



ECG biometric analysis in different physiological recording conditions

Fabienne Porée, Gaëlle Kervio, Guy Carrault

► To cite this version:

Fabienne Porée, Gaëlle Kervio, Guy Carrault. ECG biometric analysis in different physiological recording conditions. *Signal, Image and Video Processing*, Springer Verlag, 2016, 10 (2), pp.267-276. <10.1007/s11760-014-0737-1>. <hal-01099449>

HAL Id: hal-01099449

<https://hal.archives-ouvertes.fr/hal-01099449>

Submitted on 3 Jan 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ECG Biometric Analysis in Different Physiological Recording Conditions

Fabienne Porée^{a,b}, Gaëlle Kervio^c and Guy Carrault^{a,b}

Addresses:

^a Université de Rennes 1, LTSI, Rennes, France

^b INSERM, U1099, Rennes, France

^c CIC-IT INSERM 804, Rennes, France

Corresponding author:

Fabienne Porée

Laboratoire Traitement du Signal et de l'Image (LTSI), Université de Rennes 1, Campus de Beaulieu, 35042 Cedex, Rennes, France

Tel.: +33-2-23-23-73-30; Fax: +33-2-23-23-69-17

E-mail address: fabienne.poree@univ-rennes1.fr

Abstract

Biometric systems have for objective to perform identification, or verification of identity of individuals. Human electrocardiogram (ECG) has been recently proposed as an additional tool for biometric applications. Then, a set of ECG-based biometric studies has occurred in the literature but they are difficult to compare because they use various values of: the number of ECG leads, the length of the analysis window (only the QRS or more), the delays between recordings... However, they analyze nearly always the ECG in rest conditions. Here, we propose to evaluate the possibility of performing ECG-based biometry in other conditions. For this purpose, a comparative study, on three experimental conditions (supine rest, standing and exercise), has been carried out. It is based on the computing of the correlation coefficient between pairs of shapes of windowed ECG. Both verification and identification tasks are tackled. The results show that there is no advantage in comparing shapes recorded in supine rest conditions, as it is classically done, which represents an obvious benefit in biometry. Performances are evaluated as a function of the shape length. Then, different tests are performed in order to investigate how should be constructed the enrolment database when a system is devoted to work in several conditions. Last part of the paper shows how performances depend on the time and on the number of ECG leads.

Keywords: Electrocardiogram, Biometrics, Verification, Identification

1 Introduction

Biometric systems have for objective to perform identification or verification of identity of individuals using their physical and physiological characteristics. They rely on the hypothesis that more similarities exist between two recordings of a same individual than of two different individuals. The most popular biometric systems are based on fingerprint, iris-scan or also voice and are already present in real-world applications. However, other biometrical approaches are still in investigation. Human electrocardiogram (ECG) has been reported as an additional tool for biometric applications in 2001 [1], since the physiological and geometrical differences of the heart in different individuals display certain uniqueness in the ECG signals [2]. Furthermore, its major benefit, compared to other biometric modalities, leads in the fact that it is difficult to be falsified. And, even if the use of ECG for biometry may be still considered difficult for a practical use, compared to fingerprint or iris scan for example, due to the necessity to place electrodes, recent developments in wearable ECG devices using mobile phones [3] and smart textiles are now promising issues in this direction [4].

Since [1], a set of ECG-based biometric studies has occurred in the literature proposing solutions for identification [1, 5–15], or verification [16–18] or both [19–22]. A set of them, in particular the first studies, proposed systems based on the extraction of a set of fiducial temporal and amplitude features from the ECG, from the P-QRS-T [1, 5–7, 10, 11, 16] or only the QRS-T segment [8, 23], which is generally a difficult task. To bypass it, more recent approaches compute non fiducial parameters between windowed ECG into single heartbeat signals, needing only the R-peaks detection [13, 19, 21, 22], except in [9, 12] where no waveform detection is required. When the ECG is segmented, the template matching is performed between windows of various lengths: 100 ms in [19], 120 ms and 600 ms in [13], 700 ms in [21] and 750 ms in [22].

Concerning the number of leads, almost all the studies are based on one-lead ECG, since Biel [1] showed that it was possible to perform identification with only one lead. Even if it represents an interest from a technological point of view, in [14] the number of leads is equal to 2, in [19] it is equal to 3, in [23] it is equal to 12 and different numbers of leads are compared in [13] (1 and 3), in [12] (1 and 12) and in [1] (1, 6 and 12).

An important issue in biometric studies is the database, public or private (see [24] for details). On the one hand, some exploit a database containing ECG issued from only one session (e.g. MIT-BIH database [25]) and segment the recordings in two parts, one for the training set and one for the testing set. So the training set corresponds to the data recorded

in the enrolment database as reference and the testing test simulates the connection attempts. Indeed, when data come only from one date, it is obviously not realistic in relation to a real-life utilization of a biometric system. On the other hand, some studies are based on data registered at different dates, but only a minority of them contain data with several months or years between the records (e.g. PTB database [26]) [9–12, 19]. In addition, only [19] and [22] contains ECG recordings issued from more than 2 dates. In [19], the number of ECG per subject varies from 2 to 20 with most of the subjects recorded for 2 or 3 times and the average time interval between two dates is about 16 months. However, no discussion about the passing of time is done. In [22] subjects were recorded three times, the separation between recording days ranging from two weeks to six months.

However, almost all these works have a common point, since they analyze the stability of ECG in rest conditions (supine or sitting rest), which represents also a limitation for practical issues of biometric systems (e.g. access to a system controlling the entrance of a secured place just after a biking, running or going upstairs effort). Nevertheless, heart rate changes have been considered in some papers. In [7], Israel et al studies the effect of anxiety, Batchvarov et al [23] investigates the effect of the body position by comparing fiducial parameters obtained in supine rest and standing positions and non healthy subjects are included in [20] and [22]. Only a few papers deal with exercise. In [8], a fiducial approach is adopted and the exercise (stairs up and down) ECG is resampled to reduce the effects of heart rate variability. In [21] the effects of a moderate bicycle exercise, but also the recover period, are studied, using a fixed window ECG length. Intra-state and inter-state are considered and the degradation of performances is reduced by integrating resampling methods. In [27] six physiological conditions (rest, stairs up and down, ...) are studied. But, for these three studies, all ECG had been were collected on the same day.

More recently, Odinaka et al [24] summarized all these previous works and has proposed a comparative analysis. This study reports the authentication (or verification) performance of a few of the ECG biometric systems in a specific database. This study includes the cases where training and testing sets are different and come from different sessions. It clearly underlines that the degradation performance occurred with time and proposes an original training and testing procedures as an alternative procedure. Despite the interest of this work, some major questions are not addressed in this work and are studied in this communication.

The objective of this paper is not to propose a new biometric system but to study the possibility of biometric systems based on ECG recordings to perform in several experimental

conditions: supine rest, standing and exercise. For this purpose, a specific 12-lead ECG database has been acquired, with up to 4 dates by subject, and up to 16 months between dates. The methodology is based on the computation of a score between shapes of windowed ECG and the score is using the correlation coefficient (CC), largely used in matching problems. Only ECG recorded at different dates are compared.

Both verification and identification contexts are tackled. In the verification task, intra-patient case is compared to the inter-patients case and the quality of the separation is evaluated with the discrimination coefficient [28]. In the identification case, identification rates (percentage of good identification in the first rank) are computed. In addition, different temporal supports are tested, from the QRS complex duration (100 ms) to the whole beat (1 s), in order to explore if an optimal shape length exists. Then, as we study several recording conditions, we also simulate different ways to construct the enrolment database. In summary, several tests have been conducted in order to answer to three major questions:

1. Is it required to compare shape recorded in supine rest as it is classically performed?
2. Does an optimal shape length exist?
3. How to define the enrolment database?

Then additional tests evaluate the performance when the delay between records is taken into account, to verify if a degradation over time exists. With our database we can simulate performances obtained after a short delay (1 month) and after 16 months, between the construction of the enrolment database and the connection attempts. The importance of the number of leads is also investigated. Performances obtained with one, three, six and 12 leads are compared.

The paper is an extension of [15] and [18] and is organized as follows. In Section 2 some fundamental bases on ECG and biometry are reminded. Section 3 presents the database and the methodology for processing the recorded data. Then, Section 4 gives the results obtained with the different recording conditions for verification and identification tasks, in several analysis scenarios. Sections 5 presents additional tests concerning the evolution of the performances with time and the number of ECG leads. In Section 6 the conclusion is given.

2 Fundamental bases

2.1 On ECG

In this representation, we consider the description of the normal sinus beat and do not describe modifications due to pathologies such bundle branch block, Kent bundle or ST-T segment elevation. In the same manner, we consider only normal sinus rhythm either in bradycardia or tachycardia but exclude other arrhythmias (ventricular tachycardia, atrial fibrillation, ...).

A normal ECG beat contains 3 different waves: a P wave, a QRS complex and a T wave (Figure 1). The P wave corresponds to the depolarization of the right and the left atria. The QRS complex reflects the depolarization of the right and left ventricles. The T wave corresponds to the ventricular repolarization and the QT interval is dependent on the heart rate [29].

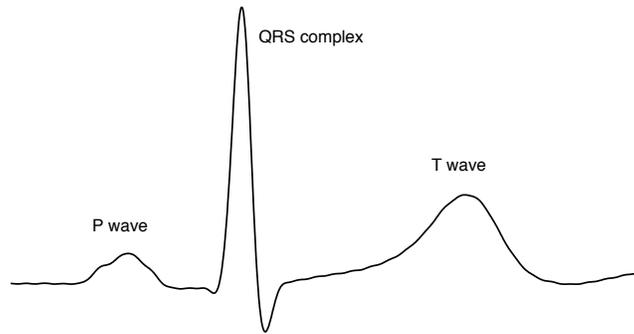


Figure 1: Example of a normal ECG beat (lead II).

Classical recording of the electrical heart activity is based on the recording of a 12-lead ECG and contains 12 channels, denoted by:

- three bipolar leads: I, II, III,
- three augmented unipolar leads: avR, avL, avF,
- and six precordial electrodes: V1 to V6.

2.2 On biometry

A biometric application is based on three steps:

1. the enrolment, where biometric information from an individual is stored,

2. the connection steps, where an individual tries to connect to a system and biometric information is detected and compared with the information stored at the time of enrolment,
3. and the decision step.

Then, two contexts exist, the identity verification and the identification.

In the context of identity verification, an identity is first announced by the subject, his signal is compared to signals owing to him (one-to-one comparison) and the decision consists in accepting or rejecting the claimed identity. When comparing two signals issued from the same subject, we talk about the genuine case, or hypothesis H_0 , or intra-subject case. When comparing two signals issued from two different subjects, we talk about the impostor case, or hypothesis H_1 , or inter-subjects case. The decision step relies on the use of a matching algorithm that decides if the subject is a genuine or an impostor, by comparing its score to a decision threshold.

To evaluate a verification system, scores are computed between all the records, two at a time, of a database, leading to two types distribution, one for hypothesis H_0 and one for hypothesis H_1 . The verification process will be all the more effective since the distributions of H_0 and H_1 are well separated. In practice, false matching rate (FMR) and false non-matching rate (FNMR) may be computed for various threshold values, allowing to plot a Detection Error Tradeoff curve. Sometimes performances are given as a function of the Equal Error Rate (EER), where FMR and FNMR are equal.

In the context of identification, the signal is compared to a biometric database (one-to-many comparison). The decision step relies on the use of a ranking algorithm, providing the identity of the subject.

To evaluate an identification system, for each record of the database, a score is computed between it and all the others. Performance may be evaluated by analyzing the nearest neighbors, and giving the rank necessary to find the same subject, or only considering the first rank (or first rank identity) which provides an identity of the subject (correct or not).

3 Database and Methodology

3.1 Database

A specific database has been built for this study. It includes fourteen healthy subjects (7 males and 7 females) with mean and standard deviation of the ages equal to 29.0 and 3.6

respectively. For each of the subjects, the recording of a 12-lead ECG has been performed using a Cardionics (Belgium) system at the frequency of 1000 Hz.

The protocol consisted in the following steps:

- 5 minutes in supine rest (not recorded)
- 3 minutes in supine rest (R)
- 3 minutes in standing (S)
- 3 minutes in exercise (bicycle effort) (E)

During the bicycle effort, the subject was asked to provide a level of effort between moderate and high but no heart rate value to attain was imposed in order to take into account the physiology of each individual and to be close to the real-life context. Furthermore, the three steps were recorded at one go and then segmented.

In order to test the long-term stability of the ECG, this protocol was repeated at 4 different dates: first date (reference date), 2 weeks after the first date, 1 month after the second, and 15 months after the third. The maximal delay, between dates 1 and 4, is equal to more than 16 months.

$$\text{Date 1} \xrightarrow{2 \text{ w}} \text{Date 2} \xrightarrow{1 \text{ m}} \text{Date 3} \xrightarrow{15 \text{ m}} \text{Date 4}$$

In the following, we will define a "record" as a 12-lead ECG signal recorded at a given date in a given condition. At a given date, we have access to three different records per subject. However, depending on each subject, the number of dates is varying in function of its availabilities, and goes from one to four, so that the number of ECG records per subject varies from three (recorded one time) to 12 (recorded 4 times). In addition, a small number records were rejected due to misplacement of electrodes or manipulation error. And finally, the total number of records is equal to $N = 106$: 35 in supine rest and standing and 36 in exercise (two files were badly recorded and removed from the database). For $k = 1, \dots, N$, each record k is characterized by $\text{Id}(k)$ the identity of the subject, $\text{Cond}(k)$ the recording condition and $\text{Date}(k)$ the date of the recording.

3.2 Methodology

3.2.1 Shape extraction

Each record is then processed as depicted in Figure 2. First, a low-pass filter is applied to each of the 12 leads with a cut-off frequency at 45 Hz. Then a beat detection procedure

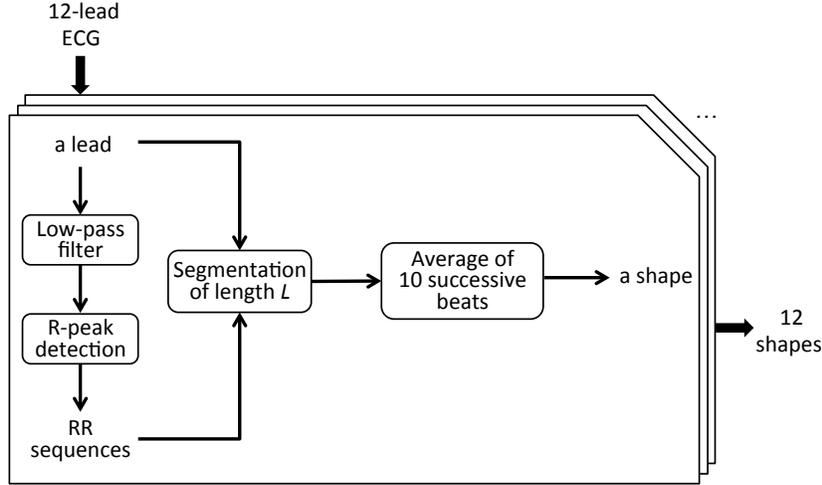


Figure 2: Shape extraction for a 12-lead ECG.

determines the R-peak positions [30] which allows to segment the ECG into successive beats of length L . The length L , centered around the R peak, is a varying parameter since different window lengths have to be tested, from the QRS complex ($L = 100$ ms) to the whole beat P-QRS-T ($L = 1000$ ms). Finally, in order to avoid a noise filtering step, the average of 10 successive beats, chosen in the second minute of the record, excluding preliminary ventricular contraction, is performed. It provides 12 shapes (a shape per lead) for each record.

Figure 3 shows the shapes extracted for a subject, on lead II, with $L = 1000$ ms. Four records per condition are available and superimposed. We observe first that only low mod-

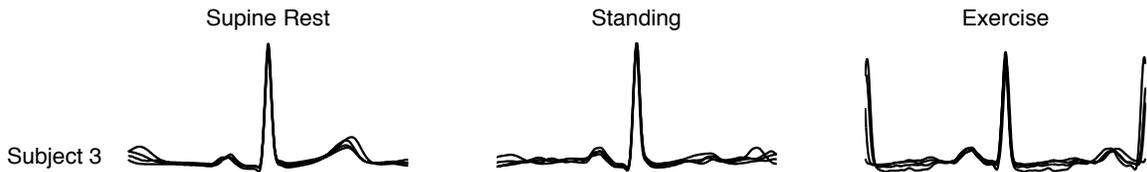


Figure 3: Extracted shapes obtained for subject 3 on lead II. For each condition (supine rest, standing and exercise), shapes issued from four different dates are superimposed.

ifications are involved on the shape morphology from date to date in the same condition. However, in exercise condition, the signal contains, not only the P-QRS-T segment, but also the end of the precedent beat and the beginning of the following beat, due to heart rate acceleration. We can also verify the influence of the autonomic nervous system that modifies T wave position and duration during exercise [29]. We also notice the presence of noise, particularly on the baseline.

By observing signals of two other subjects (Figure 4), we see that subject 9 has a morphology clearly different from the two others, contrary to subjects 3 and 5, who seem to have more similar ECG. These examples clearly exhibit why the verification and identification tasks may be sometimes difficult.

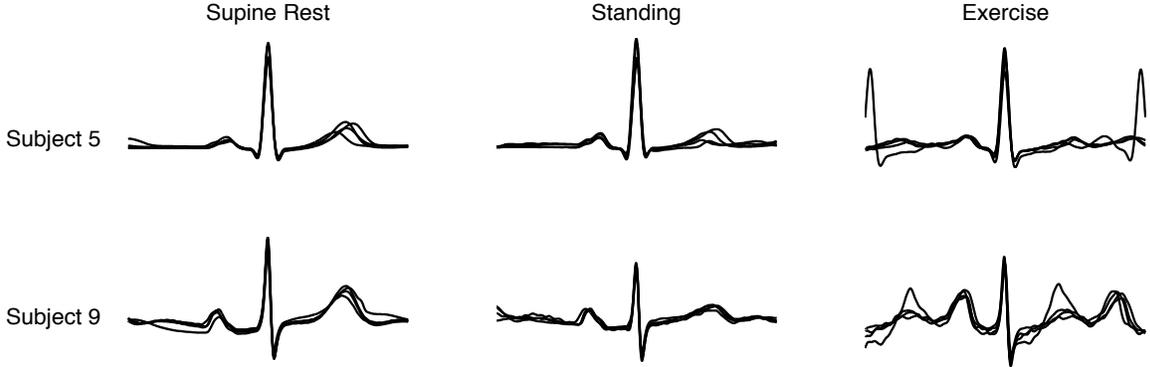


Figure 4: Extracted shapes obtained for subjects 5 and 9 on lead II. For each condition (supine rest, standing and exercise), shapes issued from four different dates are superimposed.

3.2.2 Comparison metric

The objective is to obtain a score between pairs of recorded ECG, using the extracted shapes. A preliminary step consists of applying a time-shift correction between the 12 pairs of shapes using cross-correlation. Let $s_{k,l}$ and $s_{k',l}$, two shapes of length L , issued from records k and k' respectively, on the same lead l . Their cross-correlation is computed as follows:

$$\Gamma_{kk',l}(\tau) = \begin{cases} \sum_{i=0}^{L-\tau-1} s_{k,l}[i+\tau]s_{k',l}[i] & \text{if } \tau \geq 0 \\ \Gamma_{k'k,l}(-\tau) & \text{if } \tau < 0 \end{cases} \quad (1)$$

where τ is the time lag, in the interval $[-25 \text{ ms}, \dots, 25 \text{ ms}]$. Then the delay for lead l , $\hat{\tau}_l$, is the value that maximizes $\Gamma_{kk',l}(\tau)$:

$$\hat{\tau}_l = \arg \max_{\tau} \Gamma_{kk',l}(\tau). \quad (2)$$

After computing a delay for the 12 leads, the value that is kept is the one that occurs most of the time. If several values are concurrent, we keep the smallest. If all the 12 values are different, the delay is put to 0.

Then, the correlation coefficient is used as comparison metric between shapes $s_{k,l}$ and $s_{k',l}$. It is called $r_{kk',l}$ and is equal to:

$$r_{kk',l} = \frac{1}{L} \frac{\sum_{i=0}^{L-1} (s_{k,l}[i] - \mu_{s_{k,l}})(s_{k',l}[i] - \mu_{s_{k',l}})}{\sigma_{s_{k,l}} \sigma_{s_{k',l}}}, \quad (3)$$

where $\mu_{s_{k,l}}$ and $\sigma_{s_{k,l}}$ are the mean and the standard deviation of shape $s_{k,l}$, $\mu_{s_{k',l}}$ and $\sigma_{s_{k',l}}$ are the mean and the standard deviation of shape $s_{k',l}$.

In Section 4 the number of leads is equal to 12 whereas it is a variable parameter in Section 5.2. Let n be the number of leads, CC is then computed for each lead and denoted by the vector $r_{kk'} = [r_{kk',1}, r_{kk',2}, \dots, r_{kk',n}]$. The final score between records k and k' is called $R_{kk'}$ and is the mean CC i.e. the average of vector $r_{kk'}$.

3.3 Performance analysis

3.3.1 Verification

In this section, a score $R_{kk'}$ is computed between all the possible pairs of records k and $k' \in [0, \dots, N]$, with $k \neq k'$, leading to two distributions:

- The intra-subject distribution $p_{H_0}(R_{kk'})$: the distribution of the scores $R_{kk'}$ under hypothesis H_0 , i.e. $\text{Id}(k) = \text{Id}(k')$
- The inter-subjects distribution $p_{H_1}(R_{kk'})$: the distribution of the scores $R_{kk'}$ under hypothesis H_1 , i.e. $\text{Id}(k) \neq \text{Id}(k')$.

As already said, different ways to evaluate a biometric system exist, but a verification system will be all the more effective since the distributions of H_0 and H_1 are well separated. For this reason, we propose to quantify the separation between H_0 and H_1 by the use of the discrimination coefficient (DC) proposed by Alexander et al. in [28] for speaker verification, and defined as:

$$\text{DC} = \frac{\mu_{H_0} - \mu_{H_1}}{\sigma_{H_0} + \sigma_{H_1}}, \quad (4)$$

where μ_{H_0} and σ_{H_0} are the mean and the standard deviation of distribution $p_{H_0}(R_{kk'})$, μ_{H_1} and σ_{H_1} are the mean and the standard deviation of distribution $p_{H_1}(R_{kk'})$.

According to [28], DC equal to one or below corresponds to a system with poor discrimination, if DC is between one and two the discrimination is moderate to good. Values of DC up to two imply very good discrimination.

3.3.2 Identification

For each record k of the database, a score $R_{kk'}$ is computed with all the other records k' (with k and $k' \in [0, \dots, N']$ and $k \neq k'$). Let \hat{k}_0 be the record leading to the highest score, or the nearest neighbor:

$$\hat{k}'_0 = \arg \max_{k'} R_{kk'}, \quad (5)$$

its identity $\text{Id}(\hat{k}'_0)$ provides the estimated identity of the tested individual.

If it is equal to the real identity ($\text{Id}(\hat{k}'_0) = \text{Id}(k)$), the system lead to a good identification. After applying this process to the whole database, the percentage of good identification can be computed and is called the identification rate (IR).

It is worthwhile to note that, for this task, we reduced the database to the 11 subjects recorded at least two times, which got the total number of records to $N' = 93$.

4 Results

4.1 Introduction

As already said, most of the studies only consider the supine rest condition, which represents an important limitation for an utilization of ECG-based biometric system in real-life context. Here two objectives are addressed. The first one evaluates if it is possible to envisage ECG biometric systems in other recording conditions than supine rest as it is classically done. The second objective studies how the enrolment database should be constructed in a practical use when different conditions may be encountered. In other words, should it necessary contain records of the different conditions or not?

For this purpose, three different scenarios are tested:

- The "Intra-condition" case: a score is computed only between pairs of records acquired in the same condition. For example, the case S - S represents the case where both tested ECG have been recorded in the standing condition.
- The "Inter-conditions" case: a score is computed only between pairs of records acquired in two different conditions. For example, case S - E represents the case where one record is in standing and the other is in exercise.
- The "All-Conditions" case: a score is computed between one condition and all the others. For example, in the case "All - R" we computed a score between pairs of records where at least one were in supine rest.

Performances are evaluated using $n = 12$ leads and discussed as a function of the shape length L , with $L \in [100, \dots, 1000]$ ms. Both verification and identification tasks are studied.

4.2 Verification

For the verification task, we compute scores between all pairs of records of the database, leading to two distributions, one for intra-subject values (hypothesis H_0) and one for inter-subjects values (H_1). From them, DC can be computed according to Eq. 4.

4.2.1 Intra-condition performances

Figures 5(a) and 5(b) show the mean of the scores $R_{kk'}$ of the intra-subject (i.e. μ_{H_0}) and inter-subjects (i.e. μ_{H_1}) distributions as a function of L for each of the three "Intra-condition" cases.

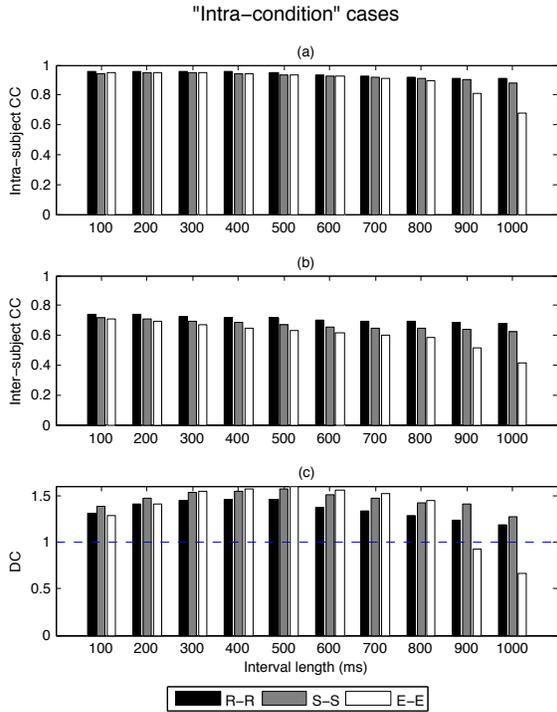


Figure 5: (a) Intra- (b) Inter-subjects mean scores and (c) Discrimination coefficient in the three "Intra-condition" cases R-R, S-S and E-E.

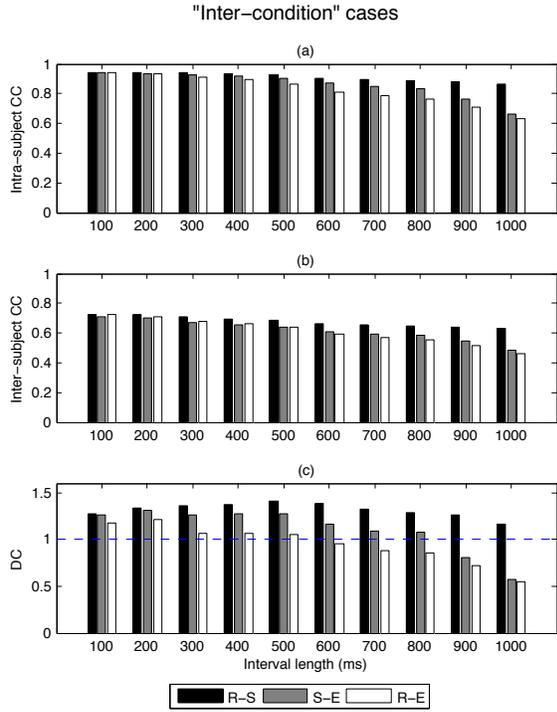


Figure 6: (a) Intra- (b) Inter-subjects mean scores and (c) Discrimination coefficient in the three "Inter-conditions" cases R-S, S-E and R-E.

Results show that up to $L = 800$ ms intra-subject mean score is high (above 0.95), quasi-

equal for the three conditions, and it is slowly decreasing when L is increasing. In addition, for both cases (intra- and inter-subjects), we observe that:

- The mean of the scores is slightly higher in supine rest, than in standing and exercise.
- In supine rest and standing conditions, results are quasi-identical whatever the shape length.
- In exercise, when $L \geq 900$ ms, the mean of the scores is breaking down. It is due to the heart rate acceleration which modifies the T wave position, as already shown in Figure 3.

The behavior of the discrimination coefficient, in order to verify if the separation capability between both hypotheses H_0 and H_1 is effective, is depicted on Figure 5(c). Results form bell shapes with maximum values when L is around 500 ms. DC is always higher to 1, except in exercise where DC is breaking down when $L \geq 900$ ms. In addition, we can observe that performance are even slightly higher in standing and exercise conditions for $300 \text{ ms} \leq L \leq 800$ ms than in supine rest.

These results are really innovative and important since they suggest that there is no requirement or advantage to compare only ECG shapes recorded in supine rest conditions, as it is classically done. Under conditions on the shape length, they show that comparing ECG shapes recorded in standing condition or in exercise provides also a good discrimination coefficient, sometimes higher, which represents an obvious benefit in biometry. However, in exercise, shape length must not be greater than 800 ms.

4.2.2 Inter-conditions performances

Figures 6(a) and 6(b) show the mean of the scores $R_{kk'}$ of the intra- and inter-subjects distributions in the three "Inter-conditions" case. Comparing with Figure 5, we especially observe lower values of the scores in the intra-subject case. Consequently, on Figure 6(c), values of DC are slightly lower than in the "Intra-condition" case. More precisely:

- In case R - S, DC is still high and is only slowly varying with the shape length.
- In cases S - E and R - E, DC is always decreasing when L is increasing and falls below 1 when $L \geq 900$ ms and $L \geq 600$ ms respectively.

To sum up, crossing the conditions provides still acceptable values of DC, provided that the shape length is lower than 600 ms.

These results show that it is also possible to compare ECG recorded in different conditions. Between rest and standing conditions, performances are high even using the whole beat (P-QRS-T). As expected, the case R - E is the most difficult. When one of the two shapes has been recorded in exercise, it is preferable to restrict the shape length to lower values (QRS or P-QRS). In other words, they show that, for example if the enrolment has been constructed in a given condition, it can be used in other conditions. It also brings flexibility for the constitution of database.

4.2.3 Construction of the enrolment database

Having observed the results of the previous section, we can now address more precisely the question of the construction of the enrolment database: is it preferable that the enrolment database contains records of these different conditions for each subject?

For this purpose, we computed performances with the database containing all the conditions ("All-Conditions" case). Results obtained for each of the three conditions are given in Figure 7.

First, we observe that once again best results are obtained with the standing condition. They are also always lower in the "Intra-conditions" case, but higher than in the "Inter-conditions".

All these results suggest that:

- If the system will work in only one condition, always the same, it is recommended to construct the enrolment database in this condition. Expected performances will be the ones of "Intra-Condition" case given in Figure 5. As already mentioned, shape length should be tuned to 500 ms.
- If the system will work in several physiological conditions, we recommend:
 - if only one condition can be recorded during the enrolment, to choose the standing condition, since it provides good performances in both "Intra-Condition" and "Inter-Conditions" cases. The standing condition appears as a compromise between the supine rest and the exercise conditions. Shape length may be set around 200 ms.
 - if it is possible, to fill the enrolment database with different physiological conditions. Expected performances are the ones of "All-Conditions" given in Figure 7. Shape length may be chosen around 200 ms. Note that, if during the connection

it is possible to know the physiological state of the user (self declaration or other), only ECG recorded in the same state will be compared and performances will be the ones depicted in Figure 5.

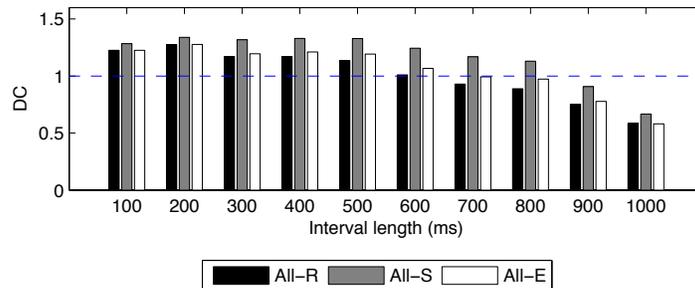


Figure 7: Discrimination coefficient computed in cases All-R, All-S and All-E.

4.3 Identification

We now consider the identification task. Identification rates are computed as presented in section 3.3.2. As for verification, we computed IR in the "Intra-conditions", "Inter-conditions" and "All-Conditions" cases. Results are presented in Table 1 as a function of the shape length.

Table 1: IR (%) computed in the three "Intra-condition" cases, in the three "Inter-conditions" cases and in the three "All-conditions" cases, for $L = 100, \dots, 1000$ ms.

	100	200	300	400	500	600	700	800	900	1000
R - R	90.3	93.6	93.6	96.8	96.8	96.8	93.6	93.6	87.1	83.9
S - S	93.6	96.8	96.8	96.8	96.8	96.8	96.8	93.6	93.6	93.6
E - E	83.9	83.9	93.6	96.8	96.8	96.8	93.6	90.3	83.9	58.1
R - S	98.4	98.4	98.4	98.4	100.0	98.4	98.4	98.4	98.4	98.4
S - E	96.8	96.8	98.4	100.0	100.0	96.8	96.8	96.8	90.3	69.4
R - E	95.2	93.6	85.5	88.7	88.7	80.5	79.0	75.8	74.2	69.4
All - R	96.8	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.8	96.8
All - S	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	96.8
All - E	96.8	96.8	96.8	100.0	100.0	100.0	100.0	96.8	90.3	67.7

In the three "Intra-conditions" cases, performances are close and high, up to $L = 900$. Again, slightly higher values of IR are obtained with the standing condition. Moreover, the optimal value of L is located between 400 and 600 ms and outside from these values, IR is

slowly decreasing in rest and standing conditions. In the exercise condition, values of IR are the lowest and are breaking down when $L = 1000$ ms.

In the three "Inter-conditions" cases, values of IR are globally decreasing when L is increasing. It is worth noticing that this loss is small, between rest and standing whatever L , and between standing and exercise up to $L = 800$ ms. However, between rest and exercise, performances are really poor as soon as $L \geq 600$ ms.

For both above cases, results obtained in the identification task are similar to those obtained with the verification. Indeed, high values of IR may be also obtained when comparing ECG shapes recorded in standing condition or in exercise. They also show that good results may be obtained if the tested shape has been recorded in a condition that is not present in the enrolment database, under constraints on the shape length and avoiding the case R - E.

However, results obtained in the "All-conditions" case are different than those obtained during the verification task since they are the highest, with most of the time 100% of good identification.

These results suggest that:

- If only one condition can be acquired during the enrolment, it should be preferable to choose the standing condition, with a shape length around 500 ms.
- However, if it is possible, it will be better to fill the enrolment database with different physiological conditions, even if the system is expected to work in only one condition. The shape length should be still tuned to 500 ms.

5 Additional tests

In this section, we propose additional tests concerning the evolution of the performances with the passing of time and the number of leads. To simplify the section, only the "Intra-condition" cases are presented, and for $L = 500$ ms, as it was the optimal length for most of the tests previously reported.

5.1 Influence of time

As pointed in the introduction, beyond studies considering data registered at different dates, only a minority of them contains data with several months or years between the records and issued from more than two dates.

Our database containing more than two dates per subject, it is possible to evaluate the effect of the time between the date of the enrolment and the dates of the connection attempts. For this purpose, we compare values of DC computed:

- between records acquired at date 1 and at date 2&3 (dates 2 and 3 merged as they were close): up to 1 month delay,
- and between records acquired date 1 and at date 4: 16 months delay.

Results are presented in Figure 8 for verification (top) and identification (bottom). We

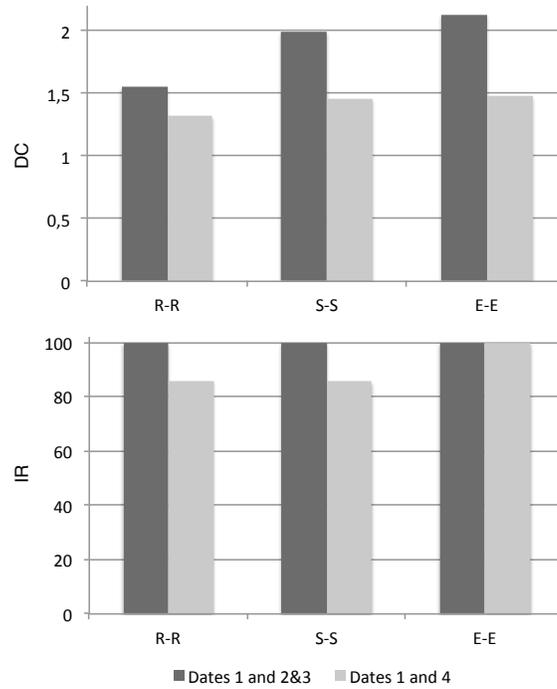


Figure 8: DC (top) and IR (bottom) computed as a function of the delay between dates.

can observe that DC and IR obtained between date 1 and dates 2&3 are always higher (or equal) than between date 1 and date 4. These results show that there may exist a degradation of the ECG stability over time. However performances are still acceptable after 16 months.

5.2 Influence of the number of leads

The majority of the studies of the literature have been led with one lead, since it represents an interest from a technological point of view. Here we propose to show performances obtained with $n = 1, 3, 6$ and 12 leads. All the possible n -combinations of 12 are tested. Results (mean

and standard deviation) are presented in Figure 9 for verification (top) and for identification (bottom).

For both DC and IR, the best performances are obtained with 12 leads and then decrease with the number of leads. With $n = 3$ and 6, results are still acceptable, with values of DC greater than 1 and values of IR greater than 90%. However for $n = 1$, where the smallest mean and the highest standard deviation are obtained, DC is always lower than 1 and IR is always lower than 90%.

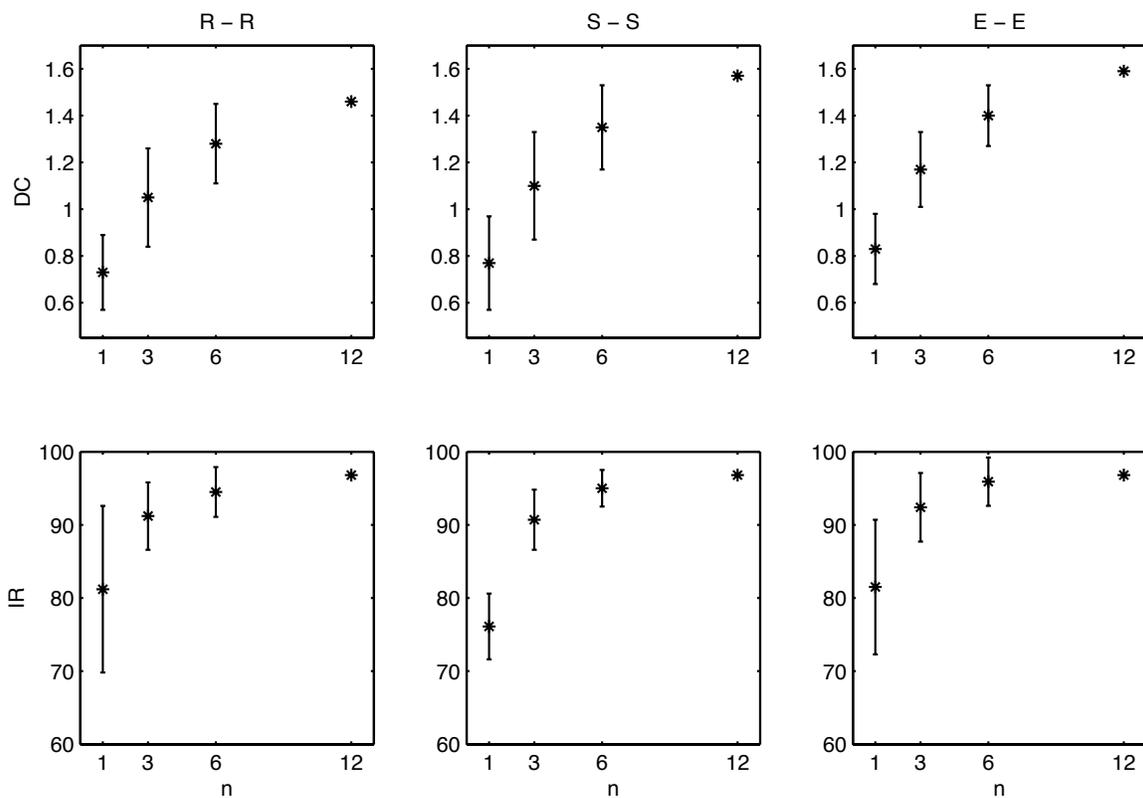


Figure 9: Mean (*) and standard deviation (-) of values of DC (top) and IR (bottom) as a function of the number of leads.

6 Conclusion

The first objective of this paper was to evaluate the possibility of performing ECG-based biometry in different physiological conditions. For this purpose, we chose to use the correlation coefficient in order to have a simple and general framework for the evaluation and to objectively compare the different conditions.

For both verification and identification tasks, we showed that there is no requirement or

advantage to restrict such studies to supine rest condition, as it is classically done. In the "Intra-Condition" case, where only ECG recorded in the same condition were considered, very high performances were obtained in the three conditions, using a shape length up to 800 ms in exercise. The best performances were encountered with the standing conditions and an shape length between 400 and 600 ms.

The second objective was to explore the different ways to construct the enrolment database. For the verification task, we showed that if the system is supposed to work in different conditions, an optimal system would consist in recording, for each individual in the enrolment database, these different conditions. However, it is not essential, since "Inter-Conditions" performances may be still high, provided that conditions on the shape length are fulfilled. If the enrolment database can contain only one condition, results suggest choosing the standing condition, that is as a compromise between the supine rest and the exercise conditions. This allows avoiding the most difficult case R - E. For the identification task, we showed that results are always higher when the database is filled with several conditions.

In the last part of the paper we studied the effect of the time and of the number of leads. Results showed that there is a degradation with time, but the results are still acceptable after 16 months. Regarding the number of leads, we obtained the best performances when it is equal to 12.

We are aware that the number of subjects is reduced but it is worthwhile to remind that the constitution of such a database is very difficult and required to follow individuals over nearly two years. Furthermore, no database, including rest, standing and exercise conditions exist elsewhere to our knowledge.

The sampling frequency has been kept at 1000 Hz, as proposed by our recording system. It may appear high but this is a value largely reported in other studies [7,9,10,12,21,22,24,27]. Decreasing this value would be possible and would probably decrease the performance but without changing the conclusions. This paper contributes to the conception of ECG-based biometric systems. It defines the usability and the impact of different physiological conditions (rest, standing and exercise in this paper), the influence of the time and the importance of the number of leads.

Acknowledgements

The authors gratefully acknowledge volunteers and technicians who participated to this work.

References

- [1] L. Biel, O. Pettersson, L. Philipson, and P. Wide. ECG analysis: A new approach in human identification. *IEEE Trans. Instrum. Meas.*, 50(3):808–12, 2001.
- [2] R. Hoekema, G.J.H. Uijen, and A. van Oosterom. Geometrical aspects of the interindividual variability of multilead ECG recordings. *IEEE Trans. Biomed. Eng.*, 48:551–9, 2001.
- [3] I. Silva, G.B. Moody, and L. Celi. Improving the quality of ECGs collected using mobile phones: The physionet/computing in cardiology challenge 2011. In *Computing in Cardiology*, pages 273–276. IEEE, 2011.
- [4] A. Pantelopoulos and N.G. Bourbakis. A survey on wearable sensor-based systems for health monitoring and prognosis. *IEEE Trans. Syst. Man Cybern. C, Appl. Rev.*, 40(1):1–12, 2010.
- [5] M. Kyoso and A. Uchiyama. Development of an ECG identification system. In *Proc. of the 23rd IEEE EMBS Conference*, volume 4, pages 3721–23, 2001.
- [6] T. Shen, W. Tompkins, and Y. Hu. One-lead QRS for identity verification. In *Proc. of the Second Joint EMBS/BMES Conference*, pages 62–3, 2002.
- [7] S.A. Israel, J.M. Irvine, A. Cheng, M.D. Wiederhold, and B.K. Wiederhold. ECG to identify individuals. *Pattern Recognition*, 38(1):133–42, 2005.
- [8] K.S. Kim, T.H. Yoon, J.W. Lee, D.J. Kim, and H.S. Koo. A robust human identification by normalized time-domain features of electrocardiogram. In *27th Annual International Conference of the Engineering in Medicine and Biology Society (EMBS)*, pages 1114–1117. IEEE, 2005.
- [9] K.N. Plataniotis, D. Hatzinakos, and J.K. Lee. ECG biometric recognition without fiducial detection. In *Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6. IEEE, 2006.
- [10] Y. Wang, K.N. Plataniotis, and D. Hatzinakos. Integrating analytic and appearance attributes for human identification from ECG signals. In *Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6. IEEE, 2006.

- [11] Y. Wang, F. Agraftoti, D. Hatzinakos, and K.N. Plataniotis. Analysis of human electrocardiogram for biometric recognition. *EURASIP journal on Advances in Signal Processing*, 2008:148658, 2008.
- [12] F. Agraftoti and D. Hatzinakos. Fusion of ECG sources for human identification. In *3rd International Symposium on Communications, Control and Signal Processing (ISCCSP)*, pages 1542–1547. IEEE, 2008.
- [13] S.C. Fang and H.L. Chan. Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space. *Pattern Recognition*, 42:1824–1831, September 2009.
- [14] C. Ye, M. T. Coimbra, and B.V.K. Vijaya Kumar. Investigation of human identification using two-lead electrocardiogram (ECG) signals. In *Fourth IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, pages 1–8. IEEE, 2010.
- [15] F. Porée, A. Gallix, and G. Carrault. Biometric identification of individuals based on the ECG. which conditions? In *Computing in Cardiology (CinC)*, pages 761–764. IEEE, 2011.
- [16] J.M. Irvine and S.A. Israel. A sequential procedure for individual identity verification using ECG. *EURASIP Journal on Advances in Signal Processing*, 2009:243215, 2009.
- [17] F. Agraftoti, 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 (BTAS)*, pages 1–6. IEEE, 2009.
- [18] F. Porée, J.Y. Bansard, G. Kervio, and G. Carrault. Stability analysis of the 12-lead ECG morphology in different physiological conditions: Of interest for biometric applications. In *Computers in Cardiology (CinC)*, pages 285–8. IEEE, 2009.
- [19] G. Wübbeler, M. Stavridis, D. Kreisler, R-D. Bousseljot, and C. Elster. Verification of humans using the electrocardiogram. *Pattern Recognition Letters*, 28:1172–75, 2007.
- [20] F. Agraftoti and D. Hatzinakos. ECG biometric analysis in cardiac irregularity conditions. *Signal, Image and Video Processing*, 3(4):329–343, 2009.
- [21] M. Chen, J.A. O’Sullivan, A.D. Kaplan, P.H. Lai, E.J. Sirevaag, and J.W. Rohrbaugh. Biometrics with physical exercise using laser doppler vibrometry measurements of the

- carotid pulse. In *International Conference on Biometrics, Identity and Security (BIDS)*, pages 1–6. IEEE, 2009.
- [22] I. Odinaka, P.H. Lai, A.D. Kaplan, J.A. O’Sullivan, E.J. Sirevaag, S.D. Kristjansson, A.K. Sheffield, and J.W. Rohrbaugh. ECG biometrics: A robust short-time frequency analysis. In *IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 1–6. IEEE, 2010.
- [23] V.N. Batchvarov, G. Bortolan, and I.I. Christov. Effect of heart rate and body position on the complexity of the QRS and T wave in healthy subjects. In *Computers in Cardiology (CinC)*, pages 225–8. IEEE, 2008.
- [24] I. Odinaka, P.H. Lai, A.D. Kaplan, J.A. O’Sullivan, E.J. Sirevaag, and J.W. Rohrbaugh. ECG biometric recognition: A comparative analysis. *IEEE Trans. Inf. Forens. Security*, 7(6):1812–1824, dec. 2012.
- [25] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, and H.E. Stanley. Physiobank, PhysioToolkit, and Physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.
- [26] M. Oeff, H. Koch, R. Bousseljot, and D. Kreiseler. The PTB diagnostic ECG database, national metrology institute of germany, 1995.
- [27] K.A. Sidek, I. Khalil, and M. Smolen. ECG biometric recognition in different physiological conditions using robust normalized QRS complexes. In *Computing in Cardiology (CinC)*, pages 97–100. IEEE, 2012.
- [28] A. Alexander, F. Botti, and A. Drygajlo. Handling mismatch in corpus-based forensic speaker recognition. In *ODYS-2004*, pages 69–74, 2004.
- [29] L. Sörnmo and P. Laguna. *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. Elsevier, Amsterdam, The Netherlands, 2005.
- [30] J. Pan and W. J. Tompkins. A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.*, 32:230–6, 1985.