Indian Journal of Science and Technology, Vol 9(30), DOI: 10.17485/ijst/2016/v9i30/97958, August 2016

ISSN (Print): 0974-6846 ISSN (Online): 0974-5645

Improvement in ECG based Biometric Systems using Wavelet Packet Decomposition (WPD) Algorithm

Zeeshan Hassan*, Syed Omer Gilani and Mohsin Jamil

Department of Robotics and Artificial Intelligence, School of Mechanical and Manufacturing Engineering (SMME), National University of Sciences and Technology (NUST), H-12 Main Campus, Islamabad, Pakistan; hzeeshan7@gmail.com, omer@smme.nust.edu.pk, mohsin@smme.nust.edu.pk

Abstract

In this paper, a non-fiducial approach using Wavelet Packet Decomposition (WPD) algorithm for repeated examination of solitary cable ECG used for individual identification is planned and tested. Multiple samples of ECG wave are extracted considering R-peak as a reference and WPD algorithm is applied for feature extraction. This feature file is fed as an input to a machine learning classifier i.e. random forest in order to classify the individuals. In this work, records from publicly available MIT/BIH arrhythmia dataset have been utilized to evaluate the proposed system. Best result relies on third level of wavelet decomposition using Daubechies wavelet to analyze the signal. Furthermore ranker search method is used in conjunction with relief attribute evaluator for feature selection and random forest classifier is applied by generating 100 trees. It is shown that the method is effective for quantifying the classification of arrhythmia ECG signals with accuracy of 92.62%.

Keywords: Biometrics, ECG, MIT-BIH Arrhythmia Database, Random Forest, Wavelet Packet Decomposition

1. Introduction

Biometrics indicates metrics associated to individual human being. These systems are also considered as a recognition tool to perform individual's identification and verification on the basis of their biological and physiological characteristics. Several identification mechanisms have been studied in the past based on physical features¹ such as face, ear, iris, finger prints, face, hand geometry and behavioral characteristics2 that include gesture, signature, voice, key stroke as well as biological signal characteristics such as ECG^{3,6}. However, physical features and behavioral characteristics can be tempered, falsified, imitated and copied. Appearance detection system can be deceived by a photograph, hand markers can be reproduced, contact lens can easily doge iris recognition system, passwords can be observed, forgotten or hacked and sound can be imitated or preprocessed4. Implementing ECG based biometric systems exhibit some unique advantages over conventional biometrics i.e. it is difficult to falsify and also provides proof for aliveness of the subject³.

ECG indicator is the explanation of the dynamic

action of the heart, which is generated by depolarization and repolarization of heart muscles during its working. Mechanism of the ECG waveform was first proposed by the Willem Einthoven in 1924 and he was presented by the Nobel Prize⁵. Almost each human being contain identical model of ECG as exposed in Figure 1, however it has reasonable difference in its detailed shape.

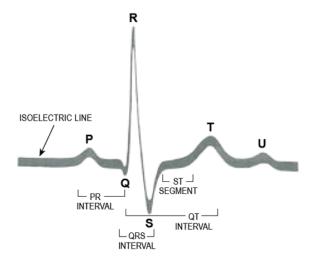


Figure 1. ECG model.

^{*} Author for correspondence

ECG signal is acquired by using single or multiple electrodes positioned externally at different locations on the skin. Hence the subject's gender, age, habits, heart diseases as well as exclusive anatomy of the heart can influence the fiducial features of the ECG waveform.

In this research we suggested a transform based technique for people recognition and verification. Wavelet Packet Decomposition algorithm is applied to extract features from multiple ECG signals. Random forest algorithm is then applied to this feature set for the reason of categorization. Given scheme is verified on publically available MIT/BIH arrhythmia dataset and results show 92.62% accuracy. Step by step diagram of proposed system is given in Figure 2.

Section 2 of this paper is concerned with literature review, part 3 will explicate the projected methodology, part 4 includes result discussion, part 5 will conclude the whole research work and finally references are specified.

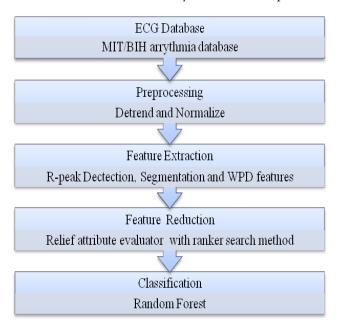


Figure 2. Flow chart of proposed system.

2. Literature Review

ECG based biometric systems fully depends on feature extraction of ECG waveform. Algorithm for feature extraction process can be divided into three main classes depending on waveform, transform and statistics⁷. Every approach has some advantages and disadvantages. Waveform based methods of feature extraction are known as fiducial detection approach while transform

and statistical based techniques are the non-fiducial detection approaches for the purpose of identification and verification. Recognition plus authentication are inherent and clear states of uniqueness8. The fiducial technique demands calculation of height and chronological length among distinctive position which are corresponding to the confined maximum as well as minimum, important elevations, low points sandwiched between climax, arrivals and balances of sole ECG waveform. Consequently, a fiducial characteristic absolutely relies on the accurate finding of concerned positions which is a big challenge for researchers as ECG is a non-stationary signal. Temporal or waveform based traits can be acquired by means of certain fiducial points for classification purpose. These algorithms are more suitable to classify regular signals, but the accuracy might not be quite good due to complex computational mechanism of non-stationary waveforms⁷.

However, non-characteristics methods dig out different instructions in large ECG data which do not include transitional attributes⁹. Therefore the former method of feature extraction enhances the inter subject variability meanwhile reducing intra subject variability. A number of transforms like, Fourier, wavelet, wavelet energy, fast approximate entropy and Discrete Cosine Transform (DCT) are commonly used transform based algorithms in literature¹⁰. Different trait mining scheme in ECG waveform is shown in Figure 3.

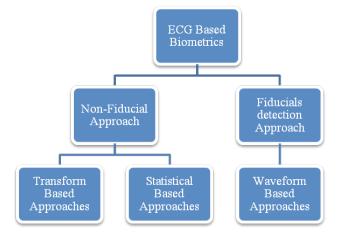


Figure 3. Feature extraction practices in ECG based biometrics.

ECG signal is a common example of non-stationary signals which exhibits a little irregularity in its shape with the passage of time. Analysis of this kind of signals

usually requires information in both time and frequency domains¹¹. Moreover, the wavelet transforms have a completely expandable window that gives further authentic and detailed explanation of signal attributes¹². Due to this reason researchers prefer to apply wavelet transform¹³ for trait mining of static signals over Fourier transform¹⁴ and DCT¹⁵. The statistical-based algorithms^{16,17} normally require short execution time, however they demands a well-designed analytical trait to provide high efficiency.

⁶Used ECG as a biometric trait and proposed that the fiducial points of the waveform are distinctive to every individual i.e. ECG statistics carries sufficient inherent information for identification purposes. ⁶Got ECG records through signal acquiring apparatus and calculated height of ECG to judge against characteristics. ⁶and ¹⁸ tested their systems on 20 subjects and obtained nearly 100% identification accuracy. Table 1 compares different techniques and related work.

3. Methodology

3.1 Pre-Processing

Pre-processing is very important step in proposed system which is applied to take out noise present ECG waveform. In the process of acquiring, a significant quantity of unwanted signal combined to the ECG as a result of dislocation of measuring apparatus. Mostly measuring apparatus moves from its actual position from the body of the subject due to the process of inhalation which causes baseline drifts and power line interfaces. Baseline drifts and power line interfaces are low and high frequency components of noise respectively. Apart from that all signal acquiring devices also have integral error in its readings. As a result, the imprecision of finding enhances owing to inception and balance of complex waves present in the signal under examination²³. Pre-processing steps involves detrend and normalization which significantly

improves the probability of R-peak detection. Equation 1 shows the way of normalization in which y presents the input signal and x presents the normalized output signal.

$$x = \frac{y - \min(y)}{\max(y) - \min(y)}$$
 (1)

3.2 Feature Extraction

After normalizing signals between zero and one in preprocessing step, segmentation is performed followed by R-peak detection. R-peak is detected by setting threshold of 0.75 for minimum peak height within find peak command in MATLAB. Only 15 peaks from each waveform are utilized in proposed system. For segmentation purpose we move 70 samples before and 140 samples after each R-peak. WPD algorithm is then applied to extract features from each segment obtained in prior step. Daubechies, Biorthogonal, Coiflets and Symlets are available wavelets which may be applied. Each wavelet posse's different shape however fundamental characteristics of these wavelets are similar. Best results are obtained by using Daubechies wavelet to analyze the signal because its shape is close to the ECG waveform²⁴. Corresponding high-pass filter or a low-pass filter obtained by multiplying Daubechies family and signal under consideration, which are respectively known as the details and the approximations of the signal.

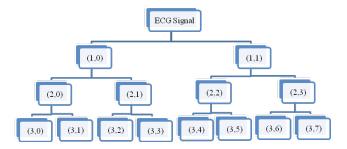


Figure 4. Wavelet decomposition tree.

The prime objective of finding features is to extract

Table 1. Assessment of methods in ECG biometrics

Writer	Algorithm	Number of people	Traits	Correctness
Biel et al. ⁶	Principle Component Analysis	20	Waveform Based Approach	100%
Shen et al.18	DBNN	20	Waveform Based Approach	99.90%
Palaniappan et al.19]	R-R interval	10	Fiducial	97.6%
Gahi ²⁰	Template Matching	16	Waveform Based Approach	99.80%
Janani et al. ²¹	K-Nearest Neighbors	17	Mix	87.90%
Abdel Raheem et al.22	VCG	22	Mix	99.30%
Chiu et al.10	Linear Discriminate Analysis	35	Transform Based Approach	99.50%
Proposed work	Wavelet Packet Decomposition	47	Non Fiducial	92.62%

and keep important data of the signal being analyzed. Each signal is decomposed at depth 3 with db1 wavelet packets using Shannon entropy. Wavelet decomposition tree of proposed system is shown in Figure 4.

3.3 Feature Reduction and Classification

Feature reduction and classification is performed in WEKA environment that is especially designed for machine learning algorithms. Relief attribute evaluator that falls under the category of supervised attribute filter is applied to select useful attributes. It is very flexible and allows a variety of search and evaluation methods to be combined. Ranker search method is used in conjunction with relief attribute evaluator to find a subset of extremely interrelated features between similar classes having small inter-class correlation. This search method ranks attributes by their individual evaluations. Random forest classifier is applied with 10 fold cross validation by generating 100 trees to evaluate best performance of proposed strategy. In this classifier percentage split of data that is used for training and testing purposes are 66% and 34% respectively.

4. Results Discussion

To determine the hit rate of the planned identification algorithm, an ample experiment was done on MIT-BIH arrhythmia database²⁵ which comprises of 48 groups. Double-lead ECG record is kept for 30 minutes, summing up to a whole of 24 hours of ECG data. There are total 47 individuals in this dataset including 25 male having age limit from 32 to 89 as well as 22 female having age limit from 23 to 89. Dataset ID 201 and 202 attain from the same body moreover it has sampling rate of 360 Hz.

In this research 15 segments are extracted from heartbeat of every individual resulting a total of 705 instances. Score of cases predicted positive which are really encouraging is called true positive while score of cases predicted positive that are in fact depressing is known as false positive. Recall is the true positive rate also referred to as sensitivity. Precision is the ratio between true positive and predicted positive and it is also referred to as Positive Predictive Value (PPV). F-Measure is a combination of precision along with recall and measured by calculating their harmonic mean. Out of 705 instances, correctly and incorrectly classified instances are 653 and 52 respectively giving 92.62% accuracy. Table 2 provides

the complete outcome of this research.

Table 2. Result summary

True Positive	False Positive	Preci- sion	Recall	F-Mea- sure	Accura- cy
Rate	Rate				
0.926	0.002	0.933	0.926	0.923	92.62%

The confusion matrix is generally known as eventuality chart. We have 47 categories which lead to 47×47 confusion matrix. The crossway of the matrix shows the sum of properly classified cases. Confusion matrix is plotted using MATLAB shown in Figure 5.

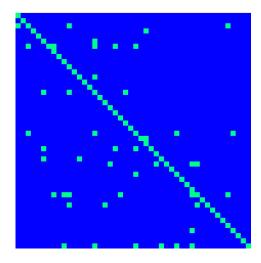


Figure 5. Plot of confusion matrix.

5. Conclusion

ECG is the most important physiological signal in human beings which can be used for various real life applications to identify individuals. Therefore the correct processing and finding useful features of ECG signal have significant importance. This research proposes robust, less intrusive, safe and highly accurate biometric approach using MIT/BIH arrhythmia dataset. It is concluded that WPD is more suitable for analyzing ECG signals and the suitability of WPD depends on the proper selection of mother wavelet. Best results are obtained by using Daubechies wavelet to analyze the signal because its shape is quite similar to the ECG waveform. Upcoming research will focus on the improvement in hybrid biometric systems depending on ECG and some other feasible traits i.e. finger print for better identification outcome.

6. References

- 1. Bansal A, Agarwal R, Sharma RK. FAR and FRR based analysis of iris recognition system. 2012 IEEE International Conference on Signal Processing, Computing and Control (ISPCC); 2012 Mar. p. 1-6.
- 2. Syed Z, Helmick J, Banerjee S, Cukic B. Effect of user posture and device size on the performance of touch-based authentication systems. 2015 IEEE 16th International Symposium on High Assurance Systems Engineering (HASE); 2015 Jan. p. 10-7.
- Silva H, Lourenco A, Lourenco R, Leite P, Coutinho D, Fred A. Study and evaluation of a single differential sensor design based on electro-textile electrodes for ECG biometrics applications. IEEE Sensors; 2011. p. 1764-7.
- 4. Phillips PJ, Martin A, Wilson CL, Przybocki M. An introduction evaluating biometric systems. Computer. 2000; 33(2):56-63.
- Ross AA, Nandakumar K, Jain AK. Handbook of multibiometrics (International Series on Biometrics). Secaucus;
- Biel L, Pettersson O, Philipson L, Wide P. ECG analysis: A new approach in human identification. IEEE Transactions on Instrumentation and Measurement. 2001; 50(3):808-12.
- 7. Zeng F, Tseng KK, Huang HN, Tu SY, Pan JS. A new statistical-based algorithm for ECG identification. 2012 IEEE Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP); 2012 Jul. p. 301-4.
- 8. Sahoo SK, Choubisa T, Prasanna SM. Multimodal biometric person authentication: A review. IETE Technical Review. 2012; 29(1):54-75.
- 9. Coutinho DP, Silva H, Gamboa H, Fred A, Figueiredo M. Novel fiducial and non-fiducial approaches to Electrocardiogram-based biometric systems. IET biometrics. 2013; 2(2):64-75.
- 10. Chiu CC, Chuang CM, Hsu CY. Discrete Wavelet Transform applied on personal identity verification with ECG signal. International Journal of Wavelets, Multiresolution and Information Processing. 2009; 7(03):341-55.
- 11. Haque AKMF, Ali MH, Kiber MA, Hasan MT. Detection of small variations of ECG features using Wavelet. ARPN Journal of Engineering and Applied Sciences. 2009; 4(6):27–30.
- 12. Addison P, Watson JN, Clegg GR, Holzer M, Sterz F, Robertson CE. Evaluating arrhythmias in ECG signals using wavelet transforms. IEEE Engineering in Medicine and Biology Magazine. 2000; 19(5):104-9.
- 13. Chan AD, Hamdy MM, Badre A, Badee V. Wavelet distance measure for person identification using Electrocardio-

- grams. IEEE Transactions on Instrumentation and Measurement. 2008; 57(2):248-53.
- 14. Chan AD, Hamdy MM, Badre A, Badee V. Person identification using Electrocardiograms. 2006 IEEE Canadian Conference on Electrical and Computer Engineering, CCECE'06; 2006 May. p. 1-4.
- 15. Ye C, Coimbra MT, Kumar BVKV. Investigation of human identification using two-lead Electrocardiogram (ECG) signals. 2010 Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS); 2010 Sep. p. 1-8.
- 16. Zhang Z, Wei D. A new ECG identification method using Bayes' Theorem. 2006 IEEE Region 10 Conference, TEN-CON; 206 Nov. p. 1-4.
- 17. Agrafioti F, Hatzinakos D. ECG based recognition using second order statistics. 2008 IEEE 6th Annual Communication Networks and Services Research Conference, CNSR; 2008 May. p. 82-7.
- 18. Shen TW, Tompkins WJ, Hu YH. One-lead ECG for identity verification. Proceedings of the Second Joint Engineering in Medicine and Biology, 2002 IEEE 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference. 2002; 1:62-3.
- 19. Palaniappan R, Krishnan SM. Identifying individuals using ECG beats. Proceedings of the International Conference on Signal Processing and Communications (SPCOM'04); 2004 Dec. p. 569-72.
- 20. Gahi Y, Lamrani M, Zoglat A, Guennoun M, Kapralos B, El-Khatib K. Biometric identification system based on Electrocardiogram data. IEEE New Technologies, Mobility and Security, NTMS'08; 2008 Nov. p. 1-5.
- 21. Sriram JC, Shin M, Choudhury T, Kotz D. Activity-aware ECG-based patient authentication for remote health monitoring. Proceedings of the 2009 International Conference on Multimodal Interfaces, ACM; 2009 Nov; p. 297-304.
- 22. Abdelraheem M, Selim H, Abdelhamid TK. Human identification using the main loop of the vector cardiogram. Am J Signal Process. 2012; 2:23-9.
- 23. Tantawi M, Revett K, Tolba MF, Salem A. A novel feature set for deployment in ECG based biometrics. 2012 IEEE Seventh International Conference on Computer Engineering and Systems (ICCES); 2012 Nov. p. 186-91.
- 24. Ogunbona PO, Milliss M, De Bore F, Fernandes M. Classification of gas metal arc welds using wavelets. Engineering Mathematics and Applications Conference; Adelaide, Australia. 1998. p. 383-6.
- 25. Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. IEEE Engineering in Medicine and Biology Magazine. 2001; 20(3):45-50.