REAL-TIME CONTINUOUS IDENTIFICATION SYSTEM USING ECG SIGNALS

Rafik Matta, Johny K. H. Lau, Foteini Agrafioti and Dimitrios Hatzinakos

The Edward S. Rogers Department of Electrical and Computer Engineering,
University of Toronto
10 Kings College Road, Toronto, ON, Canada, M5S 3G4

ABSTRACT

This paper presents a prototype biometric system that performs real time human identification using the ECG signal. Cardiac signals have been lately suggested for biometric deployment, since they carry subject specific information and at the same time allow for inherent liveness detection. The HeartID is the first system to continuously authenticate individuals who are conveniently wearing a portable, wireless ECG sensor. The AC/LDA algorithm is used for biometric template design because of its computational and performance advantages. The performance of the prototype testbed was evaluated over 10 subjects at the University of Toronto with very promising results.

Index Terms— electrocardiogram, autocorrelation, linear discriminant analysis

1. INTRODUCTION

Biometrics are physiological or behavioral characteristics of the human body that can be used to establish people's identity. Examples include the fingerprint, the face or the iris. Although the related technology is very advanced and these modalities are widely deployed in security systems, there are concerns with regard to replay or falsification attacks, which are increasingly threatening these systems.

Lately, attention has been drawn to a new set of modalities, the *medical biometrics* [1, 2, 3, 4, 5, 6, 7]. Medical biometrics are comprised of signals that are traditionally used by physicians for diagnostics such as the electrocardiogram (ECG), electroencephalogram (EEG), phonocardiogram (PPG) etc.

Although, signal processing of such signals is very advanced, since they have been thoroughly studied for diagnostics, they have only lately been suggested for biometrics. This work emphasizes the use of ECG for human recognition, however without limiting the applicability to this signal.

The reasons why ECG is herein considered for human recognition can be summarized as: *universality, uniqueness, permanence* and *robustness to attacks*. In order for a feature to be eligible for biometric recognition, it has to be universal i.e., collectable from all human subjects. This criterion is naturally satisfied by ECG since it is a vital signal.

Uniqueness is also a necessary condition for biometric characteristics, since this property guarantees subject discrimination in a population. In the ECG case, the inter-subject variability has been long known in the medical community, where the objective was to eliminate it for the establishment of universal diagnostic standards [8]. ECG depicts the cardiac electrical activity, and is directly linked to the geometrical and electrophysiological properties of the myocardium [9, 10]. The fact that the cardiac muscle anatomically differs from person to person is depicted on the signal, and supports its employment in biometric recognition.

Permanence is also an important factor for biometric characteristics, as it guarantees that a biometric signature which is designed to today will not change dramatically with time. ECG is affected by both physical and psychological activity, and this is one the major challenges in its biometric deployment. However, despite time dependency there are aspects of the signal which remain unaltered. Furthermore, ECG presents a natural shield to attacks because it provides inherent *liveness* detection i.e., any security system using the ECG signal to recognize individuals needs no extra computation to assess the originality of the reading. In addition, it is very difficult to steal or mimic someone else's signal as it is the combination of several sympathetic and parasympathetic factors of the human body.

Given the above advantages, ECG has increasingly gained acceptance from the biometrics community. Several methodologies have been proposed for feature extraction that can be roughly categorized as *fiducial points dependent* or *independent*. Fiducial points are interesting events on a heart beat such as the onsets and offsets of waves that comprise the pulse. Analogous to face recognition, where one can use for biometric classification local characteristics, like the distance between the two eyes, in ECG biometrics such features correspond to temporal or amplitude differences between interesting fiducial points. On the other hand, fiducial independent approaches treat the signal holistically and present a statistical approach to feature extraction from high dimensional data. Both solutions have advantages and disadvantages, for instance fiducial points detection may increase significantly the computational effort of the system, but guarantees to capture the finest differences of two ECG signals.

The computational load is of great interest in ECG biometrics. This is because this signal can potentially offer real time and continuous identity authentication. As opposed to static face or fingerprint images, in the ECG case, a fresh reading can be captured every couple of seconds and used for a new matching. Along these lines, this work presents a prototype ECG based identification system, the *HeartID*, with the goal of demonstrating the potential of using this biosignal in applications of high security requirements. The unobtrusive, accurate and continuous fashion by which recognition is carried out by HeartID is expected to establish the ECG in the biometric world.

2. THE HEARTID SYSTEM

This section presents the HeartID system from an algorithmic and an implementation point of view. Recognition is based on the Autocorrelation / Linear Discriminant Analysis (AC/LDA) algorithm. The AC/LDA was chosen for HeartID first, because of its performance benefits [3, 11] and second, because it offers a fiducial independent solution, which is central to the deployment of real time authentication system (fiducial points detection risks both performance and complexity).

2.1. The AC/LDA Algorithm

HeartID designs ECG biometric signatures using the AC/LDA algorithm described in [3]. The algorithm has three main steps: 1) filtering, 2) AC computation 3) LDA training. Filtering is necessary because ECG signals are affected by both high and low frequency noise. In HeartID, a butterworth bandpass filter of order 4 is used, based on empirical results. The next step is the estimation of the normalized autocorrelation according to:

$$\widehat{R}_{xx}[m] = \frac{\sum_{i=0}^{N-|m|-1} x[i]x[i+m]}{\widehat{R}_{xx}[0]} \tag{1}$$

where $x[i]$ is the windowed ECG for $i = 0, 1...(N - |m| - 1)$, $x[i + m]$ is the time shifted version of the windowed ECG with a time lag of $m = 0, 1, ...(M - 1)$; $M << N$, and N is the length of the signal. The LDA then operates on this signal for dimensionality reduction. The purpose of the LDA is to retain only those features of the $\widehat{R}_{xx}[m]$ that reduce the intra-subject variability while at the same time increase the inter-subject one i.e., alllow for better discrimination in a population.

A challenge with the AC/LDA algorithm is that LDA performs supervised learning. This limits the applicability in real life systems where the population needs to be known before the system can be used for identification. To address this problem, a pre-enrolled database of ECG signals was added to the system, so that the recognizer can learn as much of the ECG variability as possible.

At start-up the system is exposed to a large variety of ECG signals that it can be trained on to differentiate from new enrollees. This means that every time a new subject is enrolled HeartID will retrain itself and learn not only the morphologies of the enrollees that the system can identify, but also morphologies of a generic population as well.

2.2. The HeartID Testbed

The HeartID testbed was built using the CSharp programming language. The setup involves the Equivital Vital Signs Monitor, developed by Hidalgo Limited[1]. Since remote and continuous authentication is the goal of HeartID, the particular ECG sensor was chosen over others because it allows wireless transmission of the signal. A belt is conveniently worn at the chest area, in a configuration that allows two lead ECG recordings.

The device communicates with the test PC using the RF-COMM protocol over Bluetooth. Using a virtual serial port, it transmits a number of vital signals such as the ECG, the heart rate, respiration rate, body position and other. For the purpose of biometric recognition, only one lead ECG was retained for identification. At the receiver, an Intel Core 2 Duo T7300, with 2GB of RAM was used for processing of the signal and design of the biometric signature.

The system collects ECG continuously and assesses identity information every 5 seconds. There are two modes of operation in HeartID:

1. **Enrollment**. When in this mode the system *learns* the ECG morphology of a new user, extracts a biometric template according to the AC/LDA algorithm and saves it in a database along with the credentials of the particular enrollee (for example name, age, gender etc).

2. **Identification**. During the identification mode of operation the system processes continuously input ECG readings, extracts features and performs *one-to-many* matches with the database of enrollees in order to find the best match i.e., the identity of the user.

A snapshot of graphical user interface that encompasses both modes of operation is shown in Figure 1.

3. EXPERIMENTAL PERFORMANCE

The performance of HeartID was evaluated at the Biometrics Security Laboratory at the University of Toronto. 10 volunteers participated in an identification experiment after completing and signing consent forms. Every volunteer wore the ECG sensor belt, and after establishing connection with the device, identification was performed for 5 minutes. During this time, no special instructions were given to the volunteers to allow for varying mental states that would affect the heart

[1]http://www.equivital.co.uk/

Fig. 1. Snapshot of the HeartID system

rate variability. The volunteers were free to walk, exercise, or engage in conversations. With this treatment identification under real life conditions was simulated.

For every 5 sec ECG segment that was collected, HeartID outputted ranked results. More specific, for every testing ECG window the AC/LDA algorithm was used for feature extraction. The acquired feature vector was compared against the pre-enrolled ECG templates, and a vector of the respective Euclidean distances was created. This vector was consecutively sorted in ascending order and the first three identities (three best matches) were displayed to the operator.

Table 3 presents the contingency matrix of the identification experiment. Every row of this table corresponds to a testing subject and every column to a gallery (enrolled) subject. Each entry (i, j) shows the number of ECG windows of subject i that were classified as j. In a perfect system it is expected that the diagonal will show 100% for all subjects. However, this is not usually with the case with biometric systems, where the employed modalities are stochastic and thus intra-subject variability is expected. Nevertheless, for all testing subjects the majority of their ECG readings was classified as the correct identity.

Ranked results are also of great interest for ECG biometrics, as they complete the performance of the recognizer. The better the discriminative power of the selected feature space, the lower the rank of the true class in the output. In addition, one can benefit from ranked results in understanding how the system would perform under the *verification* mode of operation. Figure 2 shows the percentage of windows (across all testing subjects) for which the true class was classified first, second or third. It is clear for this Figure that even though there are cases of misidentification i.e., cases where the true

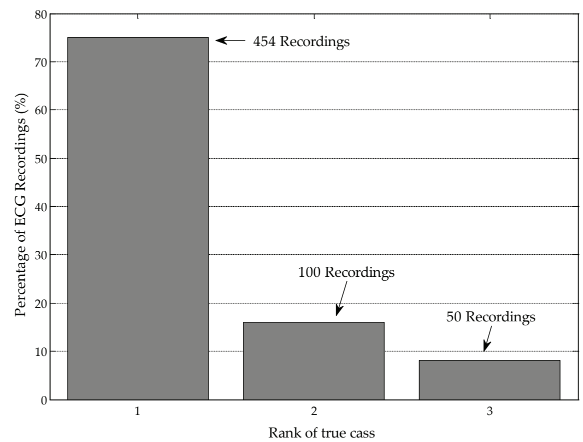

Fig. 2. Identification performance based on rank of the true class.

class wasn't ranked first, there is high probability of that being within the best three matches.

A way to increase the accuracy of the recognizer would be by allowing a decision to be made with longer (than 5 sec) ECG windows. For instance, identification upon a 15 sec segment would allow the system to vote upon three windows thereby increasing the chances of correct identification.

4. CONCLUSION

This work presented a prototype ECG biometric recognition system, the HeartID. The essence of this system is that it allows for continuous, real time, unobtrusive identification of human subjects using cardiac signals. In a monitoring setting, ECG can be used to authenticate identities continuously, each time using a fresh reading that can be acquired in as fast

	Subject 1	Subject 2	Subject 3	Subject 4	Subject 5	Subject 6	Subject 7	Subject 8	Subject 9	Subject 10
Subject 1	**59**	0	0	0	0	0	0	0	1	7
Subject 2	28	**34**	0	0	0	0	0	0	1	4
Subject 3	2	2	**57**	0	0	0	0	0	5	6
Subject 4	24	0	0	**45**	0	0	0	0	0	0
Subject 5	8	0	1	10	**33**	0	0	0	0	1
Subject 6	1	0	0	7	1	**62**	0	0	0	0
Subject 7	3	2	1	1	17	0	**40**	0	0	6
Subject 8	5	2	7	0	5	0	3	**48**	2	1
Subject 9	6	1	3	0	2	12	0	1	**46**	1
Subject 10	3	15	6	0	2	1	0	2	9	**32**

Table 1. Contingency table of the overall classification performance. Entry (i, j) shows the number of ECG windows of subject i that were classified as subject j.

as 5 seconds.

The prototype system relies on the AC/LDA algorithm to perform feature extraction and matching. The biometric signature is composed of selected features of the autocorrelated ECG. Selection takes place on the ground of increasing the within subject variability, while decreasing the between subject one. The system is equipped with a number of previously enrolled signals that participate in the LDA training along with the enrollment ECGs. By including more than the enrollee's signals in the training of the algorithm, it is expected that every projected ECG onto the LDA space will capture as much of the signal's variability as possible.

The HeartID was tested on 10 subjects using 5 minute recordings. Mental state variability was allowed during the experiment. The system provided an identity estimate continuously for every 5 seconds of the testing. The performance of the prototype recognizer is very promising however the results need to be validated with large scale testings.

5. ACKNOWLEDGEMENTS

This work has been supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

6. REFERENCES

[1] 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.

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

[3] F. Agrafioti, F.M. Bui, and D. Hatzinakos, "Medical biometrics: The perils of ignoring time dependency," in *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems, 2009. BTAS '09*, Washington, DC, USA, sept. 2009, pp. 1 –6.

[4] S. Fatemian, F. Agrafioti, and D. Hatzinakos, "Heartid: Cardiac biometric recognition," in *IEEE 4th International Conference on Biometrics: Theory, Applications, and Systems, 2010. BTAS '10*, Washington, DC, USA, sept. 2010, pp. 27 –29.

[5] G. Wübbeler, M. Stavridis, D. Kreiseler, R.D. Bousseljot, and C. Elster, "Verification of humans using the electrocardiogram," *Pattern Recogn. Lett.*, vol. 28, no. 10, pp. 1172–1175, 2007.

[6] Yongjin Wang, Foteini Agrafioti, Dimitrios Hatzinakos, and Konstantinos N. Plataniotis, "Analysis of human electrocardiogram for biometric recognition," *EURASIP J. Adv. Signal Process*, vol. 2008, no. 1, pp. 1–6, 2008.

[7] T. W. Shen, W. J. Tompkins, and Y. H. Hu, "One-lead ECG for identity verification," in *Proc. of the 2nd Conf. of the IEEE Eng. in Med. and Bio. Society and the Biomed. Eng. Society*, 2002, vol. 1, pp. 62–63.

[8] H. Draper, C. Peffer, F. Stallmann, D. Littmann, and H. Pipberger, "The corrected orthogonal electrocardiogram and vectorcardiogram in 510 normal men (frank lead system)," *Circulation*, vol. 30, pp. 853–864, 1964.

[9] R. Hoekema, G.Uijen, and A. van Oosterom, "Geometrical aspect of the interindividual variaility of multilead ECG recordings," *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 551–559, 2001.

[10] G. Kozmann, R.L. Lux, and L.S. Green, "Geometrical factors affecting the interindividual variability of the ECG and the VCG," *J. Electrocardiology*, vol. 33, pp. 219–227, 2000.

[11] F. Agrafioti, F.M. Bui, and D. Hatzinakos, "On supporting anonymity in a BAN biometric framework," in *16th International Conference on Digital Signal Processing*, Santorini, Greece, july 2009, pp. 1 –6.