Evaluating Template Uniqueness in ECG Biometrics

Carlos Carreiras, André Lourenço, Hugo Silva, Ana Fred and Rui Ferreira

Abstract Research over the past decade has demonstrated the capability of the electrocardiographic (ECG) signal to be used as a biometric trait, through which the identity of an individual can be recognized. Given its universality, intrinsic aliveness detection, continuous availability, and inherent hidden nature, the ECG is an interesting biometric modality enabling the development of novel applications, where non-intrusive and continuous authentication are critical factors. Examples include personal computers, the gaming industry, and the auto industry, especially for car sharing programs and fleet management solutions. Nonetheless, from a theoretical point of view, there are still some challenges to overcome in bringing ECG biometrics to mass markets. In particular, the issues of uniqueness (related to inter-subject variability) and permanence (related to intra-subject variability) are still largely unanswered. This work focuses on the uniqueness issue, evaluating the performance of our ECG biometric system over a database encompassing 618 subjects. Additionally, we performed tests with subsets of this population. The results cement the ECG as a viable trait to be used for identity recognition, having obtained and Equal Error Rate of 9.01 % and an Error of Identification of 15.64 % for the entire test population.

C. Carreiras (✉) · A. Lourenço · H. Silva · A. Fred
Instituto de Telecomunicações, Instituto Superior Técnico, Av. Rovisco Pais 1,
1049-001 Lisbon, Portugal
e-mail: carlos.carreiras@lx.it.pt

H. Silva
e-mail: hugo.silva@lx.it.pt

A. Fred
e-mail: afred@lx.it.pt

A. Lourenço
Instituto Superior de Engenharia de Lisboa, R. Conselheiro Emídio Navarro 1,
1959-007 Lisbon, Portugal
e-mail: arlourenco@lx.it.pt

R. Ferreira
Hospital de Santa Marta, R. de Santa Marta 50, 1169-1024 Lisbon, Portugal
e-mail: cruzferreira@mail.telepac.pt

J. Filipe et al. (eds.), *Informatics in Control, Automation and Robotics*,
Lecture Notes in Electrical Engineering 370,
DOI 10.1007/978-3-319-26453-0_7

Keywords Biometrics · Person recognition · ECG · Classification

1 Introduction

The potential of electrocardiographic (ECG) signals for identity recognition applications has been demonstrated over the past decade by numerous research groups [2, 15, 28, 33]. Due to its inherent characteristics, the ECG signal is emerging as an interesting biometric trait, given that, following the properties defined in [14], 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 [26]. 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 high-security, continuous authentication is essential, 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 the automotive industry, particularly for car sharing programs and fleet management solutions.

At the moment, the biggest challenges faced by ECG biometrics 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, inter-subject variability. Studies on the permanence question can already be found in the literature, for instance in [1, 27]. In this paper, we present a study on the *Uniqueness* question. We accomplish this by testing our recognition system on an ECG signal database with 618 subjects. We also performed tests with subsets of this population, assessing the behavior of the recognition system with a varying number of subjects.

The remainder of this paper is organized as follows: Sect. 2 provides an overview of the characteristics of the ECG signal and its use in biometric systems; Sect. 3 describes the methodology used for the biometric recognition system, including a description of the database used, feature extraction, and classification approaches; Sect. 4 summarizes the obtained experimental results; and Sect. 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 [34]. 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 [19].

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 of the person as well [4]. 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 [34]. 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 [1, 21, 30]. The former describes methods based on reference points in the signals [13, 22, 24, 29], while the latter methods rely on intrinsic information within the ECG signal, without having any particular cues as reference [6–8, 33]. Partially fiducial methods, like our approach presented in the following sections, rely on fiducial information only for ECG segmentation [3, 16, 27, 33]. We refer the reader to [1, 21, 30] and references therein for a comprehensive literature review. Figure 1 shows an example of an ECG signal trace, and Fig. 2 shows the segmented heartbeat templates for two distinct subjects, where the differences between them are clearly noticeable.

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 [26]. 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 [5]. Various research groups have used this approach [6, 16]. However, the signals obtained with this setup are harder to analyze, as they are more susceptible to noise artifacts due to unstable electrode-skin 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 Fig. 3, 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, touched a

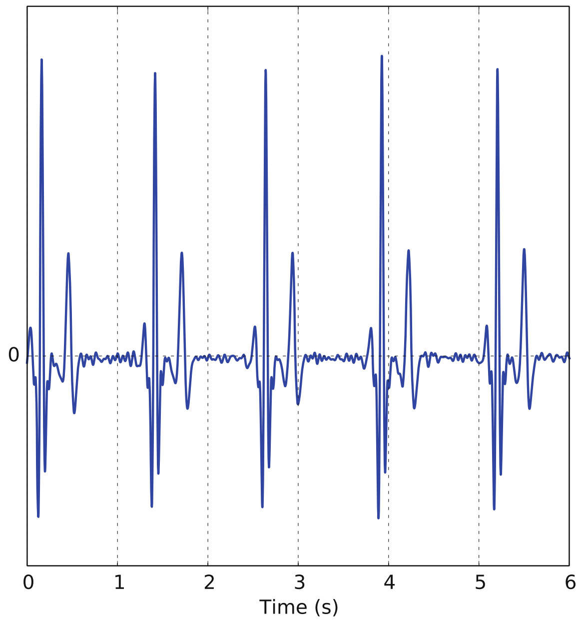


Fig. 1 Example of an ECG trace (Subject A in Fig. 2)

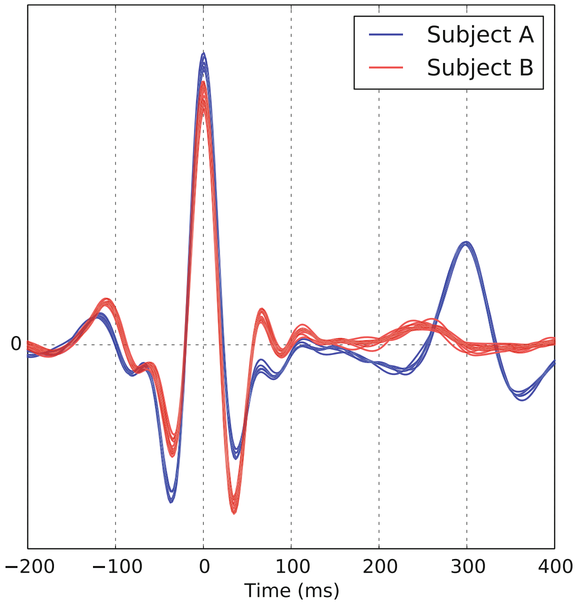


Fig. 2 Example of the heartbeat waveform templates from two distinct individuals

contacless membership card on a reader to provide an assumed identity, and perform an initial authentication on the car. This identity is then validated through the ECG signal by simply placing the hands on the steering wheel, as in a normal driving

Fig. 3 Integration of an *Off-the-Person* ECG sensor into the steering wheel of a car; the electrodes are highlighted in *red*

situation. Upon successful authentication, various user-specific configurations could be loaded, such as preferred radio stations, mirror positions, and address lists, among others. Additionally, as the system is continuously verifying the driver's identity, situations like carjacking can be prevented, as the system would lock out the intruder. Furthermore, the ECG signal acquired in this manner can also be used to monitor the health status of the driver, detecting the onset of sudden cardiac pathologies that could lead to road accidents.

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 [9]. Using the review by Odinaka et al. [21] as source, ECG biometric studies use, on average, databases of about 50 subjects. Table 1 provides

Table 1 Largest ECG biometrics studies found in the literature (adapted from [21]), with the reported authentication (AP) and identification (IP) performance

Study	Sample size	ECG Lead	AP (%)	IP (%)
Wahabi et al. [32]	1020	I (hands)	≈5 (EER)	NA
Zhang and Wei [35]	502	I	NA	85.3
		II	NA	92.0
		V1	NA	95.2
		V2	NA	97.4
Odinaka et al. [20]	269	Electrodes on lower ribcage	0.37 (EER)	99
Shen et al. [25]	168	I (hands)	NA	95.3
Safie et al. [23]	112	I	94.54 (AUR)	NA
Irvine et al. [12]	104	NA	NA	91

AUR Area Under ROC curve; *EER* Equal Error Rate; *NA* Not Available a list of the six 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 [11], 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 felt the need 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 obtained records were 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 s. 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.

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. Figure 4 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 Fig. 5. 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.

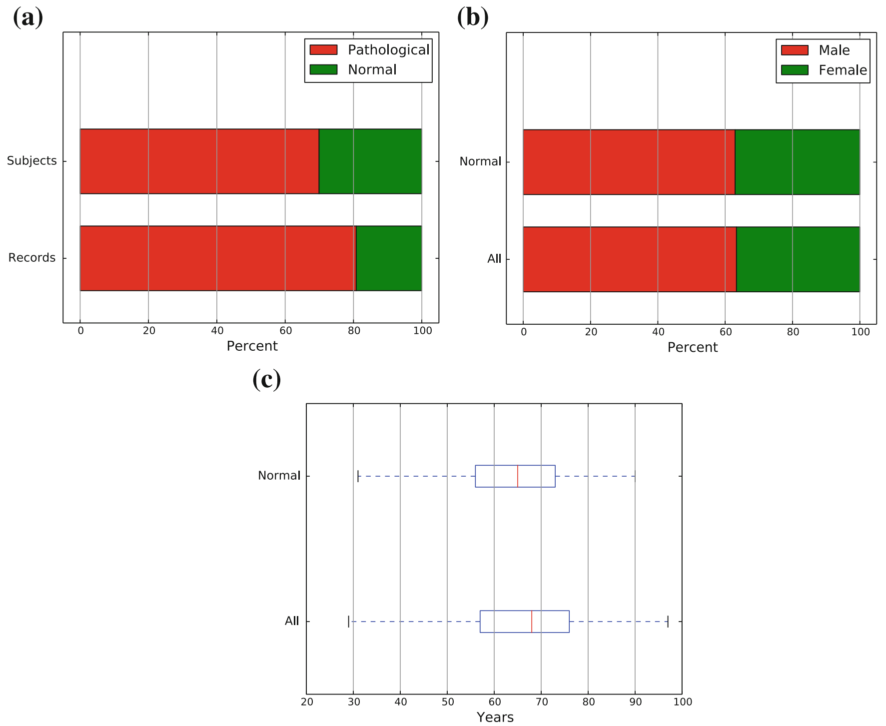


Fig. 4 Population statistics of the database, for a total of 4,332 records and 2,055 subjects, with **a** the rate of normal subjects and records, **b** subject gender, and **c** age distributions; the whiskers in the boxplots extend to the lowest and highest data points still within 1.5 times the interquartile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

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 of order 150, and cutoff frequencies $[5; 20]$ Hz) and the QRS complex detection [18]. The outputs of this block are the segmented individual heartbeats, and an 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 [17], 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.

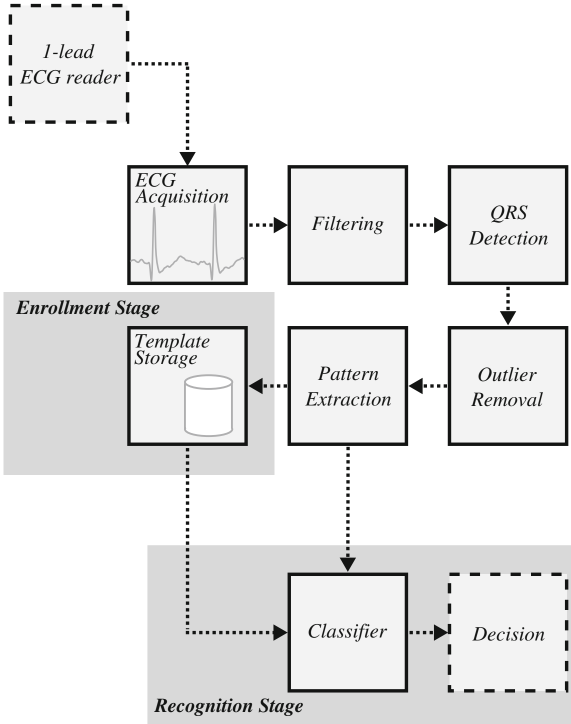


Fig. 5 Block diagram of a typical ECG biometric system

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 [31]. 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.

4 Results

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

$$\begin{aligned} FAR &= \tfrac{FP}{TN+FP}, \quad FRR = \tfrac{FN}{TP+FN}, \\ TAR &= \tfrac{TP}{TP+FP}, \quad TRR = \tfrac{TN}{TN+FP}, \end{aligned} \tag{1}$$

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 [10], 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 do 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 [35] for Lead I signals (see Table 1), with the added bonus of using a larger database.

Figure 6 plots the evolution of the FAR and FRR with the authentication distance threshold, highlighting the fact that the FAR increases more slowly than the FRR decreases with the threshold. Figure 7 shows the Receiver Operating Characteristic (ROC) curve, displaying an Area Under ROC curve (AUR) of 95.51 %, similar to the one obtained in [23].

Results for the population subsets are presented in Fig. 8, highlighting (Fig. 8a) the fact that the EER does not seem to be affected by the population size. On the other hand, Fig. 8b shows that the EID increases with the increasing number of subjects.

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

Fig. 6 Authentication FAR and FRR results for the entire population

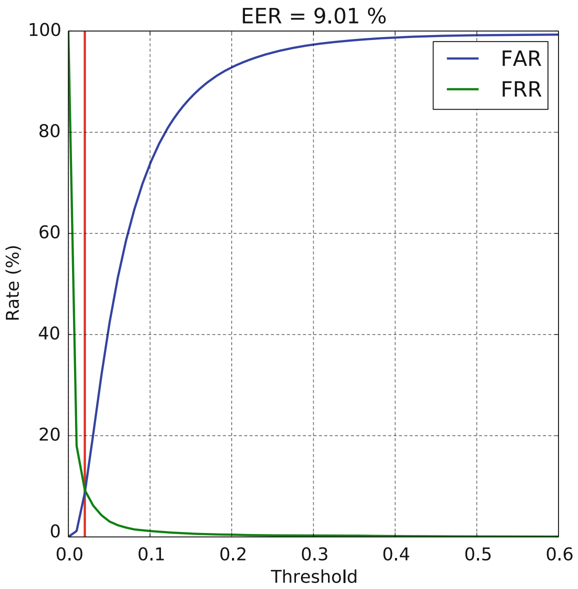
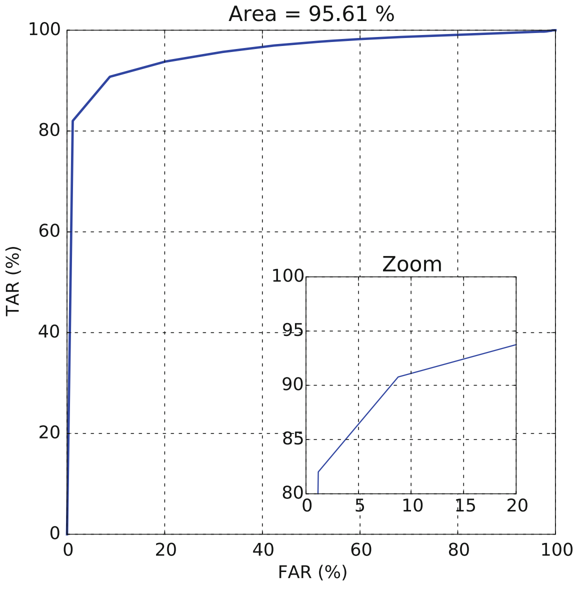


Fig. 7 Authentication ROC curve for the entire population



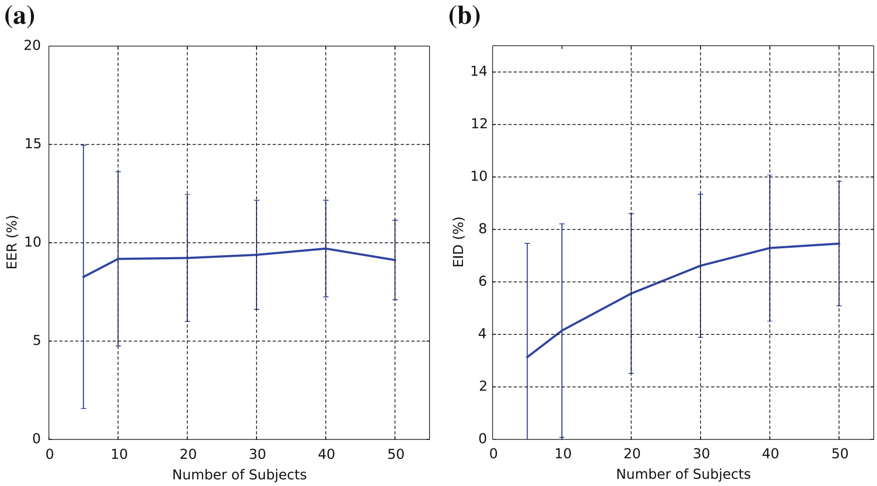


Fig. 8 Results obtained for the population subsets in the (**a**) authentication, and (**b**) identification scenarios; the *vertical bars* indicate the standard deviation

5 Conclusions

Research to date has demonstrated that the ECG signal, due to its intrinsic nature, has the potential to complement existing user recognition approaches (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 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).

Acknowledgments 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/PROTEC/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

1. Agrafioti, F., Gao, J., Hatzinakos, D.: Biometrics, chap. Theory, Methods and Applications, Biometrics. InTech, Heart Biometrics (2011)
2. Biel, L., Petterson, O., Phillipson, L., Wide, P.: ECG analysis: a new approach in human identification. IEEE Trans. Instrum. Measur. **50**(3), 808–812 (2001)
3. Carreiras, C., Lourenço, A., Silva, H., Fred, A.: A unifying approach to ECG biometric recognition using the wavelet transform. In: Proceedings of International Conference on Image Analysis and Recognition, pp. 53–62 (2013)
4. Carreiras, C., Lourenço, A., Aidos, H., Silva, H., Fred, A.: Morphological ECG analysis for attention detection. In: Proceedings of International Conference on Neural Computation Theory and Applications, pp. 381–390 (2013)
5. Carreiras, C., Lourenço, A., Silva, H., Fred, A.: Comparative study of medical-grade and off-the-person ECG systems. In: Proceedings of International Congress on Cardiovascular Technologies, pp. 115–120 (2013)
6. Chan, A., Hamdy, M., Badre, A., Badee, V.: Wavelet distance measure for person identification using electrocardiograms. IEEE Trans. Instrum. Meas. **57**(2), 248–253 (2008)
7. Chiu, C., Chuang, C., Hsu, C.: A novel personal identity verification approach using a discrete wavelet transform of the ECG signal. In: Proceedings of International Conference on Multimedia and Ubiquitous Engineering, IEEE Computer Society pp. 201–206 (2008)
8. Coutinho, D., Silva, H., Gamboa, H., Fred, A., Figueiredo, M.: Novel fiducial and non-fiducial approaches to electrocardiogram-based biometric systems. Biometrics IET **2**(2), 64–75 (2013)
9. Daugman, J.: Probing the uniqueness and randomness of iriscodes: results from 200 billion iris pair comparisons. Proc. IEEE **94**(11), 1927–1935 (2006)
10. Efron, B.: Estimating the error rate of a prediction rule: improvement on cross-validation. J. Am. Stat. Assoc. **78**(382), 316–331 (1983)
11. Goldberger, A., Amaral, L., Glass, L., Hausdorff, J., Ivanov, P., Mark, R., Mietus, J., Moody, G., Peng, C., Stanley, H.: PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. Circulation **101**(23), e215–e220 (2000)
12. Irvine, J., Israel, S., Wiederhold, M., Wiederhold, B.: A new biometric: human identification from circulatory function. In: Proceedings of Joint Statistical Meetings of the American Statistical Association, San Francisco (2003)
13. Israel, S., Irvine, J., Cheng, A., Wiederhold, M., Wiederhold, B.: ECG to identify individuals. Pattern Recogn. **38**(1), 133–142 (2005)
14. Jain, A., Bolle, R., Pankanti, S.: Biometrics: Personal Identification in Networked Society. Springer, New York (2005)
15. Kyoso, M., Uchiyama, A.: Development of an ECG identification system. In: Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 4, pp. 3721–3723 (2001)
16. Lourenço, A., Silva, H., Fred, A.: Unveiling the biometric potential of finger-based ECG signals. Comput. Intell. Neurosci. 5:1–5:8 (2011)

17. Lourenço, A., Silva, H., Carreiras, C., Fred, A.: Outlier detection in non-intrusive ECG biometric system. In: Proceedings of International Conference on Image Analysis and Recognition, pp. 43–52 (2013)
18. Lourenço, A., Silva, H., Leite, P., Lourenço, R., Fred, A.: Real time electrocardiogram segmentation for finger based ECG biometric. In: Proceedings of International Conference on Bio-inspired Signals and Signal Processing (BIOSIGNALS), pp. 49–54 (2012)
19. Neuman, M.: Biopotential amplifiers. Medical instrumentation: application and design, pp. 316–318 (1998)
20. Odinaka, I., Lai, P., Kaplan, A., O'Sullivan, J., Sirevaag, E., Kristjansson, S., Sheffield, A., Rohrbaugh, J.: ECG biometrics: a robust short-time frequency analysis. In: Proceedings of International IEEE Workshop on Information Forensics and Security (WIFS), pp. 1–6 (2010)
21. Odinaka, I., Lai, P., Kaplan, A., O'Sullivan, J., Sirevaag, E., Rohrbaugh, J.: ECG biometric recognition: a comparative analysis. IEEE Trans. Infor. Forensics Secur. **7**(6), 1812–1824 (2012)
22. Oliveira, C., Fred, A.: ECG-based authentication: bayesian vs. nearest neighbour classifiers. In: Proceedings of International Conference on Bio-inspired Systems and Signal Processing (BIOSIGNALS), pp. 163–168 (2009)
23. Safie, S., Soraghan, J., Petropoulakis, L.: Electrocardiogram (ECG) biometric authentication using pulse active ratio (PAR). IEEE Trans. Inf. Forensics Secur. **6**(4), 1315–1322 (2011)
24. Shen, T., Tompkins, W., Hu, Y.: One-lead ECG for identity verification. In: Proceedings of International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 1, pp. 62–63 (2002)
25. Shen, T., Tompkins, W., Hu, Y.: Implementation of a one-lead ECG human identification system on a normal population. J. Eng. Comput. Innovations **2**(1), 12–21 (2010)
26. Silva, H., Carreiras, C., Lourenço, A., Fred, A.: Off-the-person electrocardiography. In: Proceedings of International Congress on Cardiovasular Technologies (CARDIOTECHNIX), pp. 99–106 (2013)
27. Silva, H., Fred, A., Lourenco, A., Jain, A.: Finger ECG signal for user authentication: usability and performance. In: IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS), pp. 1–8 (2013)
28. Silva, H., Gamboa, H., Fred, A.: Applicability of lead $v2$ ECG measurements in biometrics. In: Proceedings of International eHealth, Telemedicine and Health ICT Forum (Med-e-Tel) pp. 177–180 (2007)
29. Silva, H., Gamboa, H., Fred, A.: One lead ECG based personal identification with feature subspace ensembles. In: Proceedings of 5^{th} International Conference on Machine Learning and Data Mining in Pattern Recognition, pp. 770–783. Springer, Berlin (2007)
30. Silva, H., Lourenço, A., Canento, F., Fred, A., Raposo, N.: ECG biometrics: principles and applications. In: Proceedings of 6^{th} International Conference on Bio-Inspired Systems and Signal Processing (BIOSIGNALS), pp. 215–220 (2013)
31. Silva, H., Lourenço, A., Fred, A.: In-vehicle driver recognition based on hand ECG signals. In: Proceedings of International ACM Conference on Intelligent User Interfaces, ACM pp. 25–28 (2012)
32. Wahabi, S., Pouryayevali, S., Hari, S., Hatzinakos, D.: On evaluating ECG biometric systems: Session-dependency and body posture. IEEE Trans. Inf. Forensics Secur. **9**(11), 2002–2013 (2014)
33. Wang, Y., Agrafioti, F., Hatzinakos, D., Plataniotis, K.: Analysis of human electrocardiogram for biometric recognition. EURASIP J. Adv. Signal Process (2008)
34. Webster, J.: Medical Instrumentation Application and Design, 4th edn. Wiley, New York (2009)
35. Zhang, Z., Wei, D.: A new ECG identification method using bayes' teorem. In: Proceedings of IEEE Region 10 Conference (TENCON), pp. 1–4 (2006)