ECG Signals for Biometric Applications Are we there yet?

Carlos Carreiras¹, André Lourenço^{1,2}, Ana Fred¹ and Rui Ferreira³

¹Instituto de Telecomunicações, Av. Rovisco Pais 1, Lisboa, Portugal
 ²Instituto Superior de Engenharia de Lisboa, R. Cons. Emídio Navarro 1, Lisboa, Portugal
 ³Hospital de Santa Marta, R. de Santa Marta 50, Lisboa, Portugal

Keywords: Biometrics, Person Recognition, ECG, Classification.

Abstract:

The potential of the electrocardiographic (ECG) signal as a biometric trait has been ascertained in the literature over the past decade. The inherent characteristics of the ECG make it an interesting biometric modality, given its universality, intrinsic aliveness detection, continuous availability, and inbuilt hidden nature. These properties enable the development of novel applications, where non-intrusive and continuous authentication are critical factors. Examples include, among others, electronic trading platforms, the gaming industry, and the auto industry, in particular for car sharing programs and fleet management solutions. However, there are still some challenges to overcome in order to make the ECG a widely accepted biometric. In particular, the questions of uniqueness (inter-subject variability) and permanence over time (intra-subject variability) are still largely unanswered. In this paper we focus on the uniqueness question, presenting a preliminary study of our biometric recognition system, testing it on a database encompassing 618 subjects. We also performed tests with subsets of this population. The results reinforce that the ECG is a viable trait for biometrics, having obtained an Equal Error Rate of 9.01% and an Error of Identification of 15.64% for the entire test population.

1 INTRODUCTION

Over the past decade, numerous groups have demonstrated the potential of electrocardiographic (ECG) signals for identity recognition applications (Biel et al., 2001; Kyoso and Uchiyama, 2001; Silva et al., 2007a; Wang et al., 2008). Due to its inherent characteristics, the ECG signal is emerging as an interesting biometric trait, given that, following the properties defined in (Jain et al., 1999), it can be found in all living humans (Universality), it has been shown to perform accurately for subsets of the population (Performance), and it can be easily obtained using appropriate devices (Measurability). These sensors can be designed in a non-intrusive way (Acceptability), in particular when using an Off-the-Person approach (Silva et al., 2013a). Furthermore, the ECG is not easily spoofed (Circumvention), as it does not depend on any external body traits, provides intrinsic aliveness detection, and is continuously available.

These properties of the ECG signal enable the development of novel and interesting applications, where non-intrusive and continuous authentication are critical factors. Examples of such applications include electronic trading platforms, where highsecurity, continuous authentication is essential, in the gaming industry, where the ECG sensor could be integrated into the game controller itself to identify the players in a multi-player scenario, and in the auto industry, particularly for car sharing programs and fleet management solutions.

At the moment, the biggest challenges faced by the ECG as a biometric trait relate to its Permanence and Uniqueness, and the question remains if this modality is ready for real-world applications. While Permanence deals with the question of temporal invariance of the templates, that is, intra-subject variability, Uniqueness pertains to the discernibility of the templates from different subjects, that is, intersubject variability. Studies on the permanence question can already be found in the literature, for instance in (Agrafioti et al., 2011; Silva et al., 2013c). In this paper, we present a preliminary study on the uniqueness question. We accomplish this by testing our recognition system on an ECG signal database with 618 subjects, the biggest to date to be used for biometrics, to the best of our knowledge. We also performed tests with subsets of this population, assessing

ECG Signals for Biometric Applications - Are we there yet?.

DOI: 10.5220/0005160507650772

In Proceedings of the 11th International Conference on Informatics in Control, Automation and Robotics (IVC&ITS-2014), pages 765-772 ISBN: 978-989-758-040-6

the behavior of the recognition system with a varying number of subjects.

The remainder of this paper is organized as follows: Section 2 provides an overview of the characteristics of the ECG signal and its use in biometric systems; Section 3 describes the methodology used for the biometric recognition system, including a description of the database used, feature extraction, and classification approaches; Section 4 summarizes the obtained experimental results; and Section 5 outlines the main conclusions.

2 BACKGROUND

It is widely known that the basic function of the heart is to pump blood throughout the body, demanding a highly synchronized sequence of muscular contractions. These are initiated by small electrical currents that propagate through the myocardium's cells, originating an electrical signal that can be recorded at the body surface (the ECG). These potentials can be measured by placing two electrodes on the body's surface, determining the voltage difference between them (Webster, 2009). Different electrode placements produce different perspectives of the heart, termed leads or derivations, given the spatial characteristics of the heart's electrical field and how it propagates throughout the body (Neuman, 1998).

The ECG is a semi-periodic signal, with each cycle being characterized by the typical P-QRS-T heartbeat waveform. The signal as a whole has a rich information content, being a wellbeing and health indicator, and is related with the psychophysiological state (Carreiras et al., 2013b). In order to have a coherent clinical diagnostic tool, the lead placement has been standardized. Much of the standard system is based on Einthoven's groundbreaking work, with the use of the three limb leads (I, II, and III), as the limbs are easily identified anatomical references (Webster, 2009). Additionally, the augmented leads (aVF, aVL, aVR) and the six precordial leads (V1 - V6) are also typically recorded in clinical settings.

In the context of ECG biometrics, current approaches found in the literature can be classified as either fiducial or non-fiducial (Agrafioti et al., 2011; Odinaka et al., 2012; Silva et al., 2013b). The former describes methods based on reference points in the signals (Israel et al., 2005; Shen et al., 2002; Silva et al., 2007b; Oliveira and Fred, 2009), while the latter methods rely on intrinsic information within the ECG signal, without having any particular cues as reference (Chan et al., 2008; Chiu et al., 2008; Wang et al., 2008; Coutinho et al., 2013). Partially fidu-



Figure 1: Examples of (a) of a ECG trace (Subject A), and (b) the heartbeat waveform templates for two distinct individuals.

cial methods, like our approach presented below, rely on fiducial information only for ECG segmentation (Wang et al., 2008; Lourenço et al., 2011; Carreiras et al., 2013a; Silva et al., 2013c). We refer the reader to (Agrafioti et al., 2011; Odinaka et al., 2012; Silva et al., 2013b) and references therein for a comprehensive literature review. Figure 1 shows an example of an ECG signal trace, and the segmented heartbeat templates for two distinct subjects, where the differences between them are apparent.

One significant contribution to the usefulness and acceptability of the ECG as a biometric trait is the use of an *Off-the-Person* approach for signal acquisition (Silva et al., 2013a). In this approach, only one ECG lead is used, with the signal being acquired at the hand palms or fingers, using just two (non-gelled) contact points, as opposed to multiple contact points throughout the body using gelled electrodes. The lead placement in this case is non-standard, however it has been shown to be highly correlated with the standard lead I (Carreiras et al., 2013c). Various research groups have used this approach (Chan et al., 2008; Lourenço et al., 2011). However, the signals obtained with this setup are harder to analyze, as they are more susceptible to noise artifacts due to unstable electrode-skin

· · •		· •		
Study	Sample Size	ECG Lead	AP (%)	IP (%)
		Ι	NA	85.3
(Zhang and Wei, 2006)	502	II	NA	92.0
		V1	NA	95.2
		V2	NA	97.4
(Odinaka et al., 2010)	269	Electrodes placed on lower ribcage	0.37 (EER)	99
(Shen et al., 2010)	168	I (hands)	NA	95.3
(Safie et al., 2011)	112		94.54 (AUR)	NA
(Irvine et al., 2003)	104	NA	NA	91

Table 1: Five largest ECG biometrics studies found in (Odinaka et al., 2012), with the reported authentication (AP) and identification (IP) performance; AUR: Area Under ROC curve; EER: Equal Error Rate; NA: Not Available.

contact and electromyographic (EMG) activity.

The *Off-the-Person* approach enables the seamless integration of the ECG sensor into everyday objects. One such example, as shown in Figure 2, is the integration of the ECG sensor into the steering wheel of a car using conductive textiles. In this car sharing demonstrator, the user, in order to authenticate on the system, touches the contactless member card on a reader to provide the assumed identity. This identity is then validated through the ECG signal by simply placing the hands on the steering wheel, as in a normal driving situation. On successful authentication, various user-specific configurations could be loaded, such as preferred radio stations, mirror positions, and address lists, among others.



Figure 2: Integration of an *Off-the-person* ECG sensor into the steering wheel of a car; the electrodes are highlighted in red.

Regarding the uniqueness problem, there are currently no studies assessing the performance of ECG biometric systems encompassing very large data sets, such as the work done for iris recognition by Daugman, encompassing more than 600 000 different iris patterns (Daugman, 2006). Using the review by Odinaka *et al.* (Odinaka et al., 2012) as source, ECG biometric studies use, on average, databases of about 50 subjects. Table 1 provides a list of the five largest studies, with the reported authentication and identification performances. Unfortunately, for various reasons mainly related to privacy concerns, many studies use in-house databases, which are not publicly available. Additionally, most public ECG databases, notably the ones available on Physionet (Goldberger et al., e 13), were built for research on pathophysiology, not biometrics, with most of the records having some kind of heart pathology. For this reason, as is described in the next section, we were forced to obtain our own database, in order to test our recognition methodology with a larger number of subjects.

3 METHODOLOGY

3.1 Database

Our research group entered into a collaboration with a local hospital (Hospital de Santa Marta) specialized in cardiac issues, with the goal of obtaining a large ECG database. The records thus obtained are acquired during normal hospital operation, encompassing scheduled appointments, emergency cases, and bedridden patients. Therefore, most of the records represent pathological cases.

The signals were acquired using Philips PageWriter Trim III devices, following the traditional 12 lead placement, with a sampling rate of 500 Hz and 16 bit resolution. Each record has a duration of 10 seconds. To date, we have received, over a period of 10 months, 4 332 records from 2 055 distinct subjects, whose true identities are obfuscated at the hospital.



(a) Rate of normal *vs* pathological subjects and records





(c) Age distribution

Figure 3: Population statistics of the database, for a total of 4 332 records and 2 055 subjects, with (a) the rate of normal subjects and records, subject (b) gender, and (c) age distributions; the whiskers in the boxplots extend to the lowest and highest data points still within 1.5 times the interquartile range.

As a first step, for this paper we decided to focus only on the healthy individuals. Consequently, each record had to be labeled by a specialist either as normal or pathological. Of all the records, 832 were deemed normal, corresponding to 618 subjects. This figure surpasses by about 100 the largest ECG biometrics study currently found in the literature. Figure 3 summarizes the relevant population statistics.

Note that, although our target applications follow

the *Off-the-Person* approach, such a large database takes a lot of time and effort to obtain, requiring clearance by an ethics committee, finding volunteers, signing of informed consent forms, among others. Nevertheless, if we cannot demonstrate the potential, in regards to the uniqueness question, of the ECG as a biometric in higher quality signals, then certainly that is not possible with hand signals.

3.2 ECG Biometric System

The typical block diagram of a fiducial, or partially fiducial, biometric system is depicted in Figure 4. These systems rely on the detection of notable ECG complexes for segmentation and extraction of a sequence of individual heartbeats. Typically, the QRS complex is used for that purpose.



Figure 4: Block diagram of a typical ECG biometric system.

Our ECG biometric system, designed with hand ECG signals in mind, starts with the acquisition of raw data, in this case the lead I ECG signal. The acquired signal is then submitted to a data preprocessing block, which performs a digital filtering step (band pass FIR filter order 150, and cutoff frequencies [5;20] Hz) and the QRS complex detection (Lourenço et al., 2012). The outputs of this block are segmented individual heartbeats, and a RR interval time series.

Given that segmentation algorithms are not perfect, especially for noisy signals like the ones obtained from the hands, we implement an outlier detection block, which performs detection and removal of anomalous ECG heartbeats. We follow the DMEAN approach described in (Lourenço et al., 2013), which computes the distance of all templates in a recording session to the mean template for that session, with templates being considered outliers if the computed distance is higher than an adaptive threshold.

The pattern extraction block takes the preprocessed input signals, and starts by aligning all the heartbeat waveforms by their R-peak instants, and by clipping them in the interval [-200;400]ms around that instant. In the scope of this work, we consider the features to be all the amplitudes within this interval.

Finally, a k-NN classifier (with k=3) is used together with the cosine distance metric, to produce a decision on the recognition of the individual (either in authentication or identification), as it was found to be a good compromise between performance and computational cost (Silva et al., 2012). Altogether, our biometric system is fairly simple, being computationally light and opening the possibility of integrating it into embedded systems, which have limited processing power.

-INI

4 **RESULTS**

SCIENCE

We evaluated the performance of the biometric system for both the identification and authentication scenarios. For the identification scenario, we computed the Error of Identification (EID), which corresponds to the number of incorrect identifications normalized by the total number of tests. For the authentication scenario, we computed, for each operating point of the classifier (each distance threshold), the False Acceptance Rate (*FAR*), the False Rejection Rate (*FRR*), the True Acceptance Rate (*TAR*), and the True Rejection Rate (*TRR*), given by

$$FAR = \frac{FP}{TN+FP}, \qquad FRR = \frac{FN}{TP+FN},$$

$$TAR = \frac{TP}{TP+FP}, \qquad TRR = \frac{TN}{TN+FP},$$
(1)

where TP and TN are the number of true positives and negatives, and FP and FN are the number of false positives and negatives, respectively. From these rates, we estimate the Equal Error Rate (EER), which corresponds to the operating point for which the *FAR* is equal to the *FRR*, using piecewise polynomial interpolation.

Furthermore, we used a leave-one-out (LOO) approach for cross validation (Efron, 1983), given the fact that the number of templates for some subjects was low (minimum of 4 templates), enabling us to maximize the number of templates to train the classifier, which requires at least 3 templates (3-NN). In

order to to this, we selected a random group of 4 templates for each subject, which are then partitioned with the LOO method. We repeated this procedure 10 times, computing the average authentication and identification performance across all runs.

Additionally, we assessed the behavior of the system with subsets of the population, encompassing 5, 10, 20, 30, 40, and 50 subjects. These subsets correspond to our targeted applications, ranging from a small group (*e.g.* in a multiplayer game setting, or a family sharing a car) to small businesses (*e.g.* a local distribution company). The subjects in each subgroup were randomly selected from the initial population, repeating this process 150 times, each run following the cross validation method described above.

The results obtained for the entire population (P618) are presented in Table 2, comparing them to a previous baseline experiment performed using a smaller database (63 subjects), which uses signals obtained at the hands, making obvious the costs in performance resulting from the use of hand signals. Regarding the EID, the value obtained is on par with the results presented in (Zhang and Wei, 2006) for lead I signals (see Table 1), with the added bonus of using a larger database.

Table 2: EER and EID obtained for the entire test population (P618) and the baseline experiment (63 subjects, hand ECG).

Case	EER (%)	EID (%)
P618	9.01	15.64
Baseline	13.26	36.40

Figure 5 shows the evolution of the *FAR* and *FRR* with the authentication distance threshold, as well as the Receiver Operating Characteristic (ROC) curve, which plots the the *TAR* against the *FAR*, highlighting an Area Under ROC (AUR) curve of 95.51%, similar to the one obtained in (Safie et al., 2011). Also of note in Figure 5(a) is the fact that the *FAR* increases more slowly than the *FRR* decreases with the threshold.

Results for the population subsets are presented in Figure 6. Figure 6(a) highlights the fact that the EER does not seem to be affected by the population size. On the other hand, Figure 6(b) shows that the EID increases with the increasing number of subjects.

5 CONCLUSIONS

Research to date has demonstrated that the ECG signal, due to its intrinsic nature, has the potential to complement existing person recognition approaches



Figure 5: Authentication results for the entire population, with (a) the EER determination, and (b) the ROC curve.

(a multibiometrics scenario), and, in some settings, to be used as a single modality.

However, the field is lacking a thorough examination of the limits of this modality in regards to the number of subjects, that is, we need to know if the information that we can extract from the ECG is sufficient to distinguish a large population. This paper is a contribution to that goal, assessing the performance of our ECG biometric system, which was designed for an *Off-the-Person* sensor approach, in a database with 618 subjects, examining as well the effect of the population size on the performance of the system, using subsets of the test population.

The results of our work indicate a performance of our system on par with similar studies found in the literature, with an Equal Error Rate of 9.01% and an Error of Identification of 15.64% for the entire test population. We also demonstrated that, while the authentication performance does not degrade with increasing number of subjects, the same does not happen

Figure 6: Results obtained for the population subsets in the (a) authentication, and the (b) identification scenarios.

with the identification scenario, where the error progressively increases with increasing number of subjects. Nevertheless, these results, together with the latest developments in recognition methods, template extraction and selection, and sensor devices, reinforce that the ECG is a viable trait for biometric applications.

Our future work will focus on the study of sources of intra-subject variability, in particular heart rate changes and morphological shape alterations due to pathological situations. Additionally, we will try to improve the representativeness of the test population in regards to age, and examine the performance of the system when using the other standard ECG leads, either independently or in combination (fusion of classifiers).

ACKNOWLEDGEMENTS

This work was partially funded by Fundação para a

Ciência e Tecnologia (FCT) under grants PTDC/EEI-SII/2312/2012, SFRH/BD/65248/2009 and SFRH/PR OTEC/49512/2009, whose support the authors gratefully acknowledge. We would also like to thank Joana Santos for her hard work labeling the ECG records.

REFERENCES

- Agrafioti, F., Gao, J., and Hatzinakos, D. (2011). *Biometrics*, chapter Heart Biometrics: Theory, Methods and Applications, Biometrics. InTech.
- Biel, L., Petterson, O., Phillipson, L., and Wide, P. (2001). ECG analysis: A new approach in human identification. *IEEE Trans. Instrumentation and Measurement*, 50(3):808–812.
- Carreiras, C., Lourenço, A., Silva, H., and Fred, A. L. N. (2013a). A unifying approach to ECG biometric recognition using the wavelet transform. In *Int. Conf.* on Image Analysis and Recognition, pages 53–62.
- Carreiras, C., Lourenço, A., Aidos, H., Silva, H., and Fred, A. (2013b). Morphological ecg analysis for attention detection. In *Int. Conf. on Neural Computation Theory* and Applications.
- Carreiras, C., Lourenço, A., Silva, H., and Fred, A. (2013c). Comparative study of medical-grade and offthe-person ecg systems. In Int. Congress on Cardiovascular Technologies.
- Chan, A. D. C., Hamdy, M. M., Badre, A., and Badee, V. (2008). Wavelet distance measure for person identification using electrocardiograms. *IEEE Trans. on Instrum. and Meas.*, 57(2):248–253.
- Chiu, C.-C., Chuang, C.-M., and Hsu, C.-Y. (2008). A novel personal identity verification approach using a discrete wavelet transform of the ECG signal. In *Proc. Int. Conf. on Multimedia and Ubiquitous Engineering*, pages 201–206. IEEE Computer Society.
- Coutinho, D., Silva, H., Gamboa, H., Fred, A., and Figueiredo, M. (2013). Novel fiducial and nonfiducial approaches to electrocardiogram-based biometric systems. *Biometrics*, *IET*, 2(2):64–75.
- Daugman, J. (2006). Probing the uniqueness and randomness of iriscodes: Results from 200 billion iris pair comparisons. *Proc. of the IEEE*, 94(11):1927–1935.
- Efron, B. (1983). Estimating the error rate of a prediction rule: Improvement on cross-validation. J. of the American Statistical Association, 78(382):316–331.
- Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., and Stanley, H. E. (2000 (June 13)). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220.
- Irvine, J., Israel, S., Wiederhold, M., and Wiederhold, B. (2003). A new biometric: human identification from circulatory function. In *Joint Statistical Meetings of* the American Statistical Association, San Francisco.

- Israel, S., Irvine, J., Cheng, A., Wiederhold, M., and Wiederhold, B. (2005). ECG to identify individuals. *Pattern Recognition*, 38(1):133–142.
- Jain, A. K., Bolle, R., and Pankanti, S. (1999). *Biometrics : personal identification in networked society*. Kluwer, 1st edition.
- Kyoso, M. and Uchiyama, A. (2001). Development of an ECG identification system. In *Proc. of the 23rd Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, volume 4, pages 3721–3723.
- Lourenço, A., Silva, H., Carreiras, C. C., and Fred, A. L. N. (2013). Outlier detection in non-intrusive ECG biometric system. In *Int. Conf. on Image Analysis and Recognition*, pages 43–52.
- Lourenço, A., Silva, H., and Fred, A. (2011). Unveiling the biometric potential of Finger-Based ECG signals. *Computational Intelligence and Neuroscience*.
- Lourenço, A., Silva, H., Leite, P., Lourenço, R., and Fred, A. (2012). Real time electrocardiogram segmentation for finger based ECG biometric. In *Proc. of the Int. Conf. on Bio-inspired Signals and Signal Processing* (*BIOSIGNALS*), pages 49–54.
- Neuman, M. R. (1998). Biopotential amplifiers. Medical instrumentation: application and design, pages 316– 318.
- Odinaka, I., Lai, P.-H., Kaplan, A., O'Sullivan, J., Sirevaag, E., Kristjansson, S., Sheffield, A., and Rohrbaugh, J. W. (2010). ECG biometrics: A robust short-time frequency analysis. In Proc. of the IEEE Int. Workshop on Information Forensics and Security (WIFS), pages 1–6.
- Odinaka, I., Lai, P.-H., Kaplan, A., O'Sullivan, J., Sirevaag, E., and Rohrbaugh, J. (2012). ECG biometric recognition: A comparative analysis. *IEEE Trans. on Information Forensics and Security*, 7(6):1812–1824.
- Oliveira, C. and Fred, A. L. N. (2009). ECG-based authentication: Bayesian vs. nearest neighbour classifiers. In *Proc. of the Int. Conf. on Bio-inspired Systems and Signal Processing (BIOSIGNALS)*, pages 163–168.
- Safie, S. I., Soraghan, J. J., and Petropoulakis, L. (2011). Electrocardiogram (ECG) biometric authentication using pulse active ratio (par). *IEEE Trans. on Information Forensics and Security*, 6(4):1315–1322.
- Shen, T. W., Tompkins, W. J., and Hu, Y. H. (2002). Onelead ECG for identity verification. Proc. of the Int. Conf. of the IEEE Engineering in Medicine and Biology Society, 1:62–63.
- Shen, T.-W. D., Tompkins, W. J., and Hu, Y. H. (2010). Implementation of a one-lead ECG human identification system on a normal population. *J. of Engineering and Computer Innovations*, 2(1):12–21.
- Silva, H., Carreiras, C., Lourenço, A., and Fred, A. L. N. (2013a). Off-the-person electrocardiography. In Int. Congress on Cardiovasular Technologies (CAR-DIOTECHNIX), pages 99–106.
- Silva, H., Gamboa, H., and Fred, A. (2007a). Applicability of lead v₂ ECG measurements in biometrics. In Proc. of the Int. eHealth, Telemedicine and Health ICT Forum (Med-e-Tel), pages 177–180.

y public

- Silva, H., Gamboa, H., and Fred, A. (2007b). One lead ECG based personal identification with feature subspace ensembles. In Proc. of the 5th Int. Conf. on Machine Learning and Data Mining in Pattern Recognition, pages 770–783, Berlin. Springer.
- Silva, H., Lourenço, A., Canento, F., Fred, A., and Raposo, N. (2013b). ECG biometrics: Principles and applications. In Proc. of the 6th Conf. on Bio-Inspired Systems and Signal Processing (BIOSIGNALS).
- Silva, H., Lourenço, A., and Fred, A. (2012). In-vehicle driver recognition based on hand ECG signals. In Proc. of the 2012 ACM Int. Conf. on Intelligent User Interfaces, IUI '12, New York, NY, USA. ACM.
- Silva, H. P., Fred, A., Lourenco, A., and Jain, A. K. (2013c). Finger ECG signal for user authentication: Usability and performance. In 2013 IEEE Sixth Int. Conf. on Biometrics: Theory, Applications and Systems (BTAS), pages 1–8. IEEE.
- Wang, Y., Agrafioti, F., Hatzinakos, D., and Plataniotis, K. N. (2008). Analysis of human electrocardiogram for biometric recognition. *EURASIP J. Adv. Signal Process*, 2008.
- Webster, J. G. (2009). *Medical Instrumentation Application and Design*. Wiley, 4th edition.
- Zhang, Z. and Wei, D. (2006). A new ECG identification method using bayes' teorem. In *TENCON 2006. 2006 IEEE Region 10 Conference*, pages 1–4. IEEE.