

Discriminant Analysis for Identifying Individuals of Electrocardiogram

Yogendra Narain Singh

Institute of Engineering & Technology,
Gautam Buddh Technical University Lucknow
singhyn@gmail.com

Abstract. We present a novel method to characterize the electrocardiogram (ECG) for individual recognition. The method works on analytical and appearance features of the heartbeats. The features are analyzed using the principle of Fisher's linear discriminant that produces well separated classes in a lower dimension subspace, under the presence of noise and muscle flexure. The biometric experiment is benchmarked using ECG recordings of the publically available QT database. The proposed method achieves the recognition accuracy of 98.9% on the evaluated subjects which is found optimum than the other best known methods.

Keywords: Identification, Electrocardiogram, Heartbeats, Biometrics.

1 Introduction

Many body parts, personal characteristics and signaling methods have been suggested to use as biometrics for identifying individuals. The biometric attributes that are secured and robust against falsification are the major concerns of the biometric research. The reason that the biometric attributes are not robust against falsification because they hardly ensure the liveness property [1]. The search of a new biometric modality that has an inherent property of liveness is of greater interest. In recent years, the physiological signal especially the electrocardiogram (ECG) has been proposed a novel biometric for human recognition [2]-[5].

An ECG is the graphical recording of the cardiac electric potential as a function of time. The ECG signal is persistent and repetitive. The ECG is generated from a complex self-regulatory physiological processes which makes it highly confidential and difficult to spoof. The ECG signal is an ultimate liveness metric and acceptable due to the latest advances in the sensing technologies [6]. In multi-modal framework the ECG can be combined with face and fingerprint biometrics for secure and robust individual recognition [7]. However, the ECG signal as a biometrics has notable advantages but it has limitations that must be addressed before the operational deployment of ECG biometric system [8]. The ECG data contains a relatively low information content due to low signal-to-noise ratio. The ECG signal varies with an individual emotion, excitement and stress level. There is a need of standard data representation technique and a careful transformation of signal features that preserve the discriminative information between subjects.

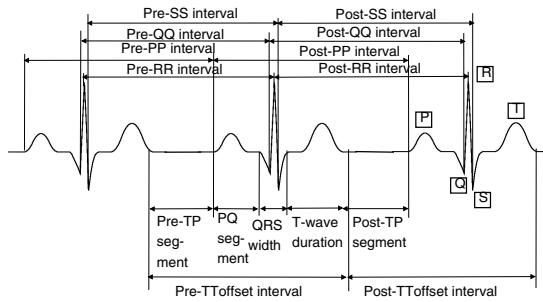


Fig. 1. Extraction of interval features from heartbeats

This paper analyzes the ECG using a novel method that is insensitive to signal variations. The method utilizes a hybrid approach consisting analytical and appearance methods of ECG characterization. The fiducial points of interest are extracted from ECG waveforms using analytical method whereas the appearance method extracts the morphological features of the heartbeats. In the ECG analysis, the analytical features are used as a complementary information with the appearance features. The separability between inter-subject ECG signals and the similarity between intra-subject ECG signals can be drawn using discriminant analysis [9] and the subject classification is performed using a nearest neighbor classifier. The method is benchmarked using ECG recordings selected from publically available QT database [10]. In the sections to follow the methodology and the experimental setup of ECG biometric system along to the results of people recognition are presented.

2 Linear Discriminant Analysis of ECG Signals

This section proposes a computational scheme for evaluating the metric of different ECG signal attributes for their discrimination potential. A normal heartbeat consists of three main components: P wave, QRS complex, and T wave (see for instance Fig. 1). The heartbeats are detected using QRS complex delineator [11]. It uses digital analysis of slope, amplitude and width information of ECG waveforms. After segmenting the heartbeats, the P wave delineator [12] and the T wave delineator [13] are employed to determine the P wave and the T wave along to their end fiducials, respectively. We propose a novel framework of ECG characterization that extracts salient features from each heartbeat, i.e., interval features and morphological features.

The heartbeat interval features contains, the QRS width, the T wave duration, the PQ segment, the pre-TP segment and the post-TP segment as shown in Fig. 1. In this class of features the temporal duration between PP, QQ, SS, TT and RR waves of successive heartbeats are also extracted. The pre-PP (post-PP) interval is the time interval between P_{onset} of a given heartbeat and the P_{onset} of the previous (following) heartbeat. The pre-QQ (post-QQ) interval is the

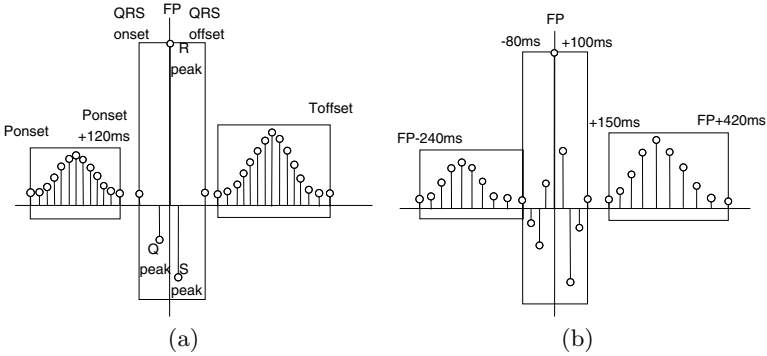


Fig. 2. Extraction of morphological features from heartbeat (a) normal and (b) scaled signals

time interval between Q_{Peak} of a given heartbeat and the Q_{peak} of the previous (following) heartbeat. The pre-SS (post-SS) interval is the time interval between S_{peak} of a given heartbeat and the S_{peak} of the previous (following) heartbeat. The pre-TT (post-TT) offset interval is the time interval between T_{offset} of a given heartbeat and the T_{offset} of the previous (following) heartbeat. Similarly, the pre-RR (post-RR) interval is defined as the RR interval between a given heartbeat and the previous (following) heartbeat. The heartbeat interval features and the interbeat interval features are shown in Fig. 1.

The morphological features contains the amplitude values of the segmented heartbeats. They are extracted from normal and scaled ECG signal whereas the signal is scaled using Pareto method [14] with an aim to reduce the sensitivity of the ECG signal both to noise and motion artifact. Thirty-two features and twenty-eight features are extracted from temporal windows that are set heuristically from the normal and scaled ECG signal, respectively using linear interpolation of uniformly distributed sample positions as shown in Fig. 2.

A gallery set composed of a relatively large group of subjects of different ECG characteristics is prepared. The database contains several examples of ECG signals for each subject in the gallery set and at least one example in the probe set. We employ linear discriminant analysis (LDA) method that uses class membership information and allows us to find eigenfeatures and therefore representations in which the variations among different ECG signals are emphasized, while the variations of the same ECG signal that are due to different physiological conditions are de-emphasized.

Let S be the representation of an original signal to be analyzed, we construct the lexicographic vector expansion R^n where n is the number of features extracted from the signal. Let us define all instances of an individual ECG as being in one class and the ECGs of different subjects as being in different classes for all subjects (N) in the gallery set. We establish a framework for performing a cluster separation analysis in the feature space. Also, having labeled all

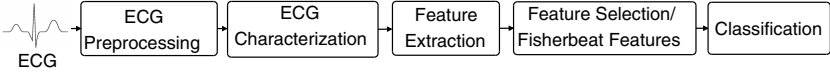


Fig. 3. Classical structure of an ECG biometric recognition system

instances in the training set and having defined all the classes, we compute the within-class and between-class scatter matrices as follows:

$$S_W^{(S)} = \sum_{i=1}^N \sum_{x_{ij} \in N_i} (x_{ij} - \bar{x}_i)(x_{ij} - \bar{x}_i)^T \text{ and } S_B^{(S)} = \sum_{i=1}^N N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$$

where $\bar{x}_i = \frac{1}{N_i} \sum_{j=1}^{N_i} x_{ij}$ is the mean vector of the class X_i and $\bar{x} \in R^n$ is the mean vector of all samples. The discriminatory power of a representation can be quantified using the separation matrix. The class separation matrix and a measure of separability, $D^{(S)}$ can be estimated by maximizing the expression, $\arg \max_{\phi} \frac{|\phi^T S_B \phi|}{|\phi^T S_W \phi|}$ and derive a set of m feature basis vectors $\{\phi\}_{k=1}^m$. We assume that the within-class scatter matrix S_W is non-singular, then the FLD finds ϕ as the k most significant generalized eigenvectors of $S_W^{-1} S_B$ those correspond to the first k largest generalized eigenvalues, thus $D^{(S)} = S_W^{-1} S_B \phi_k = \lambda_k \phi_k$. The generalized eigenvectors form the basis representation of the gallery and the probe ECG signals. For identity recognition the classification is performed using a nearest neighbour classifier.

3 Recognition Results of ECG Biometric Method

The classical structure of an ECG biometric recognition systems is shown in Fig. 3. The ECG recordings are first preprocessed that makes necessary correction of signal from noise and muscle flexure. The processed signal is then passed for ECG characterization followed by feature extraction. The Fisherbeat features are selected from heartbeat features and thus most discriminatory information is preserved. Finally the subject classification is performed in the projected domain using nearest neighbor criterion.

The performance of the ECG biometric system is tested on the ECG recordings of QT database. The first channel ECG recordings of 85 subjects are used in this experiment. The gallery and the probe datasets are prepared from ECG recorded at different time segments. The results from the subjects class specific linear projection of the ECG features demonstrate that there exist a correlation between number of Fisherbeat features and subject classification performance. The decrease in intra-subject variability and an increase of inter-subject separability, represented by first few principal components are shown in Fig. 4(a) and Fig. 4(b).

The receiver operating characteristic (ROC) curve related to better recognition rate achievable by the system, drawn for different projected dimensions are shown in Fig. 4(c). The system achieves better result at dimension ten (DIM 10) such as the values of the GAR are reported to 96% at zero FAR and 100% at 2% of FAR. The performance of the system degrades at higher dimensions and also for the dimensions below ten.

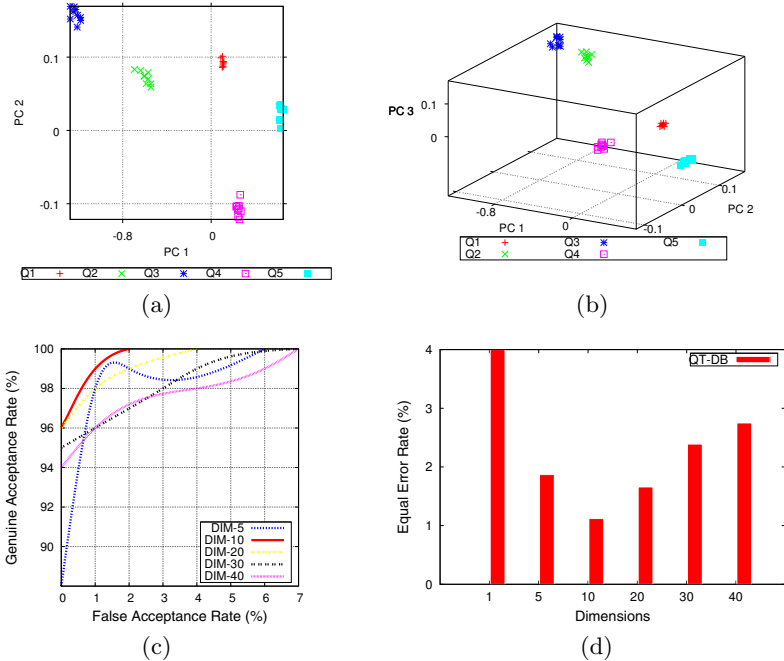


Fig. 4. Discriminability between ECG signals of five subjects represented by first (a) two and (b) three principal components. (c) ROC curve for different projected dimensions. (d) The histogram representation of recognition performance of discriminant features.

The performance of the Fisherbeat features vary with the number of principal components. The equal error rate (EER) values gradually decrease with an increase in component dimensions and found lowest to 1.1% at dimension ten. For higher dimensions of principal components, EER values raise very steadily. This trend in performance is shown in Fig. 4(d) by histogram drawn between EER and dimensions. These results signify that the discriminatory information of an ECG signal has been significantly captured by the Fisherbeat features. In particular, the Fisher-eigenspace of significantly lower dimension is capable to represent the discriminatory information of an individual heartbeat.

4 Conclusion

This study has proposed a novel method to characterize the ECG signal for individual recognition. The features of classes heartbeat interval features and waveform morphological features are derived. Using discriminant analysis the Fisherbeat features are selected. The advantages of using the discriminant analysis of the heartbeat features are the suppression of noise from the signal, reducing the complexity to access a larger attribute set and simplifying the classification process. The reported results have proved the effectiveness of the proposed ECG biometric method for individual recognition.

References

1. Singh, Y.N., Singh, S.K.: A Taxonomy of Biometric System Vulnerabilities and Defenses. *Intl. J. of Biometrics* 5(2), 137–159 (2013)
2. Biel, L., Petteesson, O., Philipson, L.: ECG Analysis: A New Approach in Human Identification. *IEEE Trans. on Instrumentation and Measurement* 50(3), 808–812 (2001)
3. Singh, Y.N., Gupta, P.: ECG to Individual Identification. In: *Proc. Biometrics: Theory, Applications and Systems*, Washington DC, USA, pp. 1–8 (2008)
4. Singh, Y.N., Gupta, P.: Biometrics Method for Human Identification Using Electrocardiogram. In: Tistarelli, M., Nixon, M.S. (eds.) *ICB 2009*. LNCS, vol. 5558, pp. 1270–1279. Springer, Heidelberg (2009)
5. Singh, Y.N., Gupta, P.: Correlation Based Classification of Heartbeats for Individual Identification. *J. Soft Computing* 15(3), 449–460 (2011)
6. Zhao, C., Wysocki, T., Agrafioti, F., Hatzinakos, D.: Securing Handheld Devices and Fingerprint Readers with ECG Biometrics. In: *Proc. IEEE Fifth Int. Conf. Biometrics: Theory, Applications and Systems*, pp. 150–155 (2012)
7. Singh, Y.N., Singh, S.K., Gupta, P.: Fusion of Electrocardiogram with Unobtrusive Biometrics: An Efficient Individual Authentication System. *Pattern Recognition Letters* 33, 1932–1941 (2012)
8. Singh, Y.N., Singh, S.K.: Evaluation of Electrocardiogram for Biometric Authentication. *Journal of Information Security* 3, 39–48 (2012)
9. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*, 2nd edn. Wiley India, New Delhi (2009)
10. Laguna, P., Mark, R.G., Goldberger, A., Moody, G.B.: A Database for Evaluation of Algorithms for Measurement of QT and Other Waveform Intervals in the ECG. *Comput. Cardiol.* 24, 673–676 (1997)
11. Pan, J., Tompkins, W.J.: A Real Time QRS Detection Algorithm. *IEEE Trans. on Biomedical Engineering* 33(3), 230–236 (1985)
12. Singh, Y.N., Gupta, P.: A Robust Delineation Approach of Electrocardiographic P Waves. In: *Proc. 2009 IEEE Symposium on Industrial Electronics and Applications*, Kuala Lumpur, Malaysia, vol. 2, pp. 846–849 (2009)
13. Singh, Y.N., Gupta, P.: A Robust and Efficient Technique of T Wave Delineation from Electrocardiogram. In: *Proc. Second Int. Conf. on Bio-inspired Systems and Signal Processing*, Porto, Portugal, pp. 146–154 (2009)
14. van den Berg, R.A., Hoefsloot, H.C.J., Westerhuis, J.A., Smilde, A.K., van der Werf, M.J.: Centering, Scaling, and Transformations: Improving the Biological Information Content of Metabolomics Data. *BMC Genomics* 7(142), 1–15 (2006)