BioMedical Engineering OnLine

REVIEW

Open Access

CrossMark

Individual identification via electrocardiogram analysis

Antonio Fratini[1*], Mario Sansone[2], Paolo Bifulco[2] and Mario Cesarelli[2]

*Correspondence:
a.fratini@aston.ac.uk
[1] School of Life and Health
Sciences, Aston University,
Aston Triangle, B4
7ET Birmingham, UK
Full list of author information
is available at the end of the
article

Abstract

Background: During last decade the use of ECG recordings in biometric recognition studies has increased. ECG characteristics made it suitable for subject identification: it is unique, present in all living individuals, and hard to forge. However, in spite of the great number of approaches found in literature, no agreement exists on the most appropriate methodology. This study aimed at providing a survey of the techniques used so far in ECG-based human identification. Specifically, a pattern recognition perspective is here proposed providing a unifying framework to appreciate previous studies and, hopefully, guide future research.

Methods: We searched for papers on the subject from the earliest available date using relevant electronic databases (Medline, IEEEXplore, Scopus, and Web of Knowledge). The following terms were used in different combinations: electrocardiogram, ECG, human identification, biometric, authentication and individual variability. The electronic sources were last searched on 1st March 2015. In our selection we included published research on peer-reviewed journals, books chapters and conferences proceedings. The search was performed for English language documents.

Results: 100 pertinent papers were found. Number of subjects involved in the journal studies ranges from 10 to 502, age from 16 to 86, male and female subjects are generally present. Number of analysed leads varies as well as the recording conditions. Identification performance differs widely as well as verification rate. Many studies refer to publicly available databases (Physionet ECG databases repository) while others rely on proprietary recordings making difficult them to compare. As a measure of overall accuracy we computed a weighted average of the identification rate and equal error rate in authentication scenarios. Identification rate resulted equal to 94.95 % while the equal error rate equal to 0.92 %.

Conclusions: Biometric recognition is a mature field of research. Nevertheless, the use of physiological signals features, such as the ECG traits, needs further improvements. ECG features have the potential to be used in daily activities such as access control and patient handling as well as in wearable electronics applications. However, some barriers still limit its growth. Further analysis should be addressed on the use of single lead recordings and the study of features which are not dependent on the recording sites (e.g. fingers, hand palms). Moreover, it is expected that new techniques will be developed using fiducials and non-fiducial based features in order to catch the best of both approaches. ECG recognition in pathological subjects is also worth of additional investigations.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 2 of 23

Background

Biometric recognition, often referred as Biometrics, is the science that uses statistical methods to uniquely identify humans by means of their physiological and behavioural characteristics. It is mostly used to solve problems of access control, providing reliable and secure alternatives to the conventional authentication methods [1–5].

Subject identification can be achieved using several human discriminants such as retinal structure, fingerprint, face, palm print, etc. However, each one of them exhibits issues related to the specific hardware to use, the practicability of the measures and the robustness against spoofing attacks. Retinal scan is a relatively quick and secure procedure (there are relatively few chances to forge it), however the technology is still expensive and more importantly the procedure is sometimes perceived as invasive and unpleasant. Fingerprint is the most widespread biometric, it has been used in forensic for about 100 years. Automated systems base their accuracy on a multispectral approach, however simple rolled fingerprint (inked impression images) can relatively easy be forged. The possibility to steal data directly from subjects and the relative ease to replicate them (e.g. silicone fingerprints, pictures or facial masks) posed serious challenges to researchers [6–10] and multimodal recognition systems have been suggested [2, 11–14].

In the last decade, the registration of the electrical activity of the heart on the body surface, namely the electrocardiogram (ECG), has been documented to be suitable for identity recognition [1, 13, 15]. Dedicated research on the ECG analysis has demonstrated its advantages in biometrics: ECG is present in all living individuals, exhibits the typical characteristics of a biometric and it is hard to forge. In addition, ECG analysis is a robust method to detect the aliveness of the subject in authentication scenarios.

To date, many different approaches to human recognition via ECG have been reported in the scientific literature but no agreement exists on the appropriate methodologies. Moreover, the use of ad-hoc signal databases makes difficult the assessment of all existing techniques [16, 17].

This study attempts to provide a survey of the techniques used so far in ECG-based human identification. Here, we present a perspective on the progresses of the last decade's research in the field and a discussion on the possible implications for future research.

Previous attempts to summarise ECG-based recognition techniques can be traced back to the work of Nasri [18], Odinaka [17] and Israel and Irvine [19].

Nasri briefly summarized the literature by 2008, Odinaka compared the performance of different algorithms testing them on a single database while Israel and Irvine suggested a sensor-based perspective. More recently, some author investigated combination of ECG-traits with other signals (voice, phonocardiography, Laser Doppler Vibrometry) to enhance the identification rate [20–22].

The aim of this survey is to provide a *pattern recognition* perspective, giving a unifying framework for interpreting previous studies and, hopefully, to guide future works. We concentrated on 'features' for ECG-based human recognition as well as classification strategies. In addition, we evaluated a weighted mean accuracy of the found journal studies to assess the overall performance of the ECG biometrics techniques used so far.

This survey is organized as follows: section two reports on the search strategy, inclusion criteria, and overall performance evaluation strategy; section three describes the

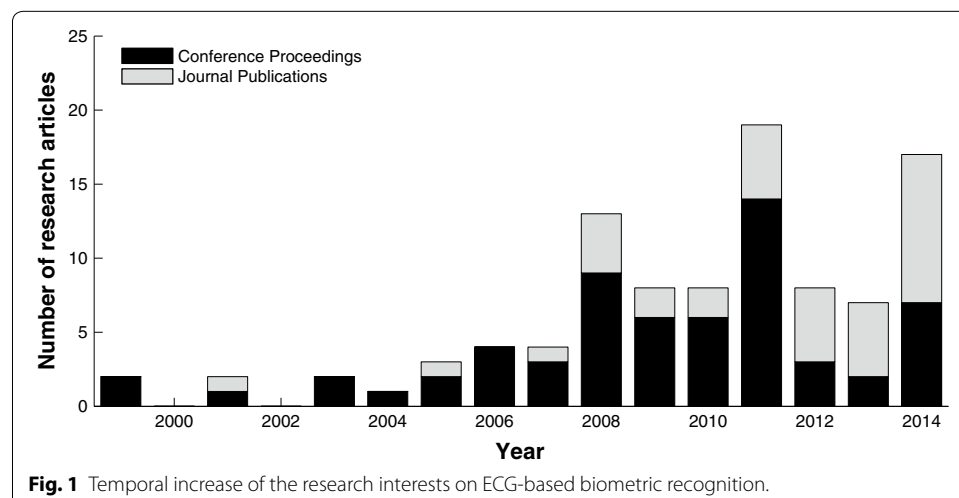Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 3 of 23

main principles underlying the use of ECG as biometrics, the most spread ECG-based features and the databases used to test algorithms performance; in section four we address the issues of feature selection and dimensionality reduction. Discussion and conclusions report the overall picture dealing with the open issues on ECG-based biometrics.

It is nevertheless worth to highlight what this survey is not aimed at. We will not review ECG pre-processing, which is an established research area and a large number of studies report on efficient methods for noise removal, power-line suppression, baseline-wandering removal etc. [8, 23, 24]. We will also not review the methods for QRS detection, although this is one of the most important issues in all algorithms for individual identification. QRS detection has been reviewed elsewhere and we refer the interested reader to specific papers such as [1, 25–27].

Search strategy

We searched for papers on the subject from the earliest available date using relevant electronic databases (Medline, IEEEXplore, Scopus, Web of Knowledge). We used the following terms in different combinations: electrocardiogram, ECG, human identification, biometric, authentication and individual variability. The electronic sources were last searched on 1st March 2015. We also performed a hand search of bibliographies of the publications that were found. In our selection we included published research on peer-reviewed journals and conferences proceedings. The search was performed for English language documents. Finally, we retrieved 100 pertinent papers that met the reported criteria. Figure 1 reports the number and the type of research publications during time.

To estimate an overall performance index for ECG biometrics we computed the weighted mean of identification rate by selecting journal publications (18 were considered appropriate due to data availability, use of ECG feature only, etc.). Weighted equal error rate for verification (authentication) scenarios was also computed. The two indexes were obtained by combining the performance of all the studies. Specifically, single study's performance were weighted (according with the number of subjects involved
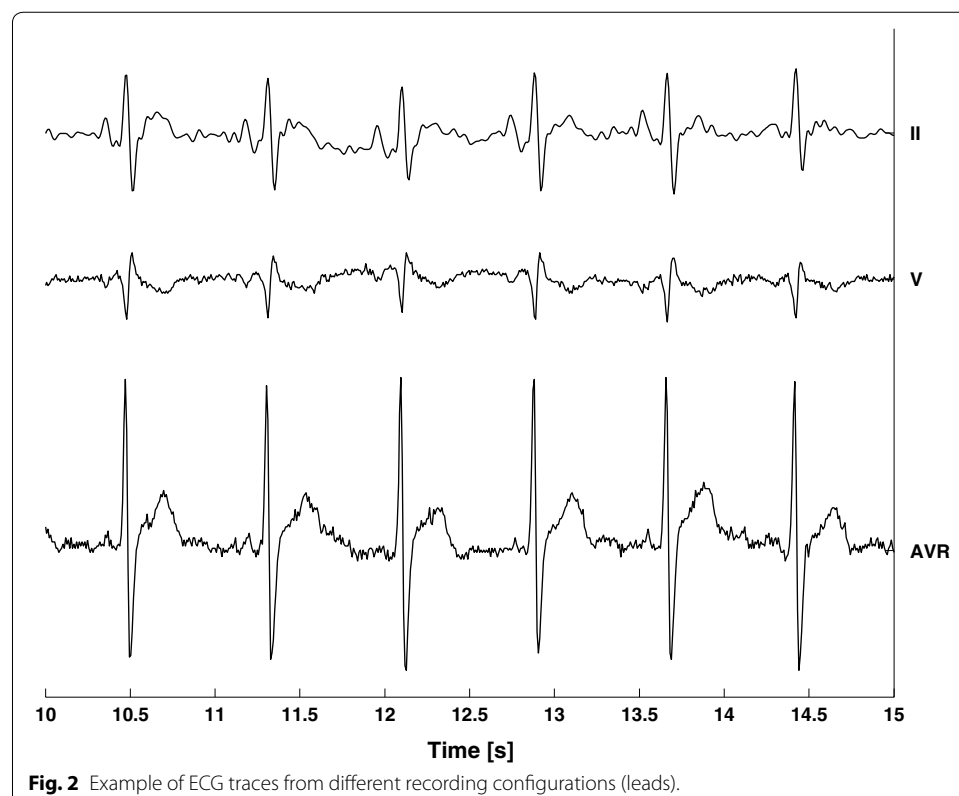

Fig. 1 Temporal increase of the research interests on ECG-based biometric recognition.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 4 of 23

with respect to the total number of subject of the selected papers) and then added to obtain mean overall performance indexes.

ECG as a biometric

ECG is the electrical activity of the heart often recorded at the chest level. During its activity, the myocardium—the heart muscle—behaves as a series of connected electric dipoles in a unique fashion called functional "*syncytium*" [28–31]. Heart's electrical activity is commonly described using an individual time-varying electromagnetic vector [28, 32–36], whose projections can be recorded onto the body surface [32]. Up to twelve specific electrodes positions (leads) are used to monitor heart functions [37, 38], and additional configurations have been proposed for specific purposes [37, 39].

Signals recorded from each lead contain different information; however, specific waves, namely P, Q, R, S, T, can be identified within a heartbeat cycle on the different leads (see Fig. 2). Nevertheless, both the time evolution of the dipole vector and its projections onto the subject's body are influenced by the electrical conduction paths inside the heart, the geometrical characteristics of the heart itself, its position within the chest and also by the inhomogeneity of the conductor volume of the thorax [38].

Therefore, ECG seems to contain enough information for subject recognition. In past studies numerous 'features'—temporal (locations and intervals among waves), amplitude (height of waves' peaks) and morphological differences (shapes, proportions, slopes and angles)—have been proposed to recognise individuals.

Fig. 2 Example of ECG traces from different recording configurations (leads).

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 5 of 23

The hypothesis of the use of ECG as personal identification attribute was suggested by Forsen [40]; however, the first study on the ECG analysis for biometric purposes was carried out by Biel and colleagues [1, 41]: the authors investigated some combinations of features in 12-leads ECG recordings on a sample of 20 subjects.

In the same years moreover, Hoekema and Van Oosterom [42, 43] highlighted and quantified, to some extent, the relevance of the geometrical characteristics in the inter-individual variability of ECG recordings.

Starting from these pioneering studies, ECG biometric literature has grown in the field of pattern recognition (see Fig. 1). Thus, the extraction of appropriate features as well as the classification procedures became both crucial issues.
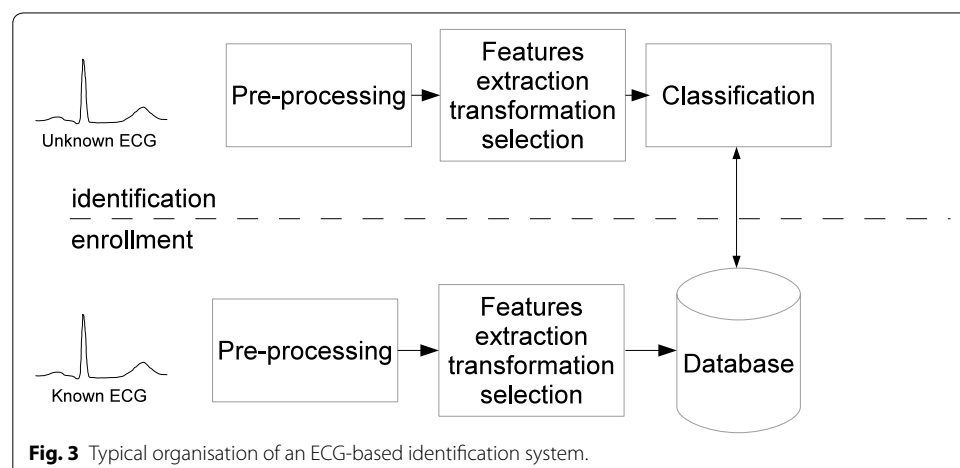
Typical realization of an ECG-based identification system

An ECG-based identification system is characterised by a well-defined workflow as depicted in Fig. 3. It firstly requires an *enrolment* phase, which serves to collect and store subject's distinctive attributes. With the enrolment, specific pre-processing, for noise and artefacts rejection, as well as feature extraction/processing are implemented before the data storage. Once the characteristics of different subjects are stored the *identification* phase can take place. During the identification, in fact, an unknown ECG is presented to the system. As in the enrolment, equal pre-processing and features extraction/transformation are performed. In addition, a specific classification algorithm assigns the extracted features to a *best matching* subject's data as stored in the database (see "Classifiers").

Clearly, ECG attributes extraction, selection and transformation, as well as the classifier structure play a fundamental role to achieve the best identification performances. The following paragraphs report the details of each of the mentioned steps.

ECG features

ECG-based recognition approaches are numerous and very different. ECG attributes (features) are intended to classify the specific subject exploiting inter-subject variability. In general, features are based on the morphology of the heartbeat, on peculiar time intervals derived from ECG waves or on specifically extracted features. The choice of the

Fig. 3 Typical organisation of an ECG-based identification system.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 6 of 23

employed features is generally driven by the complexity of the recognizer, the need of real-time identification, the specific recording device, etc.

No agreement exists on the most appropriate technique or on the type/number of features to consider. Moreover, ECG analysis is often performed on in-house databases making arduous the comparison between techniques.

For the purpose of the present survey, existing approaches have been grouped in two main categories—fiducial based and non-fiducial based—depending on the need to identify precise points in the heartbeat. Each category can be further subdivided as depicted in Fig. 4 by means of the employed features.

In addition, papers in the literature differ as regards the number of the leads used, the recording configuration and the time interval in which the recordings are collected.
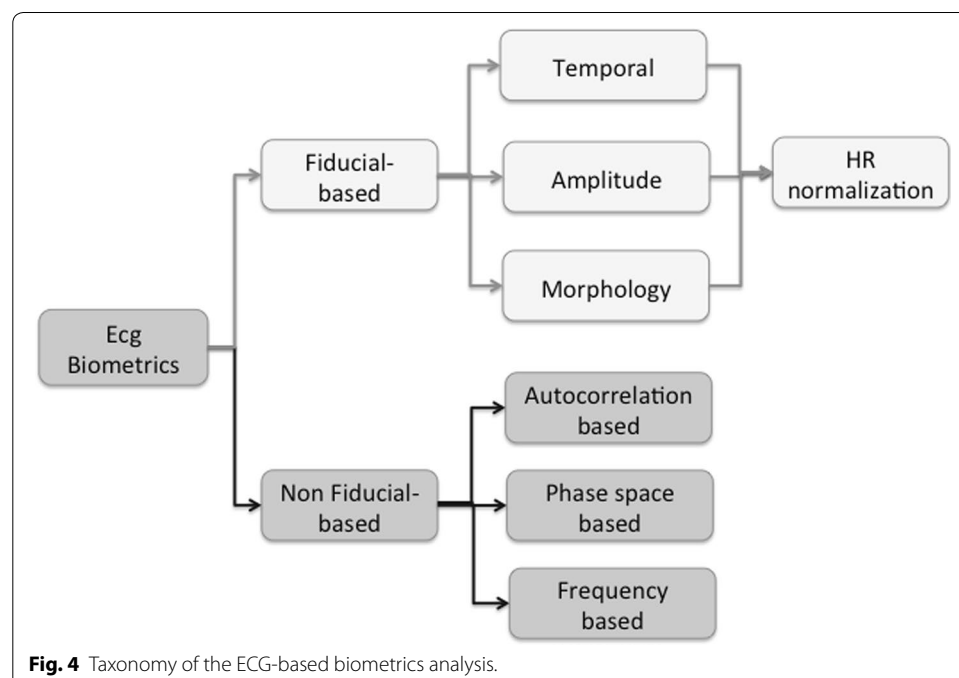
Fiducial based approaches

By locating specific anchor points on the ECG recordings, namely fiducial points or fiducials, numerous features can be extracted and used as recognizer inputs.

Wave's peaks, boundaries, slopes or other measures serve as fiducials. Detectors can use adaptive thresholds [44], Fourier synthesis [45], wavelet transform [46, 47], and other approaches as in [48].

Clearly, the extracted features are strictly influenced by the accuracy of the detection. However, in some case, researchers have limited the number of required fiducials (often to the only R peak identification) [28, 31, 35, 47–63].

Fiducials based features can be further subdivided in temporal, amplitude and morphological. Authors generally use these features in combination.

Temporal features The temporal relationships between the various ECG waves (P, QRS, and T) reflect the epochs of heart's stimulation along its electrical paths starting

Fig. 4 Taxonomy of the ECG-based biometrics analysis.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 7 of 23

from the sino-atrial node to the Purkinje fibres and can be used as biometrics discriminants.

As depicted in Fig. 5, the localisation of specific fiducials allows the computation of several temporal intervals. The most used temporal features include heartbeat wave's duration (i.e. P, QRS, T) and time intervals between them (PQ, RS, ST, etc.) [1, 2, 4, 5]. RR interval has also been used as fiducial [8].

Amplitude features   The inter-individual variability of the amplitude of the heartbeat's waves can be easily recognised in individuals [2, 13, 14]. Amplitude features capture the relative amplitude between the peaks of an ECG's wave; they are generally measured relative to the R peak.

Amplitude features also include the relative ST segment amplitude [1], the amplitude of peaks of 1st or 2nd derivatives of heartbeat [16] and ratios between them [26]. Figure 6 shows an example of some amplitude features and their location along the ECG trace.

Morphological features   Morphological attributes are those carrying information on the shape of the ECG, either as a whole or as its constituent intervals (P-QRS-T). The
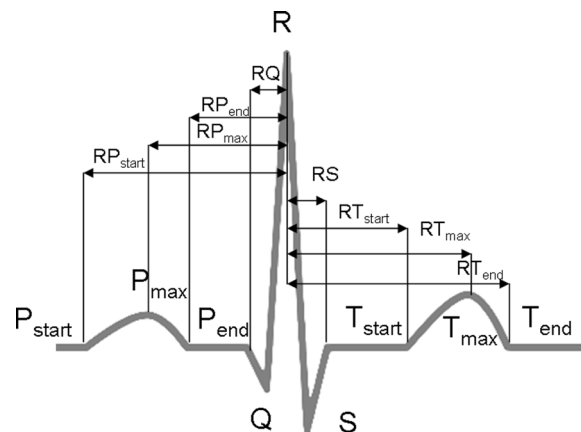

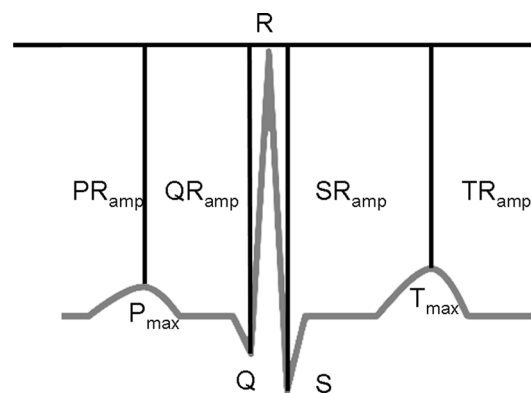**Fig. 5** Temporal features: intervals are obtained by locating specific fiducials along the heartbeat signal.

Fig. 6 Amplitude features: relative amplitudes can be obtained with respect to the R peak

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 8 of 23

simplest way to extract morphological features from a heartbeat is to perform the average of the sampled values of specific intervals (e.g. QRS) with respect to multiple aligned (i.e. centred on R peak) heartbeats [47, 48, 51, 52, 55, 57, 59, 64–67]. In this case, the morphological attributes can be considered as a simple extended set of amplitude features.

However, the study of the ECG's morphology suggested various features: researchers have primarily used slopes among waves, such as ST and RS segment slopes, and angles described by Q, R and S waves [2, 5].

Past studies also concentrate on the specific shape of the whole QRS complex. Palaniappan and Krishnan [8] introduced a peculiar attribute computed from the morphology of the QRS complex. The QRS *form factor* (FF) is defined as in Eq. 1:

$$FF = \sqrt{\frac{\mathrm{var}(x'')\big/\mathrm{var}(x')}{\mathrm{var}(x')\big/\mathrm{var}(x)}} \tag{1}$$

where x is the QRS complex waveform, x' is the first derivative of the QRS and x'' is its second derivative. Biel [1], Shen [26] and Zhang [27] evaluated the suitability of the QRS area. Wübbeler [31], Fang [28] and Sing [30] revealed the characteristic trait of the heart vector by means of studying the whole QRS on different leads.

Polynomial expansions have also been used to synthesize the heartbeat morphology. Khalil and Sufi concentrated on the discrimination feasibility of the Legendre polynomials coefficients derived by fitting the QRS complex [35, 36]. Li [68] used the Hermite interpolation coefficients, Tsao [69] the first order interpolation coefficients while Jun Shen extracted the piecewise linear representation (PLR) coefficients of the entire heartbeat [56]. Specific features (up to 98) were derived via the use of a pulsed triangular wave (Pulse Active Ratio) [70].

The morphology of ECG waves was also investigated by means of wavelet decomposition, frequency analysis and correlation coefficients. Saechia and colleagues focussed on the discriminative characteristics of frequency content of P, QRS and T waves [71]. Hou used the only QRS frequency patterns [72] Tashiro the high frequency components (40–300 Hz) of the entire heartbeat [58]. Lately, Odinaka performed a short time Fourier transform to reveal the spectrogram shape over the heartbeat cycle [54].

Wavelet analysis has been introduced by Chan [48, 49]. Wavelet coefficients were also used by Yao and Wan (Biortogonal wavelets) [47, 59], Chiu (Haar wavelets) [50] and Ye (Daubechies wavelets) [60].

Discrete cosine transform coefficients were also proposed for use in ECG biometrics by Plataniotis et al. [73]. and used by Fattah et al. [74] and Hou et al. [72].

Heart rate: based normalization All the above-described features exhibit a main drawback: the whole heartbeat wave changes with heart rate (HR). This, in turn, induces fluctuations in temporal, amplitude and morphological relationships among waves during physical activity, drug assumption or strong emotions.

Many authors therefore attempted to normalize features with respect to HR changes.

Israel and colleagues [14] assumed the existence of a linear relationship between the heart rate and the features (temporal) computed with respect to the R peak. To remove the dependence on the heart rate, they use a simple normalisation of the features to the

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 9 of 23

length of the entire P-QRS-T complex. Similar to Israel, Kim and colleagues [45] resized the entire heartbeat to a predetermined number of samples n using the inverse Fourier transform.

The approaches adopted by Israel and Kim are however partial. HR influences the timing of heart pulse but it does not homogeneously affect ECG's waves.

Shen [5, 26] and Singh [30, 75–77] utilized Bazett's formula to normalize QT interval, while Tawfik [66, 78] and Sansone [79] applied the regression based approach proposed by Sagie [80].

Fatemian and Hatzinakos addressed the heart rate ECG normalization by resampling the T wave to fit a duration of 120 ms (the typical time extension of the T-wave under rest conditions [38, 81]). After T-wave resampling they combined the obtained segment with P and QRS waves in order to compose the whole heartbeat. Finally, the entire heartbeat is resampled to fit 850 ms [64].

Non-fiducial based approaches

Temporal, amplitude and morphological features require accurate detection of fiducials and the achieved results are clearly dependent on the recognition procedure. To overcome the problem, new approaches, that do not require fiducials recognition, have been reported.

All the described techniques are based on the assumption that the ECG is a highly repetitive (quasi-periodic) signal. Scientific literature analysis revealed various approaches that we subdivide in three main categories: autocorrelation based, phase space based, and frequency based analyses.

Autocorrelation based features   The first attempt to non-fiducial approaches has been proposed by Plataniotis and Hatzinakos [73]. In order to extract valuable attributes, they randomly select 5 s from an ECG recording and estimate the normalized autocorrelation (AC) over a window of m lags (see Eq. 2). AC embeds information about ECG's peculiar characteristics: it is shift invariant and highlights non-random patterns [82]. The QRS complex, in particular, maintains a strong invariance in shape and time width. With this approach, samples that would have to be influenced by fiducials detection, are combined into a sequence of sums of products as given in Eq. 2 reported below:

$$r[m] = \frac{1}{r[0]} \sum_i s[i]s[i+m] \qquad (2)$$

In the equation, $r[m]$ is the AC, $s[i]$ is the signal at time i and m is chosen greater than the mean QRS duration (in samples).

Following this first study, Agrafioti [83–86], Wang [87, 88] and other researchers [81–83] also proposed the use of normalized autocorrelation coefficients.

Phase space analysis   As for AC based approach, the ECG signal can be characterised in a two-dimensional or even three-dimensional space by using the time-delay technique. The analysis of the phase-space trajectory in fact, can highlight unexplored peculiarities of cardiac activity.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 10 of 23

Fang [34] extended the set of features by looking at the three-dimensional vectors of single-lead, time-delayed (4–36 ms), amplitude normalised ECGs $(s(t), s(t + dt), s(t + 2dt))$.

Then, partitioning the phase-space to a $30 \times 30 \times 30$ grid, he reduced the multi loop trajectory to a coarse-grained features space lowering the computational effort and the loop variability due to noise or ECGs' irregularity.

Chen [33], and then Coutinho [51] analysed the ECG by characterising its chaotic behaviour. In detail, Chen and colleagues converted three seconds ECG trace to a phase space-plane by using a time delay equal to 20 samples. Then, they computed concise indicators like correlation dimension and Lyapunov exponents as well as the root mean square of the ECG amplitude.

Conversely, Coutinho analyses the recordings similarities using string matching and parsing algorithms. Each recording is converted by applying an 8-bit uniform quantisation, which produces a sequence of symbols (strings) from an alphabet with 2^8 symbols. Then, a Ziv-Merhav cross-parsing (ZMCP) algorithm [89] is used to evaluate the cross-complexity between different strings. The key idea behind the use of ZMCP is that the cross-complexity becomes lower when the two sequences are similar. Chen et al. explored the use of Lyapunov exponents and correlation dimension [90].

Frequency based features Other authors concentrate on the analysis of ECG's frequency characteristics [91–95].

Loong et al. [93] utilized a linear predictive technique (linear predictive coding or LPC) to model the frequency content of the ECG recordings. In detail, the model of the spectrum for each subject is obtained considering the first forty points of the linear reconstruction of the ECG spectrum using Eq. 3:

$$\hat{x}[n] = -\sum_{i=1}^{p} a_i x[n - i] \tag{3}$$

where the a_i coefficients are evaluated by minimizing the error $e[n]$ (see Eq. 4) using the Levinson–Durbin recursion [96]:

$$e[n] = x[n] - \hat{x}[n] \tag{4}$$

where $x[n]$ represent the actual value.

Kouchaki [92] and Zhao [94] used procedures similar to the Hilbert–Huang transform to obtain instantaneous ECG frequency data. The recordings set is reduced to a collection of functions (namely intrinsic mode functions or IMF) with an adaptive process called *Empirical Mode Decomposition* (EMD) [97]. Each IMF is different among individuals and it is not fixed as in Wavelet or Hermitian expansion. IMFs represent simple oscillatory modes of the system under investigation, which can be characterized by means of the Hilbert spectral analysis.

Kouchaki [92] observed that the slowest component of the EMD carries the most of the discriminative information in comparison with other IMF's.

Zhao and colleagues [94] proposed a modified version of the EMD by using averaged versions of IMFs to raise robustness with respect to noise sources [97]. A consistent

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 11 of 23

number of decompositions is collected for the same ECG by randomly adding white noise to the original signal trait. The corresponding IMFs are then averaged to become noise independent. Finally the spectrum of each IMF is evaluated by Welch analysis. The technique is named *Ensemble Empirical Mode Decomposition.*

Aghakabi [91] and Zokaee [95] use the Mel-Frequency Cepstrum Coefficients (MFCCs) [98]. More in detail, they first select a frame of N samples by using a Hamming window. Then, the Fast Fourier Transform converts the frame from the time domain into frequency domain. A triangular band pass filter bank is applied in the frequency domain to reveal the mel-frequency components [99].

Multi-lead vs single-lead features

Almost all studies reported in this review addressed the problem of subject identification via ECG using single ECG's leads. This is mainly due to the usability of ECG-based identification systems. However, since in pattern recognition problems a larger amount of information could raise the probability of a successful recognition [100], multi-lead systems have also been studied.

Biel investigated amplitude and temporal features in 12-leads ECG recordings [1, 41]. He concluded that a single lead is sufficient in assuring good recognition performances, favouring the practical application of the technique (at least three electrodes are needed).

Following the study of Biel, Zhang and colleagues compared the results obtainable with the use of different leads as a single recording [27]. Based on their outcomes they also concluded that the use of few leads is sufficient and in detail, lead V1 and lead V2 can give the best accuracy. These leads grasp a larger ECG signal with respect to other leads since electrodes are placed closer to the heart site.

Moreover, Agrafioti [83, 84] explored the use of the integration of feature extracted from all of the 12 standard measurement leads. However, he found that information integration raises the identification performance only when combining the outcomes of different classifiers at the decision level. Fang et al. [34] used three leads in his phase space trajectory analysis while obtaining comparable results also with a single lead only. Recently Raj and Hatzinakos [101] studied the feasibility of a specific single-arm single-lead on 23 subjects with discrete results (EER 4–12 %).

Databases

The wide study of ECG signal for clinical purposes favoured the research on the feasibility of ECG as biometric. ECG databases (DBs) have been utilized in the analysis of features and classification performances either public or private (see Table 1). The most used DBs for ECG biometric algorithms testing are available at the Physionet repository [102]. Many researchers used normal and pathological signals DBs: in detail, MIT/BIH's Normal Sinus Rhythm Database [102], MIT/BIH's Arrhytmia Database [103], MIT-BIH Supraventricular Arrhythmia Database [104], QT Database [105], Long Term ST Database [106], European ST-T Database [107], Paroxysmal Atrial Fibrillation Challenge Database [108], PTB Database [109]. Conversely, other papers concentrated on private DBs built by recording ECGs with specific devices [1, 14, 16, 28, 35, 45, 47, 48, 54, 59, 65, 86, 93, 110–116].

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 12 of 23

Table 1 Features of the databases used in past studies

References	NS	NL	NL$_U$	SF (Hz)	SF$_U$	NB	DU (s)	DB
Agrafioti and Hatzinakos [85]	14	12 + 3	1	1,000	1,000	16	na	PTB
Agrafioti and Hatzinakos [83]	13	1	1	128	360	16	10	MIT-BIH-NSR
	30	1		360				MIT-BIH-ARR
	13	12 + 3		1,000				PTB
Biel et al. [1]	20	12	12	na	na	na	na	na
Boumbarov et al. [120]	9	1	1	128	128	12	20	na
Chan et al. [48]	50	1	1	1,000	1,000	12	90	na
Chan et al. [49]	60	1	1	1,000	1,000	12	na	na
Chen et al. [90]	19	1	1	na	na	na	60	na
Fang and Chan [28]	100	2	2	250	250	na	30	na
Fatemian and and Hatzinakos [64]	13	1	1	128	128	16	na	MIT-BIH
	14			1,000	1,000			PTB
Israel et al. [14]	29	1	1	1,000	1,000	na	120	na
Khalil and Sufi [35]	10	1	1	na	na	na	na	na
Kim et al. [45]	10	1	1	200	200	na	30	na
Kyoso and Uchiyama [4]	9	1	1	500	500	12	200	na
Loong et al. [93]	15	1	1	256	256	na	65	na
Lourenco et al. [65]	16	1	1	1,000	1,000	na	120	na
Odinaka et al. [54]	269	1	1	10,000	1,000	na	300	na
Pathoumvanh et al. [132]	10	1	1	500	500	12	200	na
Pereira et al. [112]	77	1	1	256	na	na	600	na
Safie et al. [70]	112	1	1	1,000	1,000	na	30	PTB
Shen and Tompkins [26]	168	1	1	500	500	na	90	MIT-BIH-ARR
Silva et al. [113]	1	1	1	na	na	na	600	na
Singh [77]	50	na	na	na	na	na	na	E-ST
								MIT-BIH
								MIT-BIH-ARR
								MIT-BIH-NSR
								QT
Singh and Gupta [76]	73	na	na	na	na	na	na	E-ST
								MIT-BIH
								MIT-BIH-ARR
								MIT-BIH-NSR
								QT
Sriram et al. [114]	17	1	1	na	na	na	720	na
Tawfik and Kamal [78]	22	1	1	500	500	na	10	na
Wan and Yao [59]	38	1	1	na	na	na	240	na
Wang et al. [87]	13	12 + 3	1	1,000	na	16	na	PTB
	13			128				MIT-BIH
Wubbeler et al. [31]	74	3	2	500	na	16	10	PTB
Yao and Wan [47]	20	1	1	na	na	na	na	na
Zhang and Wei [27]	502	4	1	500	na	na	10	na
Zhao et al. [94]	28	na	na	na	250	na	na	MIT-BIH-ST
	86							Lomg-term-ST
	294							PTB

NS number of subjects, *NL* number of leads (NL$_U$ actually used), *SF* sampling frequency (SF$_U$ actually used), *NB* number of bits per sample, *DU* duration, *DB* database used, *na* indicates that information is not available or computable.

Feature selection

As described in the previous sections, the number of features extracted from the ECG analysis can be very large [1, 28, 34, 52, 60, 74, 93, 112, 117]. However, the information

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 13 of 23

obtained from an extensive set of features is generally redundant. Real-time applications moreover, require a limited number of features in order to allow faster classification. Feature selection is proven to reduce the building and testing time of a classifier by 50 % [118].

The selection of appropriate feature subsets is a critical step in pattern recognition problems; although not all the authors provide feature space dimensionality reduction for ECG based features.

In addition, extracted features are often transformed into new sets by means of linear/nonlinear operators.

Reducing the number of features means identifying the most representative attributes to describe the underlying system/phenomenon.

Israel [14] and Wang [88] used a stepwise canonical correlation [119]. The algorithm starts from one feature adding a new one per iteration; the significance of the features is evaluated by means of Wilks' Lambda distribution.

Other authors used Principal Component Analysis (PCA) [27, 47, 60, 72, 110, 120] or Linear Discriminant Analysis (LDA) [5, 83, 84, 86, 110, 120, 121] as a feature selection procedure.

PCA reduces the feature space dimensionality by performing Eigen-analysis on the covariance matrix of the original features. The covariance matrix S of a set of data x can be computed as reported in Eq. 5:

$$S = \frac{1}{N} \sum_{i=1}^{C} \sum_{j=1}^{C_i} (x_{ij} - \bar{x})(x_{ij} - \bar{x})^T \tag{5}$$

where N is the number of samples, C is the number of classes, C_i is the number of samples in the corresponding class, and $\bar{x} = \frac{1}{N} \sum_{i=1}^{C} \sum_{j=1}^{C_i} x_{ij}$ is the average of the ensemble. The eigenvectors and associated eigenvalues can be then calculated.

Sorting the associated eigenvalues from the highest to the lowest gives the components in order of significance. Thus, ignoring the components with less significance reduces the feature space dimension.

LDA is a different approach to decrease the dimensionality of a feature set. Given a set of labelled (see "Classifiers") samples $\boldsymbol{x}_1,...,\boldsymbol{x}_N$, where $\boldsymbol{x}_j = [x_{j1},...,x_{jp}]$ $j = 1,...,N$ is a vector of p features, the aim of LDA is to project them on a subspace of $M < p$ dimensions producing the best possible separation between classes maximizing the *Fisher's ratio*. In the case of K classes, maximisation of this ratio is equivalent to solving the problem of finding eigenvectors and eigenvalues of the matrix $\boldsymbol{S}_W^{-1}\boldsymbol{S}_B$ and taking the first M larger eigenvectors as the directions of the subspace. As $\boldsymbol{S}_W$ is the within-class covariance matrix, and $\boldsymbol{S}_B$ is the between-class covariance matrix, this criterion is roughly equivalent to searching for the direction from which the classes have well-separated means and small intra-class covariance. It should be emphasised that LDA must be operated on labelled data (i.e. the class of each element must be known in advance) while PCA can be used when labels are not known. Li [68] introduced a generalized LDA (HLDA) to handle heteroscedasticity of classes in mixed ECG and accelerometer data analysis.

Feature reduction has also been obtained using the information gain ratio analysis (IGR) [122–124]. IGR has been utilized in decision trees learning algorithms to select amongst feature while growing the tree [125]. IGR is based on the concept of entropy:

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 14 of 23

if the feature can assume *c* different values, then the entropy of S relative to this *c-wise* classification is defined as in Eq. 6: where p_i is the proportion of S belonging to class *i*.

$$Entropy(S) = \sum_{i=1}^{C} -p_i \log_2 p_i \tag{6}$$

Thus, IGR can be defined as the expected reduction in entropy caused by partitioning the examples according to this attribute. The formal definition of the information gain of a feature A, relative to a collection of examples S, is defined in Eq. 7.

$$IG(S, A) = Entropy(s) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v) \tag{7}$$

Values (A) is the set of all possible values for attribute *A*, and S_v is the subset of S for which attribute A has value v (i.e., $S_v = [s \; \varepsilon \; S | A(s) = v]$). Thus, the features are ranked according to their IGR. The selection algorithm begins with an empty set F of best features and then proceeds to add features from the ranked set of features until the classification accuracy begins to drop or it reaches a specific selected value.

Discrete Fourier Transform (DFT), discrete wavelet transform (DWT) or discrete cosine transform (DCT) have also been used to provide concise set of coefficients. With these approaches, the selected features are the coefficients of the transformation that result significantly different from zero [50, 73, 83, 84, 87].

Classifiers

In pattern recognition problems, classification strategies are the ways with which a vector of analysed features is assigned to a specific subject. Given a vector $\mathbf{x} = [x_1, x_2, ..., x_p]$ composed by features extracted from the ECG of an unknown subject, the aim of the classifier is to assign $\mathbf{x}$ to the correct subject. To this aim, a data-set of samples should be available. The data-set is constituted of labelled feature vectors (x_j, L) where L is the class (subject) label of the sample j = 1,...,N. Typically, a large number of samples should be available per each subject. In order to evaluate the performance of the classifier (error rate) typically the data-set is divided into a training-set and a test-set. The training-set is used for classifier design while the test-set is used for performance assessment (see "Discussion").

Distance based classification

Classifiers can be designed on the basis of several approaches. One common approach leads to assign the unknown sample to the class of the closest sample in the features space. [5, 28, 31, 35, 45, 48–50, 53, 65, 66, 71, 73, 75, 78, 83–85, 87, 88, 95, 111, 112, 114, 122–124, 126–129].

The distance between two feature vectors $\mathbf{x}_1$, $\mathbf{x}_2$ is typically measured using Euclidean norm $d(\mathbf{x}_1, \mathbf{x}_2) = ||\mathbf{x}_1 - \mathbf{x}_2||^2$. A variant of this approach involves preliminary computation of a *template* (or prototype) $\boldsymbol{\mu}_j$ per each class *j* (typically a samples average): the unknown sample is assigned to the class of the closest template (*template matching*).

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 15 of 23

However, Euclidean norm does not account for sample distributions with unequal variances in different directions. In that case, assuming that all classes share a single covariance matrix $\mathbf{\Sigma}$, a better approach could be to use the minimum Mahalanobis' distance $d(x, \mu_i) = (x - \mu_i)^T \Sigma^{-1}(x - \mu_i)$ between the unknown vector $\mathbf{x}$ and $\mathbf{\mu}_i$ is the mean (*template*) of the *ith* class [16].

Another common approach is to construct N 'discriminant functions' $g_i(\mathbf{x})$, i = 1...N, one per each class: the unknown vector is assigned to class *k* if $g_k(\mathbf{x}) > g_i(\mathbf{x})$ i = 1...N. The discriminant functions can be constructed, for example, using a Bayesian approach. As known, the Bayes formula relates the a posteriori probability that to the a priori distribution of classes and to the *likelihood* of the features, given a specified class *k* (see Eq. 8)

$$p(\omega_k|x) = \frac{p(x|\omega_k)p(\omega_k)}{p(x)}. \tag{8}$$

Using this approach with a Gaussian distribution of features, we obtain the discriminant function [27] reported in (9):

$$g_i(x) = -\frac{1}{2}(x - \mu_i)^T \sum_i^{-1}(x - \mu_i) - \frac{1}{2}\ln|\Sigma_i| \tag{9}$$

where $\mathbf{x}$ is an unknown feature vector, $\mathbf{\Sigma}_i$ is the covariance matrix of the features of *i*th class, $\mathbf{\mu}_i$ is the template of the *i*th subject.

Another comparison metric is the *maximum correlation* between two signals (vectors) [64], which exhibits the amount of similarity between two signals. The correlation is defined as in Eq. 10

$$\rho_{xy}[m] = \frac{\sum_{i=0}^{N-|m|-1} x[i]y[i+m]}{\sqrt{E_x E_y}} \tag{10}$$

where $x[i]$ and $y[i]$ $i = 1,...,P$, represent two different ECG signal windows of length *N*, $E_x = \sum_{i=0}^{N} x^2[i]$ is the energy of the signal. It achieves the maximum value for m = 0.

Israel [14] and Shen [5] used the LDA technique [130]. LDA was born as a method for dimensionality reduction; however it can also be used for classifier design.

Chan [48] used three different distances: percent residual difference (PRD) that quantifies the amount of differences between two ECG with respect to the variability contained in the unknown ECG; correlation coefficient (CCORR) that measures the least squares fitting of the two ECG to be compared and WDIST that measures the difference between the discrete wavelet coefficients of the two ECGs.

Neural networks

In the context of subject recognition by ECG processing, Neural Networks (NN) have been used in [8, 33, 58, 59, 66, 69, 71, 78, 90, 93, 120, 124, 131]. The most successful NN is the Multi Layer Perceptron (MLP). In contrast to the conventional approaches seen in the previous section, MLP is capable to solve complex non-linear classification problems.
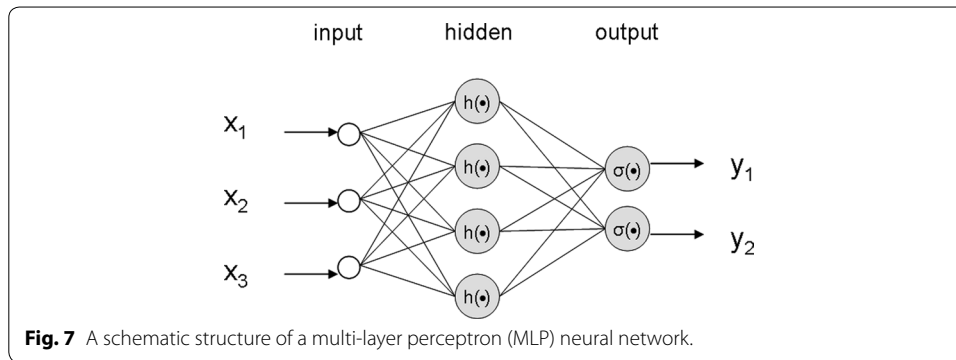
Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 16 of 23

Fig. 7 A schematic structure of a multi-layer perceptron (MLP) neural network.

Figure 7 depicts a typical MLP: each node (i.e. neuron) has a non-linear activation function acting upon the (features) inputs $\mathbf{x} = [x_1,...,x_p]$: in particular, each input has a (synaptic) weight $\mathbf{w}_j^{\mathrm{T}} = [w_{j1},...,w_{jp}]$, in order that the output z_j of the neuron j is the action of the activation function on a linear combination as $z_j = h(\mathbf{w}_j^{\mathrm{T}} \mathbf{x})$. Therefore, the output of the system is given by $z_j = \sigma(\mathbf{w}_j^{(2)\mathrm{T}} h(\mathbf{w}_j^{(1)\mathrm{T}} \mathbf{x}))$ where the superscript (1) refer to the first (hidden) layer of neurons, and the superscript (2) refer to the second (output) layer of neurons.

While a single layer NN can solve only linearly separable problems (sample are separable via a hyperplane in the features space), a hidden layer is typically added to give the MLP the capability to solve non-linearly separable problems (Fig. 7).

Non-conventional approaches

Wang et al. [87] used nearest neighbour (NN) classifiers in combination with Euclidean distance. However, they used a hierarchical approach and features integration. More importantly, they introduced a new approach in this field: the hierarchical approach that divides the problem in two sub problems (large class to small class problem): first, they used a first-level classification based on analytic features only (time + amplitude of fiducial points), then used PCA based classification module to classify subjects that can be potentially confused by the first stage!!

Singh and Singh used a different approach [30]. Per each subject j in the database they constructed a Pattern Matrix $\mathbf{P}^{(j)}$ in the following way. Per each subject they 'arbitrarily' selected a number m of heartbeats. Per each heartbeat they extracted p features $x_k^{(j)} = \left[x_{k1}^{(j)},\dots,x_{kp}^{(j)}\right]$ where $k = 1,...,m$: the latter are the rows of the Pattern Matrix. Given an unknown vector $\mathbf{x}$ the distance score measure $s^{(j)} = \frac{1}{m}\sum_{k=1}^{m} s_k^{(j)} = \frac{1}{m}\sum_{k=1}^{m}\sum_{r=1}^{p}\left|x_{kr}^{(j)} - x_r\right|$ a smaller score indicates a good match.

Discussion

The field of biometric recognition via the use of ECG characteristics is certainly engaging and results seem encouraging. According with Table 2 the weighted mean identification rate is equal to 94.95 % and the overall equal error rate (in an authentication scenario) to 0.92 %.

Our results pointed out that subject identification depends primarily on the choice of the utilised feature(s). Fiducial based approaches benefit of well-established

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 17 of 23

Table 2 Estimation of the overall performance of the use of ECG as biometric

References	IR (%)	AEER	IRW	AW	WIR (%)	WA (%)
Agrafioti and Hatzinakos [85]	100.00	na	0.010	na	0.61	na
Agrafioti and Hatzinakos [83]	96.20	0.87	0.020	0.10	2.35	0.08
Biel et al. [1]	98.00	na	0.010	na	0.86	na
Boumbarov et al. [120]	86.11	na	0.000	na	0.34	na
Chan et al. [48]	89.00	na	0.020	na	1.94	na
Chan et al. [49]	100.00	na	0.030	na	2.62	na
Chen et al. [90]	91.20	na	0.010	na	0.76	na
Fang and Chan [28]	95.00	na	0.040	na	4.15	na
Fatemian and Hatzinakos [64]	99.63	na	0.010	na	1.17	na
Israel et al. [14]	100.00	na	0.010	na	1.27	na
Khalil and Sufi [35]	na	na	na	na	na	na
Kim et al. [45]	Na	na	na	na	na	na
Kyoso and Uchiyama [4]	94.20	na	0.000	na	0.37	na
Loong et al. [93]	100.00	na	0.010	na	0.66	na
Lourenco et al. [65]	94.30	10.10	0.010	0.03	0.66	0.28
Odinaka et al. [54]	99.00	0.37	0.012	0.46	11.63	0.17
Pathoumvanh et al. [132]	97.00	na	0.000	na	0.42	na
Pereira et al. [112]	99.00	0.70	0.030	0.13	3.33	0.09
Safie et al. [70]	93.60	na	0.050	na	4.58	na
Shen and Tompkins [26]	95.30	na	0.070	na	6.99	na
Silva et al. [113]	na	na	na	na	na	
Singh [77]	82.00	0.10	0.030	0.13	2.61	0.01
Singh and Gupta [76]	99.00	na	0.020	na	2.16	na
Sriram et al. [114]	97.00	15.00	0.010	0.03	0.72	0.44
Tawfik and Kamal [78]	99.08	na	0.010	na	0.95	na
Wan and Yao [59]	100.00	na	0.020	na	1.66	na
Wang et al. [87]	100.00	na	0.010	na	1.14	na
Wubbeler et al. [31]	99.00	0.03	0.030	0.13	3.20	0.00
Yao and Wan [47]	91.48	na	0.010	na	0.80	na
Zhang and Wei [27]	97.40	na	0.220	na	21.35	na
Zhao et al. [94]	95.00	na	0.180	na	16.93	na
Total					96.22	1.16

IR identification rate, *AEER* equal error rate for authentication scenarios, *IRW* IR weights, *AW* AEER weights, *WIR* weighted IR, *WA* weighted AEEE, *na* indicates that information is not not available or not computable.

normalization algorithms to compensate for changes in ECG signal due to the heart rate variability [111, 132] but they are commonly affected by the performance of the fiducials detection algorithms.

Non fiducial based approaches, conversely, offer a promising alternative to reduce error rate and computational effort. They do not require the identification of ECG waves and have the advantage to potentially take into account *fine* features which could be lost using fiducials.

The choice of number of leads influences the recognizer outcomes. Many authors investigated the performances of multi lead systems, others have also investigated ECG related signals [20, 22, 133]; however, we believe that the feasibility of biometric identification via ECG analysis could only be obtained by limiting the number of required leads/ signals: to this regard, single ECG lead systems would be desirable.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 18 of 23

The time span in which the selected features are effective is also an important matter of research. Most of the published studies considered the capacity of a system to identify a subject only at a specific time; few papers addressed the variability of ECG features with time or with physiological (e.g. subject's aging, stress, activity, etc.) or pathological conditions [79, 134–140].

Lastly, studies generally rely on post hoc analyses, although some real time examples have been reported [136, 141, 142].

Conclusions

Four major issues must be highlighted regarding the adequacy of the studies conducted so far. First, while a great effort has been spent in feature selection and classifier design, it is not yet clear what is the best set of features and classification scheme for ECG biometrics (hierarchical, ensemble etc.). Non-fiducial based techniques can reduce the computational effort as well as the error rate due to the ECG waves recognition. Therefore, it is expected that the new techniques to be developed will use fiducials and non-fiducial based features in order to catch the best of both approaches. Further analysis should be addressed on the use of single lead recordings and the study of features which are not dependent on the recording sites (e.g. fingers, hand palms).

Second, as regards the population size, the majority of the studies have been conducted on a small population (about a few tens of subjects). Therefore, the applicability of ECG biometric recognition on a large scale (real life authentication scenario) it is not yet proven.

Third, almost all studies (except for [17] and [31]) ignored the variability of the ECG during life span (i.e. variability induced by work, ageing, iterate sport activity etc.); moreover, only few studies [57, 83, 136] considered the applicability of these techniques when subjects suffer from pathological conditions. ECG recognition in pathological subjects is another aspect worth of additional investigations.

Fourth, it must be emphasised that, while guidelines are available for ECG acquisition in the clinical scenario, there is still a lack of standardisation on ECG acquisition (number of leads and their positioning, sampling frequency, number of bits, filtering, type of electrodes, number of leads etc.) for biometrics applications. However, ECG databases for biometric recognition should ideally include recordings, at a given sampling frequency and conditions, from the same subjects in different circumstances (e.g. relaxed, during and after physical training) and along a period of several years.

If addressed, the mentioned challenges will contribute to move this promising technique from its state of adolescence to a proper daily life adoption.

Author details
[1] School of Life and Health Sciences, Aston University, Aston Triangle, B4 7ET Birmingham, UK. [2] Department of Electronic Engineering and Information Technologies, University "Federico II" of Naples, Via Claudio, 21, 80125 Naples, Italy.

Compliance with ethical guidelines

Competing interests
The authors declare that they do not have competing interests.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 19 of 23

References

1. Biel L, Pettersson O, Philipson L, Wide P. ECG analysis: a new approach in human identification. IEEE Trans Instrum Meas. 2001;50(3):808–12 **(IEEE)**.
2. Israel SA, Scruggs WT, Worek WJ, Irvine JM. Fusing face and ECG for personal identification. In: Proceedings of the 32nd applied imagery pattern recognition workshop, 2003. IEEE. 2003. p. 226–31.
3. Kim KH, Bang SW, Kim SR. Emotion recognition system using short-term monitoring of physiological signals. Med Bio Eng Comput. 2004;42(3):419–27 **(Springer)**.
4. Kyoso M, Uchiyama A (2001) Development of an ECG identification system. In: IEMBS-01, vol 4, IEEE. 2001. p. 3721–3.
5. Shen T-W, Tompkins WJ, Hu YH. Implementation of a one-lead ECG human identification system on a normal population. J Eng Comput Innov. 2011;2(1):12–21.
6. Ballard L, Lopresti D, Monrose F. Forgery quality and its implications for behavioral biometric security. IEEE Trans Syst Man Cybern Part B Cybern. 2007;37(5):1107–18 **(IEEE)**.
7. Määttä J, Hadid A, Pietikäinen M. Face spoofing detection from single images using texture and local shape analysis. Biom IET. 2012;1(1):3–10.
8. Palaniappan R, Krishnan SM. Identifying individuals using ECG beats. In: International conference on signal processing and communications, 2004 SPCOM '04, 2004. 2004. p. 569–72.
9. Singh YN, Singh SK. Vitality detection from biometrics: state-of-the-art. In: World congress on information and communication technologies (WICT), 2011. 2011. p. 106–11.
10. Van der Putte T, Keuning J. Biometrical fingerprint recognition: don't get your fingers burned. USA: Springer; 2000. p. 289–303.
11. Biggio B, Akhtar Z, Fumera G, Marcialis GL, Roli F. Security evaluation of biometric authentication systems under real spoofing attacks. Iet Biom. 2012;1(1):11–24 **(IET Digital Library)**.
12. Galbally J, Alonso-Fernandez F, Fierrez J, Ortega-Garcia J. A high performance fingerprint liveness detection method based on quality related features. Future Gener Comput Syst. 2012;28(1):311–21.
13. Hoekema R, Uijen GJ, van Oosterom A. Geometrical aspects of the interindividual variability of multilead ECG recordings. IEEE Trans Biomed Eng. 2001;48(5):551–9.
14. Israel SA, Irvine JM, Cheng A, Wiederhold MD, Wiederhold BK. ECG to identify individuals. Pattern Recognit. 2005;38(1):133–42.
15. Carreiras C, Lourenço A, Fred A, Ferreira R. ECG signals for biometric applications: are we there yet? In: ICINCO 2014—Proceedings of the 11th international conference on informatics in control, automation and robotics. 2014. p. 765–72.
16. Kyoso M. A technique for avoiding false acceptance in ECG identification. In: IEEE EMBS asian-pacific conference on biomedical engineering, 2003. 2003. p. 190–1.
17. Odinaka I, Lai P-H, Kaplan AD, O'Sullivan JA, Sirevaag EJ, Rohrbaugh JW. ECG biometric recognition: a comparative analysis. IEEE Trans Inf Forensics Secur. 2012;7(6):1812–24.
18. Nasri B, Guennoun M, El-Khatib K. Using ECG as a measure in biometric identification systems. In: Science and technology for humanity (TIC-STH), 2009 IEEE Toronto international conference. 2009. p. 28–33.
19. Israel SA, Irvine JM. Heartbeat biometrics: a sensing system perspective. IJCB. 2012;1(1):39.
20. Abo-Zahhad M, Ahmed SM, Abbas SN. Biometric authentication based on PCG and ECG signals: present status and future directions. Signal Image Video Process. 2014;8(4):739–51.
21. Bugdol MD, Mitas AW. Multimodal biometric system combining ECG and sound signals. Pattern Recognit Lett. 2014;38(1):107–12.
22. Odinaka I, O'Sullivan JA, Sirevaag EJ, Rohrbaugh JW. Cardiovascular biometrics: combining mechanical and electrical signals. IEEE Trans Inf Forensics Secur. 2015;10(1):16–27.
23. Mishra P, Singla SK. Artifact removal from biosignal using fixed point ICA algorithm for pre-processing in biometric recognition. Meas Sci Rev. 2013;13(1):7–11.
24. Sidek KA, Khalil I. Enhancement of low sampling frequency recordings for ECG biometric matching using interpolation. Computer Methods Programs Biomed. 2013;109(1):13–25 **(Elsevier)**.
25. Kohler BU, Hennig C, Orglmeister R. The principles of software QRS detection. Eng Med Biol Mag IEEE. 2002;21(1):42–57.
26. Shen T, Tompkins W. Biometric statistical study of one-lead ECG features and body mass index (BMI). Conf Proc IEEE Eng Med Biol Soc. 2005;2:1162–5.
27. Zhang Z, Wei D. A new ECG identification method using Bayes' teorem. In: TENCON 2006, 2006 IEEE region 10 conference. 2006. p. 1–4.
28. Fang S-C, Chan H-L. Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space. Pattern Recognit. 2009;42(9):1824–31.
29. Hurst JW. Naming of the waves in the ECG, with a brief account of their genesis. Circulation. 1998;98(18):1937–42.
30. Singh YN, Gupta P. Biometrics method for human identification using electrocardiogram. In: Advances in biometrics. New York: Springer; 2009. p. 1270–9.
31. Wübbeler G, Stavridis M, Kreiseler D, Bousseljot R-D, Elster C. Verification of humans using the electrocardiogram. Pattern Recognit Lett. 2007;28(10):1172–5 **(Elsevier)**.
32. Burch GE. The history of vectorcardiography. Med Hist Suppl. 1985;5:103–31.
33. Chen C-K, Lin C-L, Chiu Y-M. Individual identification based on chaotic electrocardiogram signals. In: 2011 6th IEEE conference on industrial electronics and applications (ICIEA). IEEE. 2011. p. 1771–6.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 20 of 23

34. Fang S-C, Chan H-L. QRS detection-free electrocardiogram biometrics in the reconstructed phase space. Pattern Recognit Lett. 2013;34(5):595–602.

35. Khalil I, Sufi F. Legendre polynomials based biometric authentication using QRS complex of ECG. In: International conference on intelligent sensors, sensor networks and information processing, 2008 ISSNIP 2008. IEEE. 2008. p. 297–302.

36. Sufi F, Khalil I. An automated patient authentication system for remote telecardiology. In: International conference on intelligent sensors, sensor networks and information processing, 2008 ISSNIP 2008. IEEE. 2008. p. 279–84.

37. Gargiulo GD, McEwan AL, Bifulco P, Cesarelli M, Jin C, Tapson J. Towards true unipolar ECG recording without the Wilson central terminal (preliminary results). Physiol Meas. 2013;34(9):991–1012.

38. Morris F, Brady WJ, Camm J. ABC of clinical electrocardiography. New York: Wiley; 2009. p. 112.

39. Gargiulo GD, McEwan AL, Bifulco P, Cesarelli M, Jin C, Tapson J, et al. Towards true unipolar bio-potential recording: a preliminary result for ECG. Physiol Meas. 2013;34(1):N1–7.

40. Forsen GE, Nelson MR, Staron RJJ. Personal attributes authentication techniques. 1977 (**In**).

41. Biel L, Pettersson O, Philipson L, Wide P. ECG analysis: a new approach in human identification. In: Proceedings of the 16th IEEE instrumentation and measurement technology conference, 1999 IMTC/99, vol 1. 1999. p. 557–61.

42. Hoekema R, Uijen GJH, van Oosterom A. Geometrical aspects of the inter-individual variability of multilead ECG recordings. Comput Cardiol. 1999. p. 499–502.

43. van Oosterom A, Hoekema R, Uijen GJH. Geometrical factors affecting the interindividual variability of the ECG and the VCG. J Electrocardiol. 2000;33:219–27.

44. Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Trans Biomed Eng BME. 1985;32(3):230–6.

45. Kim K-S, Yoon T-H, Lee J-W, Kim D-J, Koo H-S. A robust human identification by normalized time-domain features of electrocardiogram. In: 27th annual international conference of the engineering in medicine and biology society, 2005 IEEE-EMBS 2005. 2005. p. 1114–7.

46. Almeida R, Nez JPM, Rocha AP, Laguna P. Multilead ECG delineation using spatially projected leads from wavelet transform loops. IEEE Trans Biomed Eng. 2009;56(8):1996–2005.

47. Yao J, Wan Y. A wavelet method for biometric identification using wearable ECG sensors. In 5th international summer school and symposium on medical devices and biosensors, 2008 ISSS-MDBS 2008. 2008. p. 297–300.

48. Chan ADC, Hamdy MM, Badre A, Badee V. Wavelet distance measure for person identification using electrocardiograms. IEEE Trans Instrum Meas. 2008;57(2):248–53.

49. Chan ADC, Hamdy MM, Badre A, Badee V. Person identification using electrocardiograms. In: Canadian conference on electrical and computer engineering, 2006 CCECE '06. 2006. p. 1–4.

50. Chiu C-C, Chuang C-M, Hsu C-Y. A novel personal identity verification approach using a discrete wavelet transform of the ECG signal. In: International conference on multimedia and ubiquitous engineering, 2008 MUE. IEEE. 2008. p. 201–6.

51. Coutinho DP, Fred ALN, Figueiredo MAT. One-lead ECG-based personal identification using Ziv-Merhav cross parsing. In: 20th international conference on pattern recognition (ICPR), 2010. IEEE. 2010. p. 3858–61.

52. Fratini A, Sansone M, Bifulco P, Romano M, Pepino A, Cesarelli M, et al. Individual identification using electrocardiogram morphology. In: IEEE international symposium on medical measurements and applications proceedings (MeMeA), 2013. 2013. p. 107–10.

53. Lourenco A, Silva H, Fred A. ECG-based biometrics: a real time classification approach. In: IEEE international workshop on machine learning for signal processing (MLSP), 2012. 2012. p. 1–6.

54. Odinaka I, Lai P-H, Kaplan AD, O'Sullivan JA, Sirevaag EJ, Kristjansson SD, et al. ECG biometrics: a robust short-time frequency analysis. In: IEEE international workshop on information forensics and security (WIFS), 2010. IEEE. 2010. p. 1–6.

55. Poree F, Gallix A, Carrault G (2011) Biometric identification of individuals based on the ECG. Which conditions? Comput Cardiol 1999. p. 761–4 **(IEEE)**.

56. Shen J, Bao S-D, Yang L-C, Li Y. The PLR-DTW method for ECG based biometric identification. In: 2011 33rd annual international conference of the IEEE engineering in medicine and biology society. IEEE. 2011. p. 5248–51.

57. Sidek KA, Khalil I. Person identification in irregular cardiac conditions using electrocardiogram signals. In: Engineering in medicine and biology society, EMBC, 2011 annual international conference of the IEEE. 2011. p. 3772–5.

58. Tashiro F, Aoyama T, Shimuta T, Ishikawa H, Shimatani Y, Ishijima M, et al. Individual identification with high frequency ECG: preprocessing and classification by neural network. Conf Proc IEEE Eng Med Biol Soc. 2011;2011:2749–51.

59. Wan Y, Yao J. A neural network to identify human subjects with electrocardiogram signals. In: Proceedings of the world congress on engineering and computer science 2008, WCECS 2008. 2008.

60. Ye C, Coimbra MT, Kumar BVKV. Investigation of human identification using two-lead electrocardiogram (ECG) signals. In: Fourth IEEE international conference on biometrics: theory applications and systems (BTAS), 2010. IEEE. 2010. p. 1–8.

61. Zhao Z, Yang L. ECG identification based on matching pursuit. In: 4th international conference on biomedical engineering and informatics (BMEI), 2011, vol 2. 2011. p. 721–4.

62. Zheng G, Li Z-Y, Liu T-T, Dai M, editors. Study of human identification by electrocardiogram waveform morph. Berlin, Heidelberg: Springer; 2011.

63. Tantawi M, Salem A, Tolba MF. Fiducial based approach to ECG biometrics using limited fiducial points. Commun Comput Inf Sci. 2014. p. 199–210.

64. Fatemian SZ, Hatzinakos D. A new ECG feature extractor for biometric recognition. In: 16th international conference on digital signal processing, 2009. 2009. p. 1–6.

65. Lourenço A, Silva H, Fred A. Unveiling the biometric potential of finger-based ECG signals. Comput Intell Neurosci. 2011;1:1–8 **(Hindawi Publishing Corp.)**.

66. Tawfik MM, Selim H, Kamal T. Human identification using time normalized QT signal and the QRS complex of the ECG. In: 7th international symposium on communication systems networks and digital signal processing (CSNDSP), 2010. 2010. p. 755–9.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 21 of 23

67. Lourenco A, Silva H, Santos DP, Fred A. Towards a finger based ECG biometric system. Biosignals. 2011;2011:348–53.
68. Li M, Rozgić V, Thatte G, Lee S, Emken BA, Annavaram M, et al. Multimodal physical activity recognition by fusing temporal and cepstral information. IEEE Trans Neural Syst Rehabil Eng. 2010;18(4):369–80.
69. Tsao Y-T, Shen T-W, Ko T-F, Lin T-H. The morphology of the eectrocardiogram for evaluating ECG biometrics. In: 9th international conference on e-health networking, application and services, 2007. 2007. p. 233–5.
70. Safie SI, Soraghan JJ, Petropoulakis L. Electrocardiogram (ECG) biometric authentication using pulse active ratio (PAR). IEEE Trans Inf Forensics Secur. 2011;6(4):1315–22.
71. Saechia S, Koseeyaporn J, Wardkein P. Human identification system based ECG signal. In: TENCON 2005, 2005 IEEE Region, vol 10. 2005. p. 1–4.
72. Hou LS, Subari KS, Syahril S. QRS-complex of ECG-based biometrics in a two-level classifier. In: TENCON 2011, IEEE region 10 conference. 2011. p. 1159–63.
73. Plataniotis KN, Hatzinakos D, Lee JKM. ECG biometric recognition without fiducial detection. In: Biometrics symposium: special session on research at the biometric consortium conference, 2006. IEEE. 2006. p. 1–6.
74. Fattah SA, Jameel ASMM, Goswami R, Saha SK, Syed N, Akter S, et al. An approach for human identification based on time and frequency domain features extracted from ECG signals. In: TENCON 2011, 2011 IEEE region 10 conference. IEEE. 2011. p. 259–63.
75. Singh YN, Gupta P. ECG to individual identification. In: 2008 IEEE second international conference on biometrics: theory, applications and systems. IEEE. 2008. p. 1–8.
76. Singh YN, Gupta P. Correlation-based classification of heartbeats for individual identification. Soft Comput. 2011;15(3):449–60 **(Springer)**.
77. Singh YN. Evaluation of electrocardiogram for biometric authentication. JIS. 2012;03(01):39–48.
78. Tawfik MM, Kamal HST. Human identification using QT signal and QRS complex of the ECG. Online J Electron Electr Eng (OJEEE). 2011;3:1–5.
79. Sansone M, Fratini A, Cesarelli M, Bifulco P, Pepino A, Romano M, et al. Influence of QT correction on temporal and amplitude features for human identification via ECG. In: 2013 IEEE workshop on biometric measurements and systems for security and medical applications (BIOMS). 2013. p. 22–7.
80. Sagie A, Larson MG, Goldberg RJ, Bengtson JR, Levy D. An improved method for adjusting the QT interval for heart rate (the Framingham Heart Study). Am J Cardiol. 1992;70(7):797–801.
81. Sörnmo L, Laguna P. Bioelectrical signal processing in cardiac and neurological applications. Amsterdam: Elsevier; 2005.
82. Fratini A, Bifulco P, Romano M, Clemente F, Cesarelli M. Simulation of surface EMG for the analysis of muscle activity during whole body vibratory stimulation. Comput Methods Programs Biomed. 2013;113(1):314–22.
83. Agrafioti F, Hatzinakos D. ECG biometric analysis in cardiac irregularity conditions. SIViP. 2008;3(4):329–43.
84. Agrafioti F, Hatzinakos D. ECG based recognition using second order statistics. In: 2008 6th annual communication networks and services research conference (CNSR). IEEE. 2008. p. 82–7
85. Agrafioti F, Hatzinakos D. Fusion of ECG sources for human identification. In: 3rd international symposium on communications, control and signal processing, 2008 ISCCSP. 2008. p. 1542–7.
86. Agrafioti F, Hatzinakos D. Signal validation for cardiac biometrics. In: 2010 IEEE international conference on acoustics speech and signal processing (ICASSP). IEEE. 2010. p. 1734–7.
87. Wang Y, Agrafioti F, Hatzinakos D, Plataniotis KN. Analysis of human electrocardiogram for biometric recognition. EURASIP J Adv Signal Process. 2008;1:148658.
88. Wang Y, Plataniotis KN, Hatzinakos D. Integrating analytic and appearance attributes for human identification from ECG signals. In: Special session on research at the biometric consortium conference, 2006 biometrics symposium. 2006. p. 1–6.
89. Ziv J, Merhav N. A measure of relative entropy between individual sequences with application to universal classification. IEEE Trans Inf Theory. 1993;39(4):1270–9.
90. Chen C-K, Lin C-L, Lin S-L, Chiu Y-M, Chiang C-T. A chaotic theoretical approach to ECG-based identity recognition [application notes]. IEEE Comput Intell Mag. 2014;9(1):53–63.
91. Aghakabi A, Zokaee S. Fusing dorsal hand vein and ECG for personal identification. In: 2011 international conference on electrical and control engineering (ICECE). 2011. p. 5933–6.
92. Kouchaki S, Dehghani A, Omranian S, Boostani R. ECG-based personal identification using empirical mode decomposition and Hilbert transform. In: 2012 16th CSI international symposium on artificial intelligence and signal processing (AISP). 2012. p. 569–73.
93. Loong JLC, Subari KS, Besar R, Abdullah MK. A new approach to ECG biometric systems: a comparative study between LPC and WPD systems. World Acad Sci Eng Technol. 2010;68:759–64.
94. Zhao Z, Yang L, Chen D, Luo Y. A human ECG Identification system based on ensemble empirical mode decomposition. Sensors. 2013;13(5):6832–64.
95. Zokaee S, Faez K. Human identification based on ECG and palmprint. Int J Electr Comput Eng (IJECE). 2012;2(2):261–6.
96. Morettin PA. The Levinson algorithm and its applications in time series analysis. In: International statistical review/revue internationale de statistique, vol 52. International Statistical Institute (ISI). 1984. p. 83–92 **(vol 1)**.
97. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc A Math Phys Eng Sci. 1998;454(1971):903–95.
98. IMAI S. Cepstral analysis synthesis on the mel frequency scale. In: Acoustics, speech, and signal processing, IEEE international conference on ICASSP '83, vol 8. IEEE Press Request Permissions. 1983. p. 93–6 **(vol 3)**.
99. Stevens SS, Volkmann J, Newman EB. A scale for the measurement of the psychological magnitude of pitch. J Acoust Soc Am. 1937;8:185–90.
100. Ergin S, Uysal AK, Gunal ES, Gunal S, Gulmezoglu MB. ECG based biometric authentication using ensemble of features. In: Iberian conference on information systems and technologies, CISTI. 2014.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 22 of 23

101. Raj PS, Hatzinakos D. Feasibility of single-arm single-lead ECG biometrics. In: European signal processing conference. 2014. p. 2525–9.
102. Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG. PhysioBank, PhysioToolkit, and PhysioNet : components of a new research resource for complex physiologic signals. Circulation. 2000;101(23):e215–20.
103. Moody GB, Mark RG. The impact of the MIT-BIH arrhythmia database. Eng Med Biol Mag IEEE. 2001;20(3):45–50.
104. Greenwald SD, Patil RS, Mark RG. Improved arrhythmia detection in noisy ECGs through the use of expert systems. In: Proceedings of the computers in cardiology, 1988. 1988. p. 163–5.
105. Laguna P, Mark RG, Goldberg A, Moody GB. A database for evaluation of algorithms for measurement of QT and other waveform intervals in the ECG. Comput Cardiol. 1997;1997:673–6.
106. Jager F, Taddei A, Moody GB, Emdin M, Antolič G, Dorn R. Long-term ST database: a reference for the development and evaluation of automated ischaemia detectors and for the study of the dynamics of myocardial ischaemia. Med Bio Eng Comput. 2003;41(2):172–82 **(Springer)**.
107. Taddei A, Distante G, Emdin M, Pisani P, Moody GB, Zeelenberg C, et al. The European ST-T database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. Eur Heart J. 1992;13(9):1164–72.
108. Moody GB, Goldberger A, McClennen S, Swiryn S. Predicting the onset of paroxysmal atrial fibrillation: the computers in cardiology challenge 2001. Comput Cardiol. 2001;2001:113–6.
109. Bousseljot R, Kreiseler D, Schnabel A. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. In: Biomedizinische Technik/Biomedical Engineering. 1995. p. 317–8.
110. Boumbarov O, Velchev Y, Sokolov S. Personal biometric identification based on ECG features. J Inf Technol Control. 2008;3(4):11–8.
111. Irvine JM, Israel SA. A sequential procedure for individual identity verification using ECG. In: EURASIP J Adv Signal Process, 6th edn. 2009. p. 42–57 **(vol 5)**.
112. Coutinho D, Silva H, Gamboa H, Fred A, Figueiredo M. Novel fiducial and non-fiducial approaches to electrocardiogram-based biometric systems. Biom IET. 2013;2(2):64–75.
113. Silva H, Lourenco A, Lourenco R, Leite P, Coutinho D, Fred A. Study and evaluation of a single differential sensor design based on electro-textile electrodes for ECG biometrics applications. In: Sensors, 2011 IEEE. 2011. p. 1764–7.
114. Sriram JC, Shin M, Choudhury T, Kotz D. Activity-aware ECG-based patient authentication for remote health monitoring. In: Proceedings of the 2009 international conference on multimodal interfaces. ACM; 2009. p. 297–304.
115. Derawi M. Wireless chest-based ECG biometrics. In: Lecture notes in electrical engineering, vol 330. 2015. p. 566–79.
116. Derawi M, Voitenko I, Endrerud PE. Real-time wireless ECG biometrics with mobile devices. In: Proceedings—2014 international conference on medical biometrics, ICMB 2014. 2014. p. 151–6.
117. Zheng G, Yu T. Study of hybrid strategy for ambulatory ECG waveform clustering. J Softw. 2011;6(7):1257–64.
118. Chen Y, Li Y, Cheng X-Q, Guo L. Survey and taxonomy of feature selection algorithms in intrusion detection system. In: Information security and cryptology, vol 4318. Berlin, Heidelberg: Springer; 2006. p. 153–67 **(Chapter 13)**.
119. Morrison DF. Multivariate statistical methods. In: McGraw Hill professional. 1976.
120. Boumbarov O, Velchev Y, Sokolov S. ECG personal identification in subspaces using radial basis neural networks. In: 2009 IEEE international workshop on intelligent data acquisition and advanced computing systems: technology and applications, 2009 IDAACS. 2009. p. 446–51.
121. Zhao C, Wysocki T, Agrafioti F, Hatzinakos D. Securing handheld devices and fingerprint readers with ECG biometrics. In: 2012 IEEE fifth international conference on biometrics: theory, applications and systems (BTAS). 2012. p. 150–5.
122. Gahi Y, Lamrani M, Zoglat A, Guennoun M, Kapralos B, El-Khatib K. Biometric identification system based on electrocardiogram data. In: New technologies, mobility and security, 2008 NTMS '08. 2008. p. 1–5.
123. Guennoun M, Guennoun ZE, El-Khatib K. Efficiency in the design of a biometric identification system based on electrocardiogram data. In: International conference on computer engineering and applications. 2009. p. 1–5.
124. Tantawi MM, Revett K, Salem A, Tolba MF. Fiducial feature reduction analysis for electrocardiogram (ECG) based biometric recognition. J Intell Inf Syst. 2012;40(1):17–39.
125. Mitchell TM. Machine learning. In: McGraw-Hill, Inc. 1997. p. 432.
126. Sasikala P, Wahidauanu R. Identification of individuals using electrocardiogram. In: International journal of computer science and network security. 2010.
127. Silva H, Gamboa H, Fred A. One lead ECG based personal identification with feature subspace ensembles. In: Lecture notes in computer science, vol 4571. Berlin, Heidelberg: Springer; 2007. p. 770–83.
128. Silva H, Gamboa H, Fred A. Applicability of lead v2 ECG measurements in biometrics. In: Med-e-Tel proceedings. 2007.
129. Venkatesh N, Jayaraman S. Human electrocardiogram for biometrics using DTW and FLDA. In: 2010 20th international conference on pattern recognition (ICPR). 2010. p. 3838–41.
130. Bishop CM. Pattern recognition and machine learning. New York: Springer; 2006. p. 738.
131. Mai V, Khalil I, Meli C. ECG biometric using multilayer perceptron and radial basis function neural networks. Conf Proc IEEE Eng Med Biol Soc. 2011;2011:2745–8.
132. Pathoumvanh S, Airphaiboon S, Hamamoto K. Robustness study of ECG biometric identification in heart rate variability conditions. IEEJ Trans Electr Electron Eng. 2014;9(3):294–301.
133. Venugopalan S, Savvides M, Griofa MO, Cohen K. Analysis of low-dimensional radio-frequency impedance-based cardio-synchronous waveforms for biometric authentication. IEEE Trans Biomed Eng. 2014;61(8):2324–35.
134. Lin SL, Chen CK, Lin CL, Yang WC, Chiang CT. Individual identification based on chaotic electrocardiogram signals during muscular exercise. IET Biom. 2014;3(4):257–66.
135. Porée F, Kervio G, Carrault G. ECG biometric analysis in different physiological recording conditions. Signal, image and video processing. 2014.
136. Sidek KA, Mai V, Khalil I. Data mining in mobile ECG based biometric identification. J Netw Comput Appl. 2014;44:83–91.

Fratini *et al. BioMed Eng OnLine* (2015) 14:78

Page 23 of 23

137. Wahabi S, Pouryayevali S, Hari S, Hatzinakos D. On evaluating ECG biometric systems: session-dependence and body posture. IEEE Trans Inf Forensics Secur. 2014;9(11):2002–13.

138. Wang Z, Zhang Y. Research on ECG biometric in cardiac irregularity conditions. In: Proceedings—2014 international conference on medical biometrics, ICMB 2014. 2014. p. 157–63.

139. Sidek KA, Khalil I. Automobile driver recognition under different physiological conditions using the electrocardiogram. Comput Cardiol. 2011. p. 753–6.

140. Sidek KA, Khalil I, Jelinek HF. ECG biometric with abnormal cardiac conditions in remote monitoring system. IEEE Trans Syst Man Cybern Syst. 2014;44(11):1498–509.

141. Tan WC, Yeap HM, Chee KJ, Ramli DA. Towards real time implementation of sparse representation classifier (SRC) based heartbeat biometric system. In: Lecture notes in electrical engineering, vol 307. 2014. p. 189–202.

142. Lee S, Park SY, Kim SJ, Joeng JH, Kim SM. A study on a bio-signal biometric algorithm on the ubiquitous environments. In: Lecture notes in electrical engineering, vol 280. 2014. p. 691–7.