Human ECG for Biometrics using DTW and FLDA

Data · April 2010

2 authors:


Venkatesh N. Murthy
University of Massachusetts Amherst


Srinivasan Jayaraman

Human Electrocardiogram for Biometrics using DTW and FLDA

Venkatesh N
Embedded Systems Innovation Lab,
Tata Consultancy Services,
Bangalore, India
E-mail: venk.n@tcs.com

Srinivasan Jayaraman
Embedded Systems Innovation Lab,
Tata Consultancy Services,
Bangalore, India
srinivasa.j@tcs.com

***Abstract*— This paper proposes a new approach for person identification and novel person authentication using single lead human Electrocardiogram.**

Nine Feature parameters were extracted from ECG in spatial domain for classification. For person identification, Dynamic Time Warping (DTW) and Fisher's Linear Discriminant Analysis (FLDA) with K-Nearest Neighbor Classifier (NNC) as single stage classification yielded a recognition accuracy of 96% and 97% respectively. To further improve the performance of the system, two stage classification techniques have been adapted. In two stage classifications FLDA is used with k-NNC at the first stage followed by DTW classifier at the second stage which yielded 100% recognition accuracy.

During person authentication we adapted the QRS complex based threshold technique. The overall performance of the system was 96% for both legal and intruder situations is verified for MIT-BIH normal database size of 375 recording from 15 individual ECG.

Keywords-ECG, Biometrics, DTW, FLDA

I. INTRODUCTION

Biometrics – identification of an individual based on the physiological and/or behavioral characteristics such as face, gait, fingerprint, retina, vein and voice. Biometrics is the only authentication method that truly identifies the actual applicant as a particular individual than other conventional methods like tokens and passwords. Even though a number of biometrics methods exist, these biometrics methods are not robust against falsification. As the private biometric credentials are not secured, spoofing attack shall occur either physically or digitally. For example our finger prints may be on doors, glass and like and in addition face image could be stolen from any supermarket where we visit often. Recently, researchers [1-5] has proposed electrocardiogram (ECG) as a biometrics method for identifying an individuals. ECG has been used in clinics for the diagnosis and monitoring of heart for almost a century. It remains the best and least invasive method for the task it performs. ECG measurement systems have followed trends in technological advancement becoming more reliable and able to perform a wider range of functions and simpler to use as time has progressed. Human ECG has unique wave shape, amplitude as shown in Figure 1, due to anatomical structure of the heart and physiological conditions [6].

Due to unique feature of ECG it is more difficult to falsify. In addition to this, ECG signals can be used for aliveness detection as well. ECG is a record of time varying bio-potential generated by electrical activity of the heart.

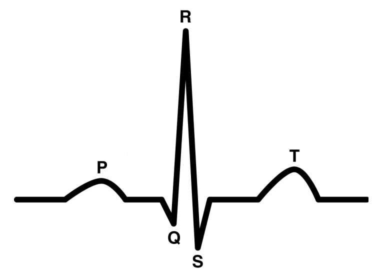

Figure 1. ECG waveform

Due to unique feature of ECG it is more difficult to falsify. In addition to this, ECG signals can be used for aliveness detection as well. ECG is a record of time varying bio-potential generated by electrical activity of the heart. ECG provides information on heart rate, rhythm and morphological details. ECG signals recording are influenced by body habits, gender, age, in long term electrode placement and pharmaceutical drug has short term influence. However inter-individual (body habitus, gender and age) differences are larger than intra-individual.

According to the author's knowledge, the research for ECG for person identification is still in its infant stage only. In previous studies on ECG biometrics Chan et al. [2] have used percent residual difference, correlation coefficient and wavelet distance using single lead (lead I) ECG configuration. In another study Wang et al. [3] have used linear discrimination analysis. The experimental result showed that correct verification rate was 100% using analytical features parameter. However features like amplitude and fiducial points depend on the electrode area and position. Shen et al. [4] combined decision based neural network and template matching method to achieve 100% verification rate. Irvine et al. [5] suggested the heart rate variability as biometric; however performance analysis presented in this paper has not separated the morphological wave features form the heart rate variability features. In few other works to increase the efficiency of the system researchers have adapted multi-modal biometrics approach

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.935

like ECG and facial image data as second source, ECG with palm recognition system and so on.

Any biometrics system can operate on any one of the two different modes

a) Person identification mode: Aims to recognize a person with a defined set of users' database.

b) Person authentication/verification mode: Aims to stop people from using a false identity.

In this paper, a complete biometric system incorporating both person identification using a two stage classification with DTW and FLDA technique to increase the efficiency of classification techniques and person authentication using QRS complex feature has been performed. The feasibility of this study is validated using MIT-BIH ECG database.

II. METHODOLOGY

ECG based biometric system consists of database establishment, data pre-processing and feature extraction, and feature classification and verification as shown in Figure 2.

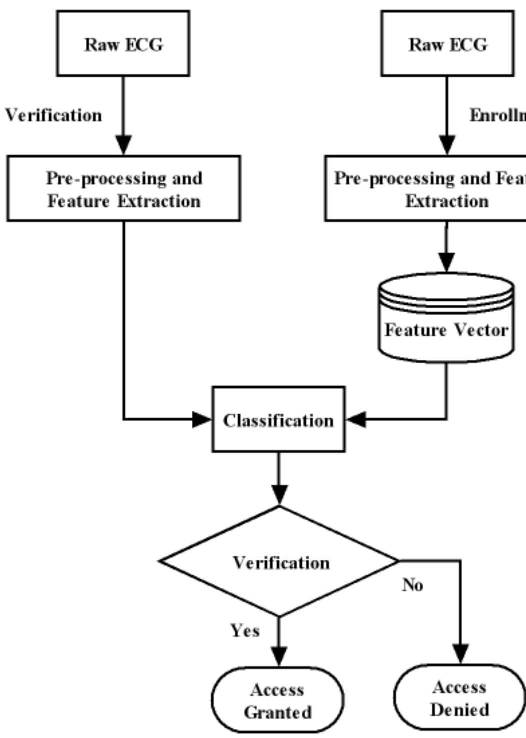

Figure 2. ECG based Biometric System

A. Database Establishment

MIT-BIH normal sinus rhythm database which consists of 18 ECG recordings from different subjects collected at Arrhythmia laboratory of Boston's Beth Israel Hospital have been used. A subset of 15 subjects was selected for our experiment with lead II ECG data set. The database offers one record of 30 minutes for each subject; we partitioned each record into 25 trails with each trail as 1 minute time interval. ECG signals were sampled at 250 Hz

B. Preprocessing and Feature extraction

The Raw ECG data is filtered using band pass filter to reduce the influence of noise, 60Hz interference, and baseline wander. The band pass filter is designed by cascading a low pass filter and a high pass filter.

Feature parameters were extracted for 1 min of filtered ECG signal. QRS complex is first detected using Tompkins method [7, 8]. Once the QRS complex is determined, RR interval, Heart rate, Standard deviation (SD) and number of peak in 1 min were calculated.

Later, zero crossing was performed on ECG trace to determine the onset and offset of P, T, Q and S wave. After identifying the onset and offset, a time window shall be fixed before and after the QRS complex to seek the P wave interval, T wave interval. Based on onset and offset, ST interval, PR interval and QT interval were calculated. From the extracted features number of R peak and Heart rate has been dropped from features vector, as they are dependent on physical conditions of the individuals and they change drastically as well. Feature vector selected were P wave interval, T wave interval. ST interval, PR interval, QRS interval and QT interval.

C. Classifier

ECG classification has been performed using Dynamic Time Warping (DTW) and Fisher's Linear Discriminant analysis (FLDA) with Nearest Neighbor Classifier (NNC) for single stage classification. Further FLDA and DTW were combined to yield better results

1) Dynamic Time Warping

The DTW classifier [9] is based on the ranking of the prototypes by the distance to the query.

Let, $F = (f_1 \ldots . f_n)$ and $G = (g_1 \ldots . g_m)$ be two time series of length n and m, respectively. To align the two sequences using DTW, we construct an n-by-m matrix whose $(i,j)^{th}$ element is the Euclidean distance d(i,j) between two points f_i and g_j. The $(i,j)^{th}$ matrix element corresponds to the alignment between the points f_i and g_j. A warping path, R is a contiguous sets of matrix elements that defines a mapping between F and G and is written as $R=\{r_1 \ldots . r_S\}$ where, $\max(m, n) \le S < m+n-1$. The warping path is typically subject to several constraints such as boundary conditions, continuity, monotonicity, and windowing [10]. The DTW algorithm finds the point-to-point correspondence between the curves, which satisfies the above constraints and yields the minimum sum of the costs associated with the matching of the data points. There are exponentially many warping paths that satisfy the above conditions. The path that minimizes the warping cost is,

$$D(F,G) = \min \sum_{s=0}^{s} r_s \tag{1}$$

The warping path can be found efficiently using dynamic programming to evaluate a recurrence relation, which defines the cumulative distance $\gamma(i, j)$ up to the element (i, j) as the sum of *d(i, j),* the cost of dissimilarity between the i^{th} and the

j^{th} points of the two sequences and the minimum of the cumulative distances up to the adjacent elements:

$$\gamma(i,j)=d(i,j)+\min\{\gamma(i-1,j),\gamma(i,j-1),\gamma(i-1,j-1)\} \quad (2)$$

2) Fisher's Linear Discriminant Analysis

FLDA [10] tries to find a set of projecting vectors W best discriminating different classes. According to the Fisher criteria, it can be achieved by maximizing the ratio of determinant of the between-class scatter matrix S_b to the determinant of the within-class scatter matrix S_w.

$$w = \arg\max \left| \frac{W^{\tau} S_b W}{W^{\tau} S_w W} \right| \quad (3)$$

S_w and S_b for L classes are defined as,

$$S_w = \sum_{i=1}^{L} \sum_{k \in c_i} (x_k - m_i)(x_k - m_i)^{\tau} \quad (4)$$

$$S_b = \sum_{i=1}^{L} n_i (m_i - m)(m_i - m)^{\tau} \quad (5)$$

where, x is the training vector, m is the mean of the total dataset and m_i is the mean ECG feature set for class C_i with n_i samples. Projection matrix W can be computed from the eigenvectors of $S_w^{-1} S_b$ [10].

III. Numerical Results

A. Classification Performance

MIT-BIH Database consisted of 12 users each having 25 trials among which 20 trials was used for training and the remaining 5 trials were used for testing.

Single stage classification was performed using DTW and FLDA with NNC as classifiers. In two stage classification, FLDA was used with k-NNC as a first stage classifier and DTW classifier for second stage classification. First stage classifications top five output (k=5) classes would be given to DTW classifier, which yielded one final output.

This experiment was performed for two situations

1. Legal situation (enrolling user/test user is present in the database) and

2. Intruder situation (enrolling user/test user is not present in the database).

TABLE I. Recognition accuracy for ECG classification

Method	**DTW**	**FLDA**	**DTW+FLDA**
Accuracy (%)	97	96	100

Table 1 presents the experimental results for single stage and two stage classifier and it is observed that by combining DTW and FLDA classifier's the accuracy of recognition rate significantly improved. The reason is, the second stage classifier is able to better distinguish the confusion among the k outputs of first stage classifier.

Classification technique will be very useful for identifying an individual in legal situations but it fails in an intruder situation because ideally intruder should not be matched with any existing users but the classifiers will always try to match the intruder (in either case) closely with the existing users in the training database. To overcome this problem we have presented a new approach for verifying the existence of an identified person in the database as explained in verification section

B. Verification performance

Two stage Classification is followed by verification. Here in addition to the above mentioned training and testing set, we have included 75 more samples of three users, each user consisting of 25 trials (which are not present in the training database) for evaluating the algorithm performance.

Since QRS interval is less variant to stress, the accuracy will not be affected. QRS interval feature is used for verifying the identified person as given by equation 6. If the QRS interval difference between the test user trial and the identified user trial exceeds experimentally found threshold 'θ' then he is a valid user.

$$\begin{array}{ll} if \ \left|C_{qrs} - T_{qrs}\right| \geq \theta & Valid\ User \\ else & Invalid\ User \end{array} \quad (6)$$

where,

C_{qrs} – QRS time interval of classifier identified person's trial and T_{qrs} – QRS time interval of test user's trial.

By varying the threshold level θ we can have a desired trade-off between false acceptance and false rejection rate.

Table 2 presents the performance of the system for different threshold level. The performance of an identification system is measured on the parameters of false rejection (FR), false acceptance (FA), false acceptance rate (*FAR*) and false rejection rate (*FRR*) reported by the system.

TABLE II. Classification and verification performance of ECG for different threshold levels

Threshold level	**#FR**	**FRR (%)**	**#FA**	**FAR (%)**	**Acc (%)**
1	25	33.3	0	0	83.3
2	13	17.3	5	6.6	88
3	3	4	8	10.6	92.6
4	0	0	9	12	94
5	0	0	10	13.3	93.3
6	0	0	10	13.3	93.3
7	0	0	12	16	92
8	0	0	12	16	92

The accuracy [6] of system is also determined using the factors FAR and FRR as given by equation 7

$$Acc(\%) = 100 - \left(\frac{FAR + FRR}{2}\right) \quad (7)$$

The overall performance of the system was 94% $\pm$ 2 for both legal and intruder situations (where the test user was not present in the database) using threshold technique adopted during verification process.

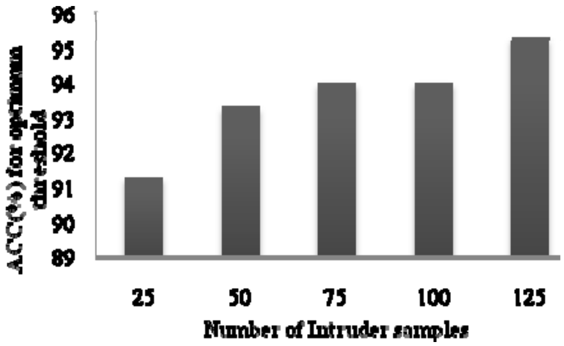

Figure 3. Performance of the system for optimum threshold θ for varying number of intruders.

Figure 3 represents the accuracy of the system with optimum threshold value θ for different set of users used for intruder testing.

IV. CONCLUSION AND FUTURE WORK

Our experimental results have presented a feasibility study of human identification and verification based on features extracted from ECG using DTW and FLDA in two stage classification techniques. Combining two classifiers have yielded comparatively better results and this would be of more significance when there are significantly large database. An idea of using QRS complex for verification process has helped us in improving the performance of authenticating the intruder.

Further experiments are being carried out to evaluate the system with large database and to make sure that database comprises of users of all age, abnormal ECG data and long span of time interval between ECG recordings.

ACKNOWLEDGMENT

The authors wish to thank Mr.Balamuralidhar.P, Head, Innovation Lab, TCS, for his continuous support.

REFERENCES

[1] W. Yongjin, A. Foteini, H. Dimitrios, and N. P Konstatantinos, “Analysis of Human Electrocardiogram for Biometric Recognaition,” EURASIP Journal of advance in signal processing, 2008. doi:10.1155/2008/148658.

[2] D. C. Chan, M. M Hamdy, A. Badre and V. Badee, “Person Identification using Electrocardiograms” CCECE apos;06. Canadian Conference on May 2006 pp:1 – 4

[3] Y. Wang, K. N. Plataniotis, and D. Hatzinakos, “Integrating analytic and appearance attributes for human identification from ECG signal,” Proc of Biometrics symposiums, September 2006.

[4] T. W. Shen, W. J. Tompkins, Y. H. Hu, “One-lead ECG for Identity Verification, Proc IEEE EMBS/BMES conf,62-63,2002.

[5] J. M. Irvine, K. B. Wiederhold, L. W. Gavshon, S. Israel, S. B. McGehee, R. Meyer, M. D. Wiederhold, “Heart rate variability: A new biometric for human identification”, Proc International Conference on Artificial Intellegent, 1106-1111,2001.

[6] N. S. Yogendra, and G. Phalguni, “ECG to Individual Identification”, IEEE Second International Conf on Biometrics: Theory, Applications and Systems (BTAS 2008), Washington DC, USA.

[7] J. Pan, and J. W. Tompkins, “A Real Time QRS Detection Algorithm,IEEE transactions on Biomedical engineering, vol33(3),1985,pp:230-236.

[8] N. Deepthi and J. Srinivasan, “Fight Or Flight Control System For Drivers Using ECG Signal,” IEEE International Conference on Electronics Design and Signal Processing (ICEDSP) 2009, INDIA

[9] J. Niranjan, G Sita, A G Ramakrishnan and M. Sriganesh,. “Tamil handwriting recognition using subspace and DTW based classifiers”, Proc. 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR-9), Tokyo, Oct 2004

[10] M. C. Bishop, “Pattern Recognition and Machine Learning”, Springer (2006).