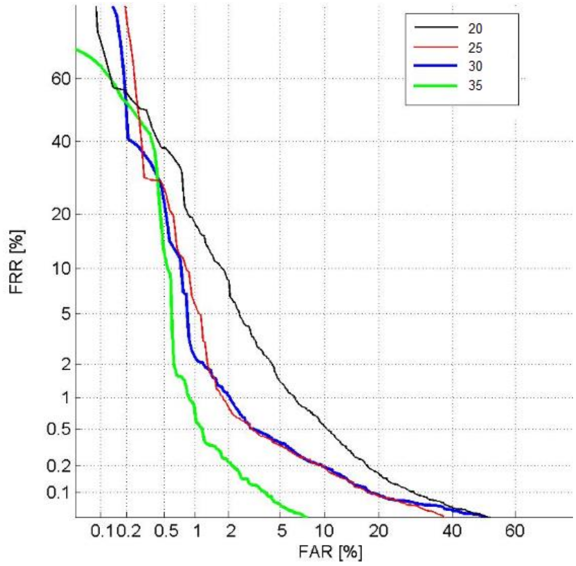


Figure 7: The comparison of DET curves of different training data

Table 2
Comparison of EER Based on Different Number of Training Data for DET Curve

Training	EER (%)
20	3.22
25	1.51
30	1.51
35	0.93

IV. Conclusion

This paper presents a system which is able to perform biometrics recognition by using ECG signals. An improvement of signal segmentation based on time domain has been employed. Subsequently, the autocorrelation method is studied in feature extraction parts and SVM is used to classify the ECG signal. Experimental results show that our system performance is promising where the EER, 0.97% and 0.93% are achieved based on the different number of training data. In conclusion, the methods for segmentation, feature extraction and classification that have been employed in this work could provide high levels of accuracy in recognizing a person based on ECG signals.

Acknowledgment

The authors would like to express their gratitude for the financial support provided by Universiti Malaysia Perlis Short Term Grant (9001-00500).

REFERENCES

[1] D. A. Ramli,., S. A Samad, and A. Hussain, (2009). "A multibiometric speaker authentication system with SVM audio reliability indicator," *IAENG International Journal of Computer Science*, vol. 36, no. 4, pp. 313-321, 2009.

[2] A. H. Mir, S. Rubab, and Z. A. Jhat, "Biometrics verification: a literature survey," *International Journal of Computing and ICT Research*, vol. 5, no. 2, pp. 67-80, 2011.

[3] E. W. K. Yih, G. Sainarayanan, and A.Chekima, "Palmprint based biometric system: a comparative study on discrete cosine transform energy, wavelet transform energy and sobelcode methods," *Biomedical Soft Computing and Human Sciences*, vol. 14, no. 1, pp. 11-19, 2009.

[4] K. A. Sidek, V. Mai, and I. Khalil, "Data mining in mobile ECG based biometric identification," *Journal of Network and Computer Applications*, vol. 44, pp. 83-91, 2014.

[5] Y. N. Singh, "Human recognition using Fisher′ s discriminant analysis of heartbeat interval features and ECG morphology," *Neurocomputing*, 2015.

[6] E. J. D. S. Luz, W. R. Schwartz, G. Cámara-Chávez, and D. Menotti, "ECG-based heartbeat classification for arrhythmia detection: A survey," *Computer methods and programs in biomedicine*, vol. 127, pp. 144-164, 2016.

[7] Y. Wang, F. Agrafioti, D. Hatzinakos, and K. N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP journal on Advances in Signal Processing*, vol. 19, 2008.

[8] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "ECG analysis: a new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808-812, 2001.

[9] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "ECG to identify individuals," *Pattern recognition*, vol. 38, no. 1, pp. 133-142, 2005.

[10] S. C. Fang, and H. L. Chan, "QRS detection-free electrocardiogram biometrics in the reconstructed phase space," *Pattern Recognition Letters*, vol. 34, no. 5, pp. 595-602, 2013.

[11] J. L. C. Loong, K. S. Subari, R. Besar, and M. K. Abdullah, "A new approach to ECG biometric systems: a comparative study between LPC and WPD systems," *World Academy of Science, Engineering and Technology*, vol. 68, pp. 759-764, 2010.

[12] A. Fratini, M. Sansone, P. Bifulco, and M. Cesarelli, "Individual identification via electrocardiogram analysis," *Biomedical engineering online*, pp. 14-78. 2015.

[13] C. Liu, D. Springer, Q. Li, B. Moody, R.A. Juan, F.J. Chorro, F. Castells, J.M. Roig, I. Silva, A.E. Johnson, and Z. Syed, "An open access database for the evaluation of heart sound algorithms," *Physiological Measurement*, vol. 37, no. 12, pp. 2181, 2016.

[14] S.H. Chen, Y.R. Luo, "Speaker verification using MFCC and support vector machine," *Proceedings of the International MultiConference of Engineers and Computer Scientists (IMECS 2009)*, vol. 1, pp. 532-535. 2009.

[15] D.A. Ramli, and H. Jaafar, "Peak Finding Algorithm to Improve Syllable Segmentation for Noisy Bioacoustic Sound Signal," *Procedia Computer Science*, vol. 96, pp.100-109, 2016.

[16] I. Guyon, S. Gunn, M. Nikravesh, and L.A. Zadeh, *Feature extraction: foundations and applications*, vol. 207, Springer Eds, 2008.

[17] A. Martin, G. R. Doddington, T. Kamm, M. Ordowski, M.A. Przybocki, "The DET curve in assessment of detection task performance," *In Proceedings of Eurospeech*, pp. 1895–1898, 1997.