*Journal of Communications Technology, Electronics and Computer Science, Issue 14, 2017*
*ISSN 2457-905X*

# Low – frequency ECG signal based Biometric Identification

Thripurna Thatipelli[1], Dr. Padmavathi Kora[2], Ambika Annavarapu[3]
[1,2,3] Department of ECE, GRIET, Hyderabad

***Abstract*— Security is the major problem facing in the modern environment relating to the information systems. This can be achieved by using various verification and identification methods with highest accuracy. In all the methods, biometric signals based method plays a pivot role. This paper presents the approach relating to the ECG based biometric identification through the experimental data.**

***Keywords: Electrocardiogram (ECG), Kolmogorov-Smirnov test, Pan-Tompkins algorithm, equal error rate (EER), false acceptance rate (FAR)***

## I. INTRODUCTION

Information about a person can be evaluated from the cardio signal. This information can be provided by the use of Electrocardiogram (ECG) of a person to examine his functional state [1]. In education systems functional state analysis can be successfully applied, e.g., building adaptive learning trajectories and smart e – learning systems. Method in [1] depends on the heart rate variability analysis and evaluating the stress express indexes and indexes of centralization in order to build an adaptive learning interface depending on ECG data. For the use of biometric identification and verification ECG signal is considered to provide the distinctive information about a person. Many systems provide data stored in their memory banks to avoid the illegal access to the sources. So many methods are been implemented now successfully in various information systems to implement the biometric identification and evaluation. The first major group of techniques relates to the cognitive ability analysis of a person, which are used to remember and reproduce information like username, password, etc. The second major group consists of verification information that a person consists of, like, USB tokens, RFID tokens, etc. The most prospective category of the identification methods is biometric identification and verification. Such algorithms are based on the analysis of biometric data that is unique for each person. In the case of applying biometric data analysis a person is not required to remember redundant information and always keep an idea of bringing smart card with him in mind: all the information about his identity is provided by a person himself. There are many types of biometric signals, e.g. ear [2], face [3] and gait [4] recognition, eye tracking [5], [6], etc., that can be used separately to identify a person or combined together in order to achieve more accurate results. Moreover, biometric signals can also be used for different purpose in other fields, e.g. eye tracking data can provide information about person's cognitive skills and functions [7] or his functional state [8], which can also be revealed by the ECG signal analysis [1] as mentioned before. ECG signal is a repeated in time signal that consists of specific fragments called QRS complexes corresponding to the depolarization of the right and left ventricles of the human heart.

The QRS complex is surrounded by P, T and U waves [9] that also provide unique information about a signal. R-peak (or R-wave) is a central part and an upward deflection of the fragment. A downward deflection Q wave is presented after the P-wave that represents atrial depolarization. An S-wave is a downward deflection after the R-peak, when T-wave and U-wave represents the repolarization of the ventricles and Purkinje fibers respectively. Authors apply different algorithms based on the ECG signal analysis in order to perform person identification using various classification methods. In [10] authors used 1000 Hz device in order to collect the data from 60 participants and applied the developed by them biometric system that uses discrete wavelet transformation to compute necessary features. The obtained results were accurate and equaled 90.8% of identification accuracy that could be improved up to 100% by changing the value of the rank level. In [9] authors applied neural networks' approach for data classification. ECG data was collected with 500 Hz device, which was then identified using different algorithms. Applying analysis of the QT interval provided 97.7% identification accuracy, while identifying datasets by QRS complexes led to 99.1%.

The authors in [11] provided results of their experimental study, where they also used neural nets algorithm called deep neural nets for the analysis. They extracted QRS complexes from the signals obtained with 500 Hz device, and tested the datasets using their own developed system that provided equal error rate value equaled 0.05%. As the technology has been developing fast during the past decades, it is currently possible to analyze various biometric signals that are collected with the use of the devices that are generally specific for each particular signal type. The existing devices provide different sampling rates and, hence, different quality of the collected measurements. Considering devices with high sampling rate, it should be mentioned that the higher the frequency of obtaining signal's data is, the more complicated hardware a device has and, hence, the more is its cost. The equipment with low frequency costs less, but the lack of measurements occurs that prevent from getting highly accurate results for most of the applied identification and verification methods. Hence, it is necessary to develop methods that could be appropriate for both high and low frequency devices and could provide accurate identification results. This work presents a biometric identification approach that was developed for its application on low-frequency biometric signals, and particularly its evaluation on the ECG data. The second section describes the approach and the two developed algorithms. The third section provides information about the conducted experimental study and the results of the algorithms' evaluation. The last section covers concluding statements.

## II. BIOMETRIC IDENTIFICATION APPROACH

The developed biometric identification approach is built based on the analysis of the repeated fragments of the biometric signal. In order to identify a specific signal it is necessary for it to include the parts that are periodically repeated. The approach considers extracting fragments that are specific for each signal and can fully identify it.

### *A. Features of the biometric signal*

At first, let us define a set of features that can identify a fragment of the signal, and, therefore, the whole signal itself. Person's cardiogram is an almost periodic biometric signal with the repeated fragments, called QRS complexes. An example of the signal is presented in Fig. 1.

Figure 1. An example of ECG signal

A QRS complex, as well as the set of these complexes, is widely used in different biometrics' studies, as it provides a unique information about person's cardiovascular system that allows using features of these complexes for the purpose of person identification. Moreover, a QRS complex is a qualifier of person's functional state [1]. The developed approach assumes a QRS complex together with surrounding P, T and U waves to form a pattern for the biometric signal that can be analyzed further. The enclosing complexes are taken into account as the approach should also be applicable for the low frequency signals. In this case, considering only QRS complexes will cause a great lack of measurements that will prevent from providing a highly accurate identification results due to the limited amount of discrete points provided by the device. The pattern of the assumed signal's fragment is depicted in Fig. 2.

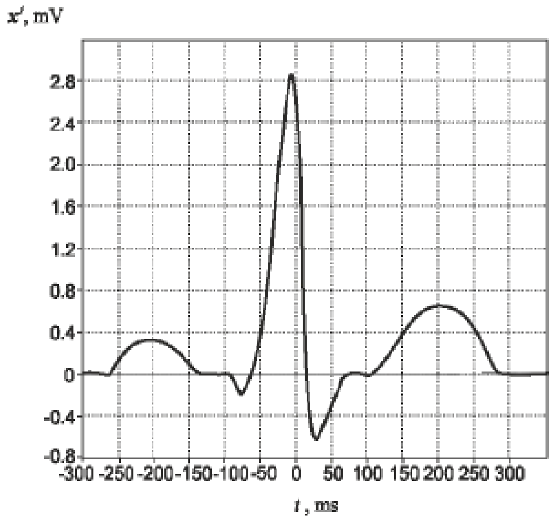


Figure 2. A pattern for the considered fragment

Each fragment is considered in a specific coordinate system in order to provide comparable results. For the ECG fragment the coordinate is built around the R-peak. At first, Pan-Tompkins [12] algorithm is used to detect R peaks in the frequency domain of the obtained signal as well as QRS complexes. After that, a fragment with QRS complex and P, T and U waves is being extracted based on the coordinate of the detected R-peak. All the fragments are assumed in coordinate systems that are built on the same principle, but are specific for each fragment. $O_j x_j y_j$ coordinate system contains a $j_{th}$ fragment, where $x_j$ is a time point in ms and $y_j$ is ECG signal's point value in mV, $A_j$ and $B_j$ are the first and the last points of the fragment respectively. The reference point or the zero time point of a fragment is an R-peak coordinate. The applied coordinate system is presented in Fig. 3.

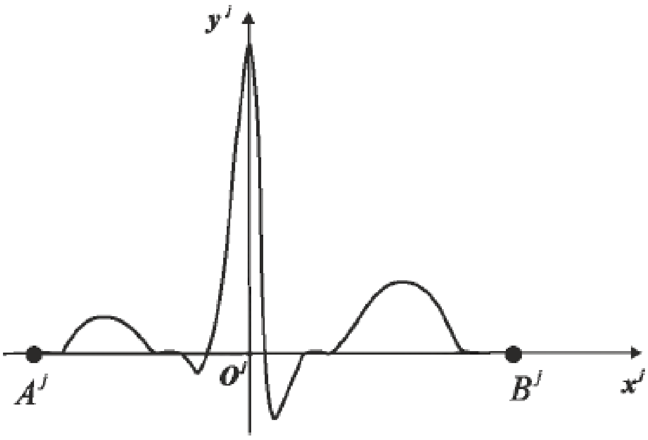


Figure 3. $O^j x^j y^j$ coordinate system of the fragment

In order to restore properties of a continuous signal, an approximation can be applied to the discrete signal obtained from the device. In our approach an approximation based on the Taylor series is used in order to get the properties of the considered fragment. The signal can be approximated by the following formula:

$$x^j(t_i+\tau)=x^j(t_i)+\frac{\left(x^j(t_i)\right)^1}{1!}\tau+\frac{\left(x^j(t_i)\right)^2}{2!}\tau^2+\cdots+\frac{\left(x^j(t_i)\right)^m}{m!}\tau^m+\cdots \qquad (1)$$

where $i=2,3,4,5,\tau=t-t_i$

$$\left(x^j(t_i)\right)^m=\frac{\nabla^m x_i^j}{\square^m}+O(\square) \qquad (2)$$

$\nabla^m x_i^j$, is a finite backward difference of the $m^{t\square}$ order, $\square>0$is a sampling interval. A set of finite differences forms a set of features that can be used for identifying an ECG signal and, hence, a person. Due to the relation between a number of the observed points in a fragment and devices' sampling rate, the size of the set (or a number of terms in Taylor series) is limited. As we assume low-frequency signals in our study, we obtain only ten features to analyze:

- The main coordinates of four considered points in fragment's coordinate system;
- Three first order differences;
- Two second order differences;
- One third order difference.

### *B. Biometric Identification Approach*

The presented approach includes two algorithms of identifying the collected biometric signals that differ from each other by the way of comparing the collected data. However, the first steps are the same for both algorithms. The scheme of the algorithm is presented in Fig. 4. As for biometric signals, the first necessary step of collecting data is a stimulus presentation that affects a person and provides a possibility to gather all required measurements. Considering electrocardiographic signals that are the main object of this paper, it can be any illustrated materials, electronic tests, texts, etc. As the stimulus is being presented to a person, ECG data is collected using a specific device and then stored in the data storage or database. A collected signal is presented as a specific dataset in storage. A set of datasets form a class pattern for a person, whom these datasets were collected for. Each $r^{th}$ dataset is related to a list of ten features $\varsigma=\{\varsigma_{r,1},\varsigma_{r,2},\dots,\varsigma_{r,10}$ that are computed for it.

#### *1) Point-to-point algorithm*

Point-to-point algorithm is based on the paired comparison of two particular datasets. It assumes computation of distances between two particular datasets. The first step is to calculate distance between datasets for a specific feature. The distance is based on the probabilities of Kolmogorov-Smirnov test that compares the distributions of the feature's values for both datasets in order to reveal their likeliness:

$$d(\varsigma_{r,t},\varsigma_{q,t})=-\ln\left[p\left[K\geq\sqrt{\frac{n_r n_q}{n_r+n_q}}D'_{r,q}\right]\right]$$

where K is a random variable with Kolmogorov distribution, and $D'_{r,q}=sup\left|F'_{r,n}(\varsigma)-F'_{q,n}(\varsigma)\right|$ is Kolmogorov- Smirnov test statistic for the $i^{t\square}$ feature of $r^{t\square}$ $and$ $q^{t\square}$ datasets f the, $F'_{r,n}(\varsigma)and\,F'_{q,n}(\varsigma)$ are empirical distribution functions for the $r^{t\square}$ $and$ $q^{t\square}$ datasets, $n_r$ $and$ $n_q$ are the number of fragments in the $r^{t\square}$ $and$ $q^{t\square}$datasets.

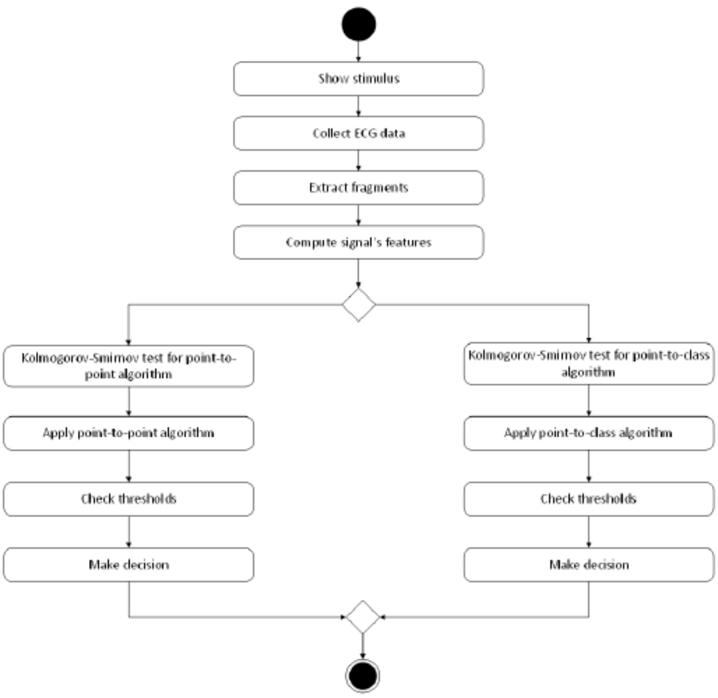


Figure 4. A scheme of identification algorithm

The next step is to calculate the distance between datasets. This value combines all the distances based on different features as following:

$$d(\varsigma_r, \varsigma_q) = \sum_{i=1}^{10} d(\varsigma_{r,t}, \varsigma_{q,t}) \qquad (4)$$

After that, a distance from a dataset to a particular class can be computed:

$$d(\varsigma_r, C_s) = \min_{q \in C_s}[d(\varsigma_r, \varsigma_q)] \qquad (5)$$

As the method should be applied in real systems, there is a possibility of impostors' access to the system. Hence, it is necessary to protect the system from such actions. Our approach assumes calculating thresholds for each of the classes presented in a database in order to limit the number of classes to compare a person's signal to. The following formula is used for this purpose:

$$d_s^+ = \max_{r \in C_s}[d(\varsigma_r, C_s/\{r\})] \qquad (6)$$

where $d_s^+$ is a threshold for an $s^{t\square}$ *class* $(C_s)$

The decision of the algorithm is class's identifier that is set according to the number of nearest datasets. When a number of datasets that are nearest to a particular class equals predefined value, the identifier of the class is assumed correct to the evaluated dataset.

### 2) *Point-to-class algorithm*

The main difference of this algorithm is a change in computing a distance between a dataset and a class. Pointto-class algorithm does not assume the distance between two datasets. Instead of this, before Kolmogorov-Smirnov test is applied it combines values of a particular feature for all the datasets from a class and computes a probability of likeliness for an observed dataset's distribution and the distribution of the class. Point-to-class algorithm was developed on the basis of Bayes theorem and, therefore, a distance from $r_{th}$ dataset to a $C_s$ class is calculated as:

$$d(\varsigma_r, C_s) = -\left[\ln\left[\frac{p(C_s)}{p(\varsigma_r)}\right] + \sum_{j=1}^{10} \ln[p(\varsigma_{r,j}|C_s)]\right] \qquad (7)$$

where $p(C_s)$ is a probability of a dataset being related to the class $C_s$, $p(\varsigma_r)$ is a probability of a dataset having a set of features $\varsigma_r$, $p(\varsigma_r)$ is a conditional probability of $\varsigma_r$, given $C_s$. Point-to-class algorithm also takes into account the possibility of impostors' access, and, hence, thresholds calculated by (6) are applied as well in the point-point algorithm of the approach.

## III. Experimental Results

The assessment of the developed biometric identification approach was held on the ECG data. Application of Pan-Tompkins filtering method [12] helped to detect QRS complexes and, after that, extract the necessary fragments containing QRS complex, P, T and U waves. The fragments being extracted, the required features of the signal are calculated for them, and then classification can be started. Our approach is aimed to be appropriate for low frequency biometric signals. Therefore, it should be mentioned that at first, signals were obtained with 125 Hz device that is of medium sampling rate. In order to evaluate the approach on low-frequency signal, the sampling rate of the signals had been artificially lowered to approximately 30 Hz by extracting each fourth point from the signal's discrete presentation. The evaluation of algorithms provided accurate identification results for both of the proposed algorithms.

(EER) values. EER is based on false rejection rate (FRR) that reflects number of falsely rejected datasets related to the number of all accesses to the system and false acceptance rate (FAR) that shows the relation of a number of falsely accepted datasets to the number of all impostors. After necessary features had been calculated for the 13 classes of the collected datasets, classification started. For all the existing classes the same algorithm of evaluation was used.

On each step of the evaluation of both algorithms one dataset was selected from each class in order to form the test set together with all datasets from other classes. The remaining datasets from a class formed class's template. The algorithms were tested for different threshold values with the lowest threshold equaled to 0 and the highest equaled to the maximum possible distance for the existing datasets and classes. Application of point-to-point algorithm provided 0.13% EER and point-to-class algorithm obtained EER equal to 0.6%.

## IV. Conclusion

Nowadays many experimental studies exploring ECG signals are accomplished in a variety of scientific fields. Person's cardiogram can make available information about his functional state that can be used in education sphere, e.g. in order to build smart systems based not only on student's answers but also on the investigation of the response of his organism and central nervous system. However, the most probable way of affecting ECG data, except for medical purposes, is biometric identification. The presented approach is suitable for many high and low frequency biometric signals, including eye tracking data, where saccades' fragments are examined in order to disclose person's identity and ECG signals. The approach encloses two algorithms, based on point-to-point and point-to-class assessment. The valuation of the approach on ECG data provides highly precise results for a low frequency signal that was used in the experimental evaluation of the algorithms. Point-to-point algorithm obtained EER equaled to 0.13%, while point-to-class algorithm showed 0.6% EER. Both results are quite comparable and, hence, we can conclude that our approach make available accurate identification results for low frequency ECG signals. As future work, we presume to insert the algorithm into a real information system, after further research on the expediency of data representation is concluded.

## REFERENCES

[1] L. Lisitsyna, A. Lyamin, A. Skshidlevsky, "Estimation of Student Functional State in Learning Management System by Heart Rate Variability Method," Frontiers in Artificial Intelligence and Applications, vol. 262, pp. 726-731, 2014.
[2] Ping Yan, K. W. Bowyer, "Biometric Recognition Using 3D Ear Shape," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 29, no. 8, pp. 1297–1308, Aug 2007.
[3] C. Herrmann, C. Qu, J. Beyerer, "Low-resolution video face recognition with face normalization and feature adaptation," *In Proc. 2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, pp. 89-94, 2015, doi: 10.1109/ICSIPA.2015.7412169.
[4] D. Ramanan, D. A. Forsyth, A. Zisserman, "Tracking People by Learning Their Appearance", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 65-81, 2007, doi: 10.1109/TPAMI.2007.250600.
[5] Y. Zhang, J. Rasku, M. Juhola, "Biometric verification of subjects using saccade eye movements," International Journal of Biometrics, vol. 4, no. 4, pp. 317–337, Oct 2012.
[6] F. Deravi, S. P. Guness, "Gaze Trajectory as a Biometric Modality," in Proc. Biosignals, 2011, pp. 335-341.
[7] Y.T. Lin, C.C. Wu, T.Y. Hou, Y.C. Lin, F.Y. Yang, C.H. Chang, "Tracking Students' Cognitive Processes During Program Debugging—An Eye-Movement Approach," IEEE Trans. Education, vol. 59, no. 3, pp. 175–186, 2016.
[8] S. Baltaci, D. Gokcay, "Role of pupil dilation and facial temperature features in stress detection," Proceedings of the 22nd Signal Processing and Communications Applications Conference (SIU), 2014.
[9] M.M. Tawfik, H. Selim, T. Kamal "Human identification using time normalized QT signal and the QRS complex of the ECG,Proceedings of the 7th International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP), pp. 755–759, 2010.
[10] A.D.C. Chan, M.M. Hamdy, A. Badre, V. Badee, "Person Identification using Electrocardiograms," Proceedings of the Canadian Conference on Electrical and Computer Engineering, pp. 1–4, 2006.
[11] A. Page, A. Kulkarni, T. Mohsenin, "Utilizing deep neural nets for an embedded ECG-based biometric authentication system,"Proceedings of the 2015 IEEE Biomedical Circuits and Systems Conference (BioCAS), pp. 1–4, 2015.
[12] J. Pan, W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.