FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Continuous Biometric Identification on the Steering Wheel

João Tiago Ribeiro Pinto

Mestrado Integrado em Bioengenharia

Internal Supervisor: Jaime dos Santos Cardoso, PhD

External Supervisor: André Ribeiro Lourenço, PhD

July 31, 2017

Continuous Biometric Identification on the Steering Wheel

João Tiago Ribeiro Pinto

Mestrado Integrado em Bioengenharia

Approved in public examination by the Jury:

President: José Alberto Peixoto Machado da Silva, PhD
Opponent: Ana Filipa Pinheiro Sequeira, PhD
Referee: Jaime dos Santos Cardoso, PhD
Referee: André Ribeiro Lourenço, PhD

July 31, 2017

Resumo

Devido à grande quantidade de tempo que as pessoas passam dentro deles, os automóveis têm evoluído de simples meios de transporte para enormes *gadgets*, recheados de tecnologia de ponta para garantir conforto, funcionalidade e segurança. A biometria tem o potencial para permitir a personalização automática de definições de segurança, conforto, e condução, tais como a posição dos bancos e espelhos, destinos favoritos no GPS, ou controlo de velocidade, segundo a identidade do condutor. A maior escala, também poderá servir grandes empresas nas tarefas de supervisão e gestão de frotas, reconhecendo continuamente os condutores, para evitar fadiga, e apurar responsabilidades em caso de acidente ou condução perigosa.

Esta dissertação visou o estudo aprofundado de novos métodos e do estado-da-arte em reconhecimento biométrico baseado em ECG, com o objetivo final de formular um sistema biométrico eficaz e fidedigno capaz de usar sinais ECG *off-the-person* altamente ruidosos, adquiridos no volante. Tal tarefa exigiu ultrapassar a evidente prevalência de ruído, incluindo frequentes períodos de saturação, que foi possível através da comparação de diversos métodos de pré-processamento, deteção de picos R, normalização, e deteção de *outliers*. De igual forma, vários conjuntos de características e métodos de reconhecimento do estado-da-arte foram explorados, de modo a garantir os melhores resultados possíveis.

Apesar dos evidentes desafios oferecidos por tais circunstâncias, descobriu-se que a aplicação de segmentação do sinal em 5s, pré-processamento com Savitzky-Golay e um filtro de média móvel, segmentação de batimentos após deteção de picos R com Trahanias, e remoção de *outliers* baseada em *clustering* e correlação cruzada, foi capaz de limpar de forma significativa o sinal. Tendo batimentos médios relativamente limpos, características extraídas por DCT, seguidas de classificação com SVM, possibilitaram a obtenção de 2.66% EER e 91.82% IDR com sinais do volante, assim como 0.60% EER e 97.23% IDR com sinais da coleção UofTDB.

Este método, com treino limitado aos primeiros trinta segundos de dados, auxiliado por uma técnica de ponderação de decisões passadas, permitiu obter 11.77% EER e 69.63% IDR com sinais do volante. Estes resultados provam que, com desenvolvimentos adicionais e a inclusão de técnicas de atualização de *templates*/modelos, será possível desenvolver um sistema biométrico fidedigno, baseado em sinais ECG adquiridos no volante.

Abstract

For the great amount of time people spend inside them, cars have been evolving from mere means of transport to supersized gadgets, filled with top-notch software technology to increase comfort, functionality, safety, and security. Biometric recognition can be the key to automatic customisation of security, comfort, and driving settings, such as seat and mirror positions, GPS favourite locations, or speed monitoring, according to the driver's identity. At a greater scale, it can also serve big companies in their fleet management and supervision tasks, continuously recognizing drivers to avoid fatigue, and apportion blame in case of reckless driving or accidents.

This dissertation aimed to perform an extensive study on prior art and novel approaches in ECG-based recognition, with the final goal of formulating a effective and reliable continuous biometric system, able to work with highly noisy, off-the-person ECG signals acquired at the steering wheel. This required overcoming the noticeable noise corruption, including frequent periods of sensor saturation, achieved through the extensive comparison of several denoising methods, R-peak detection algorithms, normalisation techniques, and outlier detection approaches. Likewise, various feature sets and recognition methods from the prior art were explored, in order to ensure the best results possible.

Despite the evident challenges brought by such settings, it was found that a succession of five-second signal segmentation, denoising with Savitzky-Golay and a Moving Average Filter, heartbeat segmentation after Trahanias R-peak detection, and outlier removal based on cross-correlation clustering, was able to significantly clean the signal. Having relatively clean ensemble heartbeats, DCT features followed by SVM classification enabled the achievement of 2.66% EER and 91.82% IDR with driving signals, as well as 0.60% EER and 97.23% IDR with signals from the public UofTDB collection.

The same method, under train settings constrained to the first thirty seconds of data, aided by a novel past score weighting procedure, was able to achieve 11.77% EER and 69.63% IDR with driving signals. These results prove that, with further improvements and the inclusion of template/model update techniques, it will be possible to develop a reliable continuous biometric system with ECG signals acquired on the steering wheel.

Acknowledgments

Ao Professor Jaime, pela atenção, pelas opiniões, pelas críticas, por toda a disponibilidade, pelo vasto conhecimento que sempre esteve pronto a partilhar, e por ser o melhor professor que tive o prazer de conhecer, um exemplo pela sua inteligência, conhecimento, simpatia, e humildade.

Ao Doutor André Lourenço, à Priscila, ao Carlos, ao Roberto, e ao Borja, todos os membros da equipa da CardioID Technologies que tive o prazer de conhecer, e que me receberam com toda a disponibilidade e simpatia, que partilharam o seu conhecimento comigo, que estiveram sempre prontos a ajudar, a dar ideias, e a esclarecer dúvidas.

Aos meus amigos, e a todos os meus colegas que me receberam de braços abertos neste curso. A muitos com os quais tive o prazer de trabalhar em projetos, e a tantos outros que me acompanharam no EMBS e no NEB. Em especial aos mais próximos, que me apoiaram sempre, e ainda mais durante esta dissertação.

Acima de tudo, aos meus pais, por aceitarem a árdua tarefa de ser meus pais, por me ensinarem quem sou, o valor da humildade, da verdade, do trabalho, e da dedicação, por me guiarem no melhor caminho e, no entanto, por sempre me garantirem a liberdade para o escolher. Também à minha irmã, aos meus padrinhos, e ao meu tio, por estarem sempre aqui, e me ensinarem o que é uma verdadeira família.

Obrigado.

The author also wishes to acknowledge the access to the UofTDB ECG Database, gracefully granted by Dr. Hatzinakos and Mr. Komeili, of the University of Toronto, that truly improved the quality of the work performed throughout this dissertation.

"You ever feel like
every time we get close to getting the answers,
somebody changes the question?"

Peter Bishop

Contents

List of Figures

List of Tables

Abbreviations

1D	Unidimensional
1DMRLPB	Unidimensional Multi-Resolution Local Binary Patterns
2D	Bidimensional
3D	Tridimensional
AC	Autocorrelation Coefficients
AC/DCT	Autocorrelation/Discrete Cosine Transform Coefficients
AHA	American Heart Association
AR	Autoregressive
AUC	Area Under the Curve
AV	Atrioventricular (node)
BPF	Bandpass Filter
BVP	Blood Volume Pressure
BW	Baseline Wander
cA	Approximation coefficients
cD	Detail coefficients
CE	Chaotic Encryption
CMC	Cumulative Match Characteristic
CNN	Convolutional Neural Network
CWT	Continuous Wavelet Transform
CYBHi	Check Your Biosignals Here initiative
DBNN	Decision-based Neural Network
DC	Direct Current
DCT	Discrete Cosine Transform
DET	Decision Error Tradeoff
DNA	Deoxyribonucleic Acid
DTW	Dynamic Time Warping
DVD	Digital Versatile Disc
DWT	Discrete Wavelet Transform
ECG	Electrocardiogram
ECU	Electronic Control Unit
EEG	Electroencephalogram
EER	Equal Error Rate
EH	Ensemble Heartbeat
EMD	Empirical Mode Decomposition
EMG	Electromyogram
FAR	False Acceptance Rate
FFT	Fast Fourier Transform
FLDA	Fisher Linear Discriminant Analysis

FMR	False Match Rate
FNIR	False Negative Identification Rate
FNMR	False Non-Match Rate
FOGD	First-Order Gaussian Differentiator
FPIR	False Positive Identification Rate
FRR	False Rejection Rate
FT1DLBP	Fuzzy Tunable Unidimensional Local Binary Patterns
GBFS	Greedy Best First Search
GBRT	Gradient Boosted Regression Trees
GMM	Gaussian Mixture Model
GMM-UBM	Gaussian Mixture Model/Universal Background Model
GPS	Global Positioning System
HLDA	Heteroscedastic Linear Discriminant Analysis
HMM	Hidden Markov Model
HPE	Hermite Polynomial Expansion
HPF	Highpass Filter
HRV	Heart Rate Variability
HTER	Half-Total Error Rate
ICA	Independent Component Analysis
ID	Identification
IDE	Integrated Development Environment
IDR	Identification Rate
IMF	Intrinsic Mode Function
kNN	k-Nearest Neighbours
KPCA	Kernel Principal Component Analysis
LBP	Local Binary Patterns
LDA	Linear Discriminant Analysis
LLR	Log-Likelihood Ratio
LMS	Least Mean Squares
Log-STFT	Log-Short-Time Fourier Transform
LPF	Lowpass Filter
LTST	Long-Term ST (collection)
MIDR	Misidentification Rate
MIT-BIH	Massachusetts Institute of Technnology - Beth Israel Hospital (collections)
MLP	Multilayer Perceptron
NB	Naïve Bayes
NB-UBM	Naïve Bayes/Universal Background Model
NCC	Normalised Cross-Correlation
NCCC	Normalised Cross-Correlation Clustering
NF	Notch Filter
NLMS	Normalised Least Mean Squares
NN	Neural Network
NSR	Normal Sinus Rhythm (MIT-BIH database)
PAR	Pulse Active Ratio
PCA	Principal Component Analysis
PCG	Phonocardiogram
PE	Peak Extraction
PIN	Personal Identification Number

PLI	Powerline Interference
PNN	Probabilistic Neural Network
PPG	Photoplethysmogram
PRD	Percent Residual Difference
PSD	Power Spectral Density
PTB	Physikalisch-Technische Bundesanstalt (collection)
PTCD	Probability of Time to Correct Decision
PVE	Peak and Valley Extraction
RAM	Random Access Memory
RBF	Radial Basis Function
RBFNN	Radial Basis Function Neural Network
RMSE	Root Mean Squared Error
RNN	Recurrent Neural Network
ROC	Receiver Operating Characteristic
RR	Reject Rate
SA	Sinoatrial (node)
SBFS	Sequential Backwards Floating Selection
SBS	Sequential Backwards Selection
SFA	Simplified Fuzzy ARTMAP
SFFS	Sequential Forward Floating Selection
SFS	Sequential Forward Selection
SG	Savitzky-Golay
SIMCA	Soft Independent Modelling of Class Analogy
SLP	Polysomnographic (MIT-BIH database)
SNR	Signal-to-Noise Ratio
SpO_2	Arterial Oxygen Saturation
STC	ST Change (MIT-BIH database)
STFT	Short-Time Fourier Transform
ST-T	European ST-T (database)
SVM	Support Vector Machine
TCD	Time to Correct Decision
TPIR	True Positive Identification Rate
UBM	Universal Background Model
UofTDB	University of Toronto ECG Database
USC	Usability-Security Characteristic
VE	Valley Extraction
WDIST	Wavelet Distance
WFDB	Waveform Database
WWPRD	Wavelet Weighted-based Percent Residual Difference

Chapter 1

Introduction

1.1 Biometric Recognition

Websites, smartphones, safes, cars, houses, buildings, banks, and airports are just a few of our society's amenities that rely on identification or authentication systems to protect and guard ourselves, our information, or our belongings. Most of them still depend on traditional systems based on extrinsic entities or knowledge like cards, keys, or passwords (Agrafioti et al., 2011; Jain et al., 2011).

In the last decades, researchers have focused on avoiding the problems of traditional systems: they can be lost, stolen, discovered, or copied (Abo-Zahhad et al., 2014). Biometrics present the perfect opportunity to achieve that goal, as they are focused on intrinsic characteristics of the person, requiring their physical presence, and minimising the probability of success of possible impostors (Jain et al., 2011; Kaur et al., 2014).

Many human traits have been proposed and studied for the purpose of identity recognition, especially face, fingerprints, voice, and iris (Kaur et al., 2014; Akhter et al., 2016). With the increasing number of applications that rely on these, the methods to circumvent them become stronger, resorting to photographic, 3D model reproductions, or sound recordings of the traits (Belgacem et al., 2012; Fratini et al., 2015), and obliging biometric systems to include deeper security measures, such as liveness detection.

More recently, a new set of biometric traits, called medical biometrics, has gained momentum (Agrafioti et al., 2012; Abo-Zahhad et al., 2014; Akhter et al., 2016). The Electrocardiogram (ECG) has proven to be the most promising of them, excelling in most of the characteristics that define the quality of a biometric trait. Its nature makes it hard to capture and inject into the system for spoofing purposes, and the inherent liveness detection ensures the biometric system is not being attacked (Li and Narayanan, 2010).

Recent studies have proven the capabilities of ECG as a biometric, and its unidimensional nature places it as a more computationally efficient alternative to image or video-based systems, especially for continuous recognition systems, highly dependant on timely decisions.

1.2 Automotive Biometrics: Challenges and Opportunities

The automotive world has, in the last decades, promoted a burst of technology integration in cars, aiming for more comfort, ease of use, functionality, safety, and security. This has been especially fueled by the fact people generally use their cars for a great amount of time each day, and cars are one of the most widespread possessions in our society (Coppola and Morisio, 2016; Huang et al., 2016).

Cars today include an array of electronic control units (ECUs) that automatically control traction, braking, stability, lane changing, lighting, collision avoidance, and many others (Coppola and Morisio, 2016; Huang et al., 2016). Besides these, other technologies for driver drowsiness, anger, and stress detection, have also been studied to avoid car crashes (Coppola and Morisio, 2016).

Biometric recognition, although less explored and studied than other technologies, also offers great advantages for the automotive world. Currently, car locks, start/stop systems, and on-loan monitoring and speed limitations are controlled by wireless key fobs, that can be lost or stolen. Biometric recognition could completely eliminate the need for such devices, and allow for automatic personalisation of car and driving settings according to the driver's identity (Lourenço et al., 2015).

ECG-based recognition could, furthermore, be a reliable way to ease remote fleet supervision and management, and associated tasks such as driver swaps and break enforcement to avoid fatigue and stress-related accidents. Continuous biometrics on the steering wheel could also, in these settings, allow managers to more reliably assess responsibility and apportion blame in case of accident or reckless driving.

The electrocardiogram is, potentially, the best trait for driving settings. It is universal, always present, easy to acquire and process, hard to counterfeit, and includes enough personal information to allow for accurate recognition (Agrafioti et al., 2011; Kaur et al., 2014). Continuous contact need is a significant setback, but its inconveniences can be effectively minimised using the recent off-the-person acquisition techniques, with dry electrodes on the steering wheel (Lourenço et al., 2015), and software solutions able to deal with contact losses and low signal quality.

1.3 Research Goals and Contributions

Medical-grade electrocardiographic signals, as those used on the first ECG biometrics studies, or even the more recent off-the-person signals, acquired during exercise on the fingers or palms, are relatively easy to preprocess. On the other hand, the ECG signals acquired on the steering wheel while driving are much more contaminated by noise artifacts, with less contained and predictable frequency ranges.

Following the advantages of ECG-based biometric recognition on the steering wheel, this dissertation aimed to provide an extensive study of the behaviour of prior art and proposed approaches in these new and more challenging settings. The main goal was, thus, to find a framework that

could recognise individuals, accurately and quickly, from the noisy signals acquired while driving, after evaluating and optimising, individually, each phase involved in the process.

This dissertation contributions begin with a complete review of the fundamentals of the electrocardiographic signal and biometric recognition, as well as a thorough survey of prior art approaches that used the former as a biometric trait for the latter. This provided a strong theoretical foundation and state-of-the-art knowledge to prepare and guide the work performed throughout the dissertation.

On signal preprocessing, several prior art and proposed denoising methods were evaluated on their ability to denoise both standard off-the-person and driving ECG signals. Afterwards, on the step of signal preparation, the Pan-Tompkins, Engelse-Zeelenberg, and Trahanias R-peak detection algorithms were explored and had their performance assessed in similar conditions. Besides this, a novel outlier detection algorithm, based on normalised cross-correlation, was devised and compared with DMEAN, one of the best prior art approaches.

Diverse prior art feature sets were extracted from ensemble heartbeats. Also, a more complete formulation of 1D Local Binary Patterns was conceived. These were compared in their ability to provide the conditions to attain high and consistent identification rates with simple classifiers, while LDA and PCA were explored for dimensionality reduction. The most promising sets were used with several decision methods, including the proposed Naïve Bayes and its UBM adaptation, and their continuous and conventional authentication and identification performances were measured and compared.

1.4 Dissertation Outline

Besides this introduction, this dissertation presents the anatomy and physiology fundamentals on the ECG signal, its variability and noise sources, and the acquisition settings, in Chapter 2. The basics on biometric recognition, its modalities and their qualities, its structure and performance evaluation metrics, as well as general system design considerations, are presented in Chapter 3.

Chapter 4 reviews some of the most commonly used ECG databases, and the CardioID collection used for this dissertation. A complete review of sixty-five prior art approaches and their acquisition, preprocessing, preparation, feature extraction, and recognition methods is presented in Chapter 5.

The experimental settings for this dissertation are defined in Chapter 6, and Chapter 7 presents several preprocessing methods and their evaluation with common off-the-person and driving ECG signals. Chapter 8 focuses on the detection of reference points, heartbeat segmentation and amplitude normalisation, outlier detection and removal, and ensemble heartbeat construction.

Chapter 9 describes various feature extraction and dimensionality reduction methods, and their evaluation with ECG signals, and Chapter 10 presents the recognition methods explored in this dissertation, and their performance in authentication and identification tasks. Finally, Chapter 11 draws some conclusions on the work performed and the results obtained, and discusses some paths to be explored in the future.

Chapter 2

Electrocardiographic Signal Characterisation

2.1 Anatomy and Physiology

In every sense of the word, the heart is a pump. Tate (2009) defines three main functions of the heart: *generating blood pressure*, through the contraction of the myocardium, in order to keep blood moving; *routing blood*, by sending venous blood to the lungs, in the pulmonary circulation, and arterial blood to the whole body, in the systemic circulation; and *regulating blood supply*, by adapting its rate and force of contraction to the current metabolic demands of the body.

The heart is able to perform these functions due to its anatomy (cf. Fig. 2.1). The heart is enveloped by the pericardium, and its wall is composed by three layers: the epicardium (the outer layer), the myocardium (the middle layer), and the endocardium (the inner layer). While the endocardium and epicardium are thin endothelium and serous layers, respectively, most of the wall is occupied by the myocardium, composed by muscle cells disposed in spiral bundles across the heart wall to allow for potent contractions (Scanlon and Sanders, 2007; Tate, 2009; Marieb and Hoehn, 2013).

Inside, the heart is divided in four chambers: the left and right atria and ventricles. The left atrium receives arterial blood from the pulmonary veins, while the right atrium receives venous blood from the caval veins (superior and inferior). The blood received will pass, upon atrial contraction, to the respective ventricle. Upon ventricle contraction, the blood on the left ventricle will be sent through the aorta, and the blood on the right ventricle will be sent through the pulmonary arteries (Scanlon and Sanders, 2007; Marieb and Hoehn, 2013).

The contraction of the heart is, thus, of the highest importance. The myocardial muscle cells contract in response to electrical currents, that cause the depolarisation of those tissues by triggering action potentials. To control the contraction, and thus channel the blood flow in the correct direction, the heart controls the depolarisation according to a defined sequence of events (cf. Fig. 2.3) (Scanlon and Sanders, 2007; Tate, 2009):

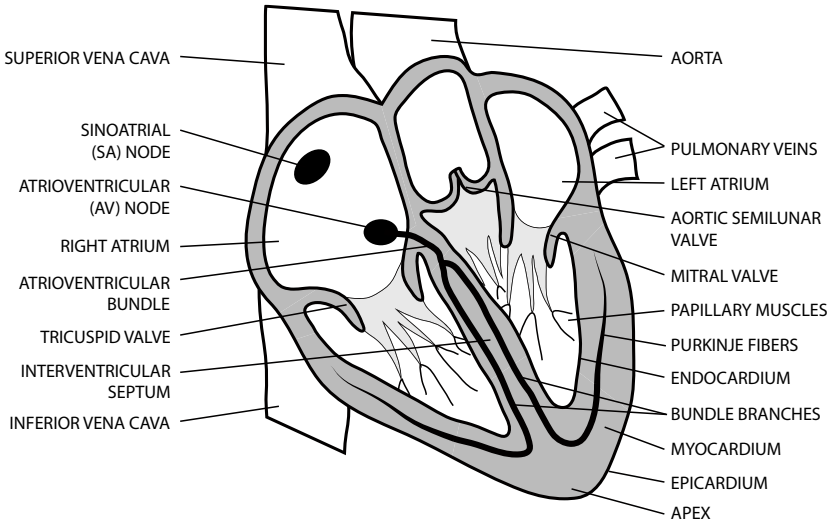

Figure 2.1: The anatomy of the heart (based on Tate (2009) and Marieb and Hoehn (2013)).

1. *Atrial depolarisation* - The depolarisation begins in the sinoatrial (SA) node, the heart's pacemaker that generates the impulse, and quickly spreads to both atria, also reaching the atrioventricular (AV) node;

2. *Atrial depolarisation complete* - The depolarisation reaches the whole atria and the atrial contraction begins. Thanks to the fibrous skeleton of the heart, a fibrous connective tissue that surrounds the four valves and insulates the ventricles from the atria, the depolarisation is unable to spread to the ventricles. The conduction to the ventricles can only be done via the AV node and the bundle branches. However, to avoid simultaneous ventricular and atrial contractions, the AV node delays the impulses for about 0.1 s before sending it down the bundle branches;

3. *Ventricular depolarisation* - The impulse delay at the AV node ends and it is sent down the bundle branches. When it reaches the apex, it starts to spread to the myocardium cells via the Purkinje fibers and begins the ventricular depolarisation. At the same time, after their contraction, the atria begin to repolarise;

4. *Ventricular depolarisation complete* - The ventricles myocardial cells contract almost immediately after being depolarised;

5. *Ventricular repolarisation* - After contraction, the ventricles begin to repolarise;

6. *Ventricular repolarisation complete* - The repolarisation of the ventricles completes, and the heart is ready to repeat the process.

These flows of depolarisation and repolarisation are nothing more than electrical currents being generated and conducted through the heart. These electrical currents can be detected and

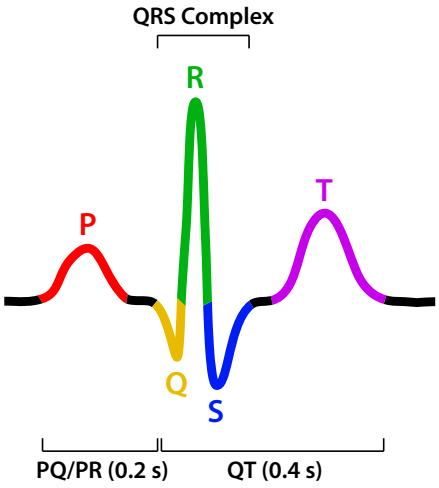

Figure 2.2: A single heartbeat in an electrocardiogram (based on Tate (2009)). The main waves P, Q, R, S, T, the QRS complex, and the important PQ/PR and QT intervals are presented. It is important to note that the presented durations of the intervals are only an example that can widely vary according to the subject and his/her condition.

measured, through electrodes placed in the body, in a process called electrocardiography. The resulting signal is called an electrocardiogram (ECG) and, in normal conditions, is a cyclic repetition of five easily recognisable waves or deflections: the P, Q, R, S, and T waves (cf. Fig. 2.2). The ECG signal is cyclic because each group of these deflections refers to a single heartbeat and, as shown in Fig. 2.3, each of these parts can be traced back to the phase that originated it (Scanlon and Sanders, 2007; Tate, 2009; Marieb and Hoehn, 2013).

2.2 Variability

As shown in the previous section, the electrocardiographic signal results of the detection and measurement of electrical current that is generated and conduced throughout the heart, to ensure its strong and timely contractions.

The ECG signal, although presenting, in normal conditions, the same deflections for all subjects at all times, is doted by a high degree of variability. Variability in the ECG can be designated as *intrasubject*, the variations between cycles (heartbeats) in the electrocardiogram of a single subject, or *intersubject*, the variations between heartbeats of different subjects. Both these variability types can have origin in several factors, and the most important of these are:

- *Heart Geometry* - Heart size, cardiac muscle thickness, and the overall shape of the heart dictate the paths the electrical current will follow inside the heart, the number of muscle cells that will depolarise, and the time it takes to do it across the whole heart. Athletes, with their high physical training, commonly have larger hearts, with thicker myocardia, which affects

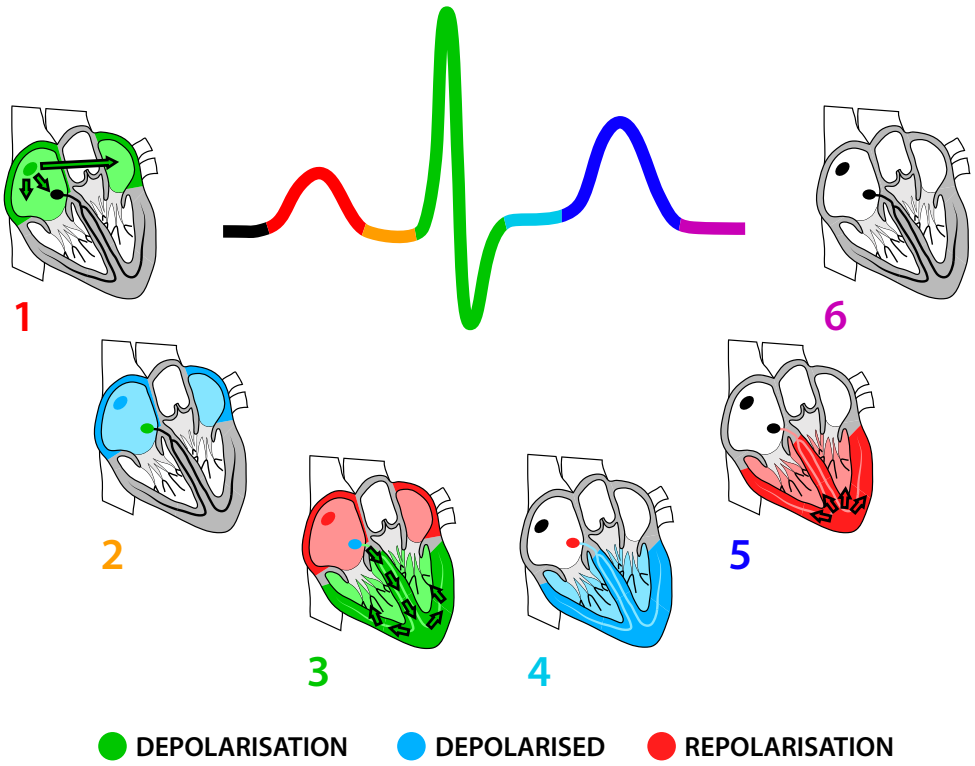

Figure 2.3: The sequence of electrical conduction events in the heart, and its relationship with different waves of a heartbeat in an ECG signal. (based on Marieb and Hoehn (2013)). The colors of the ECG trace refer to each phase 1-6.

the ECG with higher voltages in the QRS complex, and lower basal heart rates (Hoekema et al., 1999; van Oosterom et al., 2000; Hoekema et al., 2001);

- *Individual Attributes* - Age, weight, and pregnancy are some of the individual attributes that can cause shifts in the heart position and/or orientation. These shifts will change the orientation of the electrical current conduction vectors along the heart. This means that the electrodes will detect the signal in a different perspective and, thus, the ECG deflections will suffer variations both in voltage and in time (Schijvenaars, 2000);

- *Physical Exercise or Meditation* - The duration of, and intervals between the different deflections of the heartbeats in an ECG signal, vary with the heart rate. These changes are especially visible on the interval between the QRS complex and the T wave in situations of tachycardia (higher heart rates) or bradycardia (lower heart rates). Changes in the heart rate caused by physical exercise or meditation do, effectively, affect the electrocardiogram (Agrafioti et al., 2012);

- *Cardiac Conditions* - Medical conditions of the heart can also interfere in the dynamics of the electrical pulse conduction and generate variability. In the scope of biometrics, one

of the most studied conditions is Arrhythmia, that causes wide variations in the heart rate across time and, as reported by several researchers, can consistently shrink the performance of ECG-based biometric systems (Ye et al., 2010; Agrafioti et al., 2011; Safie et al., 2011).

- *Posture* - Postures like standing or laying down differ widely on the position and shape of internal organs. The heart also suffers with this, changing its position in the thorax, and thus its position in reference with the electrode placement, which will cause variations in the collected ECG signal (Schijvenaars, 2000);

- *Emotions* - The sympathetic and parasympathetic systems, of the autonomous nervous system, work to, respectively, increase or reduce the heart rate. These systems are under direct influence of psychological states and thus, under stress, fear, or other strong emotions, the heart rate and the ECG signal can be affected (Agrafioti et al., 2012);

- *Electrode characteristics and placement* - The type and size of the electrodes, and whether they use gel or not, is a main origin of noise and, thus, variability. The mispositioning of electrodes, variant placements, reversal of leads, are also sources of variability, as they change the perspective of detection of the electrocardiographic signal (Hoekema et al., 1999; Schijvenaars, 2000).

All the previously presented factors reflect on the morphology of the electrocardiographic signals acquired from an individual. Thus, when considering the acquisition of ECG, both for medical and biometric recognition purposes, it is of the utmost importance to consider all of these and the way they can ease or difficult the task at hand. In the next section, the ways to measure electrocardiographic signals of an individual are presented.

2.3 Acquisition

Medical diagnostic and biometric recognition are very different tasks with very different purposes and requirements that, thus, differ on the most commonly used techniques to measure the ECG signals. Besides that, common biometric systems also differ from the more specific settings that are studied in this dissertation - the driving environment. Hence, in this section, the electrocardiogram acquisition methods are presented in three groups: the standard medical acquisition settings, the common biometric acquisition settings, and the specific acquisition settings for driver monitoring.

2.3.1 Standard Medical Acquisition Settings

For medical purposes, there are a few defined and established configurations of electrodes for the measurement of electrocardiogram signals in standard, comparable formats that ease the diagnostic of cardiac conditions. Here, two of the most common are presented: the standard 12-lead configuration, and the corrected orthogonal ECG configuration.

2.3.1.1 Standard 12-Lead Configuration

The Standard 12-Lead Configuration, as the name suggests, allows the acquisition of an ECG signal in 12 leads (or channels). Each lead corresponds to a different electrical perspective of the heart, and can deliver important information for diagnostic. The 12 leads can be grouped in three categories: three bipolar limb leads, three monopolar limb leads, and six monopolar precordial leads (Biel et al., 2001).

The bipolar limb leads originated in the work of Willem Einthoven, the creator of the first electrocardiograph. These leads (cf. Fig. 2.4) compose the sides of a equilateral triangle, whose vertices are the left (LA) and right arm (RA) wrists and the left leg (LL) ankle (the position of the three electrodes), that serve to capture a projection of the cardiac dipole in each of these directions. The right leg serves as ground (N). The signals are obtained by measuring the electric potentials between RA and LA (Lead I), RA and LL (Lead II), and between LA and LL (Lead III) (Biel et al., 2001; MacLeod and Birchler, 2014)

The three limb electrodes also serve to capture the monopolar limb leads, that consist on the measurement of the cardiac dipole in the direction of each limb electrode. In the direction of LA we obtain the aVL lead, for RA we obtain aVR, and with LL we obtain aVF. Artificial references for all these are obtained with a combination of signals at various electrodes (Biel et al., 2001). Finally, the six monopolar precordial leads (V1 to V6) are placed on the chest, as shown on Fig. 2.4, and allow the collection of signals in directions on the axial plane, while the limb leads are for the frontal plane.

2.3.1.2 Corrected Orthogonal Configuration (Frank Leads)

Instead of aiming for the collection of signals on several directions in the axial and frontal anatomical planes (as the standard 12-lead configuration does), the commonly designated Frank Leads aim to acquire the electrocardiogram in just three channels, each corresponding to an orthogonal direction x, y, or z, thus capturing the heart dipole in three dimensions (MacLeod and Birchler, 2014).

The corrected orthogonal configuration involves the use of seven electrodes: I, E, C, A, M, F, and H (cf. Fig. 2.5). The first five electrodes are are on the same plane, the axial, with I and A below the right and left axillae, respectively. E and M on the intersection of the axial and sagittal planes, E is placed anterior to the subject and M is placed posterior. C is halfway between E and A, and makes a 45 degree angle with each, with center on the referential origin. F is placed on the left ankle, and H is placed on the back of the neck. Processing the signals obtained from all these electrodes, as shown in Fig. 2.5, allows the collection of the three orthogonal leads, P_x, P_y, and P_z (MacLeod and Birchler, 2014).

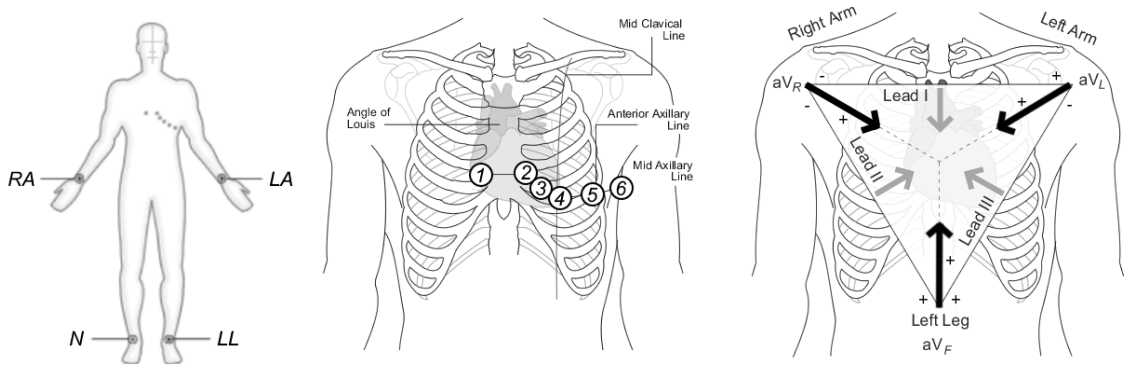

Figure 2.4: Leads and electrode placement of the standard 12-lead configuration (adapted from Moran and Fecowycz. Left: limb electrode placement; centre: precordial electrode placement; right: monopolar and bipolar limb leads).

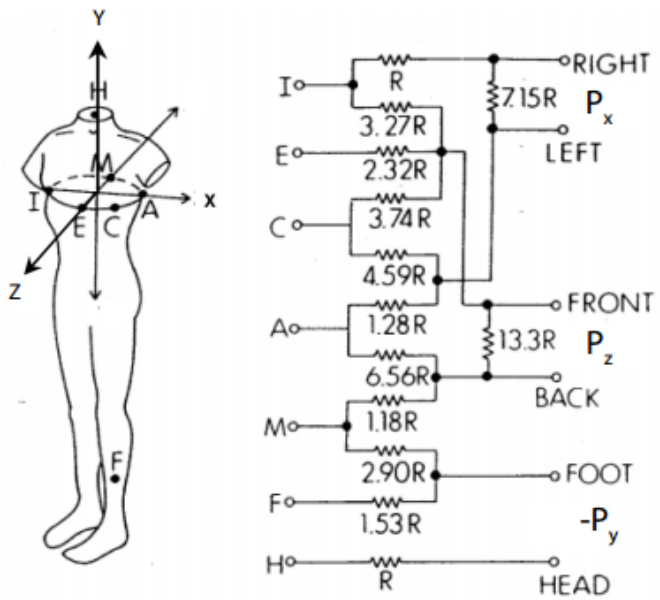

Figure 2.5: Leads and electrode placement of the corrected orthogonal configuration (from MacLeod and Birchler (2014). Left: electrode placement; right: lead measurement computation).

2.3.2 Common Biometric Acquisition Settings

2.3.2.1 Adaptation of Medical Configurations

In ECG-based biometrics, researchers have commonly used recordings from standard 12-lead ECG configuration or Frank leads as basis for the development and evaluation of their biometric systems (Plataniotis et al., 2006; Wübbeler et al., 2007; Ghofrani and Bostani, 2010). Even more common has been the selective use of certain leads of these configurations, especially the Lead I (Palaniappan and Krishnan, 2004; Zhang and Wei, 2006; Molina et al., 2007), because of its higher acceptability due to the electrode placement on the wrists, but also Lead II (Kyoso and Uchiyama, 2001a,b; N and Jayaraman, 2010; Pathoumvanh et al., 2014), and chest leads (Kyoso et al., 2000; Fang and Chan, 2009; Ye et al., 2010).

2.3.2.2 Off-the-Person Acquisition

More recently, biometrics researchers have been focusing on the so-called *off-the-person* approaches. Unlike *on-the-person approaches*, that are characterised by somewhat invasive acquisition techniques, with gelled electrodes placed on the limbs and chest, *off-the-person* acquisition focuses on minimally intrusive acquisition, closer to simple real biometric applications (Silva et al., 2013; Carreiras et al., 2016).

Off-the-person approaches have clear advantages on what concerns the acceptability, simplicity, and unobtrusiveness of the acquisition. However, the farther placement of the electrodes and the fact they are commonly ungelled makes the acquisition more vulnerable to contamination, decreasing the signal-to-noise ratio (SNR), and thus deteriorating the quality of the signal (Belgacem et al., 2012; Silva et al., 2014).

To the best of our knowledge, the first research work to focus on off-the-person acquisition was conducted by Chan et al. (2008), that collected the ECG signals using button Ag-AgCl electrodes held by the subjects (cf. Fig. 2.6). Coutinho et al. (2010, 2011) followed, acquiring signals at the palms during computer tasks (Silva et al. (2013) used a similar configuration).

Later on, Shen et al. (2011) recorded signals from both palms using two small metallic rods that were grabbed by the subjects during acquisition (Belgacem et al. (2012, 2013) and Lin et al. (2014) used similar electrodes). Lourenço et al. (2011b) mounted the three electrodes on a plaque, positioned to contact with the index finger of the left hand and the thumb of the right hand.

Silva et al. (2014) used four electrodes mounted on a cloth on a table, with one Ag-AgCl electrode in contact with each palm, and one electrolycra strip (conductive polymeric fabric) in the index and middle fingers of each hand. Lourenço et al. (2014) and Matos et al. (2014) used only two Ag-AgCl electrodes (with virtual ground) to acquire ECG at the index fingers.

2.3.3 Specific Driver Monitoring Settings

For the specific field of driver monitoring, be it for stress, attentiveness, fatigue detection, or biometric recognition, there are two ways to conduct the collection of the ECG signals. The first

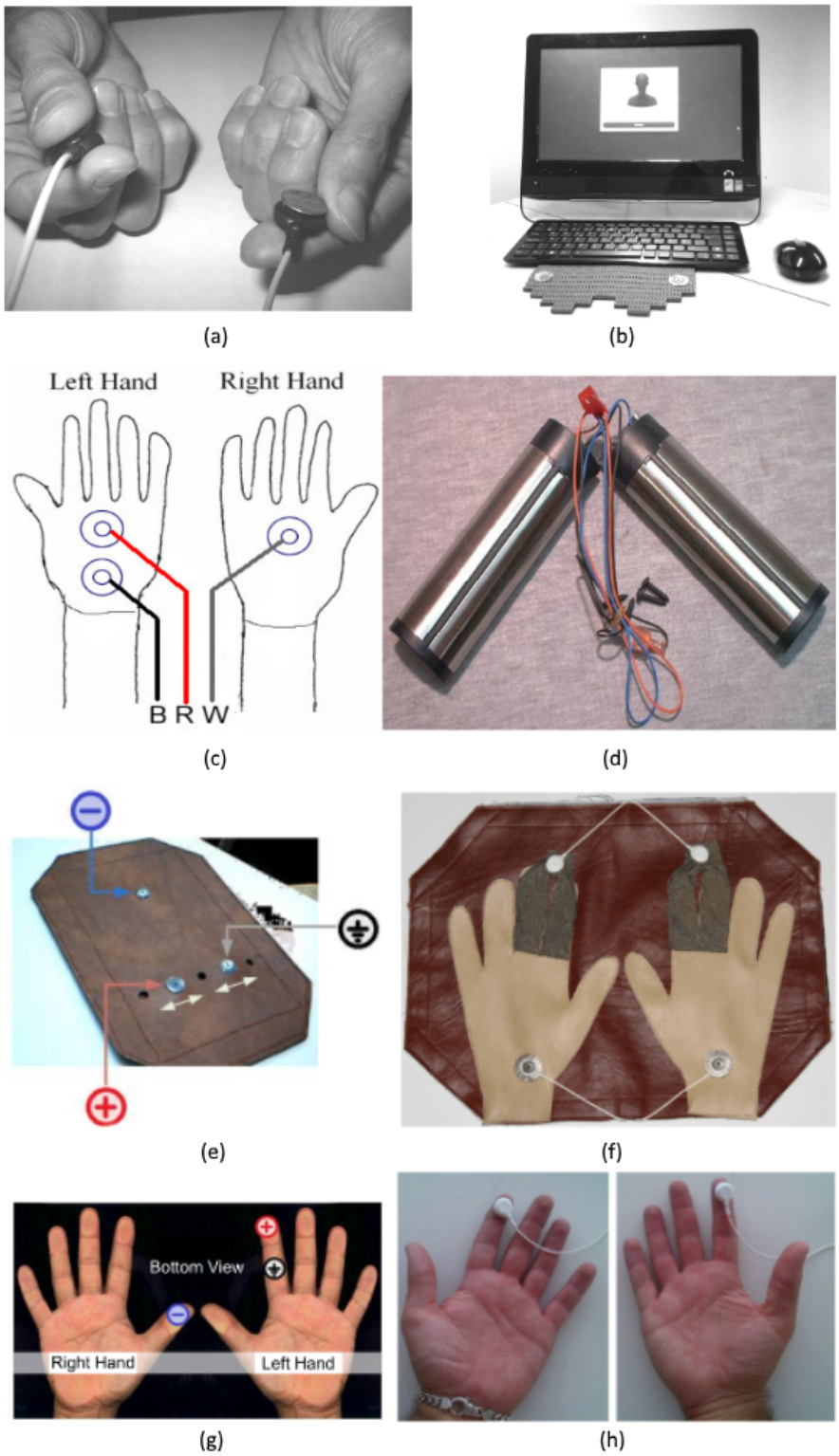

Figure 2.6: Electrode configurations used on off-the-person acquisition settings (a - Chan et al. (2008); b - Silva et al. (2013); c and d - Shen et al. (2011); e and g - Lourenço et al. (2011b); f - Silva et al. (2014); h - Matos et al. (2014)).

Figure 2.7: Steering-wheel-based ECG acquisition configurations (left: mock-up from Gomez-Clapers and Casanella (2012); center: CardioWheel mockup; right: CardioWheel integrated in car (Lourenço et al., 2015)).

consists in following configurations similar to other settings, as did Healey and Picard (2005), using adapted electrode placements fixed on the patient's chest and limbs.

The second way is to fix the electrodes on the wheel. It allows for a complete fit with the off-the-person goals, minimising the obtrusiveness of the acquisition procedure, and eliminates the need for intervention when the driver enters or exits the car.

In 2012, Gomez-Clapers and Casanella (2010, 2012) studied a wireless steering wheel prototype with embedded electrodes (cf. Fig. 2.7), that could be mimicked in real cars' steering wheels, and found that the signals acquired had enough quality and stability to allow for the monitoring of the driver. Lourenço et al. (2015), of CardioID Technologies[1], developed the CardioWheel (cf. Fig. 2.7) and, in a partnership with Ceiia[2], got to adapt it and install it in real automobiles, allowing for acquisition and testing in real settings.

Besides the scientific research efforts, many car manufacturers (especially the largest ones) have already begun to invest in the development and exploration of steering wheel ECG-based monitoring technologies, aiming for higher competitiveness, as stated by Choi et al. (2016). D'Angelo et al. (2010), in partnership with BMW, integrated ECG electrodes on the steering wheel, gear selector, and armrest, to monitor the heart rate of the driver. Along with ECG, Daimler AG's Mercedes-Benz also installed skin conductivity, SpO_2, and peripheral temperature sensors on the steering wheel (Heuer et al., 2010). Ford and Denso opted to acquire ECG using seat-mounted electrodes (Wartzek et al., 2011; Sakai et al., 2013).

2.4 Noise Contamination

2.4.1 Standard Noise Sources

The acquisition of electrocardiographic signals is highly susceptible to be corrupted by noise. Depending on the electrode characteristics and placement, the amplitude of the waveforms obtained can vary but, in ideal conditions (chest leads in medical settings), the QRS complex only reaches

[1]CardioID Technologies: `https://www.cardio-id.com/` (visited on 30/01/2017).
[2]Ceiia: `https://www.ceiia.com/` (visited on 30/01/2017).

$2-3$mV, the largest amplitude of the whole cyclic beat (Fatemian and Hatzinakos, 2009). This means that the farthest the location of the electrodes, the weaker the signal and the more dominant the noise.

Denoising is a key aspect in both diagnostic and biometric systems, and constitutes the focus of the signal preprocessing block. ECG noise can originate in many sources, most commonly:

- *Powerline inteference* (PLI) - Powerline interference is the designation given to the noise generated by the sinusoidal alternating current, used as energy source by the acquisition equipment. It reflects on the acquired signal as a high-frequency noise, with frequency equal to that of the current (60 Hz in the United States and other American countries, and 50 Hz in Europe, Asia, and most other countries) (Singh et al., 2015; Fatemian and Hatzinakos, 2009);

- *Baseline wander* (BW) - Baseline wander is caused by breathing movements. Thus, it reflects on the acquired signal as a low-frequency ondulation of the signal's baseline, normally below 1 Hz (Singh et al., 2015; Fatemian and Hatzinakos, 2009);

- *Elecromyographic (EMG) interference* - Just like the cardiac muscle, the other muscles in the body also use electric impulses to contract. While capturing ECG, it is easy to also capture the electromyographic signals of the contraction of nearby muscles, that will reflect in the signal as high-frequency, high-amplitude, short-term bursts (Singh et al., 2015; Fatemian and Hatzinakos, 2009);

- *Electrode movement* - These artifacts result from skin impedance changes around the electrode, caused by movements of the subject, and reflect as high amplitude waveforms in the signal (Fatemian and Hatzinakos, 2009);

- *Lead reversal* - The reversal of leads will cause the incorrect measurement of potentials. The acquired signal heartbeats will, thus, normally present some or all of its characteristic waveforms reversed in amplitude (Singh et al., 2015);

- *Pacemaker interference* - The signals from artificial pacemakers, installed in individuals with heart conditions, can be captured along with the ECG signal. These interferences appear as short spikes before the S wave (Singh et al., 2015).

When we consider off-the-person approaches, using ungelled electrodes in the palms or fingers, it is reasonable to expect a considerable increase in the noise influence, with lower SNR. The ability to capture the ECG signal weakens, so the amplitude of the ECG components is smaller, when compared with chest leads, and movement artifacts are much more frequent and dominant (Lourenço et al., 2014; Matos et al., 2014; Silva et al., 2014).

2.4.2 Noise Sources in Driving Settings

If, as said, noise is more influential on the acquired signal when we use off-the-person approaches, when using steering wheel acquisition in driving settings this influence becomes even higher.

Driving requires frequent movement, and drivers should be expected to move on their seat and move their hands, to respond to noises and the surroundings, to steer the car, to brake and to accelerate (Healey and Picard, 2005). The electrocardiographic signals acquired are, thus, much more prone to be contaminated with both electromyographic and movement artifacts (Gomez-Clapers and Casanella, 2012). The way the drivers hold the wheel, the frequency with which they take their hands off to change gears or steer, will also deteriorate the quality of the obtained signals, as both hands are required to be placed on the wheel to acquire the signals (Gomez-Clapers and Casanella, 2012; Lourenço et al., 2015; Choi et al., 2016).

The only source of noise that is not expected in a driving environment, when compared with common acquisition settings, is the powerline interference, as the equipments will not use alternating current energy. Despite this, the many other sources and their enhanced contaminations of the signals must be addressed properly with effective preprocessing techniques.

2.5 Biometrics Application: Challenges and Opportunities

The electrocardiogram is a cyclic signal that varies according to many factors, including heart geometry and individual attributes, that confer to it information about the person. The way cardiac conditions model the ECG has been extensively studied and used for diagnostic, including automatic tools. However, as shown, the acquisition techniques widely used for medical diagnostic can also be adapted or reformulated to allow for the compliance with real usable biometric systems, using the trending off-the-person acquisition approaches.

The off-the-person approaches, despite the great challenges it has to overcome, pave the way for a great new opportunity. For a highly dynamic environment such as driving, where movement is almost always present, but driver monitoring would still render very beneficial information, off-the-person ECG acquisition (such as on the steering wheel) poses itself as the way to easily, unobtrusively, and effectively monitor the driver. This monitoring would be advantageous not only for stress, fatigue, or attentiveness measurement, but also for identity recognition.

In the next chapter, the fundamental concepts on biometric recognition are presented. It includes a description and comparison of the most commonly explored biometric modalities, including the electrocardiogram and the characteristics that make it stand out on the task of identity recognition.

Chapter 3

Biometric Recognition: Fundamental Concepts

3.1 Biometric Systems: The Basics

Identification and authentication systems are, nowadays, nearly ubiquitous. Such systems are used almost everywhere, in an effort to more effectively and completely protect people, their identities, their information, and their possessions. Many of these systems use the so-called *entity-based/token-based* or *knowledge-based* methods, that are based, respectively, in something the user has (such as keys or identification cards) or something the user knows (such as passwords or PIN codes) (Agrafioti et al., 2011; Jain et al., 2011; Abo-Zahhad et al., 2014).

However, these traditional systems have fundamental flaws that put in risk everything they were built to protect. A password or PIN code can easily be lost, stolen, shared, forgotten, or discovered by a third-party (Abo-Zahhad et al., 2014). Keys or cards are also easily forgotten, and can be copied or counterfeit.

All these breaches originate in the fact that these methods evaluate the identity of a person based on extrinsic entities, and so anyone can, in possession of such entities, assume the identity of another person. This is not true for biometric traits, that require the physical presence of the person, in order to proceed with the identification (Jain et al., 2011; Kaur et al., 2014).

Biometric traits can be defined as human characteristics, whether anatomical, physiological, or behavioural, that include enough personal information to reliably serve as means of recognition of a certain individual (Agrafioti et al., 2011; Kaur et al., 2014). Physiological biometric traits are those that originate from direct measurements of a part of the human body, while behavioural biometric traits originate from the measurement of a person's actions (Abo-Zahhad et al., 2014).

Biometric recognition systems are, thus, systems that use statistical methods based on these biometric traits in order to identify or validate the identity of a human being, effectively using its intrinsical features as input (Sufi et al., 2010; Kaur et al., 2014; Fratini et al., 2015).

Biometric systems, just like traditional identification systems, are most needed for access control in critical, sensitive, and private settings, including civil contexts like research laboratories,

medical restricted areas, airport security, or entrepreneurial environments; military or police contexts such as the protection of crime evidences, witnesses information, and control of prison entrances and exits (Dar et al., 2015b); or even personal security settings like the control of access to a house, a car, or even a personal laptop or smartphone.

Biometric traits present the great advantages of providing high levels of security in all of these applications (Belgacem et al., 2012), being difficult to counterfeit or steal when compared with traditional credentials, and providing more ease of use and accessibility, as they do not require the use or transport of anything other than our bodies (Agrafioti et al., 2011).

3.2 Biometric Modalities

Various biometric modalities have been proposed over the last decades, following research findings proving certain human characteristics include, in fact, enough unique personal information (cf. Fig. 3.1).

Bolle et al. (2004), Sufi et al. (2010), Abo-Zahhad et al. (2014), and Kaur et al. (2014) listed several physiological biometric modalities already explored, like DNA, ear shape, face, facial thermogram, fingerprints, skin reflectance, hand geometry, hand vein thermogram, iris, odor, palm print, and retina, as well as behavioural traits like gait, keystroke, signature, voice, and lip motion.

Some of these are still under-developed for biometric purposes (mostly behavioural traits like gait or keystroke), and do not present enough recognition accuracy yet (Belgacem et al., 2012). On the other hand, face, fingerprints, iris, and voice have been the most studied and are currently the most used traits (cf. Fig. 3.2) (Kaur et al., 2014; Akhter et al., 2016). However, these have recently seen a growth of successful *spoofing* methods (methods of counterfeiting a certain person's feature to gain access through biometric systems).

Faces can usually be replicated by printing, in 2D or 3D, a replica of the person's face, shot by any medium to high-quality camera. Fingerprints can be easily copied using latex, silicone, or playdough. Voice can be recorded for posterior use, and iris and retina systems can be spoofed using contact lenses (Belgacem et al., 2012; Fratini et al., 2015).

To increase resistance to spoofing methods, it has become more and more common for biometric systems to integrate *liveness detection* (Prabhakar et al., 2003; Irvine et al., 2008). If, when acquiring an image of the subject's face, the system requires him to mimic a certain expression or movement, it will be easier to verify if the acquisition is from a living person or from a photo or pre-recording.

To overcome these difficulties, the so-called *medical biometric traits* have gained momentum. They are mainly biosignals such as the electrocardiogram (ECG), the electroencephalogram (EEG), electromyography (EMG), and other signals and features like phonocardiogram (PCG), photoplethysmogram (PPG), heart rate variability (HRV), blood volume pressure (BVP), or the brain response to stimuli (Agrafioti et al., 2012; Abo-Zahhad et al., 2014; Akhter et al., 2016). Due to their "hidden" nature, unlike face, fingerprints, or voice, these traits are harder to capture

or collect for spoofing purposes and, as they are linked to critical physiological mechanisms like the heart beat or the blood flow, they provide inherent liveness detection (Li and Narayanan, 2010).

Despite all advantages that human traits bring to biometric systems, the main disadvantage is their fuzzy nature. While traditional ID cards and keys either grant access or not, and passwords and PIN codes either are correct or wrong, and neither presents a third option, biometric traits are so variable, due to anatomical, physiological, and environmental factors, that no instance is equal to another (Agrafioti et al., 2011), and thus there is never a certainty of identification, only a fuzzy degree of confidence.

To achieve better performance, biometric systems can be developed to work with more than one biometric trait (e.g. face and fingerprint), being thus designated *multimodal biometric systems*. Another common option is to include both biometric traits and traditional credentials (e.g. face and PIN code), forming the so-called *hybrid systems*.

3.3 Qualities of a Biometric Modality

In order to consider a certain physiological or behavioural trait as fitting to be used as foundation for a biometric system, a few conditions are desirable (or even required), in order to guarantee the performance and reliability of the feature for human identification or verification.

Jain et al. (1999) defined seven desirable qualities that, when verified for a certain feature, attest its potential as a biometric modality. These qualities are:

1. *Universality*: the trait should be present in all subjects using the system;
2. *Uniqueness*: the trait should include enough personal information in order to present differences between all subjects, and thus allow their identification;
3. *Permanence*: despite the inter-subject variability desired (uniqueness), the trait should be sufficiently stable over time (reduced intra-subject variability) so as to allow the identification through the comparison of measurements in different instances;
4. *Measurability*: the trait should be easily and comfortably acquired and digitised, and its representation should allow easy processing and measurement;
5. *Performance*: a system based on such a trait should meet or exceed the recognition accuracy requirements, set by the context in which it will be applied;
6. *Acceptability*: there should be no foreseeable reservations that could make the subjects unwilling to allow the trait acquisition;
7. *Circumvention*: the trait should be as hard as possible to mimic or counterfeit, in any way, in order to prevent spoofing of the biometric system.

As stated before, these qualities are desirable but not compulsory. In fact, most of the biometric features used today do not fulfill all these requirements, at least completely. This originates in the

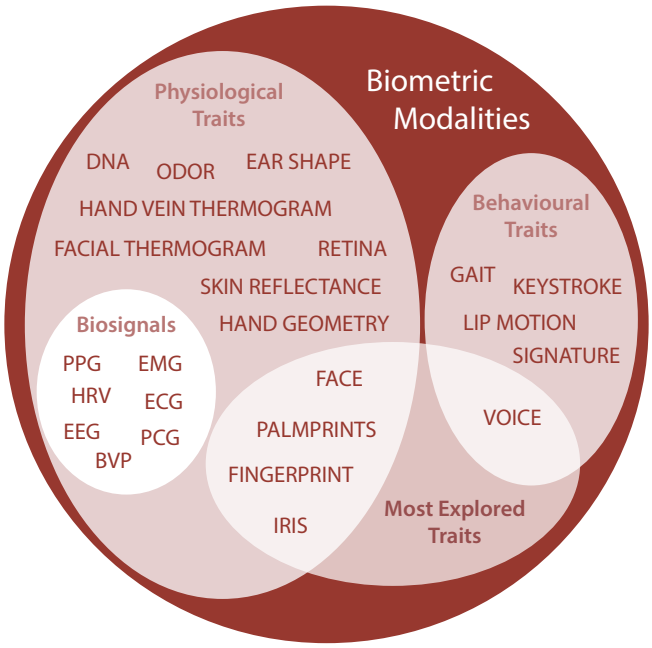

Figure 3.1: The world of biometrics: currently used traits for biometric recognition (based on information from Bolle et al. (2004); Sufi et al. (2010); Agrafioti et al. (2012); Kaur et al. (2014); Abo-Zahhad et al. (2014); and Akhter et al. (2016)).

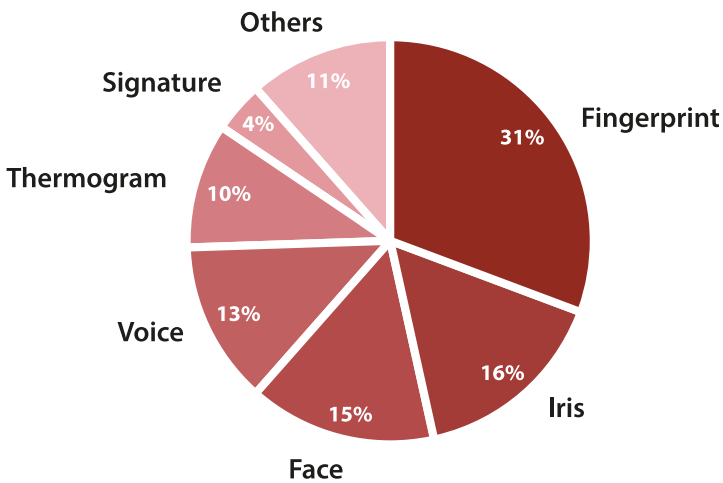

Figure 3.2: Share of the 2015 biometrics market by modality (adapted from Mani and Nadeski (2015)).

fact that each trait has a different nature, and so presents unique characteristics, that makes each trait more adapted for some application settings, and less for others (Agrafioti et al., 2011).

Faced with the seven qualities proposed by Jain et al. (1999), and considering this diversity between features, Abo-Zahhad et al. (2014) compiled two tables comparing 16 physiological and behavioural traits, according to the completeness with which they verify each quality (L - low, M - medium, and H - high).

In those tables, it is possible to verify that, generally, the lowest overall quality traits are the behavioural ones (gait, keystroke, signature, and voice) and phonocardiogram (heart sounds), with almost all of these presenting low performance, permanence, and distinctiveness. For the behavioural traits, the circumvention and universality are also downsides.

The traits with reported highest overall quality are the DNA, facial thermogram, fingerprint, iris, palm print, and ECG. DNA has the best results in universality, distinctiveness, permanence, performance, and circumvention. However, DNA performs very badly in collectability and acceptability. Facial thermogram is highly universal, distinctive, collectable, acceptable, and hard to circumvent, however, its permanence is low. Fingerprint performs very well in terms of distinctiveness, permanence, and performance, however it is only average in universality, collectability, acceptability, and circumvention avoidance. Iris presents very good scores in universality, permanence, distinctiveness, and performance, but lacks in acceptability and circumention. Palm print is average in all qualities but distinctiveness, permanence, and performance, where it excels. At last, ECG, the trait used in this work, excels in all qualities, with the exception of the average scores in collectability and acceptability.

3.4 Common Structure of a Biometric System

Biometric systems, independently of the trait considered, are tools that use pattern recognition algorithms to evaluate the similarity between biometric data acquired in the moment and biometric data previously stored (designated as *templates*) (Prabhakar et al., 2003; Jain et al., 2011). According to the result of such evaluations, the system then decides on executing (or not) an action, similar to the actions performed by traditional non-biometric identification systems, usually consisting in granting access to information, restricted areas, among others. To achieve this goal, biometric systems are composed by an array of modules with defined purposes, and are able to operate on different modes, both of which will be described in detail below.

3.4.1 System Modules

To ensure the proper functioning of the recognition system, a biometric system is composed by four modules (cf. Fig. 3.3) (Bolle et al., 2004; Jain et al., 2011), specifically:

- *Sensor* - The sensor constitutes the interface between the system and the subject and, thus, has an important role in the correct capture of the biometric trait data. The sensor must be designed in order to not only fit the specific requirements of the context on which the system

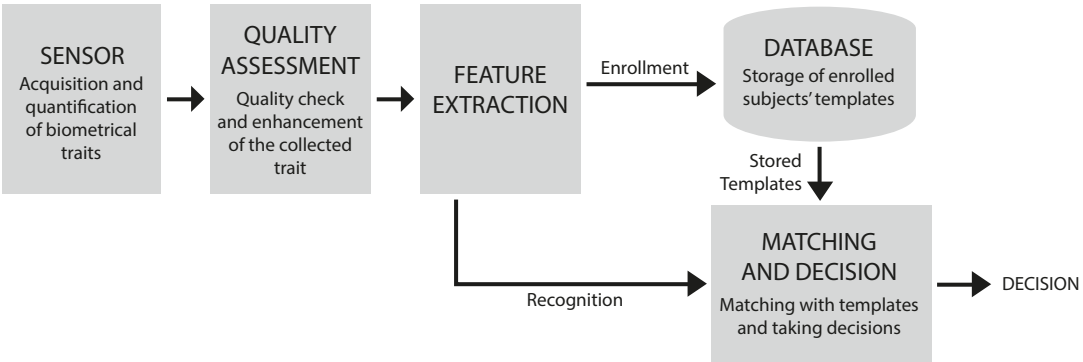

Figure 3.3: Schematic of the modules of a biometric system, their main functions, and the interactions between them (based on Prabhakar et al. (2003); Bolle et al. (2004); and Jain et al. (2011)).

will be installed, but also to minimize the possibilities of corruption of the data. Noise and external variability sources must be dealt with in order to guarantee that the signal is as pure as possible;

- *Quality assessment and enhancement* - The quality assessment module has the goal of guaranteeing the quality of the data acquired by the sensor module by, if necessary, subjecting it to enhancement procedures or rejecting it, according to its noise corruption;
- *Feature extraction* - The module of feature extraction consists in the processing and/or measurement of the acquired and enhanced data, in order to extract a set of features that provides high inter-subject discrimination power, and low intra-subject variability;
- *Matching and decision* - The matching and decision module is the part of the system where the comparison between collected data and stored templates occurs. From simple similarity measures to advanced pattern recognition algorithms, many options exist for this comparison, and its result, depending on the operation mode (described in subsection 3.4.2), is used to identify, validate, or reject an identity;
- *Database* - The database stores the enrolled individuals' biometric templates, and their identities, to be used by the other modules. For multimodal biometric systems, the database can contain several types of trait feature templates for each enrolled individual, and, for hybrid systems, the database can also store other data, such as codes or passwords.

3.4.2 Operation Modes

To adapt to the conditions and requirements of its context of application, a biometric system can operate in one of two different modes (cf. Fig. 3.4) (Prabhakar et al., 2003; Bolle et al., 2004; Agrafioti, 2011; Jain et al., 2011; Abo-Zahhad et al., 2014):

- *Authentication mode* - Also commonly designated as verification, in this mode, along with the collected biometric trait, the system is also fed an identity claim. Having the collected data and a stored template of the claimed identity, the system performs a *one-vs-one* comparison and decides, accordingly, whether or not the subject has the claimed identity;

- *Identification mode* - This mode is similar to authentication, however, no identity claim is given to the system. In this case, the system will have to compare the collected trait measurement with the data from each single identity stored in the system, thus performing a *one-vs-all* comparison. According to the results of all comparisons, the system can choose to assume the identity with the strongest match, or refuse to output an identity if no match is strong enough.

Generally, only one of these modes is chosen for a given biometric system to operate on throughout its lifetime. However, in order to correctly perform its function, the systems are able to operate in a different way, designated as the *Enrollment phase* (cf. Fig. 3.4). This phase works as a registration process, where a biometric trait of a new subject, previously unknown to the system, is collected, measured, and associated with an identity. The collected data usually undergoes quality check procedures and storage, to allow for future comparison in the above-mentioned modes (Prabhakar et al., 2003; Agrafioti, 2011).

3.4.3 Conventional vs. Continuous Biometrics

Biometric systems can also be divided into two categories: conventional and continuous systems (cf. Fig. 3.5). *Conventional systems* were the first to be developed, and are still the most common. In these systems, the acquisition of the biometric data occurs in one single period, after which the signal is given to the recognition system for identification or authentication, and a decision process outputs a result (Sim et al., 2007; Niinuma et al., 2010).

These steps occur non-simultaneously and each of them in a single moment in time, meaning that after the decision is taken, if access is granted, it is maintained throughout the lifetime of the session (Sim et al., 2007; Guennoun et al., 2009; Matta et al., 2011). In certain settings, this may be a great disadvantage: the user may forget to terminate the session before leaving the premises, and the system becomes vulnerable to intruders (Sim et al., 2007; Niinuma et al., 2010).

One of the techniques to avoid this is to expire the session after a predefined period of idleness. However, when users are, for example, reading a document or watching something on the screen, they can be locked out and have to frequently log themselves in, causing inconvenience and loss of productivity (Guennoun et al., 2009; Niinuma et al., 2010).

Continuous biometrics were developed to address this problem. The biometric data is continuously acquired by the sensor module, and the most recent segments of it are fed to the recognition process. This recognition process, instead of happening a single time, happens frequently throughout a session lifetime, in order to ensure the user is still there and an intruder has not replaced him (Sim et al., 2007; Niinuma et al., 2010).

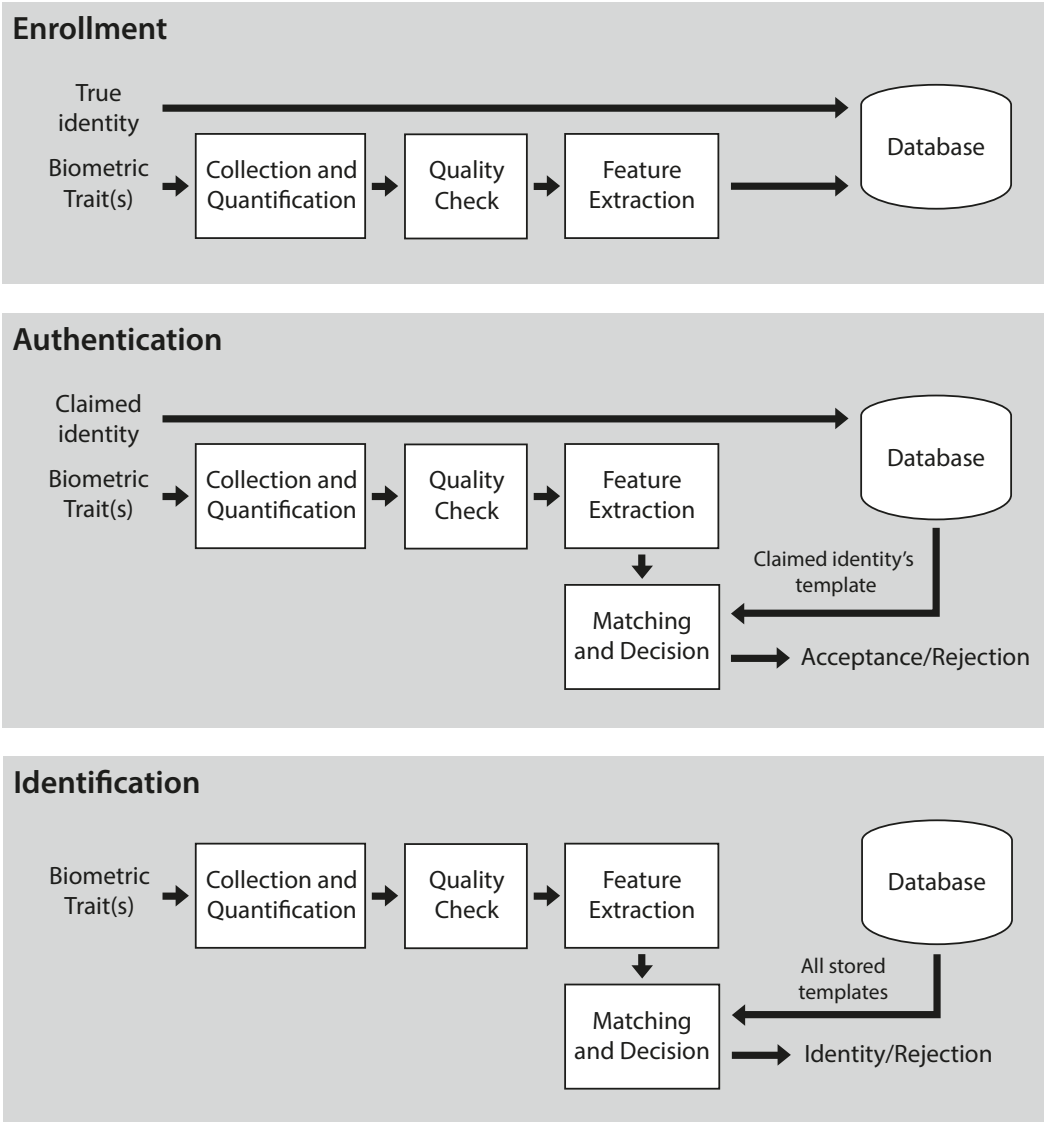

Figure 3.4: Schematics of the functioning of the identification and authentication modes, and the enrollment phase, in a biometric system (adapted from Prabhakar et al. (2003); Sufi et al. (2010); Jain et al. (2011); and Fratini et al. (2015)).

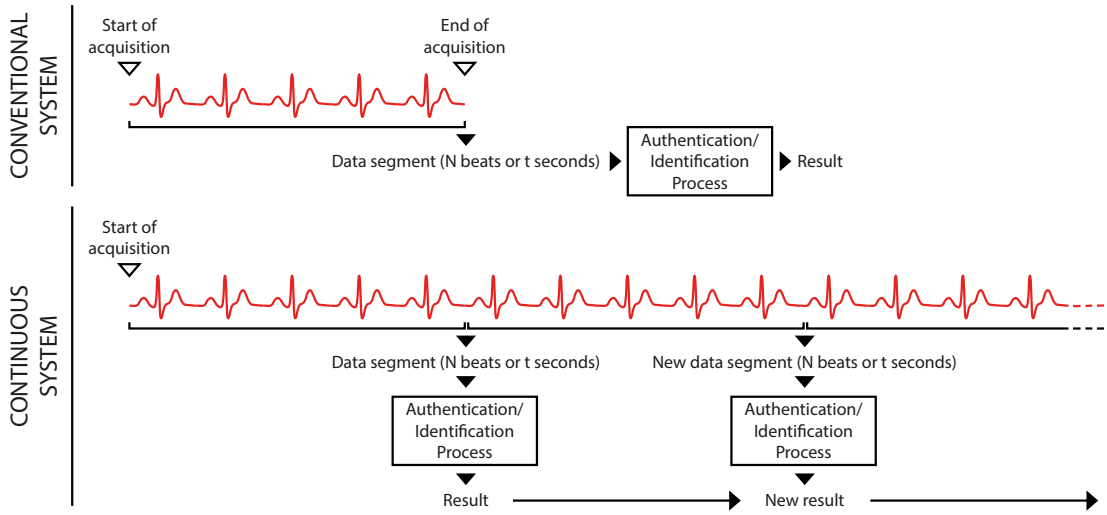

Figure 3.5: Illustration of the operation difference between conventional and continuous biometric systems (based on Guennoun et al. (2009); Matos et al. (2014); and Louis et al. (2016)). In this figure, in harmony with the focus of this document, an ECG signal is used as biometric, but the operation methods illustrated are similar for the other biometric traits.

New decision results will be made available frequently and, ideally, the session will be terminated if the user is deemed to have abandoned the system, effectively protecting against intruder attacks (Matta et al., 2011). If the user is still there, the biometric system will recognise him and maintain the session open, thus avoiding any inconveniences.

It is important to state that not all biometric traits are viable for an efficient continuous biometric system. Using fingerprints, the user would have to keep his finger still on the sensor throughout the session, and with face images, although continuous contact is not required, it would only take the user to turn his head to, possibly, make the system close the session (Niinuma et al., 2010). In that sense, the electrocardiographic signal, despite the need for contact during acquisition, is one of the most convincing traits, as the sensors can easily be designed to allow for maximum mobility and comfort of the user, without triggering recognition errors.

3.5 Performance Evaluation in Biometric Recognition

3.5.1 Conventional Biometrics

For any biometric modality, a robust performance evaluation under appropriate validation settings is the key to accurately predict the potential of the considered biometric system in real-life applications. This is also required if we want to reliably and objectively compare two different approaches.

Performance assessment in biometrics is reflected on a handful of metrics, used by most researchers on the field, selectively based on whether the problem is of identification or authentica-

tion. To understand those metrics, first we need to analyse the possible outcomes of the recognition procedure. In identification problems, the possible outcomes are:

1. The subject is enrolled, and the system correctly identifies him/her, granting access;
2. The subject is enrolled, but the system mistakes his/her identity for another enrolled subject, granting access under the identity of the latter;
3. The subject is enrolled, and the system fails to identify him/her with any enrolled subject, and rejects access;
4. The subject is not enrolled, and the system correctly rejects access;
5. The subject is not enrolled, but the system erroneously identifies him/her as one of the enrolled subjects, and grants access under this identity.

Three of these outcomes are system errors, and they can be divided in two categories (Grother et al., 2010). Type I errors comprise the situations in which the system fails to match the subject with the respective stored template (above, situation 3), and type II errors are those where the system erroneously matches the subject with a different person's stored template (situations 2 and 5).

It is easily understandable that, to achieve the best possible biometric system, the frequency of the situations 1 and 3 must be maximized, while all the type I and type II errors should be avoided. The easiest way to fine tune the behaviour of a method towards this goal is to vary the value of a similarity threshold T, that will define the minimum similarity (between acquisition and stored template) required for an identity to be attributed. Setting higher values for T will likely lead to a drop in the unwanted occurrence of situation 5 and, possibly, situation 2 as well, but will also increase the possibility of situation 3 occurring.

3.5.1.1 Identification Error Metrics

Based on the situations exposed above, for identification problems, performance metrics are based on their frequency. The most relevant error metrics were compiled by many authors, including Bolle et al. (2004), Grother et al. (2010), and Agrafioti (2011), and can be divided into two groups. The first group of metrics focuses on situations where the subject being identified is enrolled in the system:

- *True-Positive Identification Rate* (TPIR or Hit Rate) - For a total number of trials (identification procedures) where the subject is enrolled in the system, TPIR corresponds to the fraction of those where one of the system's R strongest predictions (among the list of enrolled candidates L) corresponds to the true subject identity:

$$TPIR(R,T,L) = \frac{\text{No. of trials where, among } L\text{, one of the strongest } R \text{ predictions above } T \text{ is correct}}{\text{Total number of trials}} \tag{3.1}$$

As most other metrics, TPIR depends on the defined threshold T, the selected number of top ranks R, and the list of enrolled candidates L (in identification, each enrolled subject is considered a candidate, and TPIR, as most metrics, varies not only with the size of L, but also with the variety and individual characteristics of each subject in L);

- *Identification Rate* (IDR or Accuracy) - If we take TPIR and consider only the single highest ranking identity prediction ($R = 1$) above the threshold T, we get the identification rate metric. It gives us the fraction of the total trials where the true identity was simultaneously above the threshold T and the method's strongest prediction. In the literature, this is one of the most used metrics;

$$IDR(T,L) = TPIR(1,T,L) \tag{3.2}$$

- *Reliability* - If we take TPIR and choose $R = N$ (where N is the number of enrolled subjects), we get the reliability metric. Reliability allows us to assess the fraction of times the true identity satisfies the minimum threshold constraint, regardless of its ranking;

$$Reliability = TPIR(N,T,L) \tag{3.3}$$

- *False-Negative Identification Rate* (FNIR or Miss Rate) - Represents the fraction of trials where the true identity does not correspond to one of the R strongest predictions above the threshold T;

$$FNIR(R,T,L) = \frac{\text{No. of trials where the true identity prediction is below } T}{\text{Total number of trials}} \tag{3.4}$$

- *Reject Rate* (RR) - This metric pertains to the very specific situations where all identity predictions stand below the defined threshold T, and the system has no choice but to reject to identify;

$$RR(T,L) = \frac{\text{No. of trials where all identities are below } T}{\text{Total number of trials}} = 1 - TPIR - FNIR \tag{3.5}$$

- *Misidentification Rate* (MIDR or, commonly, Misclassification or Error of Identification) - It is the complement of the Identification Rate, equivalent to FNIR with $R = 1$, and measures the fraction of total trials where, with the subject enrolled in the system, the true identity is not the system's top ranking prediction above T.

$$MIDR(T,L) = 1 - IDR(T,L) \tag{3.6}$$

The second group of metrics focuses on situations where the subject is not enrolled in the system:

- *False Positive Identification Rate* (FPIR) - It is the fraction of trials, using an unenrolled subject, where at least one of the system's identification predictions stands above the threshold T, thus being able to grant access to the unenrolled subject;

$$FPIR(T,L) = \frac{\text{No. of trials with unenrolled subjects where there is one or more predictions above } T}{\text{Total number of trials}} \tag{3.7}$$

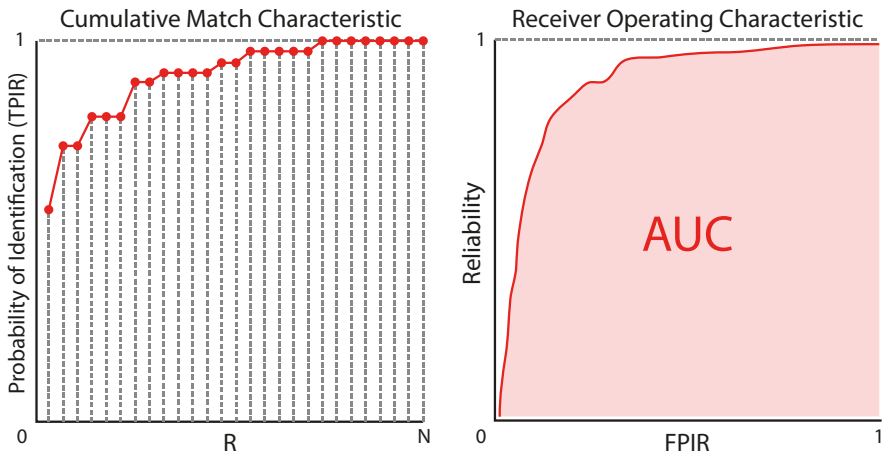

Figure 3.6: Examples of a Cumulative Match Characteristic (CMC) curve, and a Receiver Operating Characteristic (ROC) curve for an identification system. The Area Under the Curve (AUC) representation is worthy of notice.

- *Selectivity* - Similar to FPIR, but it counts the average number of predictions above the threshold T across all trials conducted with unenrolled subjects.

$$Selectivity(T,L) = \frac{\text{No. predictions above } T \text{ across all trials}}{\text{Total number of trials}} \tag{3.8}$$

To help in the fine-tuning of variables like the threshold T or the ranking R, and thus the development of more robust systems, these metrics can be varied or combined to build two-dimensional plots to visualise the evolution of two error metrics. These combinations are called Performance Characteristics (cf. Fig. 3.6), and the most common are:

- *Cumulative Match Characteristic* (CMC) - Plots $TPIR(R,T,L)$ (designated as probability of identification) with $T = 0$ against R, varying R between 1 and the total number N of enrolled subjects in list L;

- *Receiver Operating Characteristic* (ROC) - Plots Reliability against FPIR, for various T values. A point of equal error can be located, as a way to minimize both rejection of enrolled individuals and identification of intruders.

From both CMC and ROC curves, we can compute the also common metric *Area Under the Curve* (AUC), that serves as a global measure of performance of a method (the higher its value is, the better the method will, in general, perform).

3.5.1.2 Authentication Error Metrics

For authentication problems, due to having a single candidate identity to analyse instead of all enrolled subjects, possible outcomes differ from identification. If the subject is enrolled and the candidate identity matches his/hers, the system either grants or rejects access, and similarly if the subject is not enrolled or selects other identity instead of his own, resulting in four possible

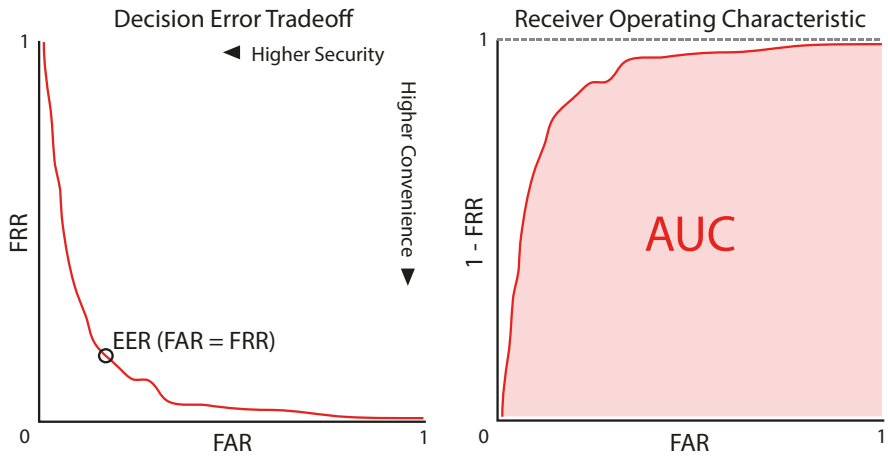

Figure 3.7: Examples of a Decision Error Tradeoff (DET) characteristic curve, and a Receiver Operating Characteristic (ROC) curve for an authentication system. The EER point illustration and the Area Under the Curve (AUC) representation are both worthy of notice.

outcomes. The metrics are, thus, distinct from those used for identification. They are (Bolle et al., 2004; Grother et al., 2010; Agrafioti, 2011):

- *False Acceptance Rate* (FAR or, commonly, False Match Rate, FMR) - The fraction of trials where the true and candidate identities do not match, but the system erroneously accepts it because the prediction score is above the threshold T (it thus depends on the value of T);

$$FAR(T) = \frac{\text{Number of trials were prediction score was above } T}{\text{Total number of trials}} \tag{3.9}$$

- *False Rejection Rate* (FRR or, commonly, False Non-Match Rate, FNMR) - Fraction of trials where the true and candidate identities match, but the system rejects access due to prediction scores below the set threshold (depends on T).

$$FRR(T) = \frac{\text{Number of trials were prediction score was below } T}{\text{Total number of trials}} \tag{3.10}$$

FAR and FRR both depend on the threshold defined, as illustrated on Fig. 3.8. Based on these metrics, two others can be defined (Bolle et al., 2004):

- *Convenience* - The complement of the FRR, the higher it is, the lesser the probability of an enrolled user to be incorrectly rejected access (the higher the better);

$$Convenience = 1 - FRR \tag{3.11}$$

- *Security* - The complement of the FAR, the higher it is, the lesser the probability of an impostor being unrightfully granted access (the higher the better).

$$Security = 1 - FAR \tag{3.12}$$

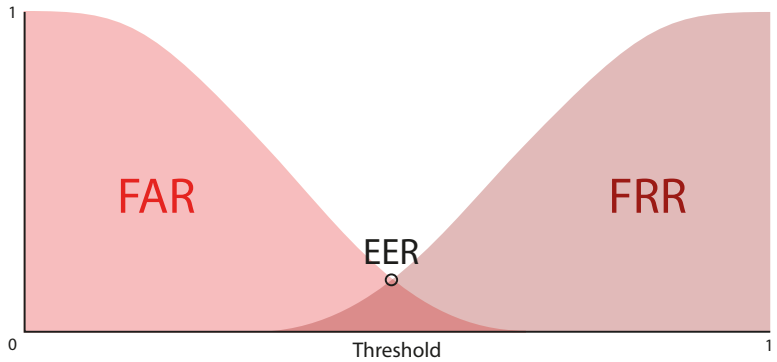

Figure 3.8: Evolution of False Acceptance Rate and False Rejection Rate with the threshold defined (adapted from Syris (2004)).

Similarly to the identification case, there are also some commonly used performance characteristic for authentication systems (cf. Fig. 3.7):

- *Decision Error Tradeoff* (DET) - Plots FRR versus FAR for varying values of T;
- *Receiver Operating Characteristic* (ROC) - For authentication, it plots $1 - FRR$ versus FAR for varying values of T. As with identification, AUC is usually used with ROC in authentication.

From these performance characteristic curves, two metrics are usually extracted. One of them is the *Equal Error Rate* (EER), that represents the point of the DET curve where $FAR = FRR$, and provides an equilibrium between convenience and security. The other is the *Area Under the Curve* (AUC), that has meaning equivalent to its identification use.

3.5.2 Continuous Biometrics

Subsection 3.5.1 of this document addressed the outcomes that are susceptible to be observed in conventional identification and authentication biometric systems, and the metrics that try to quantify them, in the quest to accurately measure the performance of such systems.

Despite the variety of metrics presented, all of them fall short when considering continuous biometrics. While conventional biometric systems are generally composed of a single identification/authentication procedure after a single period of collection, in continuous biometric systems *time* is of the essence (Sim et al., 2007).

In continuous biometrics, there is a sequence of identification/authentication procedures that feed from a single continuous collection of biometric data. Unlike conventional biometrics, in this continuous collection there are new possibilities that must be addressed: at some point in time the subject may leave the premises and lose contact with the acquisition system and, in those situations, an impostor may seize the opportunity to unrightfully access the system.

Time constraints are usually overlooked in conventional biometrics. In continuous biometrics, however, they are of utmost importance, as a good continuous biometric system must be able to

quickly detect the user absence or substitution, in order to terminate the session, and minimize the periods of vulnerability to intruders.

Sim et al. (2007), in the process of developing a new continuous biometric system, have designed criteria to ensure a continuous system reliability. Although these are specifically fitted to authentication multimodal systems, they can be reformulated to adapt to identification as well:

1. *Different modalities must have decision weights adapted to their relative reliability.* For multi-modality systems, each modality must contribute to the final decision proportionally to their reliability in identifying individuals;

2. *Older data influence on the decision must be discounted.* Although the same data can be used in several successive decision instances, the older data must always have their weight reduced, due to the uncertainty of whether or not those conditions still apply;

3. *Recognition certainty assessment must be possible at any point in time.* The system must be able to decide on whether or not the user is still there, or if it is the same user, at any moment, regardless of the availability of a biometric trait to be measured. Even if, for example, no ECG signal is collected by the sensor, or no face is detected by the camera, the system must be able to proceed with a decision anyway.

Recognizing the singular impact time has on continuous systems, Sim et al. (2007) considered that FAR or FRR (described in subsection 3.5.1) are not fitted to evaluate these tools in a complete fashion. Thus, they have also proposed three new metrics, *Time to Correct Reject*, *Probability of Time to Correct Reject*, and *Usability*, to evaluate the completeness with which an authentication system meets the specific criteria of continuous systems. Likewise, if these metrics are adapted to identification as well, we get:

- *Time to Correct Decision* (TCD) - TCD measures the time from the moment the logged user is substituted, by another enrolled individual or an impostor, and the moment the system detects this and takes an appropriate decision. For an ideal system, this should be zero, but that is virtually impossible to achieve, thus Sim et al. (2007) state that it should be sufficient for this to be always lower than W, called *Window of Vulnerability*, that is the minimum access time required for the impostor or wrong individual to cause any kind of damage;

$$TCD = \text{Seconds it took for the system to correct a wrong decision} \tag{3.13}$$

- *Probability of Time to Correct Decision* (PTCD) - Having TCD, and a defined W, PTCD measures the statistical probability of TCD being lower than W. The closer this metric value is to 1, the lesser the probability of an impostor having time to cause damage before the system acts;

$$PTCD = \frac{\text{Number of trials where TCD was lower than } W}{\text{Total number of trials}} \tag{3.14}$$

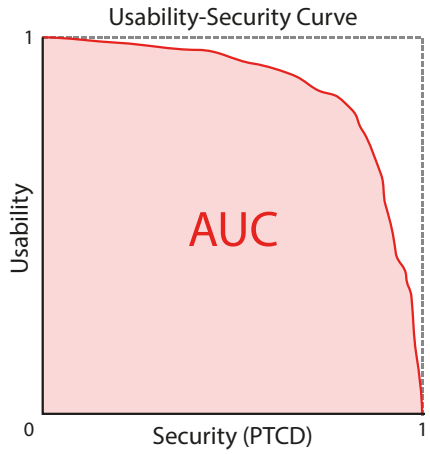

Figure 3.9: Example of a Usability-Security characteristic curve (adapted from Sim et al. (2007)).

- *Usability* - In a normal continuous biometric system, it is normal that among the several successive recognition procedures that take place in a single session, some of the decisions will be incorrect. In a t seconds session of a legitimate user, Usability measures the fraction of t where the user is deprived of access due to wrong decisions of the system. For any biometric system, this should be as close to zero as possible.

$$Usability = \frac{\text{Total sum of time where the user is granted access}}{\text{Total session time } (t)} \tag{3.15}$$

From these metrics, the authors also define the *Usability-Security curve* (USC), a new characteristic curve that plots Usability vs. PTCD for a varying threshold T (cf. Fig. 3.9). USC is described as an analog of the ROC curve and, thus, AUC can also be computed and considered as a good metric for evaluation of the global capabilities of a biometric system.

3.5.3 The Biometric Menagerie

Throughout this section, many ways of assessing and quantifying the performance of a biometric system have been presented. However, they all base themselves in the response of the system, and this response is not only reliant on the intrinsic characteristics of the system, but also on the population used to train the decision method.

Moreover, the population has power over the system's performance, not only because of the amount of individuals it includes, but also due to the personal characteristics of each, that will define their biometric traits. This means that a system can perform quite well with a certain population set, and err frequently with another, just because some populations (or some of the individuals inside them) offer more difficulty to the recognition process than others.

Thus, to accurately and objectively evaluate a biometric system solely based on its intrinsic characteristics, it is first necessary to evaluate the difficulty level of the population set used.

This was the objective of Doddington et al. (1998), that first defined the concept of *Biometric Menagerie*.

Biometric Menagerie is composed by a set of classes that are attributed to each individual in the population of a biometric system, according to their relative recognisability. All individuals are expected to be unique, however some can blend in with the population more easily than others, thus being less recognisable.

The categories defined by Doddington et al. (1998) were defined according to the two types of errors discussed in subsection 3.5.1. Following type I errors, where enrolled users are erroneously rejected access, we have:

- *Sheep* - Sheep compose, generally, the greatest part of any population and are usually recognized without any problem, matching well with their own templates and poorly with others', leading to low FRR and high IDR;

- *Goats* - Goats are those individuals that, mostly due to high intra-subject variability, frequently match poorly with their own templates, leading to low IDR and high FRR.

Following type II errors, where an impostor is incorrectly matched to an enrolled individual and granted access, two categories were defined:

- *Lambs* - Lambs consists of enrolled individuals that, because they have very "common" traits, make it easy for the system to match an impostor with their template. Because they open up the system to false acceptance or mis-identifications, they must be regarded as a vulnerability to the system;

- *Wolves* - Wolves are individuals that, for reasons similar to those of lambs, are easily matched by the system with the template of an enrolled individual (normally a lamb), and can thus easily gain access in their name.

The animal names used refer to the perceived nature of these animals: while sheep are usually calm and in order, goats are generally problematic, and wolves usually take advantage of the weakest, the lambs, to feast. This group of animals is also the origin of the classification name, menagerie.

About a decade later, Yager and Dunstone (2010) proposed a whole new set of classes (cf. Fig. 3.10). While the original four members were distinguished by the genuine/impostor scores (their recognisability), the new four differ in the relationship they present between genuine and impostor match scores.

Genuine match scores represent the rank of matches between the subject and his own template (the higher the better), while impostor match scores represent the rank of matches between the subject and another's template (the lower the better). These new categories were:

- *Chameleons* - Chameleons are individuals that generally match well with themselves (high genuine match scores), but present the setback of also matching well with others (high impostor scores), thus leading to false acceptances (higher FAR);

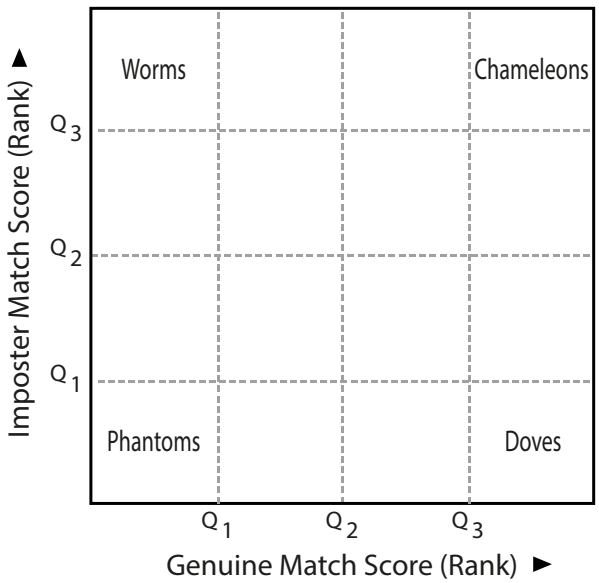

Figure 3.10: Illustration of the new classes of Biometric Menagerie (adapted from Yager and Dunstone (2010)).

- *Phantoms* - Unlike chameleons, phantoms are not only hard to match with other individuals (low impostor match scores), but also match poorly with their own templates, leading to false rejections (higher FRR);
- *Doves* - These are the best individuals for biometric recognition. They match well with themselves and poorly with others, thus presenting high genuine match scores and low impostor match scores;
- *Worms* - The complete opposite of doves, worms are the worst possible individuals with whom biometric recognition has to deal. They not only match poorly with themselves but are also easy to match with others, leading to a joint increase in both false acceptances and false rejections.

Taking into account the population used in the development and evaluation of a biometric system is, thus, of the highest relevance. Regardless of the biometric system, accurately assessing the distribution of each of these categories (both from Doddington et al. (1998) and Yager and Dunstone (2010)), will allow the weighing of the performance results with the difficulty of the population used.

3.5.4 Ideal Conditions for a Thorough Performance Assessment

Considering the information collected and presented in detail throughout this section on performance evaluation, it is possible to define some standards for the thorough performance assessment

of a biometric system.

First, to allow for the benchmarking with other existing methods, it is important to evaluate performance on both identification and authentication modes, regardless of which is the needed mode for the expected application settings. Benchmarking will also benefit from the use of popular publicly available databases, and this specific point is be further explained on Chapter 4.

For identification, Identification Rate (or Accuracy) is the most commonly used performance metric, while for authentication, the most common are the False Acceptance Rate, False Rejection Rate, and Equal Error Rate. All these should be used in the interest of benchmarking.

The evaluation of the population used (the signal collection), proven to be influential on the system's performance, should also be thoroughly performed, based on the biometric menagerie proposed by Doddington et al. (1998) and Yager and Dunstone (2010). Finally, in the specific case of continuous biometric systems, regardless of the modality, the metrics presented, based on the original authentication metrics proposed by Sim et al. (2007), should be used to address the characteristic time-dependency of this type of systems.

3.6 System Design Considerations and Concerns

System performance, addressed in section 3.5, should not be the only concern when studying a new biometric recognition approach or designing a biometric system. Bolle et al. (2004) advised biometric system designers to always be aware of six central aspects, of which the last two are the hardest to define due to the high sensitivity of biometric data:

- *System accuracy* - System accuracy, or system performance, measures the frequency with which the biometric system make a correct decision. For authentication, accepting a true identity claim, and rejecting false identity claims or impostors, and for identification, correctly matching the acquired trait to a stored template, and refusing to identify when the subject is not enrolled, are all central for the quality of the biometric system;

- *Computational speed* - Computational speed deals with the time required between the moment the user contacts with the system and the moment a decision is made. It depends on the time needed for trait acquisition, quality assurance, feature extraction, matching, and decision, and should be as high as possible, to improve usability. This aspect is especially important for continuous recognition systems;

- *Exception handling* - For some biometric traits, acquisition and/or enrollment is not always possible, or possible for all individuals that will use the system. As ECG is universal, this is not a common problem, but, e.g., with fingerprints or palmprints, people missing fingers or hands will not be able to enroll, and thus will be unable to use the system. A biometric system should be able to allow manual matching, to handle these exceptions;

- *System cost* - The system cost includes all expenses related to acquisition and processing equipment needed, algorithm development and implementation, routine maintenance, and operational costs;
- *Security* - Biometric systems decisions can serve as proof of the actions of a certain individual, and this, in some settings, can escalate to serious legal implications. Thus, it is extremely important to minimise the possibility of decision flaws that allow impostors to act under the identity of an authorised person;
- *Privacy* - Biometric systems require the storage of templates, consisting on discriminating information about each of the enrolled individuals. That information, in the interest of anonymity and security, should be kept as safe as possible, resorting to encryption techniques that allow matching but minimise the possibility of reconstructing the original acquired trait data.

As Bolle et al. (2004) stated, these considerations can sometimes seem contradictory. System accuracy decreases as the needs of computational speed and system cost are addressed. Working towards security can sometimes mean accepting higher false rejection rates, reducing system accuracy. Improving on all fronts usually has a great impact on system cost. All of these should be carefully weighed and combined in order to build a system that is affordable, efficient, accurate, secure, and usable.

3.7 Summary and Conclusions

Biometric identity recognition offers clear advantages over traditional techniques. While the keys and cards used in token-based methods are easy to clone, lose, or steal, and the PIN codes and passwords used in knowledge-based methods are easily forgotten or discovered by impostors, biometric systems ensure the relation between the identity and person requesting access, by using intrinsic characteristics of the individual.

The individual information content of many human characteristics has been studied, to prove their capabilities as biometric modalities, most commonly the face, fingerprints, iris, and voice. However, in recent years, the electrocardiogram has surged among many others due to its outstanding performance in the set of seven desirable qualities of biometric modalities.

However, the advantages of the electrocardiogram have to translate into accurate systems, able to compete with the remaining traits, currently much more developed. And that requires the development of ECG-based recognition approaches, integrated in carefully designed biometric systems, to follow the considerations exposed in this chapter. After the characterisation and presentation of some publicly available ECG collections on Chapter 4, some of these approaches, found in the literature, are presented on Chapter 5.

Chapter 4

ECG Signal Collections

4.1 The Significance of a Well-Structured Signal Database

As in all biometric systems and, in general, all machine learning problems, the development of an electrocardiogram-based biometric system requires a collection of signals, to guide the implementation, simulation and evaluation of the performance of the system, not only at the end, but throughout all phases of its development.

Having a well-structured signal collection is, thus, key to guide the development towards the exploitation of the best possibilities for the system, and accurately predicting its performance upon real-life application. To achieve such a complete collection, a few aspects have to be considered:

- *Electrode number* - Some researchers have studied the influence of these factors on the performance: Fang and Chan (2009) showed that the performance of their method with just one lead was 6% worse than with three, and Porée et al. (2016) used their method with 1, 2, 3, 6, and 12 leads, and the performance consistently decreased with lesser number of leads.

- *Electrode placement* - Zhang and Wei (2006) studied the use of two chest and two limbs leads (V1, V2, I, and II), individually, and found that both placement and orientation had influence on the system's performance, as both results with limbs leads were worse than with chest leads, V2 rendered better results than V1, and the same happened between leads II and I. It can be concluded that, the farthest from the heart the electrodes are placed, the most negative impact on the performance;

- *Sampling frequency* - Sampling is a necessary step in the conversion of the analogic, natural heart electric signal into the digital ECG. In this process, there is always the loss of fine details, that, as stated by Porée et al. (2016), can have influence on the recognition process. The lower the sampling frequency, the larger the amount of details that can be lost, and systems must be prepared to deal with diverse sampling frequencies. Beyond this, EMG interference is characterised by a wide range of frequencies, from 20 Hz up to 1 kHz. If low sampling frequencies are used for acquisition, the risk of aliasing increases, and the ECG range of frequencies (approximately 1-40 Hz) could be contaminated;

- *Subject posture and activity* - Porée et al. (2016), Wahabi et al. (2014), and Pathoumvanh et al. (2014) have studied the effect of varying conditions (supine rest, sitting, standing, tripod, exercise, and after exercise). They concluded that the database should follow the system's expected operation context, but it is important to reflect various conditions in order to ensure the systems capabilities, as intense activities usually have a negative effect on the recognition performance if the systems have not been trained accordingly;

- *Subject health* - As stated on Section 2.2, there are many factors that confer variability to the electrocardiographic signal. Some health issues, mainly arrhythmia, can generate intrasubject signal variability, over time, and encumber the recognition process (Safie et al., 2011; Dar et al., 2015a,b). A way to make systems robust against this is to include subjects with heart conditions, such as arrhythmia, in the datasets used during development and validation of the methods, effectively designing the system around the constraints that are imposed by such conditions;

- *Number of subjects* - The number of subjects of the population used to guide the development and assess the performance of the biometric recognition methods is a very important aspect to take into account. As previously exposed, in subsection 3.5.3, the diversity of individuals and their own characteristics may ease or difficult the job of the biometric systems, and the use of a signal collection with a great number of subjects ensures the presence of subject diversity, and thus increases the thoroughness of the performance assessment. Odinaka et al. (2012) have evaluated prior art approaches with a 265-subject database, and have concluded that most were not prepared for such a large set of individuals, and their results deteriorated quickly;

- *Acquisition sessions* - Biometric systems shouldn't be made to work perfectly in a one-time recognition instance. In reality, biometric systems will be used frequently to identify or authenticate individuals, and have to be prepared to deal with the variability over time of the ECG signals. Labati et al. (2013, 2014) have proven, using a Holter ECG database, that even in a small 24 hour period, the ECG signal varies enough to cause recognition errors in most biometric systems. Other researchers, like Silva et al. (2013) and Porée et al. (2016) showed that, between sessions weeks and even months apart, the performance of the systems quickly erodes;

- *Data origin* - Considering the previous criteria, the ideal signal collection is very specific of each method, according to its expected real application settings. Thus, researchers can opt to use publicly available collections, build their own private databases, or opt to use both. The main advantages of public databases are their generally larger size and availability of annotations, as well as the possibility of accurate performance benchmarking with other methods evaluated with the same databases. Private acquisitions are more useful for very

specific application settings (such as driving, heavy exercise, or unique electrode placement/configurations), as the researchers can adjust the acquisition protocols to fit the expected conditions. Generally, the use of both, when possible, is the best option, as it benefits from the advantages of both.

Summing up, all these factors can have an impact on the performance of an ECG-based biometric system. In order to correctly assess the capabilities of such systems, it is of the highest relevance to not only build a database that fits the system's expected application context, but also one that reflects all possibilities mentioned above, in order to study the use of the same biometric system in a wider set of contexts.

4.2 Publicly Available ECG Databases

Many researchers, when working with ECG signals, for biometric recognition purposes or for automatic diagnosis of medical cardiac conditions, opt for an own acquisition of data for private use. However, as the needs grew for complete signal collections, with more subjects, more medical conditions, on more sessions, spread across wider time frames, under different posture and activity conditions, researchers became more aware of the importance of public signal collections (Silva et al., 2014).

Physionet[1] (Goldberger et al., 2000) lead this effort by storing a wide variety of databases recorded both in hospitals, focused in medical applications (like the popular MIT-BIH collections), and in biometric recognition contexts (such as the ECG-ID database) (Silva et al., 2014). Other databases, resulting from private research works, are generally gently ceded by their owners to other researchers (like UofTDB or CYBHi). In this section, the most used and relevant publicly available ECG databases are presented, accompanied by short descriptions of their technical specificities. Finally, in Table 4.1, they are compared according to the criteria defined in section 4.1, and based on the work of Molina et al. (2007), Pouryayevali et al. (2014), Silva et al. (2014), Wahabi et al. (2014), and Merone et al. (2017).

AHA

The American Heart Association began in 1977 an effort to build a database for the guided training of health professionals on the diagnosis of arrhythmias. From this effort resulted the AHA ECG database[2], that includes 154 ECG recordings from actual patients, donated by various institutions.

The recordings are composed by 2 lead signals, sampled at 250 Hz, with 12 bit precision. Each recording is three hours long, of which the final 30 minutes' beats have been classified by experts, according to the presence or absence of seven types of arrhythmia. These records were made publicly available, requestable in the form of a DVD.

[1] Physionet ECG databases. Available on: `https://www.physionet.org/physiobank/database/#ecg` (visited on 19/01/2017).

[2] American Heart Association ECG database. Available on: `https://www.ecri.org/components/Pages/AHA_ECG_DVD.aspx` (visited on 20/01/2017).

CYBHi

The Check Your Biosignals Here initiative (CYBHi) was led by researchers from the Instituto Superior Técnico, the Instituto Superior de Engenharia de Lisboa, and the Escola Superior de Saúde da Cruz Vermelha Portuguesa, in Lisbon, Portugal, and was described by Silva et al. (2014).

The collection of ECG signals differed from common medical settings by following an *off-the-person*, biometric-focused approach, using minimally invasive electrodes: two dry (non-gelled) Ag/AgCl electrodes at the palms, and two electrolycras at the middle and index fingers.

Acquisition was performed at a sampling rate of 1 kHz, with 12 bit resolution, for two datasets: a short-term dataset, with single-session recordings of 65 volunteers; and a long-term dataset, where 63 subjects were recorded in two-sessions, three months apart. All sessions had an average length of 5 minutes, while the subjects, with their hands placed in contact with the electrodes, were exposed to a prepared video designed to cause emotional reactions.

ECG-ID

The ECG-ID is a database whose construction was entirely focused on biometrics. Lugovaya (2005), in the scope of her master thesis, collected short-duration ECG recordings from 90 subjects, with the purpose of individual identification. The final collection was also described on Nemirko and Lugovaya (2005), and was contributed to Physionet, where it is currently available.

For each subject, the ECG-ID database has between 2 and 20 different recordings (a total of 310), 20 seconds long, collected over a six month period. The signals were from Lead I (RA-LA), and were acquired at a sampling frequency of 500 Hz, with 12 bit precision on a nominal ± 10 mV range, using simple and highly acceptable limb-clamp electrodes at the wrists.

E-HOL 24h Holter

The E-HOL-03-0202-003 database (commonly designated as E-HOL 24h Holter) is one of many ECG databases created and made publicly available by the University of Rochester[3]. Unlike most other databases on their Telemetric and Holter ECG Warehouse that are mainly focused on medical diagnosis, this database was focused on biometrics.

A total of 203 healthy subjects were enrolled, and their ECG signals were recorded using a Holter monitor, during 24 hours. The acquisition was performed at a sampling rate of 200 Hz, with four electrodes placed on the chest, from 3 leads following a pseudo-orthogonal configuration.

European ST-T

The European ST-T database was created by Taddei et al. (1992), and made available for research purposes via the Physionet repository. The collection was intended for the development of algorithms for the analysis of ST and T-wave changes, but has also been used in research works on biometrics.

[3]University of Rochester Medical Center, Telemetric and Holter ECG Warehouse. Database E-HOL-03-0202-003. Available on: `http://thew-project.org/Database/E-HOL-03-0202-003.html` (visited on 20/01/2017).

Table 4.1: Summary of the technical specificities of the most commonly used publicly available ECG collections and the CardioID database.

Database	Origin	No. of Subjects	Sampling Rate (Hz)	Electrode Placement	Leads/ Electrodes	Health Conditions	Activity/Posture Settings	Sessions
AHA	American Heart Association[2]	154	250	Chest	2 / -	Various cardiac conditions	-	3 hour recordings
CYBHi	Silva et al. (2014)	128	1000	Palms + Fingers	2 / 4	None	Reactions triggered by sound and video	5 min. single-sess. and two sessions 3 months apart
ECG-ID[4]	Lugovaya (2005); Nemirko and Lugovaya (2005)	90	500	Wrists	1 / -	-	Sitting, unrestrained movement	Various 20 s rec. per subject during 6 months
E-HOL 24h	University of Rochester[3]	203	200	Chest	3 / 4	None	Ambulatory recordings	24 hour sessions
European ST-T [4]	Taddei et al. (1992)	79	250	Chest	2 / -	Various cardiac conditions	Ambulatory recordings	2 hour sessions
LTST [4]	Jager et al. (2003)	80	250	Chest	2/3 / -	Arrhythmia and ischaemia	Ambulatory recordings	21-24 h sessions
MIT-BIH Arrhythmia[4]	Mark et al. (1982); Moody and Mark (1990)	47	360	Chest	2 / -	None	Ambulatory recordings	30 min. sessions
MIT-BIH NSR[4]	Mark et al. (1982); Moody and Mark (1990)	18	360	Chest	2 / -	None	Ambulatory recordings	30 min. sessions
PTB[4]	Bousseljot et al. (1995)	290	1000	Chest + Limbs	15 / -	Various cardiac conditions	At rest only	1 to 5 per patient, 38.4-104.2 s
QT[4]	Laguna et al. (1997)	105	250	Chest	- / -	Various cardiac conditions	Rest and exercise	15 min. recordings
DriveDB[4]	Healey and Picard (2005)	9	456	Chest	1 / -	-	Rest, highway, and city driving	Sessions from 50 min. to 1.5 h
UofTDB	Wahabi et al. (2014)	1019	200	Fingers	1 / 2	None	Sit, stand, supine, exercise, and tripod	Up to six 2-5 min. recordings over 6 months
CardioID	CardioID Technologies[5]	Still underway	1000	Palms	1 / -	None	Coach bus and simulated driving settings	Continuous acquisition during 17 days

[4]Database available on Physionet: `https://www.physionet.org/physiobank/database/#ecg` (visited on 19/01/2017).
[5]CardioID Technologies: `https://www.cardio-id.com/` (visited on 30/01/2017).

The database is composed by 90 annotated excerpts of recordings from 79 subjects, from 2 leads, sampled at 250 Hz with 12 bit resolution, on a 20 mV nominal input range. Each excerpt has a duration of two hours, and were chosen in order to have a representative selection of abnormalities with origin in myocardial ischaemia, hypertension, ventricular dyskinesia, and effects of medication.

Long-Term ST

The Long-Term ST (LTST) database, available on Physionet and described on Jager et al. (2003), was created to include a variety of ST segment changes from various ischaemic and non-ischaemic origins, for the development of algorithms for the diagnosis of myocardial ischaemia.

This database includes 86 records from 80 subjects. The ambulatory records have duration between 21 and 24 hours, and the signals were acquired from two and three leads, with sampling frequency of 250 Hz, and resolution of 12 bit on a range of ± 10 mV.

MIT-BIH Arrhythmia

The MIT-BIH Arrhythmia database is one of the most used databases in ECG-based biometrics research. It was completed and made available in 1980, and was described by Mark et al. (1982) and Moody and Mark (1990). As many other databases available at the Physionet repository, it was constructed as a joint effort from the Massachusetts Institute of Technology (MIT) and the Beth Israel Hospital (BIH, now Beth Israel Deaconess Medical Center).

The database is composed by a total of 48 signals from 47 subjects. Each signal, with duration of 30 minutes, is an excerpt from ambulatory two-lead recordings, performed at a sampling rate of 360 Hz with 11 bit resolution, on a 10 mV range. Subjects were selected in order to obtain a representation of a wide variety of clinically significant arrhythmias.

MIT-BIH Normal Sinus Rhythm

The MIT-BIH Normal Sinus Rhythm database resulted from a joint effort of the MIT and the Beth Israel Hospital. Its origin is the MIT-BIH Arrhythmia database (Mark et al., 1982; Moody and Mark, 1990), presented above, and it includes excerpts deemed to be completely free from arrhythmias or other abnormalities. It includes records from 18 subjects (originally 20, 2 were removed due to the presence of ectopic beats), and is available on Physionet.

PTB

The PTB Diagnostic ECG database was created by the Physikalisch-Technische Bundesanstalt (German Federal Metrology Institute) and decribed in Bousseljot et al. (1995) and Kreiseler and Bousseljot (1995). Healthy subjects and individuals with various cardiac conditions (such as myocardial infarction, dysrhythmia, hypertrophy, or heart failure) were recorded, at rest, at a sampling frequency of 1 kHz, with 16 bit resolution over a range of ± 16 mV.

In total, the database includes 549 recordings from 290 subjects (1-5 per subject), with duration between 38.4 and 104.2 seconds. The recordings include simultaneously acquired signals from all 12 standard leads (I, II, III, aVR, aVL, aVF, V1, V2, V3, V4, V5, and V6) and 3 Frank leads (vx, vy, and vz).

QT

The QT database is available on Physionet, and was developed to aid the development of automatic methods of measurement of QT waveforms, as described in Laguna et al. (1997). This collection is a compilation of records from other public databases (such as the MIT-BIH Arrhythmia, the MIT-BIH Long Term, and the European ST-T), whose most relevant signals where selected, and extracts of 15 minutes were uniformly resampled to 250 Hz. In total, the database includes 105 recordings.

Stress Recognition in Automobile Drivers

The Stress Recognition in Automobile Drivers (DriveDB) database was created and described by Healey and Picard (2005), and had as purpose the collection of signals to allow the monitoring of stress in drivers, thus being the only database publicly available on Physionet composed of signals recorded in a driving context.

Various physiological parameters (electrocardiogram, electromyogram, and skin conductivity) were recorded from 9 subjects in a total of 18 driving sessions, with duration from 50 minutes to 1 hour and 30 minutes, including periods of rest (lower stress levels), highway driving, and city driving (higher stress levels). ECG signals were recorded from a modified Lead II with electrodes placed on the chest, and sampled at 496 Hz.

UofTDB

The University of Toronto ECG Database (UofTDB), described by Wahabi et al. (2014), was recently created by a group of the University of Toronto, the Bio.Sec Lab, specifically focused on biometric recognition and security. Thus, it was highly oriented towards the coverage of all criteria that make a database suitable for a thorough evaluation of the performance of a biometric system.

The ECG signals were captured from Lead I, using only two small ungelled electrodes at both thumbs of a total of 1019 subjects. For each subject, the database includes up to six recordings over a period of six months, each with duration between 2 and 5 minutes, sampled at 200 Hz. Subjects were recorded in various postures: supine, tripod, exercise, sitting, and standing.

4.3 The CardioID ECG Collection

CardioID Technologies[6] is a technology company based in Lisbon, Portugal, that aims to explore the use of the electrocardiogram for biometric recognition, for health and wellbeing monitoring, and for fatigue and drowsiness detection. Its foundation followed academic research work related to physiological computing, the automatic analysis of biosignals using signal processing and pattern recognition techniques.

The main current focus of CardioID is the use of ECG in driving settings. The CardioWheel, already mentioned in subsection 2.3.3 and presented in detail by Lourenço et al. (2015), is used by CardioID to acquire the ECG signals unobtrusively and continuously from the drivers, that will be used for their identification and monitoring.

For the development and evaluation of their monitoring and recognition algorithms, CardioID conducted the formation of a private signal collection. This collection includes signals from three separate initiatives. The first was SteeringWheel v1, with signals from ten subjects, acquired along 17 days during coach bus travels. These were unconstrained acquisitions, with periods of urban and highway travel (various levels of stress), and from long travels (various levels of fatigue).

SteeringWheel v2 was the second acquisition initiative, currently underway, that follows the same settings as v1. Finally, SteeringWheel Simulator consists of acquisitions in simulated driving settings, using a gaming steering wheel and a driving simulator software.

In all these initiatives, the leather dry electrodes were mounted on the steering wheel, and the sampling rate was of 1 kHz. Accelerometer data was also acquired, at 100 Hz, as well as lead-on sensor contact detection (at 1kHz), aiming to allow the posterior rejection of segments with lower quality. A fourth, smaller, acquisition initiative, RPiDemo, was conducted using stainless steel electrodes mounted in a Raspberry Pi case, at the fingers of 8 subjects, on several days, at 1 kHz. Albeit less noisy, these signals were similar to the signals obtained in driving conditions.

4.4 Summary and Conclusions

Electrode number, placement, sampling frequency, number of subjects, subject posture and activity during recordings, health conditions, and the duration, number, and time span of acquisition sessions, are all factors that influence the thoroughness that a ECG signal collection ensures in the assessment of the performance of a biometric system. They all should be carefully considered when choosing a public database or creating a private collection, before the development of a new ECG-based biometric system.

For this dissertation, CardioID Technologies provided the CardioID ECG database, that was expanded through own acquisition sessions during the dissertation, and served the purpose of testing the developed methods in the expected application context, the driving environment. However, as it is a private database, public databases were be necessary to both assess the performance of the method in different settings, and perform a benchmarking against state-of-the-art methods.

[6]CardioID Technologies: `https://www.cardio-id.com/` (visited on 30/01/2017).

Chapter 5

ECG Biometric Recognition: Prior Art

5.1 Brief History Overview

Since the beginning of the XX century, after the development of the first electrocardiogram by Willem Einthoven, the electrocardiogram has been evolving, by providing more accuracy, higher noise resistance, and the need for less electrodes while guaranteeing equal reliability. The main use was, and still is, aiding physicians in the task of diagnosing medical cardiac conditions (Biel et al., 2001; Agrafioti et al., 2011).

The use of ECG as a biometric trait is, when compared with other modalities like face or fingerprint, fairly recent, beginning to gain traction in the research community only after the beginning of this millennium.

To the best of our knowledge, the first study that hypothesised the use of ECG for individual recognition was a military technical report from the United States Department of Defense, dated from 1977 (Forsen et al., 1977). In this report, ECG was one of thirty-three personal attributes studied to improve the performance of authentication systems.

The ECG acquisition was surprisingly similar to today's most advanced systems, with the signal collection being performed using only two dry electrodes placed at the index fingers of the subject. Even at that time, with lesser technologies, among the thirty-three attributes, ECG deserved some of the highest scores in six out of seven evaluation criteria used, giving a glimpse into the results obtained decades later by Abo-Zahhad et al. (2014), already presented in section 3.3.

The first researchers to address exclusively the ECG as a biometric trait were Biel et al. (1999), from the Örebro University, at Sweden, using features directly output by an ECG acquisition device, conducing the matching and decision using Principal Component Analysis (PCA) for Soft Independent Modelling of Class Analogy (SIMCA). They obtained an identification rate of 100% with a reduced number of 20 enrolled subjects. Their research was also the first to propose the use of a single lead signal (Lead I), and was later complemented on Biel et al. (2001).

Around the same time, Kyoso et al. (2000), from the Kanagawa Institute of Technology and

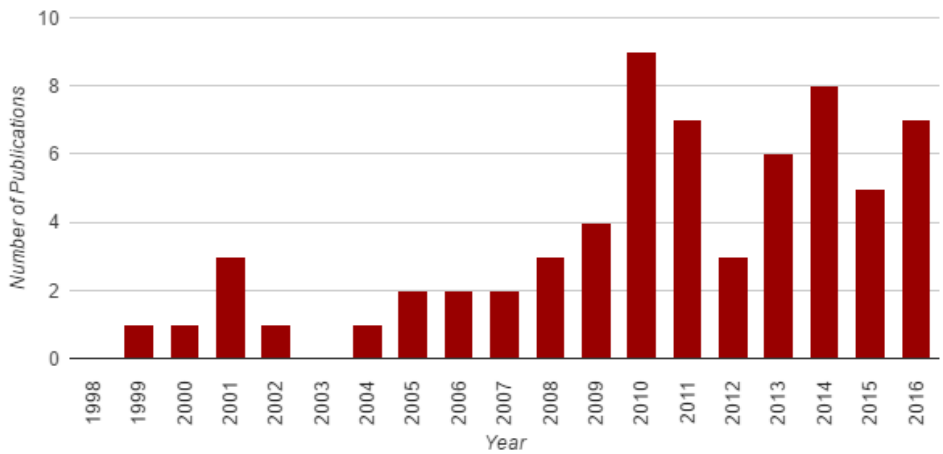

Figure 5.1: Distribution per year of the surveyed publications about biometric recognition using the ECG.

the Waseda University in Japan, were also pioneers in the study of biometrics with ECG. They extracted fiducial latency features from segmented heartbeats on ECG signals, and tried to recognise a group of 3 individuals using Mahalanobis distance and Discriminant Analysis, obtaining a final identification rate of 99.5%. Their work was shortly after restructured in Kyoso and Uchiyama (2001a) and Kyoso and Uchiyama (2001b) for new fiducial features and a larger set of 9 individuals.

The turn of the millennium was, indeed, a key moment for ECG as a biometric trait. Along with the first focused research publications on this theme, Hoekema et al. (1999) and van Oosterom et al. (2000) were also the first to analyse in detail the inter-subject variability of the electrocardiographic signal, essential for the task of identification.

Compelled by the promising results of these researchers, a great number of works have been published since then, including the research of Israel et al. (2003), that, to the extent of our knowledge, was the first to propose a multimodality biometric recognition method using ECG (in this case, accompanied by face).

Since then, and especially after 2009, this research field has gained traction, and ever more advanced and daring methods have been proposed, backed by ever more complete and thorough validation procedures. Of these, sixty-five original research publications (cf. Fig. 5.1) about biometric recognition using ECG were surveyed and, following the results, the panorama on the the current methods is presented in the next sections, in what concerns acquisition settings, preprocessing methods, signal preparation, features, matching, decision, and continuous analysis. For a general overview of the surveyed methods, please confer Table B.1.

5.2 Acquisition Settings

This section presents an overview of the surveyed methods according to the acquisition settings, i.e., the characteristics of the signal databases used for their performance evaluations. As already presented in section 4.1, there are some criteria that are crucial to define the quality of a signal database and, consequentially, the thoroughness they impose on the performance assessment. Following those criteria, this section presents an overview of the surveyed methods (cf. Table 5.1), highlighting some of the research works that stand out and can provide useful insight for future research.

5.2.1 Electrode Number and Placement

Most analysed publications refer only the leads of the ECG data used, and not the electrode number, but the latter can be inferred through the former. The vast majority of surveyed research works have used ECG signals from a single lead, and most of these from Lead I, following the conclusions of Biel et al. (1999, 2001) that recognition can be performed accurately using only one lead, and due to the much higher acceptability of acquisition using three electrodes (or two with virtual ground) on the wrists, hands, or fingers, rather than on the chest.

More unusual leads were the CM5, used by Kyoso et al. (2000), and the Lead II, used by Kyoso and Uchiyama (2001a,b) and Pathoumvanh et al. (2014). As already mentioned in section 4.1, Porée et al. (2016) used all 12 classical leads, in order to study the impact of the number of leads in performance, and Zhang and Wei (2006) used limbs and chest leads separately, for similar purposes.

As for the placement of electrodes, among the publications that specify such detail, sixteen have chosen the chest (including Kyoso et al. (2000); Fang and Chan (2009); Li and Narayanan (2010); Odinaka et al. (2010); Ye et al. (2010); and Agrafioti et al. (2012)), while other sixteen have chosen the limbs (such as Kyoso and Uchiyama (2001a,b); Palaniappan and Krishnan (2004); Molina et al. (2007); Wübbeler et al. (2007); Fatemian et al. (2010); N and Jayaraman (2010); and Wang et al. (2013)).

Most of those that have acquired the signal at the limbs used Lead I. For this lead, some fewer papers have chosen to acquire the signals at the palms instead of the wrists (Coutinho et al., 2010, 2011; Shen et al., 2011; Belgacem et al., 2012, 2013; Lourenço et al., 2012a; Lin et al., 2014), or even the fingers (Chan et al., 2008; Lourenço et al., 2011a, 2014; Silva et al., 2013; Matos et al., 2014; Louis et al., 2016), coining the term *off-the-person* acquisition (in opposition to *on-the-person* for acquisition at the chest or limbs), due to its much higher versatility and acceptability.

5.2.2 Sampling Frequency

As already explained in section 4.1, the sampling frequency at which the acquisition is performed can affect the quality of the signal by possibly losing some of the individual information it includes. Almost half of the surveyed publications that disclose this detail have used high sampling

Table 5.1: Summary of the acquisition of the data used by the surveyed approaches (ordered ascendingly by year, and alphabetically by the last name of the first author). Dashes denote the respective information could not be obtained, and publications where none of the details were specified were omitted from this table. Legend: column NS - Number of Subjects; column SF - Sampling Frequency (Hz); column EP - electrode placement (C - chest, L - limbs, P - palms, F - fingers); column L - leads (name, number, or S - several leads used); column HC - includes health conditions; column A - activity (R - rest, M - mild, H - heavy); column AS - various acquisition sessions.

Researchers	**Databases**	**NS**	**SF**	**EP**	**L**	**HC**	**A**	**AS**
Biel et al. (1999, 2001)	Private	20	-	Limbs	I		R	
Kyoso et al. (2000)	Private	3	250	Chest	CM5		R	
Kyoso and Uchiyama (2001a,b)	Private	9	500	Limbs	II		R	
Shen et al. (2002)	MIT-BIH NSR	20	-	-	1		R	
Palaniappan and Krishnan (2004)	MIT-BIH NSR	10	-	Limbs	I		R	
Israel et al. (2005)	Private	49	1000	-	-		M	
Plataniotis et al. (2006)	PTB	14	-	-	-		R	
Zhang and Wei (2006) [1]	Private	502	500	Limbs	I		R	
Zhang and Wei (2006)[1]	Private	502	500	Limbs	II		R	
Zhang and Wei (2006)[1]	Private	502	500	Chest	V1		R	
Zhang and Wei (2006)[1]	Private	502	500	Chest	V2		R	
Molina et al. (2007)	Private	10	128	Limbs	I		R	
Wübbeler et al. (2007)	PTB	74	500	Limbs	S		R	✓
Agrafioti and Hatzinakos (2008)	PTB + NSR	27	128	Limbs	I		R	
Chan et al. (2008)	Private	50	1000	Fingers	I		R	
Irvine et al. (2008)	Private	39	-	-	-		M	
Boumbarov et al. (2009)	Private	9	128	Limbs	I		R	
Fang and Chan (2009)	Private	100	250	Chest	S		R	
Fatemian and Hatzinakos (2009)	PTB + NSR	27	-	-	S		R	
Guennoun et al. (2009)	Private	16	300	-	-		R	✓
Coutinho et al. (2010)	Private	19	-	Palms	I		M	
Fatemian et al. (2010)	Private	21	200	Limbs	I		R	
Ghofrani and Bostani (2010)	PTB	12	1000	L. + C.	S		R	✓
Li and Narayanan (2010)	MIT-BIH NSR	18	128	Chest	-		R	
N and Jayaraman (2010)	MIT-BIH NSR	15	250	Limbs	II		R	
Odinaka et al. (2010)	Private	269	1000	Chest	1	✓	R	
Sasikala and Wahidabanu (2010)	MIT-BIH Arrh.	10	360	-	-	-	R	
Tawfik et al. (2010)	Private	22	500	Limbs	I		M	
Ye et al. (2010)[1]	Arrh.	47	360	Chest	2	✓	R	
Ye et al. (2010)[1]	NSR (short)	18	360	Chest	2		R	
Ye et al. (2010)[1]	Long Term	65	360	Chest	2	✓	R	✓
Ye et al. (2010)[1]	NSR (long)	18	360	Chest	2		R	✓

[1] Used various sets of data separately.

Researchers	Databases	NS	SF	EP	L	HC	A	AS
Coutinho et al. (2011)	Private	19	-	Palms	I		M	
Lourenço et al. (2011a)	Private	16	1000	Fingers	I		R	
Matta et al. (2011)	Private	10	-	Chest	2		M	✓
Safie et al. (2011)	PTB	112	1000	Limbs	I	✓	R	
Shen et al. (2011)	Private	168	500	Palms	I		R	
Sufi et al. (2011)	MIT-BIH Arrh.	-	-	-	-	✓	R	
Agrafioti et al. (2012)	Private	42	256	Chest	-		M	
Belgacem et al. (2012)	Arrh. + ST-T + NSR + PTB + Private	80	250	Palms	I		R	
Lourenço et al. (2012a)	Private	32	-	Palms	I		R	
Singh and Singh (2012)	ST-T + Arrh. + NSR + QT	73	-	-	-	-	R	
Belgacem et al. (2013)	Arrh. + ST-T + NSR + PTB + Private	80	250	Palms	I		R	
Coutinho et al. (2013)[1]	PTB	51	256	Chest	1		M	
Coutinho et al. (2013)[1]	Private	26	256	Chest	1		M	
Labati et al. (2013)	E-HOL 24h	185	200	Chest	S		M	✓
Matos et al. (2013)	Private	27	256	Chest	1		M	
Silva et al. (2013)	Private	63	1000	Fingers	I		R	✓
Wang et al. (2013)	PTB	100	500	Limbs	I		R	
Ergin et al. (2014)	MIT-BIH NSR	18	-	-	-		R	
Iqbal et al. (2014)	Private	30	1000	Chest	1		H	
Labati et al. (2014)	E-HOL 24h	185	200	Chest	S		M	✓
Lin et al. (2014)	Private	26	-	Palms	I		H	
Lourenço et al. (2014)	Private	63	1000	Fingers	I		R	✓
Matos et al. (2014)	Private	10	1000	Fingers	I		R	
Pathoumvanh et al. (2014)	Private	10	500	Limbs	II		H	
Zhou et al. (2014)	Private	20	250	Chest	-		R	✓
Brás and Pinho (2015)	PTB	52	500	-	-		R	
Choudhary and Manikandan (2015)	Arrh. + STC + QT + NSR + SLP	127	Many	-	-	✓	R	
Dar et al. (2015a,b)[1]	MIT-BIH Arrh.	47	-	-	1	✓	R	
Dar et al. (2015a,b)[1]	MIT-BIH NSR	18	-	-	1		R	
Dar et al. (2015a,b)[1]	ECG-ID	90	-	-	1		R	
Jahiruzzaman and Hossain (2015)	MIT-BIH Arrh.	11	360	Chest	2	✓	M	
Carreiras et al. (2016)	Private	618	500	Limbs	I		R	
Chun (2016)	ECG-ID	89	500	-	1		R	
Hejazi et al. (2016)	Private	52	200	Limbs	I		R	
Louis et al. (2016)	UofTDB	1012	200	Fingers	I		R	
Porée et al. (2016)	Private	14	1000	L. + C.	S		H	
Rezgui and Lachiri (2016)	NSR + Arrh.	48	360	-	-	✓	R	
Waili et al. (2016a)	PTB	14	-	-	-	-	R	✓

frequency values: ten used signals sampled at 500 Hz (such as Kyoso and Uchiyama (2001a,b); Zhang and Wei (2006); Wübbeler et al. (2007); and Tawfik et al. (2010)), and eleven at 1000 Hz (including Israel et al. (2005); Chan et al. (2008); Ghofrani and Bostani (2010); and Odinaka et al. (2010)).

Most of the remaining research works have used between 200 and 300 Hz, with the exception of four at 360 Hz (Sasikala and Wahidabanu, 2010; Ye et al., 2010; Jahiruzzaman and Hossain, 2015; Rezgui and Lachiri, 2016), and another four at 128 Hz, the lowest value found (Molina et al., 2007; Agrafioti and Hatzinakos, 2008; Boumbarov et al., 2009; Li and Narayanan, 2010). Some research groups, using various sets of data, have chosen to not uniformise their sampling frequencies, in order to test their methods' performance under these varying conditions (Wübbeler et al., 2007; Fang and Chan, 2009; Fatemian and Hatzinakos, 2009; Labati et al., 2013, 2014; Porée et al., 2016).

5.2.3 Subject Posture and Activity

Subject posture and activity can alter the heart rate, generate noise through movement, and alter the position of electrodes relative to the heart, thus influencing the acquired signal and the performance of the biometric system. The first to acknowledge this were Israel et al. (2005), by inducing anxiety with a set of seven two-minute tasks. They were followed by many other researchers, including Irvine et al. (2008), Coutinho et al. (2010), and Tawfik et al. (2010), who have opted not to restrain subjects to a resting posture while acquiring signals, or have used computer tasks to trigger emotions.

Iqbal et al. (2014) went even further and studied heavy activity settings, by acquiring signals from subjects walking up and down stairs, lying, resting, working, and walking normally. Lin et al. (2014) acquired signals from subjects exercising on a bicycle (as well as Porée et al. (2016)), and Pathoumvanh et al. (2014) opted to collect signals before and after jogging, to study the effects of increased HRV.

5.2.4 Subject Health

Health conditions can play a major role in undermining the performance of a biometric system, and it is desirable to prevent this by including subjects with health conditions, such as arrhythmia, in the signal database used. A great majority of the research works surveyed overlooked this aspect, especially in the first years. In 2010, Odinaka et al. (2010) and Ye et al. (2010) were the first researchers to include non-healthy individuals in their datasets, and proved the feasibility of such versatile systems by obtaining identification rates between 99% and 100%. Safie et al. (2011), Sufi et al. (2011), Choudhary and Manikandan (2015), and Dar et al. (2015a,b) were some who followed their steps.

5.2.5 Number of Subjects

In what concerns the number of subjects of the collections used, the analysed publications vary widely. The first works to use ECG for identity recognition, Biel et al. (1999, 2001), Kyoso et al. (2000), Kyoso and Uchiyama (2001a,b), Shen et al. (2002), and Palaniappan and Krishnan (2004), all relied on small databases from 3-20 subjects. This trend was maintained for a few years, as Zhang and Wei (2006) were, seemingly, the first researchers to use more than 50 subjects, with a collection of 502 individuals.

Since then, as is possible to analyse in Fig. 5.2, most research works have used between 10 and 50 subjects, with some exceptions, among which the most remarkable are Odinaka et al. (2010), Carreiras et al. (2016), and Louis et al. (2016), that used collections with signals of, respectively, 269, 618, and 1012 subjects.

5.2.6 Acquisition Sessions

A first step towards continuous biometrics, the focus of this work, is to acknowledge the variations the ECG signal suffers through time, and that a simple repetition of a conventional biometric recognition process does not suffice. Labati et al. (2013, 2014) performed a complete analysis of the permanence of ECG signals in a 24-hour period, using Holter acquisitions, and concluded that even in such a short period of time, the signal undergoes variations that are significant enough to influence the recognition process.

Acknowledging the need for the acquisition of signals widely spread in time, Wübbeler et al. (2007) was the first to work with multiple acquisition sessions, using signals from the PTB database (as well as, later, Ghofrani and Bostani (2010) and Waili et al. (2016b)). Shortly after, Chan et al. (2008) worked with acquisitions performed in three different days. Guennoun et al. (2009) used 15-minutes-long acquisitions for the first continuous approach (as well as Matta et al. (2011)), and Ye et al. (2010) studied their method's performance with MIT-BIH Long Term signals with and without arrhythmias. Lourenço et al. (2014), worked with two acquisitions, three months apart.

5.2.7 Data Origin

The surveyed research works can be categorised into two main groups according to the origin of the electrocardiographic signals used. Some researchers choose to use publicly available databases, such as the MIT-BIH Normal Sinus Rhythm (MIT-BIH NSR) database (Shen et al., 2002; Palaniappan and Krishnan, 2004), the MIT-BIH Arrhythmia database (Sufi et al., 2010; Jahiruzzaman and Hossain, 2015), the Physikalisch-Technische Bundesanstalt (PTB) database (Plataniotis et al., 2006; Wübbeler et al., 2007; Safie et al., 2011; Wang et al., 2013; Waili et al., 2016a), the E-HOL 24h Holter database (Labati et al., 2013, 2014), the University of Toronto (UofTDB) database (Louis et al., 2016), the ECG-ID (Chun, 2016), and some decided to use several (Ye et al., 2010; Choudhary and Manikandan, 2015; Dar et al., 2015a,b). These databases were previously described in more detail in Chapter 4.

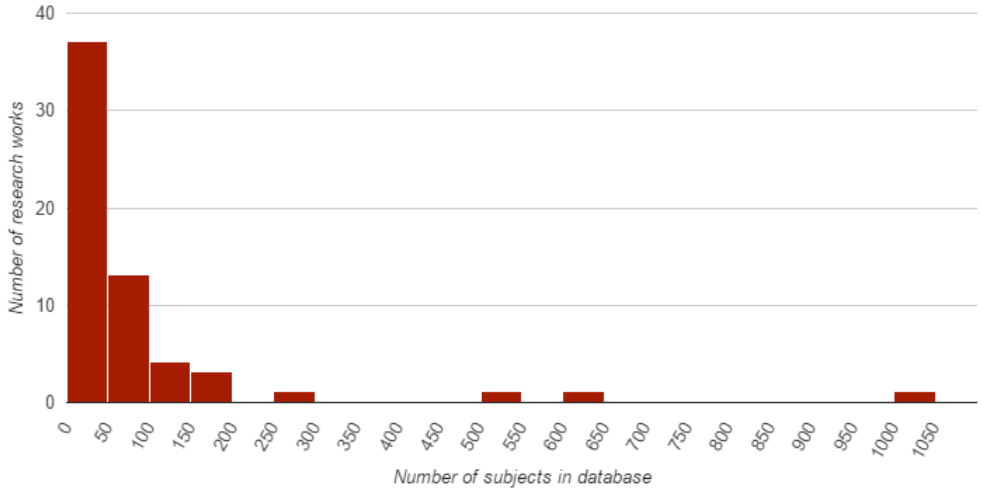

Figure 5.2: Histogram showing the number of analysed research works in each range of number of subjects in the database. For approaches evaluated with more than one database separately, only the largest was considered, and for those evaluated with a joint group of subjects' signals from more than one database, the total number of subjects was considered.

An approximately equal number of researchers opted for own acquisitions of data, following protocols defined by them, resulting in private collections of data, including Biel et al. (1999, 2001); Kyoso and Uchiyama (2001a); Israel et al. (2005); Molina et al. (2007); Chan et al. (2008); Boumbarov et al. (2009); Odinaka et al. (2010); Shen et al. (2011); Agrafioti et al. (2012); Matos et al. (2013); Lourenço et al. (2014); and Carreiras et al. (2016). A few researchers have resorted to the use of both publicly available databases and own acquisitions of ECG signals, thus making the most of the advantages of both (Belgacem et al., 2012; Coutinho et al., 2013).

5.3 Preprocessing Methods

As Rezgui and Lachiri (2016) state, and as already discussed in section 2.4, electrocardiogram signals are prone to be contaminated by noise from several sources during acquisition. The most relevant sources of noise are powerline interference from alternating current, electromyographic signals from muscles, and baseline interference from breathing movement.

To ensure the noise does not affect the recognition performance, the first step in most biometric systems is the preprocessing, that aims to denoise the signal and enhance its quality (Wahabi et al., 2014). In what concerns the process of denoising the acquired ECG signals, most of the current approaches fit into one of three categories: *filters*, *transforms*, or *line fitting*. A summary of the methods used in the analysed approaches is presented in Table 5.2.

Table 5.2: Summary of preprocessing methods used on the surveyed approaches (N.A. - number of approaches; HPF - highpass filter; LPF - lowpass filter; BPF - bandpass filter; NF - notch filter).

Type	N.A.	Specificities	N.A.	Approaches
Filters	35	BPF 1-40 Hz	5	Agrafioti and Hatzinakos (2008); Agrafioti et al. (2012); Belgacem et al. (2012, 2013); Louis et al. (2016)
		HPF 0.06 Hz + LPF 60 Hz + NF 50 Hz	3	Kyoso et al. (2000); Kyoso and Uchiyama (2001a,b)
		BPF 2-30 Hz	3	Coutinho et al. (2010, 2011, 2013)
		BPF 2-40 Hz	3	Israel et al. (2005); Safie et al. (2011); Rezgui and Lachiri (2016)
		LPF 30 Hz	2	Palaniappan and Krishnan (2004); Guennoun et al. (2009)
		BPF 0.5-40 Hz	2	Plataniotis et al. (2006); Zhou et al. (2014)
		HPF 0.5 Hz + NF 50 Hz	2	Labati et al. (2013, 2014)
		BPF 5-20 Hz	2	Silva et al. (2013); Carreiras et al. (2016)
		NF 60 Hz	1	Chan et al. (2008)
		BPF 0.05-60 Hz	1	Irvine et al. (2008)
		BPF 2-50 Hz	1	Fang and Chan (2009)
		BPF 0.5-150 Hz + NF 50 Hz	1	Ghofrani and Bostani (2010)
		BPF 0.05-40 Hz	1	Li and Narayanan (2010)
		HPF 0.5 Hz + LPF 500 Hz + NF 60 Hz	1	Odinaka et al. (2010)
		HPF 1Hz + LPF 40 Hz	1	Tawfik et al. (2010)
		BPF 0.5-30 Hz	1	Lourenço et al. (2011a)
		BPF 1-50 Hz	1	Shen et al. (2011)
		BPF 1-30 Hz	1	Lourenço et al. (2012a)
		HPF 0.5 Hz + NF 50 and 100 Hz	1	Matos et al. (2013)
		LPF 50 Hz	1	Matos et al. (2014)
		BPF 0.4-40 Hz	1	Pathoumvanh et al. (2014)
		BPF 0.5-45 Hz	1	Jahiruzzaman and Hossain (2015)
		LPF 45 Hz	1	Porée et al. (2016)
		HPF 0.05 Hz + LPF 40 Hz	1	Waili et al. (2016a)
		BPF (unspecified frequencies)	3	N and Jayaraman (2010); Ye et al. (2010); Matta et al. (2011)
Line fitting	3	6th order polynomial line	2	Dar et al. (2015a,b)
		Savitzky-Golay	1	Molina et al. (2007)
Transforms	3	DWT-based 3rd scale reconstruction	3	Fatemian et al. (2010); Chun (2016); Hejazi et al. (2016)
		DCT-based filtering	1	Choudhary and Manikandan (2015)
Others	4	Moving median filter + LPF 75 Hz	1	Wübbeler et al. (2007)
		HPF 0.05 Hz + DWT coefficients soft thresholding	1	Boumbarov et al. (2009)
		DWT-based 3rd scale reconstruction + Moving avg. filter	1	Fatemian and Hatzinakos (2009)
		Moving median filter + DWT	1	Sasikala and Wahidabanu (2010)
		NF 50 Hz + Moving avg. filter + LPF 40 Hz	1	Brás and Pinho (2015)

Filters

The vast majority of the surveyed approaches fall into this category. Despite the simplicity and low computational requirements of filters, they report great efficiency in the denoising of ECG signals. Three of the first works, Kyoso et al. (2000) and Kyoso and Uchiyama (2001a,b), used a cascaded bandpass filter (composed by a lowpass and a highpass filter) with cut-off frequencies of 0.06 Hz and 60 Hz, and a notch filter at 50 Hz. Simple bandpass filters were also commonly used, with bands between $1-40$ Hz (Agrafioti and Hatzinakos, 2008; Agrafioti et al., 2012; Belgacem et al., 2012; Louis et al., 2016), and $2-30$ Hz (Coutinho et al., 2010, 2011, 2013).

Transforms

A few research works, especially in the last couple of years, have used transforms in order to cancel the noise of the acquired ECG signal. Choudhary and Manikandan (2015) proposed an approach based on the Discrete Cosine Transform (DCT), and compared it with the results of bandpass filters to conclude it conferred less distortion to the signal. Fatemian et al. (2010) and Hejazi et al. (2016) used an approach of signal decomposition and reconstruction based on the Discrete Wavelet Transform (DWT) but, unlike the previous researchers, performed no comparison with other commonly used methods.

Line Fitting

Some of the prior art approaches analysed resorted to line fitting for signal denoising. Dar et al. (2015a,b) proposed a method of reducing the baseline wander and powerline interference noise that consists on fitting a sixth order polynomial curve to the ECG signal and detrending it.

Molina et al. (2007) used a Savitzky-Golay filter (Savitzky and Golay, 1964) to smooth the signal and, thus, reduce high-frequency noises like powerline interference. This type of filter fits a polynomial curve to the neighbourhood of each sample point in the signal, and adjusts its value accordingly.

Other Methods

Some of the researchers have opted for totally different methods for noise attentuation. These methods can be a combination of various common techniques, or totally different approaches.

Wübbeler et al. (2007) combined the use of a moving median filter, with 1 second width, and a lowpass filter with 75 Hz cut-off frequency. The subtraction of the filtered signal to the original signal reduced the baseline wander interference, while the lowpass filter attenuates the high frequency noise.

Other combination of techniques was proposed by Boumbarov et al. (2009). These researchers used a highpass filter with 0.5 Hz cut-off frequency, to attenuate baseline wander, and soft thresholding of Wavelet Transform coefficients is used for high frequency denoising.

Fatemian and Hatzinakos (2009) performed a Discrete Wavelet Transform reconstruction at the third scale to preserve most of the energy of the ECG signal, while reducing high frequency noise, and then applied a moving average filter for further smoothing. Brás and Pinho (2015) used a notch filter of 50 Hz cut-off frequency and a low pass filter of 40 Hz for high-frequency noise attenuation, and a moving average filter for baseline wander.

5.4 Signal Preparation Techniques

Frequently, prior biometric recognition methods using ECG, formulated over the last decades and surveyed in the scope of this dissertation, resort to the application of processing operations over the acquired ECG signal, besides denoising. These operations have the main goal to prepare the signal for the feature extraction phase, in order to maximise the performance of the system.

Table 5.3 summarises their uses in the reviewed approaches. Reference point detection and signal segmentation have been widely implemented in prior research. Amplitude normalisation, although much less popular than the previous two, was still more used than time normalisation. The most common situation verified was for researchers to choose the use of reference point detection and signal segmentation, and overlook the normalisations. Below, some of the most relevant examples of each, found in the literature, are presented.

5.4.1 Reference Point Detection

In order to aid posterior processes, such as signal segmentation, the preparation of the signal for recognition can include a step of detection of reference points. A great majority of the surveyed research works have used this technique, varying in the points detected and the methods used.

The Pan-Tompkins algorithm (Pan and Tompkins, 1985), developed for real-time QRS detection in ECG signals, was chosen by a great part of the surveyed prior art works, including Palaniappan and Krishnan (2004); N and Jayaraman (2010); Shen et al. (2011); Louis et al. (2016) and Waili et al. (2016a).

Molina et al. (2007) opted for the use of another method of QRS detection, developed by Trahanias (1993), to pinpoint the R peaks to posteriorly break the signal into R-R segments. The Engelse-Zeelenberg algorithm (Engelse and Zeelenberg, 1979) was the option of Lourenço et al. (2011a), for the detection of the QRS complexes and, for Lourenço et al. (2012a), the R detection was performed using a Steep-slope thresholding approach devised by Christov (2004).

Other methods specified in the surveyed approaches for detection of the R-peaks were a First order Gaussian Differentiator (FOGD) based technique proposed by (Choudhary and Manikandan, 2015), followed by peak correction stages, and a thresholding of local maxima used by Dar et al. (2015a,b).

Table 5.3: Summary of the use of signal preparation techniques by the surveyed approaches (ordered ascendingly by year, and alphabetically by the last name of the first author). Checkmarks and crosses denote, respectively, the approach used or did not use the technique, and dashes denote the information could not be obtained.

Researchers	**Ref. point det.**	**Signal segm.**	**Amp. norm.**	**Time norm.**
Biel et al. (1999, 2001)	-	-	-	-
Kyoso et al. (2000)	✓	✓	×	×
Kyoso and Uchiyama (2001a,b)	✓	✓	×	×
Shen et al. (2002)	✓	✓	✓	✓
Palaniappan and Krishnan (2004)	✓	✓	×	×
Israel et al. (2005)	✓	✓	×	×
Saechia et al. (2005)	✓	✓	×	✓
Plataniotis et al. (2006)	×	×	×	×
Zhang and Wei (2006)	✓	✓	×	×
Molina et al. (2007)	✓	✓	×	×
Wübbeler et al. (2007)	✓	✓	×	×
Agrafioti and Hatzinakos (2008)	×	✓	×	×
Chan et al. (2008)	✓	✓	×	×
Irvine et al. (2008)	×	✓	✓	✓
Boumbarov et al. (2009)	×	✓	×	×
Fang and Chan (2009)	✓	✓	✓	×
Fatemian and Hatzinakos (2009)	✓	✓	×	✓
Guennoun et al. (2009)	✓	×	×	×
Coutinho et al. (2010)	✓	✓	×	×
Fatemian et al. (2010)	×	✓	×	✓
Ghofrani and Bostani (2010)	×	-	×	×
Li and Narayanan (2010)	-	✓	✓	✓
N and Jayaraman (2010)	✓	×	×	×
Odinaka et al. (2010)	✓	✓	✓	×
Sasikala and Wahidabanu (2010)	✓	✓	×	×
Tawfik et al. (2010)	✓	✓	✓	✓
Ye et al. (2010)	✓	✓	×	×
Coutinho et al. (2011)	✓	✓	×	×
Lourenço et al. (2011a)	✓	✓	✓	✓
Matta et al. (2011)	×	✓	×	×
Safie et al. (2011)	×	✓	✓	×

Researchers	Ref. point det.	Signal segm.	Amp. norm.	Time norm.
Shen et al. (2011)	✓	✓	✓	✓
Sufi et al. (2011)	✓	✓	×	×
Agrafioti et al. (2012)	×	✓	✓	×
Belgacem et al. (2012)	✓	✓	✓	×
Lourenço et al. (2012a)	✓	✓	✓	×
Singh and Singh (2012)	✓	✓	✓	✓
Belgacem et al. (2013)	✓	✓	✓	×
Coutinho et al. (2013)	✓	✓	✓	×
Labati et al. (2013)	✓	✓	×	×
Matos et al. (2013)	✓	✓	×	×
Silva et al. (2013)	✓	✓	×	×
Wang et al. (2013)	×	✓	✓	×
Ergin et al. (2014)	✓	✓	×	×
Iqbal et al. (2014)	✓	✓	×	×
Labati et al. (2014)	✓	✓	×	×
Lin et al. (2014)	×	×	×	×
Lourenço et al. (2014)	✓	✓	×	×
Matos et al. (2014)	✓	✓	×	×
Pathoumvanh et al. (2014)	✓	✓	×	×
Zhou et al. (2014)	✓	✓	×	×
Brás and Pinho (2015)	×	×	✓	×
Choudhary and Manikandan (2015)	✓	✓	✓	✓
Dar et al. (2015a)	✓	✓	✓	×
Dar et al. (2015b)	✓	✓	✓	×
Jahiruzzaman and Hossain (2015)	×	×	×	×
Carreiras et al. (2016)	✓	✓	×	×
Chun (2016)	✓	✓	×	×
Hejazi et al. (2016)	×	×	×	×
Louis et al. (2016)	✓	✓	×	×
Porée et al. (2016)	✓	✓	×	×
Rezgui and Lachiri (2016)	✓	✓	×	×
Waili et al. (2016a)	✓	✓	✓	×
Overall usage	49 in 61	54 in 61	21 in 62	11 in 62

5.4.2 Signal Segmentation

As is possible to extract from Table 5.3, the signal segmentation is the most commonly used signal preparation technique among the surveyed approaches. It is used to limit the signal span for feature extraction, or to set a fixed size to ease template matching when the feature is the signal itself.

In some cases, the segmentation follows the reference point location and consists on the cropping of the QRS complex and/or other waveforms (Tawfik et al., 2010; Sufi et al., 2011; Waili et al., 2016a), or is meant to include the whole heartbeat (or a majority of it) and is thus performed at fixed distances before and after detected R-peaks or QRS complexes (Zhou et al., 2014; Carreiras et al., 2016; Louis et al., 2016). Other research works, such as Odinaka et al. (2010); Matta et al. (2011); and Ergin et al. (2014), have performed segmentation of the signal using sliding windows, with or without overlap, regardless of the completeness of the heartbeat cycles inside it.

Related to the signal segmentation is also the alignment and averaging of various signal segments. The alignment is generally performed using the R-peak as reference after its location, or it's performed through correlation. It usually serves as a way to guarantee the template and the collected signal aren't affected by rare variabilities, that distort the personal information the signal contains, and could threaten the recognition task. This approach accompanied the signal segmentation in several of the analysed research works, such as Belgacem et al. (2012); Coutinho et al. (2013); Lourenço et al. (2014); and Choudhary and Manikandan (2015).

5.4.3 Amplitude Normalisation

The amplitude of the signal and its DC offset depends highly on the acquisition equipment used, and a versatile and reliable biometric system must be able to protect from equipment variability (Irvine et al., 2008). Thus, amplitude normalisation seeks to rescale all signals to the same amplitude range, and to align their baseline to the same DC offset (usually zero).

Irvine et al. (2008), one of the first groups to apply this technique, defined an expression for the amplitude normalisation of a heartbeat segment, that sets the maximum value to 1 and the minimum to 0 (Z and U denote, respectively, the normalised and the original segments, h denotes the subject, i denotes the heartbeat, and t denotes each individual sample):

$$Z_{t,i}^{h} = \frac{U_{t,i}^{h} - \min(U_{t,i}^{h})}{\max(U_{t,i}^{h}) - \min(U_{t,i}^{h})} \tag{5.1}$$

This same expression was used by some posterior approaches that applied amplitude normalisation, such as Fang and Chan (2009); Li and Narayanan (2010); and Safie et al. (2011). Odinaka et al. (2010) opted to normalise heartbeat segments by subtracting the signal mean and dividing by the standard deviation, and Tawfik et al. (2010) and Lourenço et al. (2011a) simply divided the entire beats by the R-peak amplitude value.

5.4.4 Time Normalisation

Time normalisation techniques aim to reduce the impact of heart rate variability on the electrocardiogram's heartbeats. The most noticeable impact is the total length of the heartbeat and, so, most of the surveyed approaches performed normalisation through the shrinking of the segmented signal to a predefined length, usually through signal resampling (Saechia et al., 2005; Li and Narayanan, 2010; Lourenço et al., 2011a).

Tawfik et al. (2010) normalised only the QT waveform, more prone to variations from the heart rate, using the Framingham study formula to compute the linearly corrected QT length, through:

$$QT_{LC} = QT + 0.154(1 - RR). \tag{5.2}$$

One unique and interesting approach for time normalisation was proposed by Fatemian and Hatzinakos (2009). The researchers broke each ECG heartbeat into its key waveforms (P, QRS, and T), which are individually resampled and then pierced back together, with regulated intervals between them. This approach allowed for an efficient reduction of heart rate variability impact on the signal, while avoiding the distortion of the individual waveforms of typical normalisation approaches.

5.5 Feature Extraction Methods

The variety of features extracted from the ECG signal for individual recognition in prior art approaches is impressively wide. These can vary in their type - *fiducials*, or *non-fiducials* (Matta et al., 2011; Matos et al., 2013) - and in the domain they are extracted - *time domain*, *frequency domain*, or others (Hassan et al., 2016). Extracted features can, additionally, suffer dimensionality reduction for performance enhancement (Wahabi et al., 2014). In the next subsections, these aspects are presented in more detail, and the most relevant state-of-the-art examples are exposed.

5.5.1 Approach Types

ECG-based biometric systems can be categorised in three groups, according to the features they use: *fiducial-based*, *non-fiducial*, or *hybrid* (Matta et al., 2011; Lourenço et al., 2011a; Matos et al., 2013).

Fiducial-Based Approaches

Fiducial-based approaches are those that use fiducials (specific local maxima or minima points such as P, Q, R, S, or T) as features, measuring their amplitude and time, and the amplitude difference or time intervals between two or more of these (Matta et al., 2011; Lourenço et al., 2011a; Matos et al., 2013). Fiducial approaches were a remarkable majority through the first years of this area of research. Seven of the ten earliest publications used exclusively fiducial measurements as features, and the first to reject this type of features was only published in 2005.

Table 5.4: Summary of the methods of feature extraction and dimensionality reduction used on the surveyed approaches (ordered ascendingly by year, and alphabetically by the last name of the first author). The dashes denote the respective method did not apply a dimensionality reduction technique. Legend: Column T - approach type: F - fiducial-based; N - non-fiducial; and H - hybrid. Column D - feature domain: T - time; F - frequency; and O - others.

Researchers	**T**	**D**	**Feature extraction**	**Dim. Reduction**
Biel et al. (1999, 2001)	F	T	10 Lead I fiducial features	Correlation matrix
Kyoso et al. (2000)	F	T	PQ and QT times	-
Kyoso and Uchiyama (2001a,b)	F	T	QRS duration and QT time	-
Shen et al. (2002)	F	T	RQ, RS, and ST amplitudes, QS and QT intervals, RS slope, and QRS triangle area	-
Palaniappan and Krishnan (2004)	H	T	Amplitudes of R, QR, RS, width of QRS, R-R interval, and beat form factor	-
Israel et al. (2005)	F	T	RQ, RS, RP, RL, RP', RT, RS', RT', P width, T width, ST, PQ, PT, LQ, ST'	Wilkes' lambda stepwise correlation
Saechia et al. (2005)	N	T	Fourier transform of PQRST (whole), P, QRS, and T	-
Plataniotis et al. (2006)	N	T	Autocorrelation coefficients	Discrete Cosine Transform
Zhang and Wei (2006)	F	T	Fiducial amplitudes, durations, intervals, levels, and areas	Principal Component Analysis
Molina et al. (2007)	N	T	R-R signal segments	-
Wübbeler et al. (2007)	N	T	2D QRS segments (combination of leads I, II, and III)	-
Agrafioti and Hatzinakos (2008)	N	T	Normalised autocorrelation	DCT/LDA
Chan et al. (2008)	N	T	Signal-averaged ECG	-
Irvine et al. (2008)	N	T	Covariance matrix eigenvectors	-
Boumbarov et al. (2009)	N	T	Cardiac cycle vector matrix	PCA and LDA
Fang and Chan (2009)	N	O	Avg. beat phase space portrait	-
Fatemian and Hatzinakos (2009)	N	T	Heart-rate-normalized single heartbeat template construction	-
Guennoun et al. (2009)	F	T	Fiducial amplitude and time features	Physiological-state-independent feature selection
Coutinho et al. (2010)	N	T	Uniformly quantised avg. beats	-
Fatemian et al. (2010)	N	T	Averaged ensemble heartbeat	-
Ghofrani and Bostani (2010)	N	T	AR coeff.; PSD; Lyapunov exponents; Approximation Entropy; Higuchi Fractal Dimension; Shannon Entropy	-
Li and Narayanan (2010)	N	O	Hermite poly. expansion; and Cepstral features	Heteroscedastic Linear Discriminant Analysis
N and Jayaraman (2010)	F	T	P, T, ST, PR, QRS and QT intervals	-
Odinaka et al. (2010)	N	F	Log-STFT spectrogram Gaussian models	Bin selection
Sasikala and Wahidabanu (2010)	F	T	Fid. amplitudes and differences	-
Tawfik et al. (2010)	N	F	QRS DCT coefficients	-
Ye et al. (2010)	N	F	Daubechies DWT and ICA	-
Coutinho et al. (2011)	N	T	User-tuned Lloyd-Max quantised averaged ensemble heartbeats	-

Researchers	T	D	Feature extraction	Dim. Reduction
Lourenço et al. (2011a)	N	T	Average normalised beat	-
Matta et al. (2011)	N	T	Autocorrelation coefficients	LDA
Safie et al. (2011)	N	O	Pulse Active Ratio	-
Shen et al. (2011)	F	T	Amplitudes, durations, slopes, angles, and QRS area	-
Sufi et al. (2011)	N	O	Cardioid graph centroid, extremas, area, and perimeter	-
Agrafioti et al. (2012)	N	T	Autocorrelation coefficients	LDA
Belgacem et al. (2012)	N	F	Average DWT coefficients	-
Lourenço et al. (2012a)	N	T	Segmented heartbeats	-
Singh and Singh (2012)	F	T	Interval, angle, and amplitude fiducial features	-
Belgacem et al. (2013)	N	F	Avg. beat Daubechies DWT	-
Coutinho et al. (2013)[2]	F	T	Fid. latency and amplitude from mean waveform subsampling	-
Coutinho et al. (2013)[2]	N	T	User-tuned Lloyd-Max quantised beats	-
Labati et al. (2013)	N	T	QRS segment set templates	-
Matos et al. (2013)	N	F	STFT spectrogram Gaussian models and Spectral zoom	Bin selection
Silva et al. (2013)	N	F	Mean and median ensemble beats	-
Wang et al. (2013)	N	T	Max-pooling representation elements	-
Ergin et al. (2014)	H	O	Fusion of QRS fid., time domain, wavelet transform, and PSD features	-
Iqbal et al. (2014)	N	O	QRS cardioid graph coordinates	-
Labati et al. (2014)	N	T	QRS segments	-
Lin et al. (2014)	N	T	Correlation dimension, Lyapunov exponents, RMSE	-
Lourenço et al. (2014)	N	T	Mean ensemble beats	-
Matos et al. (2014)	N	F	STFT window features	Kullback-Leibler
Pathoumvanh et al. (2014)	N	F	Continuous Wavelet Transform	FLDA
Zhou et al. (2014)	N	T	Signal segment between 3 consecutive R peaks	-
Brás and Pinho (2015)	N	O	Kolmogorov-based normalised relative compression	-
Choudhary and Manikandan (2015)	N	T	Averaged ensemble heartbeat	-
Dar et al. (2015a)	N	F	Haar transform	Greedy Best First Search
Dar et al. (2015b)	H	F	Haar transform and HRV	Greedy Best First Search
Jahiruzzaman and Hossain (2015)	N	F	CWT and Chaotic Encryption	-
Carreiras et al. (2016)	N	T	Segmented heartbeat	-
Chun (2016)	N	T	Guided filtering avg. beat	-
Hejazi et al. (2016)	N	T	Autocorrelation coefficients	KPCA
Louis et al. (2016)	N	T	1D Multi-res. LBP	-
Porée et al. (2016)	N	T	10 beat average ensemble	-
Rezgui and Lachiri (2016)	F	T	Amplitudes, areas, time intervals and slopes of fiducials	-
Waili et al. (2016a)	F	T	12 QRS fid. amplitudes	-

[2]Proposed two different approaches: one fiducial and one non-fiducial.

Kyoso et al. (2000) used the PQ and QT time intervals as features, while Kyoso and Uchiyama (2001a,b) used the QT time interval and the QRS duration. Israel et al. (2005) extracted fifteen features, including the P and T wave widths, and the time intervals between the R-peak and S, P, T, among others. The most recent fiducial approach was proposed by Waili et al. (2016a), who have extracted a total of 36 features of Q, R, and S amplitudes, of twelve different segmented QRS complexes.

Non-Fiducial Approaches

Non-fiducial approaches are those that use the entirety of the signal (or segments of it), holistically, to extract features related to the waveform morphology (Matta et al., 2011; Lourenço et al., 2011a; Matos et al., 2013). The first non-fiducial approach was published in 2005 (Saechia et al., 2005) but, one year earlier, Palaniappan and Krishnan (2004) has also extracted non-fiducial features for its hybrid approach. Non-fiducial approaches, not being constrained to the limited variety of fiducial points, offer a wider variety of possibilities, and have thus become the most popular.

Saechia et al. (2005), already mentioned, pioneered by applying a Fourier transform of whole PQRST segments and P, QRS, and T waves, separately. Fourier, cosine, or wavelet transforms were used by many other researchers, including Odinaka et al. (2010); Tawfik et al. (2010); Ye et al. (2010); Belgacem et al. (2012); and Matos et al. (2013, 2014). Other researchers have opted for the use of signal segments as templates, such as whole heartbeats Lourenço et al. (2012a); Labati et al. (2014); Carreiras et al. (2016), or segments between R peaks (Molina et al., 2007; Zhou et al., 2014), or averaged ensemble heartbeats (Fatemian et al., 2010; Lourenço et al., 2011a, 2014; Choudhary and Manikandan, 2015).

Other popular technique was that of Plataniotis et al. (2006), Agrafioti and Hatzinakos (2008), Agrafioti et al. (2012), and Hejazi et al. (2016), that used autocorrelation coefficients, of sliding-window-selected signal segments, as features. Two groups of researchers transformed the ECG signal segments into cardioid graphs, and used as features its centroid, area, perimeter, and extremas (Sufi et al., 2011), or all its x and y coordinates (Iqbal et al., 2014).

More singular approaches include Coutinho et al. (2011, 2013), that used user-tuned Lloyd-Max heartbeat segment quantisation to build the feature vector. Brás and Pinho (2015) opted for a Kolmogorov-based normalised relative compression of the heartbeats. Irvine et al. (2008) used eigenvectors from the data covariance matrix of segmented heartbeats, while Fang and Chan (2009) used phase space portraits of averaged heartbeats. Louis et al. (2016) computed one-dimensional multi-resolution Local Binary Patterns (1DMRLBP) on 1 second segmented heartbeat waveforms.

Hybrid Approaches

Hybrid approaches are those that use features from both origins, fiducials and holistic signal measurements. In the performed prior art analysis, these proved to be much more uncommon than the other two types of approaches.

Palaniappan and Krishnan (2004) combined the common fiducial features R amplitude, QR interval, RS interval, QRS width, and RR interval, with a non-fiducial QRS complex form factor, computed using the segment and its first and second derivatives. Ergin et al. (2014) proposed the fusion of QRS fiducials, with time domain, wavelet transform, and power spectral density features, computed across 2 second sliding windows. Dar et al. (2015b) opted for the extraction of a total of 46 features from Haar transform and heart-rate-variable RR intervals.

5.5.2 Features' Domain

Features can also be extracted from the electrocardiographic signal on various domains. The most common options are to perform the extraction either on the *time* domain or in the *frequency* domain (Hassan et al., 2016), but many researchers opt to perform other transformations on the original signal, in order to obtain different features.

Time Domain

Most of the surveyed approaches propose the extraction of features on the time domain. This group includes all fiducial-based approaches, such as those from Biel et al. (1999, 2001); Kyoso and Uchiyama (2001a,b); Shen et al. (2002); Guennoun et al. (2009); N and Jayaraman (2010); Shen et al. (2011); and Coutinho et al. (2013). Approaches that performed feature extraction from autocorrelation (Plataniotis et al., 2006; Agrafioti and Hatzinakos, 2008; Agrafioti et al., 2012; Hejazi et al., 2016), also worked on the time domain. Other methods include those of Coutinho et al. (2010, 2011, 2013), that proposed the quantisation of averaged heartbeat segments, and Louis et al. (2016), that performed a 1D multi-resolution Local Binary Pattern extraction in the time-domain.

Frequency Domain

For the extraction of features on the frequency domain, the signal must be first transformed into this domain. Following the results of the prior art survey performed, some used techniques to achieve this were the Fast Fourier Transform (FFT) (Saechia et al., 2005), the Short-Time Fourier Transform (STFT) (Matos et al., 2013, 2014), the Log-STFT (Odinaka et al., 2010), the Discrete Cosine Transform (DCT) (Tawfik et al., 2010), the Continuous Wavelet Transform (Pathoumvanh et al., 2014; Jahiruzzaman and Hossain, 2015), and the Discrete Wavelet Transform (DWT) (Belgacem et al., 2012), and more specifically the Daubechies Wavelets (Ye et al., 2010), and the Haar Transform (Dar et al., 2015a,b).

Other Transformations

This category includes those approaches whose feature extraction takes place on both the time domain and the frequency domain, and the approaches where the signal suffers other types of transformations before the features are extracted.

One of these was proposed by Fang and Chan (2009), that built a Phase-Space representation of a heartbeat segment, that served as feature vector for recognition. Li and Narayanan (2010) combined time domain and cepstral information by modeling heartbeats with Hermite polynomial expansion (HPE) and extracting cepstral features with cepstral mean subtraction and cepstral variance normalisation.

A transformation of the heartbeat into a pulse representation was proposed by Safie et al. (2011), building a Pulse Active Ratio (PAR) pulse by superimposing a triangular waveform with the segmented signal. Sufi et al. (2011) and Iqbal et al. (2014) transformed each ECG heartbeat into a cardioid curve, from which the features for recognition were then extracted.

5.5.3 Dimensionality Reduction

In the quest to entirely capture the wide variety of individual information stored by the electrocardiographic signal, the number of features extracted by biometric systems can easily become too high for a time-efficient and reliable recognition process Fratini et al. (2015). Thus, dimensionality reduction serves the purpose of transforming or selecting the extracted features, in order to reduce its number to a more computationally viable number, while keeping the maximum discriminant power to ensure the maintenance or improvement of the system's recognition performance (Wahabi et al., 2014).

Biel et al. (1999, 2001), the pioneers of ECG biometric recognition, applied dimensionality reduction using correlation matrices, to select 10 of 30 features from Lead I, of a grand total of 360 features extracted from all 12 leads. Israel et al. (2005) selected 12 of 15 extracted fiducial features using Wilkes' lambda stepwise correlation. Plataniotis et al. (2006) used the Discrete Cosine Transform (DCT) to reduce the features extracted from windowed autocorrelation.

Later, Agrafioti and Hatzinakos (2008) obtained better performance with dimensionality reduction using Linear Discriminant Analysis (LDA) than with DCT, also for autocorrelation features. LDA was also the choice of Boumbarov et al. (2009); Matta et al. (2011); Agrafioti et al. (2012). Li and Narayanan (2010) opted for an extension of LDA, the Heteroscedastic Linear Discriminant Analysis (HLDA), while Pathoumvanh et al. (2014) used the less general Fisher Linear Discriminant Analysis (FLDA).

Matos et al. (2013, 2014) used the symmetric Kullback-Leibler divergence for bin selection after computing Gaussian models of heartbeat segment STFT spectrogram, and for feature selection between 350 extracted STFT window features. Dar et al. (2015a,b) applied the Greedy Best First Search (GBFS) algorithm for selection of Haar transform features. Hejazi et al. (2016) used, besides LDA, Principal Component Analysis (PCA) and Kernel Principal Component Analysis (KPCA) for dimensionality reduction, concluding that KPCA rendered the best performance results.

5.6 Recognition Approaches

After the acquisition, preprocessing, preparation, and feature extraction, ECG-based biometric systems can perform the recognition of the individual being analysed, based on the stored templates of enrolled subjects. Be it an authentication or identification problem, generally recognition can be divided into two phases, the *matching* and the *decision*.

Decision refers to the methods used to obtain a decision on the acceptance or identity of the subject, based on the previously extracted features and the stored templates. In some cases, it is necessary or beneficial to compute separate comparison measures between these two sets of features, and this composes the matching phase. Along with these, thresholding (cf. subsection 3.5.1) is usually applied to accept or reject an individual based on its claimed identity's template, or for rejecting any identification, in case the similarity score isn't high enough to make a confident decision (Matos et al., 2013).

In the next subsections, the most common and most relevant matching and decision methods of prior art approaches found in the surveyed publications are presented, grouped by categories that can be defined according to the similarities between them. A summary on the essential aspects of recognition of the analysed recognition can be found in Table 5.5.

5.6.1 Matching Methods

Distance Metrics

A substantial fraction of the research works that apply matching methods have opted for the computation of distance metrics. These metrics perform a comparative analysis between the feature vectors, measuring the difference between them. The most popular distance metric was, by far, the Euclidean distance, used by Plataniotis et al. (2006); Lourenço et al. (2011a); Matta et al. (2011); Safie et al. (2011); Agrafioti et al. (2012); Singh and Singh (2012); Coutinho et al. (2013); Silva et al. (2013); Pathoumvanh et al. (2014); and Chun (2016).

Mahalanobis distance was also used various times, in three of the first publications on ECG-based biometrics (Kyoso et al., 2000; Kyoso and Uchiyama, 2001a,b), as well as the first continuous biometrics research work (Guennoun et al., 2009). Choudhary and Manikandan (2015) tested the performance of four different distance metrics, the Root Mean Square Error (RMSE), the Percent Residual Difference (PRD), the Wavelet Weighted-based Percent Residual Difference (WWPRD), and the Wavelet Distance (WDIST). WDIST and PRD were also used by Chan et al. (2008).

Correlation

The correlation coefficient, unlike distance metrics, serves to measure the statistical similarity between two signals or feature vectors Agrafioti and Hatzinakos (2008). Its use as a matching method was first proposed by Shen et al. (2002). They were followed by Agrafioti and Hatzinakos (2008); Chan et al. (2008); Fatemian and Hatzinakos (2009); Fatemian et al. (2010); Sasikala and

Wahidabanu (2010); and Shen et al. (2011). Choudhary and Manikandan (2015) opted for the Normalised Cross-Correlation (NCC), and Fang and Chan (2009), that used as feature vector a 3D average beat phase space portrait, used spatial correlation.

Labati et al. (2013, 2014) proposed the computation of similarity matrices, through cross-correlation, between the two templates, and produced a final similarity score by fusing matrix measures such as the mean, median, percentiles, and maximum.

Others

Besides distance metrics and similarity metrics based on correlation, other techniques were found in the surveyed publications. The log-likelihood ratio (LLR), that returns a similarity score between two feature vectors, that can be then used for identification or authentication, was used by Odinaka et al. (2010) and Matos et al. (2013, 2014).

Molina et al. (2007) was the first to propose the Dynamic Time Warping (DTW) paths as matching method. DTW has previously been used for speaker recognition using out of time scale or unsynchronised signals, not needing signal alignment. N and Jayaraman (2010) and Zhou et al. (2014) were two other research groups that used DTW.

Another matching method worthy of mention was used by Coutinho et al. (2010, 2011, 2013). The Ziv-Merhav cross parsing algorithm, originally used with symbol sequences for data compression, was modified to compare two quantised heartbeat segments and output two measures, similarity and relative entropy.

5.6.2 Decision Methods

Nearest Neighbour

Nearest neighbour classifiers, commonly k-Nearest-Neighbour (kNN), takes the feature vector being classified and those of the stored templates and, in the feature space, computes the distance between the former and each one of the others. The feature vector is then attributed the most verified class among the k closest template vectors. kNN is most commonly used for identification, with $k = 1$, meaning the feature vector is given the identity of the template vector that is closest to it (Wang et al., 2013). It can also be used in authentication, along with a thresholding phase, to verify a claimed identity.

Nearest neighbour classifiers were used on a large majority of the surveyed approaches, including Molina et al. (2007); Wübbeler et al. (2007); Chan et al. (2008); Ghofrani and Bostani (2010); N and Jayaraman (2010); Agrafioti et al. (2012); Matos et al. (2013); Wang et al. (2013); Labati et al. (2014); Brás and Pinho (2015); Dar et al. (2015a); Carreiras et al. (2016); and Porée et al. (2016). These classifiers have the main advantage of, unlike most other techniques, not requiring previous training with all subjects signals (Chun, 2016).

Neural Networks

Neural Network classifiers mimic the function of their biology homonyms, that consist of webs of interconnected neurons that receive inputs, analyse and modify them, and pass them along until it reaches a target organ or tissue (Scanlon and Sanders, 2007). The classifiers are also composed by neurons (or nodes), arranged in a varying number of layers, and connected between them. The first layer receives the inputs (feature vectors), the nodes have activation functions, and their connections are weighted to guide the final classification, output by the last node layer (Iqbal et al., 2014; Fratini et al., 2015).

Neural Networks are especially useful in non-linear classification problems (Fratini et al., 2015). Various types of these classifiers were used in the surveyed approaches, namely the Multilayer Perceptron (MLP) (Palaniappan and Krishnan, 2004; Ghofrani and Bostani, 2010; Iqbal et al., 2014; Waili et al., 2016b), the Decision-based Neural Network (DBNN) (Shen et al., 2002), the Simplified Fuzzy ARTMAP (SFA) (Palaniappan and Krishnan, 2004), the Radial Basis Function Neural Network (RBFNN) (Boumbarov et al., 2009), and the Probabilistic Neural Network (PNN) (Ghofrani and Bostani, 2010). Most of these vary on the node activation functions used, and have been trained using backpropagation.

Support Vector Machines

Support Vector Machines (SVM) are classifiers that, based on a given set of training data and a defined kernel function, compute an optimal classification boundary between classes (Rezgui and Lachiri, 2016).

SVM were used in ECG-based recognition by Li and Narayanan (2010); Ye et al. (2010); Lourenço et al. (2012a, 2014); Silva et al. (2013); Lin et al. (2014); Hejazi et al. (2016); and Rezgui and Lachiri (2016). In what concerns the kernel functions, Gaussian radial basis function and linear and non-linear polynomial have been explored in the analysed approaches.

Discriminant and Component Analysis

Discriminant or Component Analysis methods are more commonly used for dimensionality reduction, as they allow the transformation of the feature space in order to minimise intraclass variance and increase discrimination between classes (Biel et al., 2001). However, this transformation of the feature space can also be used for the classification of new samples, for biometric recognition.

Among the analysed research works, some methods have applied different algorithms of discriminant or component analysis for identification or authentication. Kyoso et al. (2000) and Kyoso and Uchiyama (2001a,b) have used the Linear Discriminant Analysis (LDA). Biel et al. (1999, 2001) opted for SIMCA, a commercial PCA-based data analysis algorithm. Irvine et al. (2008) and Chun (2016) used Principal Component Analysis (PCA).

Table 5.5: Summary of the matching and decision methods used on the surveyed approaches (ordered ascendingly by year, and alphabetically by the last name of the first author). The dashes denote the respective method did not apply a matching phase, and the features extracted were directly used in the decision phase.

Researchers	**Matching**	**Decision**
Biel et al. (1999, 2001)	PCA	SIMCA
Kyoso et al. (2000)	Mahalanobis distance	Linear Discriminant Analysis
Kyoso and Uchiyama (2001a,b)	Mahalanobis distance	Linear Discriminant Analysis
Shen et al. (2002)	Correlation coefficient	Decision-based Neural Network
Palaniappan and Krishnan (2004)	-	MLP and SFA
Israel et al. (2005)	LDA	Contingency Matrix Majority Voting
Saechia et al. (2005)	-	Neural Networks
Plataniotis et al. (2006)	Normalised Euclidean distance	Normalised Gaussian Log-Likelihood
Zhang and Wei (2006)	-	Bayes minimum-error-rate class.
Molina et al. (2007)	Dynamic Time Warping path	Nearest Neighbour
Wübbeler et al. (2007)	Temporal derivatives distance	Nearest Neighbour
Agrafioti and Hatzinakos (2008)	Correlation coefficient	Nearest Neighbour
Chan et al. (2008)	Percent Residual Difference; Correlation coefficient; and Wavelet distance	Nearest Neighbour
Irvine et al. (2008)	-	PCA
Boumbarov et al. (2009)	-	Radial Basis Function N. N.
Fang and Chan (2009)	Spatial correlation; and Mutual Nearest Point Distance	Nearest Neighbour
Fatemian and Hatzinakos (2009)	Correlation coefficient	Nearest Neighbour
Guennoun et al. (2009)	Mahalanobis distance	Thresholding and Voting
Coutinho et al. (2010)	Ziv-Merhav relative entropy	Nearest Neighbour
Fatemian et al. (2010)	Correlation coefficient	Nearest Neighbour
Ghofrani and Bostani (2010)	-	Nearest Neighbour; MLP; PNN
Li and Narayanan (2010)	-	Fusion of SVM and GMM-UBM Supervector
N and Jayaraman (2010)	FLDA and DTW	Nearest Neighbour
Odinaka et al. (2010)	Log-likelihood ratio	Nearest Neighbour
Sasikala and Wahidabanu (2010)	Correlation coefficient	Maximisation and mean dist.
Tawfik et al. (2010)	-	Neural Network
Ye et al. (2010)	-	Gaussian RBF SVM
Coutinho et al. (2011)	Ziv-Merhav cross parsing similarity	Nearest Neighbour
Lourenço et al. (2011a)	Euclidean distance	Nearest Neighbour
Matta et al. (2011)	Euclidean distance	Nearest Neighbour
Safie et al. (2011)	Euclidean distance	Nearest Neighbour
Shen et al. (2011)	Correlation coefficient and LDA	Nearest Neighbour

Researchers	Matching	Decision
Sufi et al. (2011)	Straight line and percentage dist.	Nearest Neighbour
Agrafioti et al. (2012)	Euclidean distance	Nearest Neighbour
Belgacem et al. (2012)	-	Random Forest Ensemble
Lourenço et al. (2012a)	-	kNN and SVM
Singh and Singh (2012)	Euclidean distance	Nearest Neighbour
Belgacem et al. (2013)	-	Random Forest Ensemble
Coutinho et al. (2013)[3]	Euclidean distance	Nearest Neighbour
Coutinho et al. (2013)[3]	Ziv-Merhav cross parsing similarity	Nearest Neighbour
Labati et al. (2013)	Cross-correlation similarity matrix and score fusion	Nearest Neighbour
Matos et al. (2013)	Log-likelihood ratio	Nearest Neighbour
Silva et al. (2013)	Euclidean and cosine distances	kNN and SVM
Wang et al. (2013)	-	Nearest Neighbour
Ergin et al. (2014)	-	C4.5 Decision Tree and Bayesian Network
Iqbal et al. (2014)	-	Multilayer Perceptron
Labati et al. (2014)	Cross-correlation similarity matrix and score fusion	Nearest Neighbour
Lin et al. (2014)	-	Support Vector Machine
Lourenço et al. (2014)	-	Support Vector Machine
Matos et al. (2014)	Log-likelihood ratio	Nearest Neighbour
Pathoumvanh et al. (2014)	Euclidean distance	Nearest Neighbour
Zhou et al. (2014)	Dynamic Time Warping path	Nearest Neighbour
Brás and Pinho (2015)	-	Nearest Neighbour
Choudhary and Manikandan (2015)	RMSE; PRD; NCC; WWPRD; and WDIST	Nearest Neighbour
Dar et al. (2015a)	-	Nearest Neighbour
Dar et al. (2015b)	-	Random Forest Ensemble
Jahiruzzaman and Hossain (2015)	-	Unique CE sequences identification
Carreiras et al. (2016)	-	Nearest Neighbour
Chun (2016)[3]	Dynamic Time Warping	Nearest Neighbour
Chun (2016)[3]	Euclidean distance	Nearest Neighbour
Chun (2016)[3]	-	PCA
Hejazi et al. (2016)	-	Support Vector Machine
Louis et al. (2016)	-	Bagging
Porée et al. (2016)[3]	-	Discrimination coefficient
Porée et al. (2016)[3]	-	Nearest Neighbour
Rezgui and Lachiri (2016)	-	Support Vector Machine
Waili et al. (2016a)	-	Multilayer Perceptron

[3]Proposed more than one recognition approach.

Others

Besides the already mentioned decision methods, some researchers have opted for unusual alternatives. Zhang and Wei (2006) built a classification method based on Bayes theory error minimisation, while Ergin et al. (2014) used Bayesian Networks along with C4.5 Decision Trees.

A few approaches used ensemble decision methods, like Random Forests, proposed by Belgacem et al. (2012, 2013) and Dar et al. (2015b), and Bagging, used by Louis et al. (2016). Jahiruzzaman and Hossain (2015), after computing Chaotic Encryption (CE) features, based its decision approach on the identification of unique CE sequences for each subject.

5.7 Conventional vs. Continuous Biometrics

The vast majority of the surveyed publications were focused on conventional biometrics, but a few research groups have already begun to address the specific problem of continuous or real-time biometrics. The first step is to acknowledge the need for signals widely spread across time, due to the variable nature of the ECG.

The first continuous approach was proposed by Guennoun et al. (2009), that frequently generated a comparison score to assess if the authenticated user was, still, the one using the system. Louis et al. (2016) devised a similar approach, by allowing for three different results in each cycle: blocking out any impostors (reject), maintaining the system open for the authenticated user (accept), or, in doubt, simply delay the decision for the next cycle (continue). Matta et al. (2011) was the first to focus on continuous identification, performing it every five seconds. Matos et al. (2014) performed a real-time recognition process on 100ms sliding windows over a continuously received ECG signal.

Other researchers have devised methods for *template update*, a reliable way to ensure the system maintains the capability of recognising enrolled subjects, in order to keep up with their constant variability, by frequently updating the respective templates stored in the database. These researchers included Coutinho et al. (2011, 2013) and Agrafioti et al. (2012).

5.8 Performance Achievements

As explained on subsection 3.5.3, the population that composes the database used for the evaluation of a biometric system can have a significant effect on the performance results obtained. As visible on Table 5.1, there is a large variety of data used by the surveyed prior art research works. This means that, as the testing environment of the methods is not the same, trying to compare them based on the metrics disclosed by the researchers is futile.

This is the core of the need for public signal collections. As the populations that compose them is the same through all research works, and the databases are freely available to all researchers, the testing environment is the same and makes benchmarking possible. Although some researchers

Table 5.6: Results of approaches evaluated with the PTB database (ordered by total number of subjects - NS).

Researchers	NS	Type	Results	
Safie et al. (2011)	112	Authentication	EER	19.2%
Wang et al. (2013)	100	Identification	IDR	99.5%
Wübbeler et al. (2007)	74	Both	IDR	98.1%
			EER	2.8%
Brás and Pinho (2015)	52	Identification	IDR	99.91%
Coutinho et al. (2013)	51	Both	IDR	99.9%
			EER	0.01%
Plataniotis et al. (2006)	14	Both	IDR	100%
			FAR	0.02%
Waili et al. (2016a)	14	Identification	IDR	96%
Ghofrani and Bostani (2010)	12	Identification	IDR	98.6%

Table 5.7: Results of approaches evaluated with the MIT-BIH Normal Sinus Rhythm database (ordered by total number of subjects - NS).

Researchers	NS	Type	Results	
Shen et al. (2002)	20	Identification	IDR	100%
Li and Narayanan (2010)	18	Both	IDR	98.3%
			EER	0.5%
Ye et al. (2010)	18	Identification	IDR	99.3%
			FPIR	26.9%
Ergin et al. (2014)	18	Identification	F-score	0.97%
Dar et al. (2015a)	18	Both	IDR	99.4%
Dar et al. (2015b)	18	Both	IDR	100%
			EER	0%
N and Jayaraman (2010)	15	Both	IDR	96%
Palaniappan and Krishnan (2004)	10	Identification	IDR	96.2%

Table 5.8: Results of approaches evaluated with the MIT-BIH Arrhythmia database (ordered by total number of subjects - NS).

Researchers	NS	Type	Results	
Ye et al. (2010)	47	Identification	IDR	99.6%
			FPIR	12.3%
Dar et al. (2015a)	47	Both	IDR	93.1%
Dar et al. (2015b)	47	Both	IDR	95.85%
			FAR	4.1%
			FRR	0.1%
Jahiruzzaman and Hossain (2015)	11	Identification	IDR	96.9%
Sasikala and Wahidabanu (2010)	10	Identification	IDR	62.7%
Sufi et al. (2010)	-	Both	MIDR	1%
			EER	0.5%

chose to only use subsets of subjects in public databases, in Tables 5.6, 5.7, and 5.8, the performance results of approaches using the PTB, the MIT-BIH Normal Sinus Rhythm, and the MIT-BIH Arrhythmia databases, respectively, are presented. The remaining research works, unable to undergo accurate benchmarking, have their results presented in Table B.1, in the appendix.

5.9 Summary and Conclusions

Through the analysis of sixty-five collected prior art original research publications, it was possible to extract some conclusions and guidelines for the work to be conducted in the dissertation. In what concerns the acquisition settings, prior art research have already covered a great amount of research possibilities, on almost everything that can affect the performance of a biometric system. However, none has gone as far as to use signals as noisy as those acquired on the steering wheel whilst driving.

It is important, though, to keep in mind the need for performance assessment on both public, benchmarking-allowing databases, and private signal collections that fit, in the best possible way, the system's expected application settings. Variety in subjects health, positioning, activity, acquisition sessions over time, and more acceptable number and placement of electrodes, should all guide the collection of signals for both private and publicly available databases for biometric performance evaluation.

As for the preprocessing methods, it would be desirable to conduct a thorough evaluation of the several denoising techniques proposed, using ECG signals acquired under various conditions, especially to assess their performance in driving signals. For signal preparation techniques, amplitude normalisation and time normalisation, besides being underexplored, also lack a complete evaluation of their true impact in a system's performance. It would also be interesting to devise new approaches that lack the need for the often unreliable reference point detections and signal segmentations.

When it comes to features, both fiducial and non-fiducial approaches have been widely explored, although their applicability to driving off-the-person signals is still unknown. On the domains, however, frequency domain and other transformations of the original time domain signals still present some room for future research. The same applies for dimensionality reduction, where a comparison of methods, using the same original set of features, would be desirable.

Finally, matching and decision have also been widely explored. However, string matching methods similar to the Ziv-Merhav cross parsing and the Kolmogorov-based compression are still in their infancy and, while presenting great results, leave room for improvement. Other possibilities, not yet explored, would include the use of deep learning neural networks, Markov models, or ensemble algorithms other than Bagging and Random Forests.

The highest focus, however, should be given to the integration of these in continuous systems, still in an early stage of research maturity. Especially, the aim should be to assess their capabilities when using the highly noisy signals acquired on the steering wheel whilst driving, and to ensure the systems could offer accurate and reliable decisions, in a timely fashion.

Chapter 6

Experimental Settings

6.1 Introduction

Just as all ECG-based biometric recognition methods and systems, the work performed for this dissertation used a combination of specialised hardware, acquired data, and software, for the effective recognition of individuals based on signals acquired on the steering wheel.

The data used for the evaluation of the approaches explored for this dissertation, and the hardware and software used for their development are, in this chapter, presented in detail. The settings for the evaluation of the performance of the methods are also here described.

6.2 Software and Hardware

Along the following chapters, the run times of some operations are presented. For reference, the computer used to obtain such measurements was an Asus X555LJ, with 8 Gb of RAM, and a processor Intel Core i7-5500U, with two cores, with a base frequency of 2.40 GHz. The run times are just meant to allow the comparison of methods, as in embedded settings (with much less processing power) these will, likely, be much higher.

However, the methods that prove to be faster in these better settings, should, reasonably, be also faster in embedded systems. The results presented should, thus, serve as a guide for future embedded implementations.

As for the software, Python was the programming language chosen for the scientific work performed during the dissertation, running on Microsoft Windows 10. The Spyder IDE was installed for this purpose, through the Anaconda data science platform (Python 2.7 version).

Several Python libraries related to machine learning and data science were installed, including NumPy, Matplotlib, scikit-learn, scikit-image, pandas, Seaborn, Keras, Theano, and Jupyter. One last library, BioSPPy[1], developed and maintained by Instituto de Telecomunicações, was installed, and its functionalities focused on physiological signal processing and analysis were used during the dissertation.

[1] BioSPPy. Available on: `https://pypi.python.org/pypi/biosppy/0.2.0` (last visited on 20/05/2017).

Intervals		15 s	20 s	15 s	25 s	20 s	25 s	15 s	15 s	20 s	25 s	20 s
Subjects	S0	S1	S3	S2	S3	S4	S0	S4	S2	S5	S1	S5
Duration	173 s	327 s	284 s	302 s	315 s	297 s	298 s	169 s	287 s	311 s	297 s	40 s
Ensembles	N=59	N=250	N=128	N=88	N=138	N=100	N=196	N=33	N=162	N=259	N=184	N=31
Day	1	2	3	3	3	5	3	5	3	4	2	4

Figure 6.1: Structure of the driving dataset built from the continuous recordings by CardioID (the duration includes occasional periods without contact with the steering wheel, interval refers to the time interval used to simulate quick changes of drivers, ensembles refers to the number of ensemble heartbeats it was possible to extract, and the day of recording is also specified).

6.3 Signal Databases and Datasets

From all available ECG collections (cf. Chapter 4 for more detailed information), the work performed in this dissertation used signals from the University of Toronto ECG database (UofTDB, whose access was gracefully granted by the responsible researchers), and driver ECG signals acquired at the steering wheel of a coach bus by CardioID Technologies (a week-long recording from the SteeringWheel v1 collection, with a total of six drivers).

The continuously acquired signals from the CardioID SteeringWheel v1 collection were chosen as they provided the most realistic settings for performance evaluation, considering the main goal of this dissertation to recognise individuals continuously, on the steering wheel, whilst driving. As visible in Fig. 6.2, the signals from this database also posed the greatest challenges, due to the greater influence of low frequency noise, and especially the frequent sensor saturation periods that cause the complete loss of ECG information.

The UofTDB collection was selected because it provided a greater amount of subjects, and its acquisition settings allowed benchmarking against prior art approaches. Although its signals are less affected by noise than those recorded in the steering wheel, the fact they were recorded on the fingers, during several activities and postures, makes it more challenging than on-the-person ECG collections.

From the CardioID SteeringWheel v1 six-day recording, continuous excerpts of varying durations were extracted (two for each driver, preferentially from different days). These composed a small, approximately one-hour-long recording (cf. Fig. 6.1), designed to simulate a quick succession of drivers, with short intervals between them.

The contact with the steering wheel (and the sensors) is not guaranteed nor expected at all times while driving. Hence, the lead-on detection data was used to reject any samples acquired during periods without contact. Among the remaining samples, only those that were part of sequences of continuous contact longer than 5 seconds were considered.

Each continuous contact sequence was broken into five-second-long segments (with time steps of 1 s, cf. Fig. 6.3) to allow for posterior processes of template extraction, outlier detection, and heartbeat ensemble, while ensuring a quick first decision after the contact with the sensor starts, and very frequent decision renewals. These segments made up the CardioID real driving dataset.

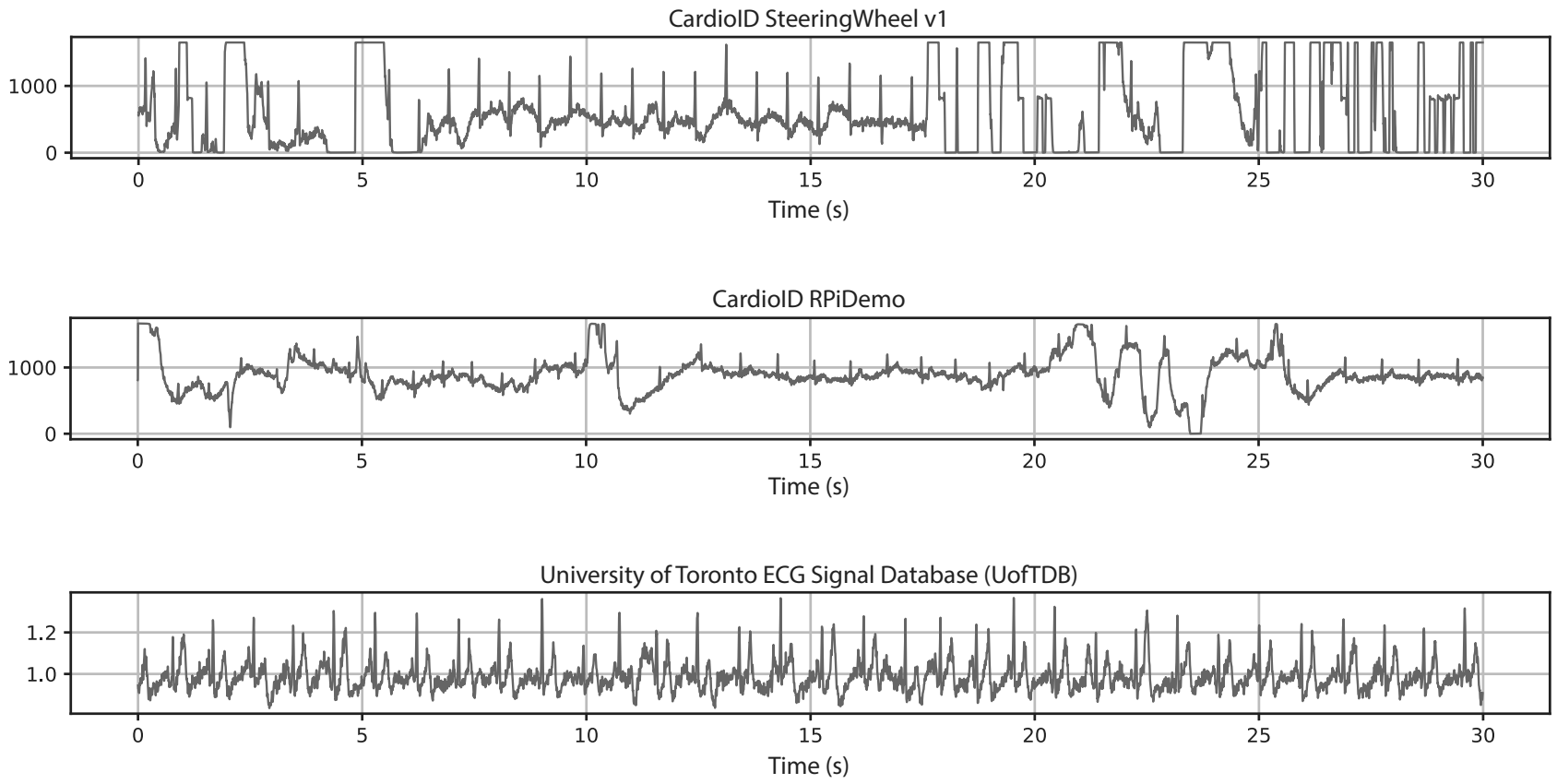

Figure 6.2: Example excerpts from the ECG databases used in this dissertation.

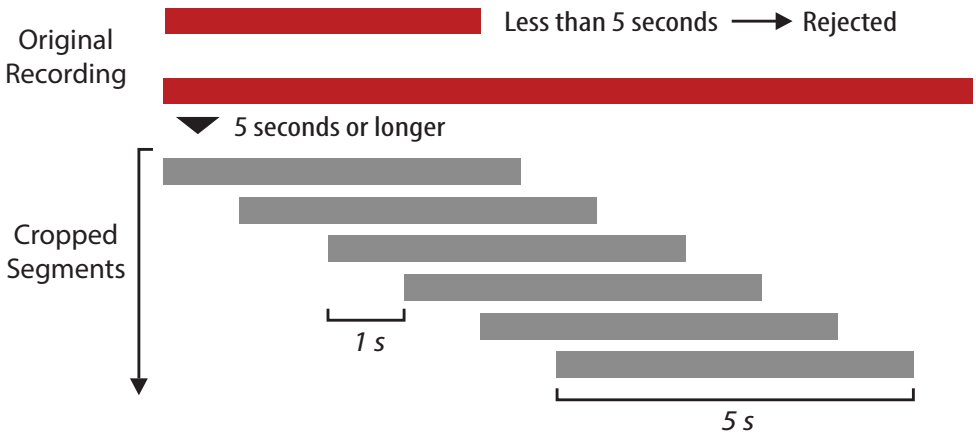

Figure 6.3: Illustration of the process of signal recording cropping into small, overlapping segments.

The same operations were used to build a dataset from the RPiDemo recordings, that was mainly used as separate data to train UBM models of Naïve Bayes and Gaussian Mixtures, due to the similarity with the driving signals, despite the evidently lower noise influence (cf. Fig. 6.2).

As for the signals of the University of Toronto ECG database, while they were acquired at different points in time and are, thus, not continuous, the available recordings were also broken into five-second-long segments with 1s steps. These segments composed the UofTDB dataset used throughout the dissertation. From the 1019 subjects of the collection, the dataset is composed by a continuous set of 100 subjects (725 to 824), chosen due to the more frequent acquisitions in several weeks and conditions (cf. Table 6.1).

6.4 Performance Evaluation

The performance evaluation of the recognition methods is presented in section 10.5, and was performed in both operation modes possible for a biometric system: identification and authentication.

For authentication, performance was evaluated based on Receiver Operating Characteristic (ROC) curves, traced through the computation of false acceptance (FAR) and false rejection rates

Table 6.1: Status of the hundred subjects selected from the UofTDB collection (The subjects selected were 725 through 824. NS - number of subjects).

Recordings	**NS**	**Subjects**
10 (on 6 weeks and 5 conditions)	8	725, 761, 781, 788, 793, 800, 818, and 821
9 (on 5 weeks and 5 conditions)	2	792, and 816
8 (on 5 weeks and 5 conditions)	4	736, 748, 792, and 804
6 (on 3 weeks and 4 conditions)	1	805
2 (on 2 weeks and 2 conditions)	1	802
only 1	84	

(FRR) for several thresholds (cf. Fig. 3.7). The metrics presented in this document are the equal error rates (EER, the operation point where $FAR = FRR$) and the area under the ROC curve (AUC). In the case of identification, the performance of the approaches was based on the most commonly used metric, the identification rate (IDR).

Although these are the most common performance metrics among related prior art works, this dissertation also focused in a step-by-step evaluation of the approaches on the modules prior to decision. This was motivated by the fact that each of these steps influences the final performance of the recognition methods, and their individual optimisation is paramount in such challenging settings.

Preprocessing methods were evaluated through the measurement of the root mean square error (RMSE) between simulated signals and their denoised versions (cf. section 8.3). R-peak detection methods were evaluated by counting the peaks correctly detected, as well as possible false peaks (cf. subsection 8.3.1).

Outlier detection performance was assessed through template set standard deviation evolution, and number of templates rejected (cf. subsection 8.3.2). At last, feature sets were evaluated through simple identification procedures, through ten-fold cross-validation using SVM and kNN classifiers, of which the identification rates were measured (cf. section 9.4).

6.5 Summary and Conclusions

In conclusion, the experimental settings defined and described in this chapter should be enough to ensure a complete and adequate development and evaluation of the algorithms in this dissertation. The data, from both the CardioID SteeringWheel v1 collection and the University of Toronto, enable the development and performance assessment of the algorithms bearing in mind wide applications in off-the-person systems, with diverse signal characteristics and population sizes.

The performance evaluation structure, with both authentication and identification modes, and the metrics selected, allow for a complete assessment of the capabilities of the explored methods, both with driving and standard off-the-person ECG signals. The run time measurements of time-sensitive processes is also paramount to ensure the best performance in continuous settings.

Chapter 7

Signal Preprocessing

7.1 Problem Statement

One of the main concerns of every study related to the use of the electrocardiogram or other biosignals is the noise and the way it may impact the reliability of the signal for the desired goal. In the specific case of biometrics, noise can present serious challenges on the accurate recognition of individuals, as it distorts the signals and introduces undesirable variability.

The influence of such noise, presented in section 2.4, is usually easy to cancel by the use of simple approaches (like bandpass filters). So, the great majority of the prior art approaches, analysed and described in Chapter 5, did not give much emphasis to the preprocessing stage of their proposed algorithms.

However, as systems evolve towards the use of ECG signals acquired on off-the-person settings, noise becomes much more dominant, and the need for more robust denoising methods quickly increases. This need becomes even greater when considering signals acquired on the steering wheel whilst driving, highly corrupted by noise, as those used on this dissertation.

This chapter aims to address the important concern of the denoising of the ECG signals for biometric recognition, through the evaluation of prior art and preprocessing techniques to simulated and real signals, contaminated by standard off-the-person and driving noise sources. Own proposals were also explored: Savitzky-Golay followed by a Moving Average Filter or a Highpass Filter, and DWT Soft Thresholding combined with a Moving Average Filter.

7.2 Selected and Proposed Methods

7.2.1 Filters

Filters are among the most simple and commonly used methods for denoising ECG and other biosignals. Oppenheim et al. (1999) described a filter as any system that modifies certain frequency components relative to others. In this broad set of denoising techniques, this dissertation focuses on frequency-selective filters: bandpass, lowpass, and highpass filters.

These three filters are characterised by allowing or rejecting certain frequency ranges of the signals to which they are applied. Highpass filters ideally reject all frequency components below a defined cut-off frequency value, while lowpass filters reject all frequency components lower than that threshold. Bandpass filters fuse the two and reject any frequency components but those inside a range defined by two cut-off frequency limits.

Although the behaviour of filters is not of total rejection of frequency components, but more of attenuation, if the cut-off frequencies are correctly defined, the denoising process should present acceptable results, allowing for the cancellation of both high and low frequency noise.

Among the denoising methods selected for this dissertation, the ones based on filters were the bandpass filters with frequency bands 1-40 Hz (Agrafioti and Hatzinakos, 2008; Agrafioti et al., 2012; Belgacem et al., 2012, 2013; Louis et al., 2016), 2-40 Hz (Israel et al., 2005; Safie et al., 2011; Rezgui and Lachiri, 2016), 2-30 Hz (Coutinho et al., 2010, 2011, 2013), and 1-30 Hz (Lourenço et al., 2012a); as well as the combination of a highpass filter of 1 Hz and a notch filter at the signal's PLI frequency (adapted from Labati et al. (2013, 2014)).

Considering the specific noises on driving settings, the last method (HPF + NF) is no longer suitable for correct filtering, as the notch filter is prepared to only attenuate the signal components at a single frequency (PLI), and not at the wide range of frequencies of EMG interference. Thus, this specific method was discarded for driving settings.

7.2.2 Moving Average and Median Filters

Moving average and median filters are two other very common techniques for the denoising of signals. These are especially fit for baseline wander correction, and thus are almost always used in combination with another technique, that is able to cancel high frequency noise.

These two filters operate based on a window, of predetermined width, that will be centered on each sample of the signal to be filtered. A new version of the signal, at that index, will take the average (or median) value of the original signal samples inside that window.

Using large window widths (around one second long) allows for the capture of the baseline of the signal on the filter's output. The posterior subtraction of the filter's output from the original signal allows for the removal of the baseline wander.

7.2.3 Savitzky-Golay

The Savitzky-Golay (SG) method was proposed by Savitzky and Golay (1964) as a smoothing technique for continuously acquired data from chemistry experimentations. The formulation of this method followed the detection of flaws on the common moving average smoothing technique (described above), of which the most undesirable was the degradation of the intensity of peaks in the signal.

The authors were inspired by the common perception of data smoothing: tracing a smooth line among the data points so that it fits them with the least overall error possible. Hence, the Savitzky-Golay method is based on least-squares error minimisation that, according to the researchers,

fulfills the purpose of noise reduction without distorting the shape and height of peaks. As stated by Schafer (2011), this advantage is beneficial not only in analytic chemistry but also in signal processing and, specifically, ECG processing.

Considering a signal $x[n]$, the SG method considers a neighbourhood of M samples on each side of a center sample (total width of $2M+1$), and fits a polynomial line of degree $n < 2M+1$, so that the mean square error is minimised. The center point of the neighbourhood will take the value resulting from the evaluation of the fitted polynomial at the respective abscissa, and the same process is then applied to all other data points of the signal.

One other advantage of the Savitzky-Golay smoothing technique is the nonexistence of a symmetry requirement for the neighbourhood relative to the center point, meaning that the procedure can be applied to the start and end of finite sequences, such as a signal recording (Schafer, 2011).

The Savitzky-Golay method was explored in this dissertation for the denoising of ECG signals, using sixth order polynomial lines (as used by Dar et al. (2015a,b)), and a total neighbourhood width of ten percent the sampling frequency of the signal. Following the results obtained and discussed in the next section, that prove this method is unable to filter low frequency noise, two adaptations are proposed in this dissertation: Savitzky-Golay combined with a Moving Average Filter with 1s window, and SG combined with a highpass filter with cut-off frequency of 1Hz.

7.2.4 DWT-based Denoising

The Discrete Wavelet Transform (DWT) is, like the Discrete Fourier Transform (DFT), a filter that slides over a signal and decomposes it into frequency components. As described by Devasahayam (2012), the DWT differs from the DFT in its fundamental assumption that high frequency components can be resolved with smaller time windows (as they complete their cycles in shorter periods) than those needed to resolve the signal components of lower frequencies. Thus, the DWT filters' bandwidth is higher for higher frequencies, while for the DFT it is kept the same.

Considering a bank of K filters, the Discrete Wavelet Transform can be defined by:

$$W[k,t] = x[t] * \psi\left[\frac{t}{a^k}\right] \tag{7.1}$$

In this equation, $x[t]$ is the signal being filtered and $k = 1,...,K$ is the number of the filter, while ψ refers to the filter function, that is scaled by a factor a^k, where generally $a = 2$, so that the frequencies and the filter bandwidth increase geometrically. While DFT uses sinusoids and so has a single filter function, DWT has several wavelets that can be used to decompose the signal.

The Discrete Wavelet Transform is commonly used in the form of hierarchical filter banks in order to decompose the signal in several scales. Following equation 7.1, the signal is sequentially passed through the filters $k = 1,...,N$ for N levels (cf. Fig. 7.1). The first level will decompose the signal in two scales, the lowest from 0 to $f_n/2$, and the highest from $f_n/2$ to f_n, where f_n is the Nyquist frequency of the signal. The highpass filter outputs the scale coefficients cD (or detail coefficients), while the lowpass output will move to the second level, where will, itself, be decomposed in two scales (0 to $f_n/4$, and $f_n/4$ to $f_n/2$), and so on for each of the N levels. The

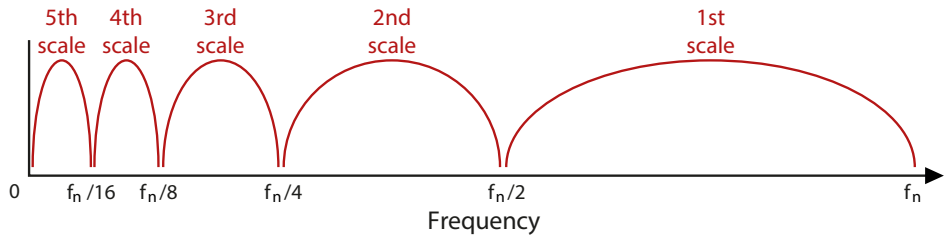

Figure 7.1: Illustration of the DWT scales for four levels (based on Devasahayam (2012)). The relationship between the frequencies and the bandwidth of the scales is worthy of notice.

last level, along with the detail coefficients of its scale, outputs the approximation coefficients cA, that correspond to the lowest of all scales.

Ergo, each layer will effectively be composed of a lowpass and a highpass filter and, as visible on Fig. 7.2, before stepping into the next layer, the scale output by the lowpass filter suffers downsampling, following the reasoning that while finer details require higher sampling frequencies, the components with lower frequencies do not require them.

After the decomposition of the signal, following the prior art approaches, the denoising of the filter can be attained through two different techniques. As Fatemian and Hatzinakos (2009) proposed, it is possible to reconstruct the signal at a certain scale, deemed to include most of the signal's information and the least noise possible. For the signals used by Fatemian and Hatzinakos (2009) and Fatemian et al. (2010), sampled at 250 Hz, this would be the third scale (capturing frequencies on the range from 15.6 to 31.3Hz).

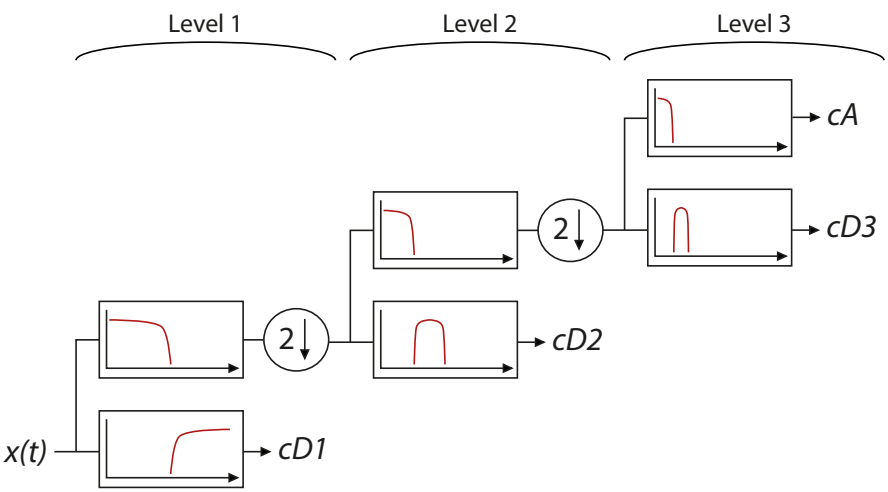

Figure 7.2: Scheme of DWT deconstruction for three levels (based on Devasahayam (2012)). In each level, a highpass and a lowpass are depicted, along with a downsampling after the following level. $cD1$, $cD2$, and $cD3$ refer to the detail coefficients of scales 1 to 3, respectively, and cA refers to the approximation coefficients output by the last level.

It is also possible to resort to signal reconstruction using all scales, instead performing a soft threshold of the various level detail coefficients, based on the relationship between the absolute value of the coefficients and the root mean square value of all coefficients of the respective level (as used by Boumbarov et al. (2009); Sasikala and Wahidabanu (2010); and Hejazi et al. (2016)).

Both these techniques were explored and evaluated. However, based on the results shown on the following section, and some of the prior art publications, these are not always suitable for the cleanse of low frequency noise. Hence, the appropriate scale reconstruction was also combined with a moving average filter (as used by Fatemian and Hatzinakos (2009)), and the soft threshold denoising was combined with an highpass filter of 1Hz, and with a moving median filter (based on the approaches of Boumbarov et al. (2009) and Sasikala and Wahidabanu (2010), respectively). Soft threshold was also, by our own proposal, combined with a moving average filter. Both MAF used here windows of 1s width, and MMF used windows half-a-second wide.

7.2.5 DCT-based Filtering

The Discrete Cosine Transform (DCT) was proposed by Ahmed et al. (1974) as a tool for pattern recognition dimensionality reduction, and for Wiener filtering of data sequences, reporting near-optimal performance. For the purpose of dimensionality reduction, DCT serves as an orthogonal transform that maps the original feature space onto one with fewer dimensions, reducing the computational load upon classification or regression processes.

Reduced computational load is the main advantage of DCT for Wiener filtering. As stated by Ahmed et al. (1974), discrete Wiener filters are commonly represented by a $M \times M$ matrix G, (where M is the length of the data sequence X), and the estimate $\hat{X}$ of the data sequence can be computed through $\hat{X} = GZ$, with $Z = X + N$ (where N is the noise vector). This means that a filtering process requires approximately $2M^2$ computations. With DCT, however, being an orthogonal transform means that many of the elements of G will be approximately equal to zero and can be annulled, thus reducing the number of required computations.

The vector G_x of Discrete Cosine Transform coefficients of a signal X of total length M can be computed through 7.2 and 7.3.

$$G_x[0] = \frac{\sqrt{2}}{M} \sum_{m=0}^{M-1} X[m] \tag{7.2}$$

$$G_x[k] = \frac{2}{M} \sum_{m=0}^{M-1} X[m] \cos \frac{(2m+1)k\pi}{2M}, k = 1, 2, ..., (M-1) \tag{7.3}$$

Each of the M coefficients of G_x corresponds to a frequency bin from 0 Hz at $G_x(0)$ to the signal's Nyquist frequency at $G_x(M-1)$. Based on this, Choudhary and Manikandan (2015) used G_x to denoise ECG signals corrupted with baseline wander and powerline interference, by setting

to zero the DCT coefficients corresponding to 0.8 Hz (BW) and to the range between 48-52 Hz (PLI of 50 Hz). The denoised signal is obtained through the inverse DCT:

$$X[m] = \frac{1}{\sqrt{2}} G_x[0] + \sum_{k=0}^{M-1} G_x[k] \cos \frac{(2m+1)k\pi}{2M}, m = 1, 2, ..., (M-1) \tag{7.4}$$

Unlike the work of Choudhary and Manikandan (2015), that tested their approach using ECG signals from standard collections using on-the-person acquisition settings, this dissertation focuses on off-the-person acquisitions, more heavily corrupted by noise. Hence, the DCT-based filtering approach was adapted, and for standard settings all coefficients corresponding to the frequency ranges $[0, 1]$ Hz and $[PLI - 5, PLI + 5]$ Hz were set to zero. For driving settings, the last range was replaced by the range of frequencies between 30 Hz and the Nyquist frequency, to fit the wider range of high frequency noise.

7.2.6 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) was proposed by Huang et al. (1998) and used for ECG signal denoising by Blanco-Velasco et al. (2008) and, later on, for the specific purpose of biometric identification, by Zhao et al. (2013). Just like Fourier transforms decompose a signal into a set of sine waves of different frequencies, EMD decomposes the signal into intrinsic mode functions (IMF), that satisfy the following conditions: the number of extremes and zero-crossings must be equal or differ by no more than one, and the mean value of the envelopes defined by the local maxima and minima must be always zero.

The process of decomposing a signal into N IMFs (c_n), as described in detail by Zhao et al. (2013), consists on extracting local extrema of the signal, building the upper and lower envelopes, calculating their mean, and assessing the compliance with the aforementioned conditions to consider it an IMF. If true, the process is repeated using the original signal removed of the IMFs already found, until this signal presents less than two extrema, at which point it is considered a residue (r) of the original signal, that can be expressed through:

$$x(t) = \sum_{n=1}^{N} c_n + r \tag{7.5}$$

A relationship between the IMFs and the frequency components of the signal can be traced (cf. Fig. 7.3). The first IMFs correspond to the highest frequencies, and the last IMFs to the lowest, while the residue usually corresponds to near zero frequency ondulations (or the DC offset of the signal).

As discernible from this definition of Empirical Mode Decomposition, the number of IMFs obtained for a signal, depending on the number of extrema detected, will vary with the length of the signal, and, in the case of ECG, also with the influence of noise that can generate false local extrema in the signal. Hence, the process of denoising based on Empirical Mode Decomposition consists on selecting of the IMFs with useful information and summing them to obtained a

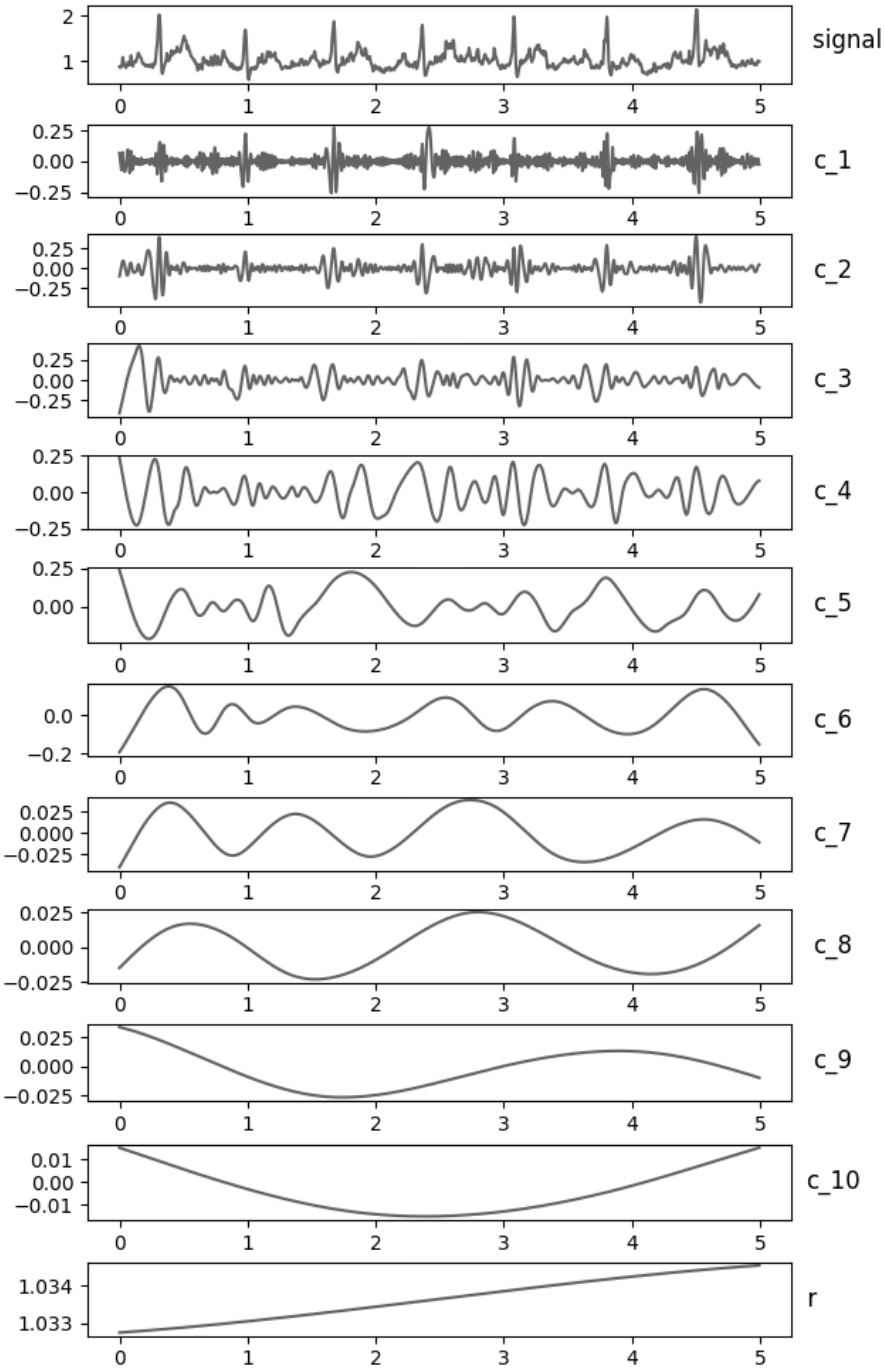

Figure 7.3: Example of the EMD's IMF set of an ECG signal excerpt (from the top, the original five-second ECG excerpt of a recording from the University of Toronto collection, the ten IMFs from Empirical Mode Decomposition, and the residue).

Table 7.1: Summary of the explored preprocessing methods.

Method	**Parameters**	**Settings**
Bandpass filter	1-30 Hz	Standard and driving
	1-40 Hz	Standard and driving
	2-30 Hz	Standard and driving
	2-40 Hz	Standard and driving
HPF + NF	1 Hz; PLI frequency	Standard
MAF + NF + LPF	1 s window; PLI freq.; 40 Hz	Standard
MAF + LPF	1 s window; 30 Hz	Driving
SG	6th degree polynomial	Standard and driving
SG + MAF	6th deg. poly.; 1 s window	Standard and driving
SG + HPF	6th deg. poly.; 1 Hz	Standard and driving
DCT	Annull coeffs. 0-1 Hz and PLI-5 to PLI+5 Hz	Standard
	Annull coeffs. 0-1 Hz and $\geqslant$30 Hz	Driving
DWT Rec.	Selective scale rec.	Standard and Driving
DWT Rec. + MAF	Selective scale rec.; 1 s window	Standard and Driving
DWT S.T.	Coeff. soft thresh.	Standard and Driving
DWT S.T. + MAF	Coeff. soft thresh.; 1 s window	Standard and Driving
DWT S.T. + MMF	Coeff. soft thresh.; 0.5 s window	Standard and Driving
DWT S.T. + HPF	Coeff. soft thresh.; 1 Hz	Standard and Driving
EMD	IMF Selection	Standard and Driving

denoised signal, discarding those that correspond to high frequency noise (the first ones) or low frequency noise (the last).

7.3 Results and Discussion

7.3.1 Objective Assessment with Simulated Signals

In order to correctly evaluate the performance of a preprocessing technique, two signals must be available for comparison: the signal after the denoising process, and the signal before contamination. Considering this, the use of signals from real collections, regardless of the acquisition settings, is not suitable for a thorough evaluation, as the signal before contamination is never available.

Thus, to fulfill these needs, the ECGSYN[1] generator was used. ECGSYN was described by McSharry et al. (2003), and consists on an algorithm based on three coupled differential equations that allow for the generation of realistic ECG signals, with highly tunable amplitude, duration, heart rate variability, and LF/HF ratios.

Using ECGSYN, an ECG signal (cf. Fig. 7.4) was generated with the following parameters: 200 Hz sampling frequency and internal frequency, 200 heartbeats, 0 mV additive noise (to ensure a clean signal), mean heart rate of 75 bpm (beats per minute) with standard deviation of 5 bpm

[1]ECGSYN. Available on: `https://www.physionet.org/physiotools/ecgsyn/` (visited on 07/02/2017).

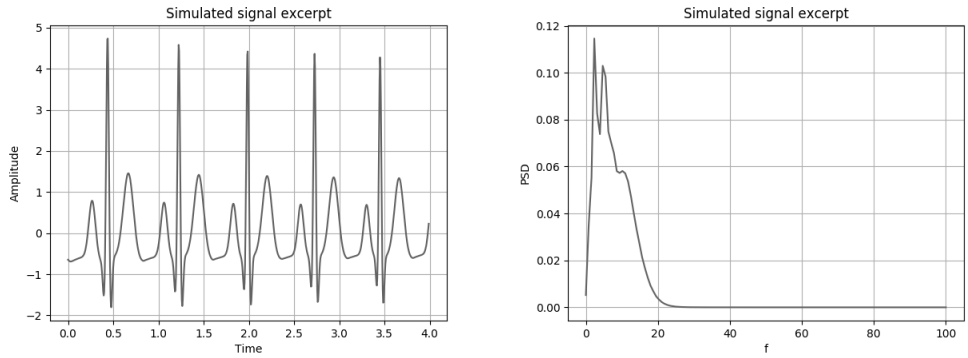

Figure 7.4: Excerpt and frequency content of the noise-free signal generated with ECGSYN (left: excerpt of the simulated signal; right: Power Spectral Density plot (Welch method)).

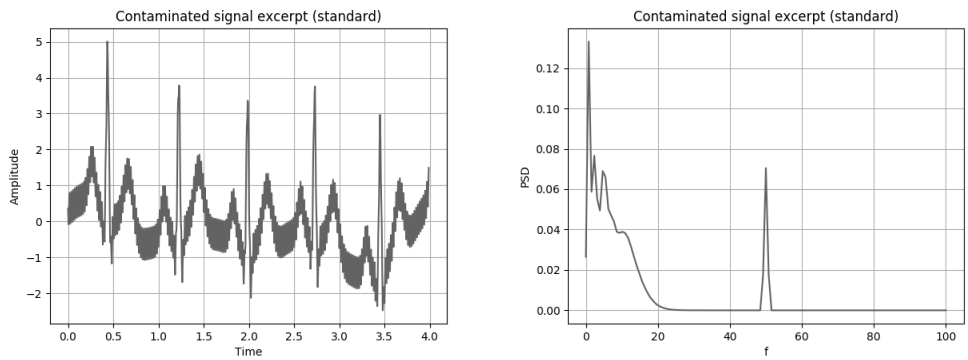

Figure 7.5: Excerpt and frequency content of the simulated signal contaminated with noise from standard off-the-person noise sources (left: excerpt of the simulated signal; right: Power Spectral Density plot (Welch method)).

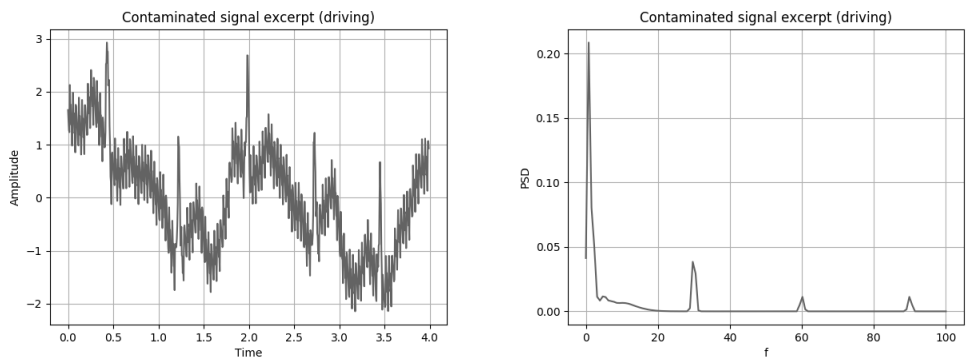

Figure 7.6: Excerpt and frequency content of the simulated signal contaminated with driving settings' noise (left: excerpt of the simulated signal; right: Power Spectral Density plot (Welch method)).

(to include high variability), LF/HF ratio of 0.5, and default angles and z-positions of extrema and Gaussian width of peaks.

This noise-free signal was then contaminated with two noise sets. For standard off-the-person noise sources (cf. Fig. 7.5), the signal was contaminated with a 50 Hz sinusoidal signal (powerline interference), and 0.25, 0.5, and 0.75 Hz signals (simulating low frequency noise and moderate baseline wander). For driving settings (cf. Fig. 7.6), baseline wander simulation was combined with motion artifacts, ridding the signal with 0.5, 1.0 and 2.0 Hz interference signals, and the PLI-simulating signal was replaced by 30, 60, 90 Hz signals (simulating the wide frequency range of EMG interference, from 20 to 1000 Hz).

The contaminated signal was then segmented into excerpts of 1, 2, 5, and 10 s, to evaluate the aforementioned methods' average performance with segments with different lengths. The denoised segments were compared with the original signal using root mean square error (RMSE).

7.3.1.1 Standard Off-the-Person Noise Sources

The results of the evaluation of the various explored methods based on the difference between the denoised and the original simulated signals, using root mean square error, can be conferred on Fig. 7.7. Through their analysis, it can be concluded that, independently of the length of the signal being denoised, the Discrete Cosine Transform filtering is the best option, confirming the statements of Choudhary and Manikandan (2015).

A visual comparison of the original signal and the denoising results proves that DCT preprocessing is fit to eliminate both high and low frequency noise without causing major distortions to the signal. For signals longer than two seconds, the combinations of a notch filter or Savitzky-Golay with a highpass filter also return very desirable results, as well as DWT soft thresholding with a moving median filter for shorter signals.

It is relevant to remark that the bandpass filters with higher cut-off frequency of 40 Hz still preserve some high-frequency noise, and those with lower cut-off frequency of 2Hz show some distortion of the signal, denoting the loss of ECG information. Among these filters, the 1-30 Hz proposed by Lourenço et al. (2012a) would be the best choice.

DWT reconstruction at defined scales proved to be completely inadequate, as it clearly causes the loss of useful waveform information left outside the chosen scale. The soft thresholding of DWT coefficients presented much better results, and its combination with an highpass filter allowed for the best attenuation of baseline wander.

Empirical Mode Decomposition presented good results in eliminating both low and high frequency noise. However, its dependency on the length of the signal and the influence of noise on it means EMD is unstable and such results may not be observed with real signals in less controlled settings.

The combination of HPF with a notch filter failed to eliminate completely the high frequency noise at 50 Hz. Moving average filter followed by a notch and a lowpass filter, similarly to the bandpass filters with 2Hz lower cut-off frequencies, caused the loss of useful, low frequency information, distorting the signal.

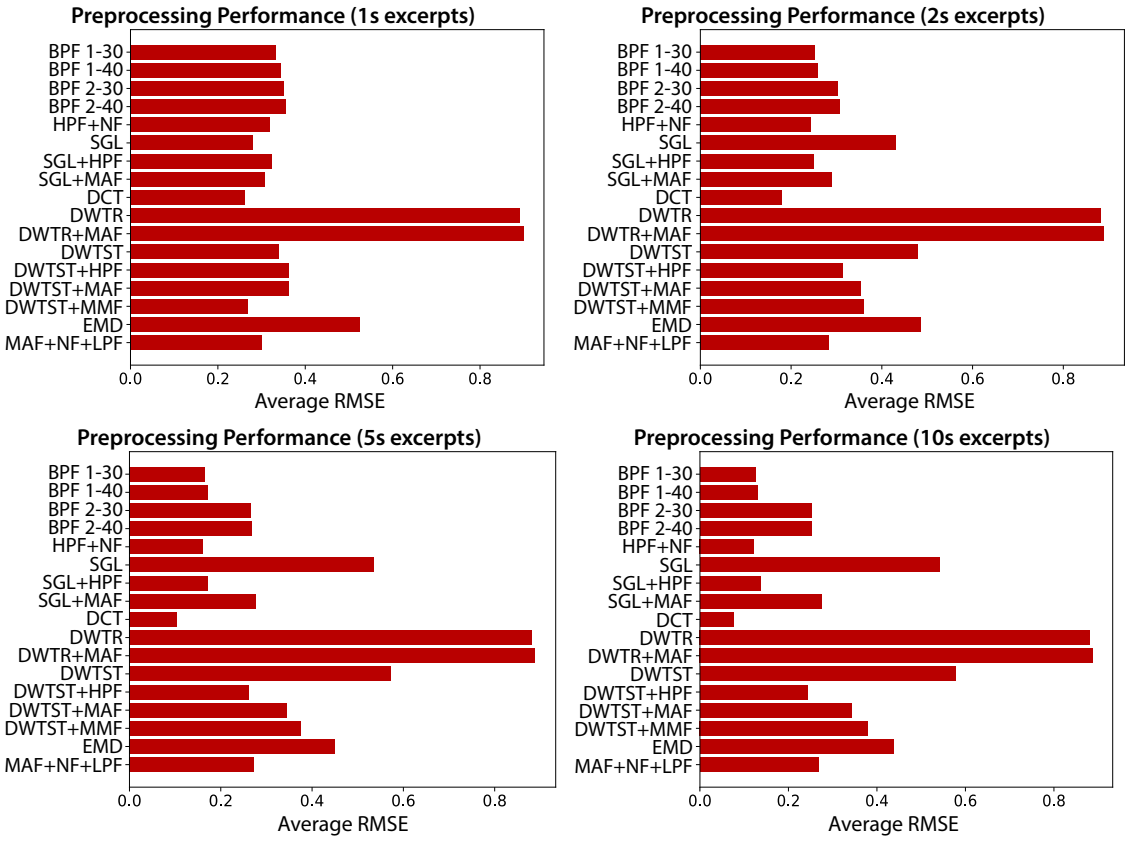

Figure 7.7: Average results of denoising using the explored methods on excerpts of various lengths, contaminated with noise from standard off-the-person settings.

The Savitzky-Golay filter proved unsuitable to eliminate baseline wander, so it worked better when combined with HPF or MAF. Still, even with these additions, some of the high frequency noise managed to remain in the denoised signal.

7.3.1.2 Driving Acquisition Settings

The average root mean square error between varying-length excerpts of the simulated signal (contaminated with noise from driving settings) and the result of denoising, for the various methods considered, is shown in Fig. 7.8. Based on this, DCT filtering, the best method for standard settings noise, was no longer the frontrunner, as it is unable to deal with the wider range of low frequencies of noise.

The best results for driving settings were obtained by the combinations of moving average filter with Savitzky-Golay or a lowpass filter, with the latter performing better for shorter signals (1 or 2 s excerpts). Nevertheless, the effect of harsher conditions was clearly felt on the results, as even these methods cannot output a clean signal, and still presented both high and low frequency noise, along with some signal distortion.

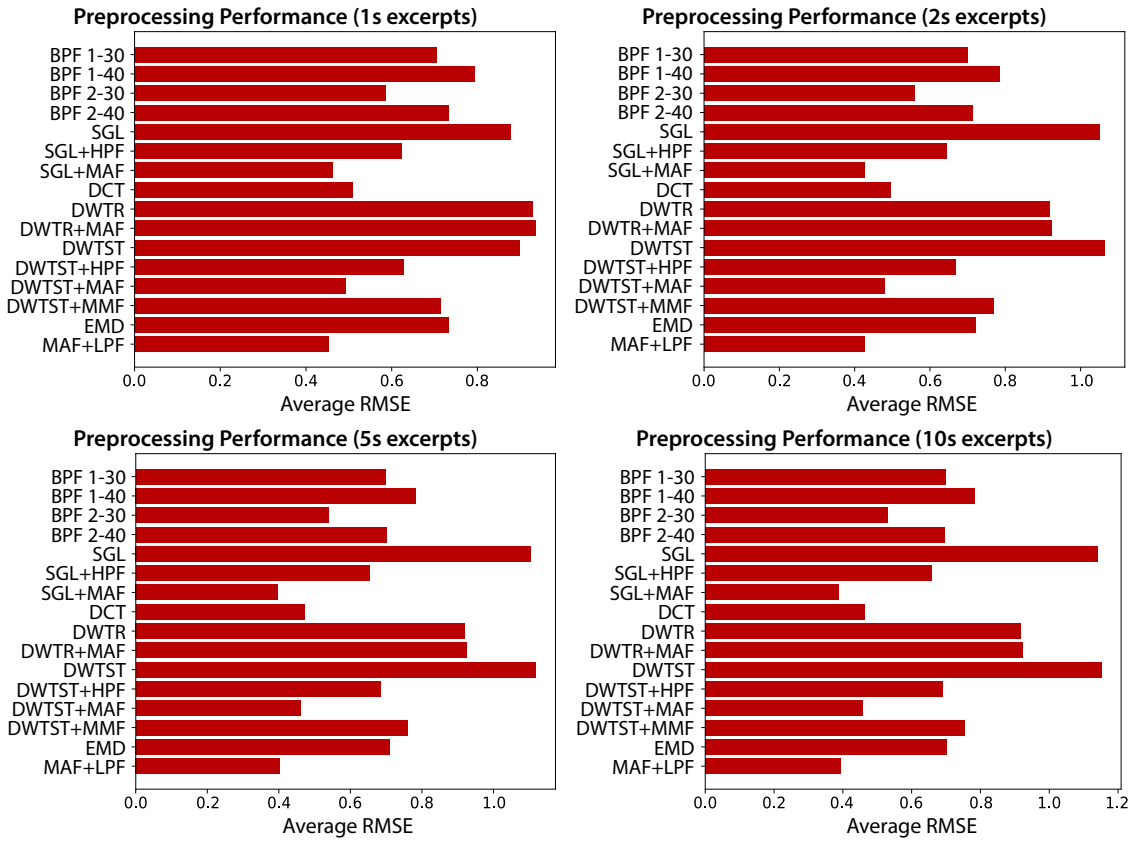

Figure 7.8: Average results of denoising using the explored methods on excerpts of various lengths, contaminated with noise from driving settings.

All bandpass filters used proved unable to filter the wide range of noise frequencies, as did Empirical Mode Decomposition. Scale reconstruction using DWT presented poor results, similar to those on standard settings. Likewise, DWT soft thresolding presented the incapability to deal with baseline wander. When combined with HPF, MMF, or, especially, MAF, this handicap was reduced, but some distortion was unveiled.

7.3.2 Subjective Comparison using Real Signals

Although simulated signals should offer a reasonably thorough evaluation of the capabilities of the different preprocessing techniques, testing with real signals is always necessary, as the results obtained will offer a uniquely accurate glimpse into the behaviour the method will show upon real application.

However, real signals are acquired already corrupted by noise of various noise sources (cf. section 2.4). Hence, the comparison of the output of the preprocessing methods with a clean, uncorrupted version of the signal is impossible. The output can only be compared with outputs of other techniques, subjectively.

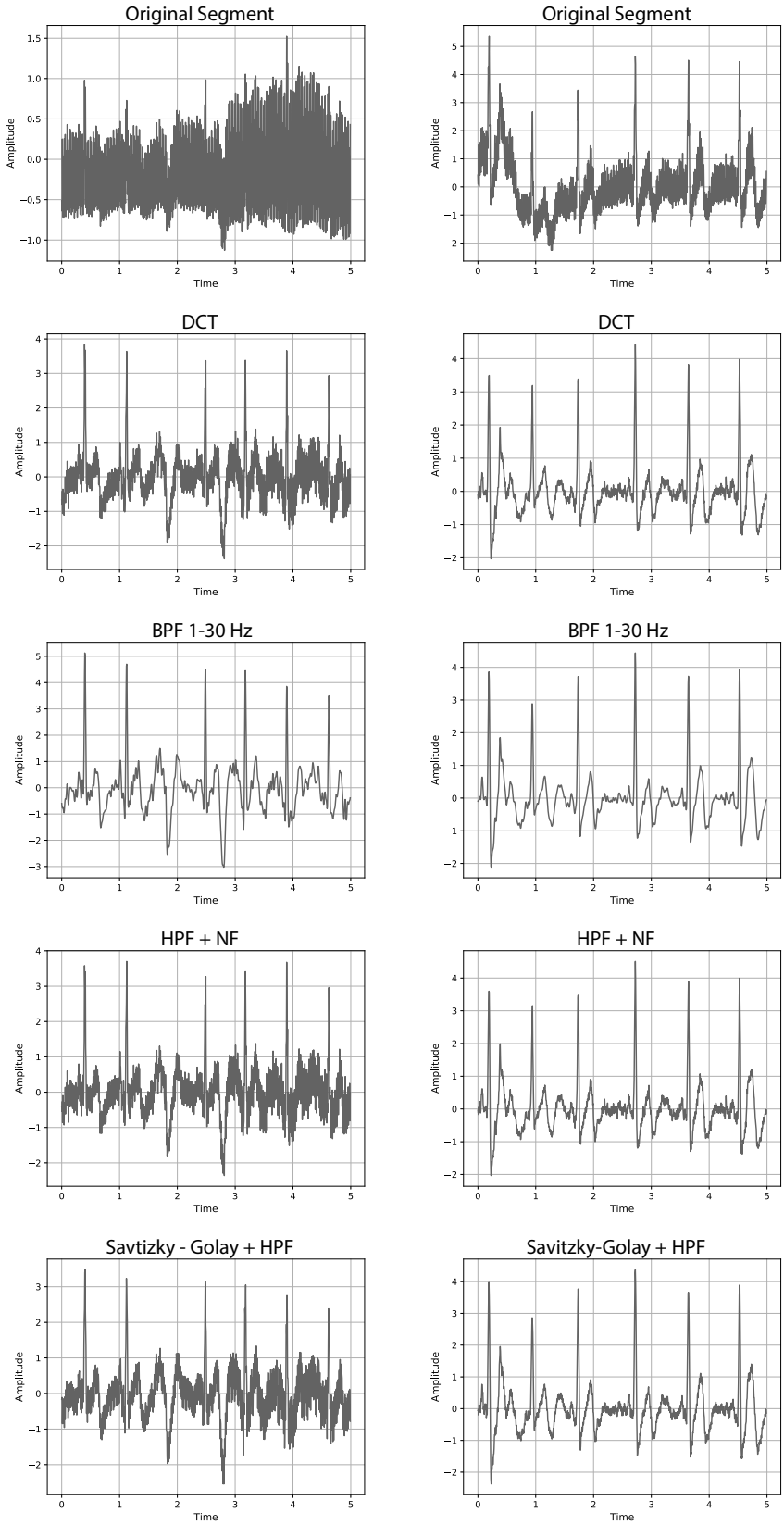

Figure 7.9: Example of the results of the preprocessing methods on UofTDB signals (standard off-the-person acquisition).

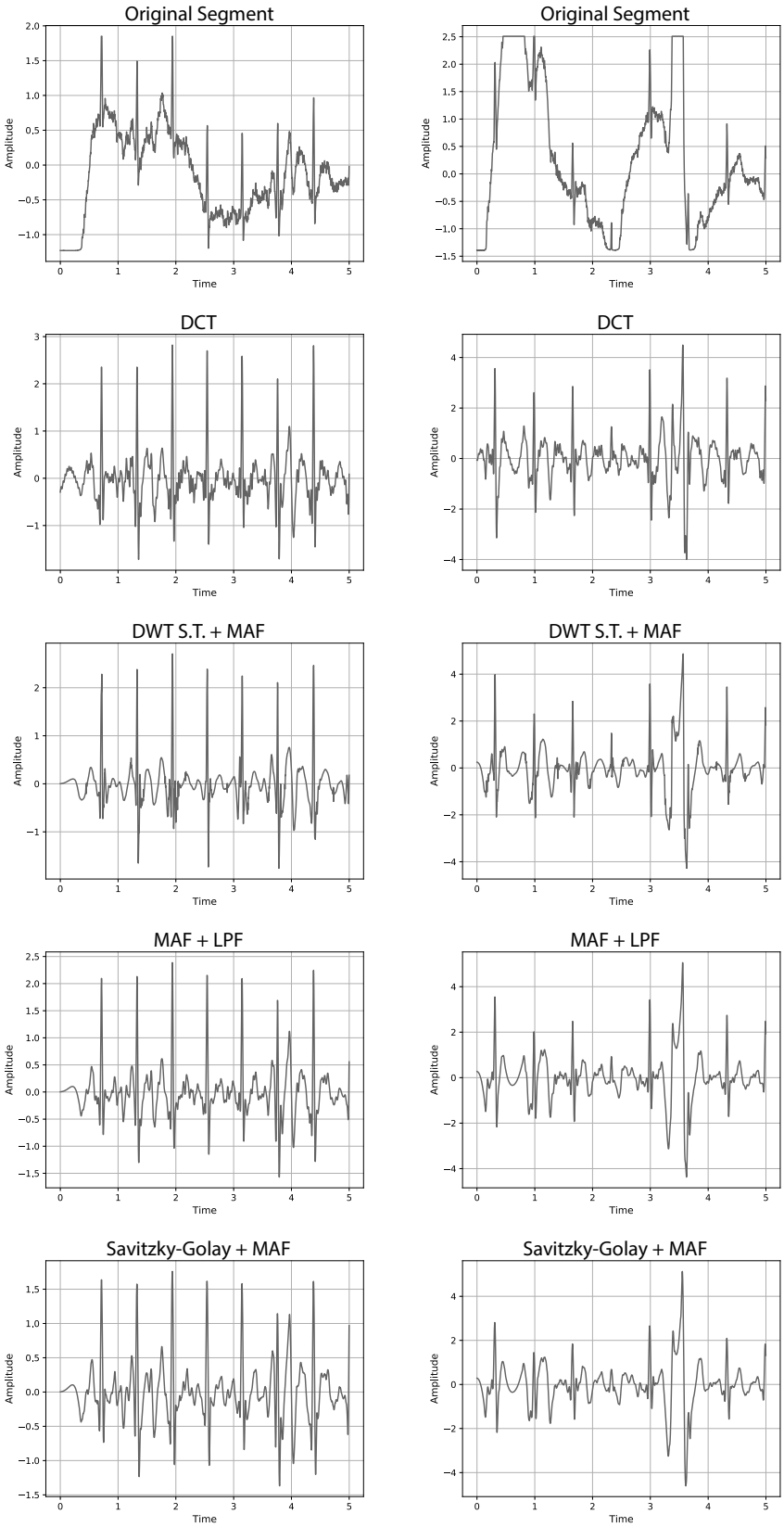

Figure 7.10: Example of the results of the preprocessing methods on real driving signals (from CardioID SteeringWheel v1 collection).

Thus, the four overall best preprocessing methods for standard off-the-person settings (DCT, HPF + NF, BPF 1-30 Hz, and SG + HPF) were compared, according to their apparent capabilities of reducing both high frequency and low frequency noise (baseline wander), using signals from the University of Toronto ECG database (UofTDB). The four best denoising methods for driving settings (DCT, MAF + LPF, DWT S.T. + MAF, and SG + MAF) were compared using signal excerpts from the CardioID collection (from the dataset of real coach bus driving settings).

Two examples of the results obtained with real, standard settings' signals (from UofTDB) are presented in Fig. 7.9. Unlike the results obtained with the simulated signal, DCT did not present itself as the best preprocessing method, with the Bandpass filter 1-30 Hz taking its place. While all methods seem to adequately correct the baseline wander on the original signals, only the bandpass filter can attenuate most of the high frequency noise.

Considering the UofTDB Canadian origins (where 60 Hz alternating current is used) and the sampling frequency of 200 Hz, the remaining high frequency noise would likely be the second harmonic frequency of PLI (120 Hz), that suffered aliasing and introduces a noise component at 80 Hz, only effectively attenuated by the bandpass filter.

A quick analysis of the frequency spectrum of UofTDB signals proved that, indeed, there is a contamination by 80 Hz noise. This aspect is quite important as it reminds of the utmost importance of adequate sampling frequency. Here, the results differ from the simulated signals but, in different signals, this problem is not expected to be observed and DCT should be the best method.

As for the driving signals, two examples of the results of the application of the selected preprocessing methods are presented in Fig. 7.10. As expected, the proposed method Savitzky-Golay with a moving average filter presents the best results, cleaning the signal of both high and low frequency noise. Unlike the proposed method, DCT and MAF+LPF maintain some high frequency noise, while DWT soft thresholding with a MAF causes the distortion of the shape of the ECG waveforms.

An analysis of the processing times required for each denoising method was performed (cf. Table 7.2). DCT and DWT Reconstruction were the fastest methods, while EMD and DWT Soft Thresholding with a Moving Median Filter took the longest to complete preprocessing. Once again, DCT proved the superior suitability for standard off-the-person applications. Savitzky-Golay with MAF, although not one of the fastest methods, presented very fast run times, compatible with continuous applications.

7.4 Summary and Conclusions

In conclusion, the evaluation of the several preprocessing methods described in this chapter, allowed for the determination of the best methods to be used for the denoising of off-the-person signals. While, for near-medical-grade signals (only minimally contaminated by noise), bandpass

Table 7.2: Average run times of each preprocessing method explored, when applied to signal segments of different durations (times presented in milliseconds).

	Segment length			
	1 s	**2 s**	**5 s**	**10 s**
BPF 1-30	0.587	0.610	0.583	0.626
BPF 1-40	0.689	0.610	0.578	0.675
BPF 2-30	0.740	0.648	0.571	0.635
BPF 2-40	0.560	0.587	0.579	0.634
DCT (std.)	0.029	0.049	0.054	0.082
DCT (drive)	0.035	0.035	0.057	0.086
DWT Rec.	0.147	0.173	0.173	0.208
DWT Rec. + MAF	0.396	0.589	0.628	0.863
DWT S.T.	0.340	0.490	0.658	1.110
DWT S.T. + HPF	0.816	1.740	1.170	1.520
DWT S.T. + MAF	0.635	1.530	1.120	1.730
DWT S.T. + MMF	3.920	11.100	21.300	43.200
EMD	29.400	52.400	56.100	79.100
HPF + NF	1.100	1.270	1.050	1.130
MAF + NF + LPF (std.)	1.180	1.510	1.460	1.800
MAF + LPF (drive)	0.627	0.690	0.850	1.200
SGL	0.604	0.756	0.636	0.638
SGL + HPF	1.070	1.270	1.130	1.160
SGL + MAF	0.856	1.230	1.070	1.320

filters are more than sufficient to obtain near-perfect signals, this study proved that there are better methods when using off-the-person signals, be them from standard acquisition procedures or driving settings.

For standard off-the-person settings signals, such as those of the University of Toronto, DCT was deemed to be the most robust method for the attenuation of noise, allowing for fair reconstruction of the original simulated signals, with much distortion. Although bandpass filters have proven to be more adequate to signals like those of UofTDB, where aliasing has taken place, DCT deserves to be the recommended choice for the denoising of signals on standard settings.

As for the driving signals, the proposed combination of Savitzky-Golay with a moving average filter presented the best results both on simulated and real signals, proving to be fit to clean the signals without distorting them. This is likely due to the approach's fewer assumptions on the frequency ranges of noise artifacts, that allows it to adapt to unforeseeable circumstances.

Nevertheless, it is understandable that all preprocessing techniques are flawed and some noise is expected to survive them. Specifically for driving signals, sensor saturation periods have shown to cause false peaks that could be mistaken by R-peaks or other ECG waveforms. It is, thus, paramount to prepare the signal and correct these situations before the feature extraction and recognition takes place. The next chapter, 8, presents the efforts to address these concerns.

Chapter 8

Signal Preparation

8.1 Problem Statement

Following the results reported in the prior art, for ECG signals acquired on standard settings, the need for robust preprocessing is low, as the noise that disturbs the signal is generally confined to very specific frequency ranges, rarely directly contaminating the useful frequency ranges of the ECG. Hence, while more refined algorithms can be used, such as the DCT-based filtering, the simple and widely used bandpass filters fulfill their purpose to a very satisfying extent.

However, as presented in the previous chapter, even after denoising with the best preprocessing methods, the signals acquired in an off-the-person driving context were still markedly distorted by high frequency noise and occasional sharp variations caused by sensor saturation. This will, expectedly, degrade the performance of the recognition process, as it injects unwanted variability into the signal.

Therefore, in this step of signal preparation, before the extraction of features that would otherwise carry this undesirable noise, it is of utmost importance to detect and discard the portions of the signal that have poor quality. In this dissertation, as further explained in the next section, this goal was achieved through the detection of R-peaks on the signal, the segmentation of heartbeats on fixed widths around them, the detection and removal of outliers through a novel method based on cross-correlation clustering, and the generation of heartbeat ensembles.

8.2 Proposed Approach

8.2.1 General Overview

On the quest to discard the noisiest parts of the acquired signals, the proposed approach focuses, mainly, on the detection of R-peaks for identification and segmentation of heartbeat segments, and the removal of outliers (false heartbeats caused by sensor saturation or heartbeats highly corrupted by noise) in order to avoid intrusions that, later on, will degrade the performance of the recognition algorithms.

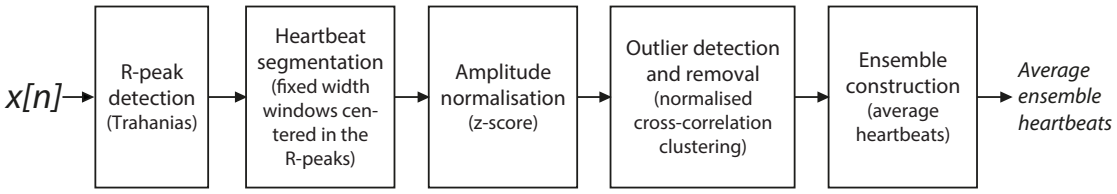

Figure 8.1: General overview schema of the entire signal preparation process ($x[n]$ denotes the preprocessed signal; this schema includes the chosen individual methods for each step, selected based on the results presented on this chapter).

The structure of the signal preparation process is presented in the next subsections, and a general schema, already specifying the individual methods, selected upon the results obtained in this part of the dissertation, can be conferred in Fig. 8.1.

8.2.2 Reference Point Detection

The first step towards the rejection of noisy parts of the signal was the detection of the R-peaks on the signal. Locating these reference points allows the posterior segmentation of heartbeats and, afterwards, the rejection of those which deviate more than acceptable from the normal.

For the detection of R-peaks, three of the most commonly used prior art methods were explored: Pan-Tompkins, proposed by Pan and Tompkins (1985); Trahanias, proposed by Trahanias (1993); and Engelse-Zeelenberg, proposed by Engelse and Zeelenberg (1979). These are, here, described in detail, and the comparative results of their application on the noisy off-the-person signals used for this dissertation are presented in section 8.3.

8.2.2.1 Pan-Tompkins

The Pan-Tompkins QRS detection method is one of the most commonly used in prior art research on ECG-based biometric recognition. Proposed by Pan and Tompkins (1985), this method is composed by five steps, oriented towards the goal of emphasising the characteristic sharpness of R-peaks to ease their detection (cf. Fig. 8.2).

The first step consists on a bandpass filtering process. Pan and Tompkins (1985) proposed the use of a cascaded bandpass filter, composed by a second-order lowpass filter with 11 Hz cutoff frequency, followed by a highpass filter with cutoff frequency of 5 Hz. This filtering process aims to reduce to a minimum any existing baseline wander or electromyographic interference.

The filtered signal is then differentiated to obtain information on the slopes of the QRS complex. The third step is the squaring of the value of each sample of the filtered and differentiated signal. The squared signal then suffers the action of a moving window integrator, where the window width was set to 150 ms.

The fifth and last step consists on a process of adjustable thresholding. A set of local maxima on the signal is determined and each is considered a candidate for R-peak. Signal peak and noise

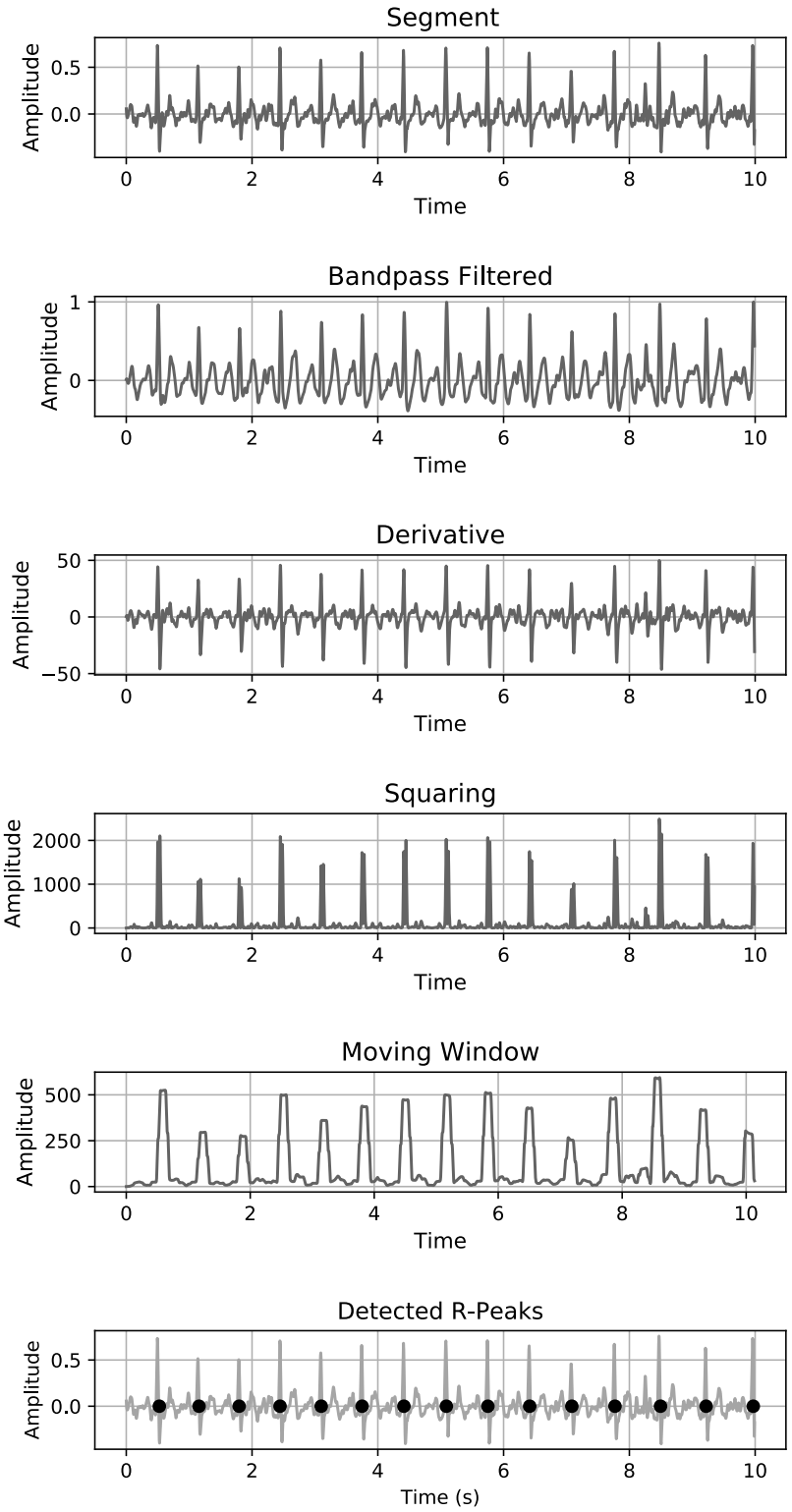

Figure 8.2: Example of the application and result of the application of the Pan-Tompkins method for R-peak detection on an ECG signal (the signal used was an excerpt from the UofTDB collection).

peak amplitude estimates are computed based on this set of maxima, and two thresholds are then defined according to these estimates. The amplitude of each R-peak candidate relative to these thresholds will determine if it is considered a signal or noise peak, and the thresholds are updated accordingly, so that they float above the noise, while following possible signal fluctuations.

The Pan-Tompkins method also defines and maintains two different reference R-R interval average measurements to avoid mistakenly detecting peaks too close or too far apart, but nevertheless allowing heart rate variability. These reference average values are also updated, and take into account the latest eight detected R-peaks. At last, the set of detected R-peaks is corrected for delays caused by the filtering, differentiation, and integration steps.

8.2.2.2 Trahanias

The Trahanias QRS Complex Detector was proposed by Trahanias (1993), and has since been used by researchers on ECG-based recognition works, including Molina et al. (2007).

This method is almost entirely based on one-dimensional morphological operations over the signal, namely, *open* and *close*, in order to act as a peak-valley extractor. The morphological operations open and close are combinations of two simpler operations, *dilation* ($\oplus$) and *erosion* ($\ominus$), that can be described by the equations:

$$(x \oplus s)[t] = \sup_{y \in S} (x[t] + b[t-y]) \tag{8.1}$$

$$(x \ominus s)[t] = \inf_{y \in S} (x[t+y] - b[t]) \tag{8.2}$$

Where $x[t]$ denotes the signal being dilated/eroded, $s[t]$ represents the structural element used for the operation (defined in the space S), and *sup* and *inf* denote, respectively, the infimum and the supremum of a set of values. Open ($\circ$) and close ($\bullet$) can be, thus, described through:

$$x \circ s = (x \ominus s) \oplus s \tag{8.3}$$

$$x \bullet s = (x \oplus s) \ominus s \tag{8.4}$$

Trahanias makes use of these simple morphological operations to enhance the sharp peaks in the signal that are, characteristically, the R-peaks. This process is composed by a filtering phase and a peak and valley extraction (*PVE*) phase (confer Fig. 8.3 for the overview schema of the process, and Fig. 8.4 for an example of its application).

The filtering phase is composed by two branches. In the first branch, the original signal suffers an operation of opening, followed by closing. In the second branch, the original signal suffers closing followed by opening. The filtering process concludes by computing the average between the resulting signals from both branches. This results in the suppression of signal noise, while avoiding deterioration of the ECG waveforms.

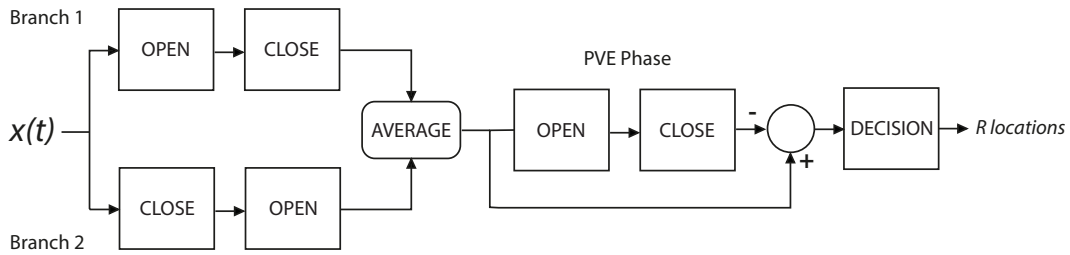

Figure 8.3: Structure of the Trahanias algorithm for R-peak detection.

The signal output by the filtering phase is then used on the *PVE* phase. This phase is a combination of a peak extraction (*PE*) and a valley extraction (*VE*) operations, that maintains the R-peaks, while effectively suppressing the remaining ECG waveforms that could interfere in the R-peak detection. *PVE* can be defined by the equation:

$$PVE(x) = x - ((x \circ s) \bullet s) \tag{8.5}$$

A decision rule is then applied to the result of the peak and valley extraction, in order to locate the R-peaks on the signal. In this case, a single initial threshold is computed based on local positive peaks found in the PVE signal (considered R-peak candidates). Then, each candidate is compared with the threshold to determine if it is a true R-peak or not, and at each affirmative decision the threshold is adapted to fit possible long-term variations on the signal.

8.2.2.3 Engelse-Zeelenberg

This algorithm was proposed by Engelse and Zeelenberg (1979), and works on a differentiated version of a filtered signal to detect the location of R-peaks. In this dissertation, a real-time adaptation proposed by Lourenço et al. (2012b) is used, as its development was directly focused on ECG signals acquired at the fingers.

The first step of this method consists on the differentiation and filtering of the input signal. First, the signal is differentiated, following the equation:

$$y_1[n] = x[n] - x[n-4] \tag{8.6}$$

Then, the differentiated signal passes through a lowpass filter, described by the expression:

$$y_2[n] = \sum_{i=0}^{4} c_i . y_1[n-i], c_i = [1,4,6,4,1] \tag{8.7}$$

An illustration of the application of this step on an example signal can be analysed in Fig. 8.5. After this brief processing phase, the detection of R-peaks on the signal is based on the recognition of the two lobes created on the differentiated signal by the abrupt slopes between Q-R and R-S.

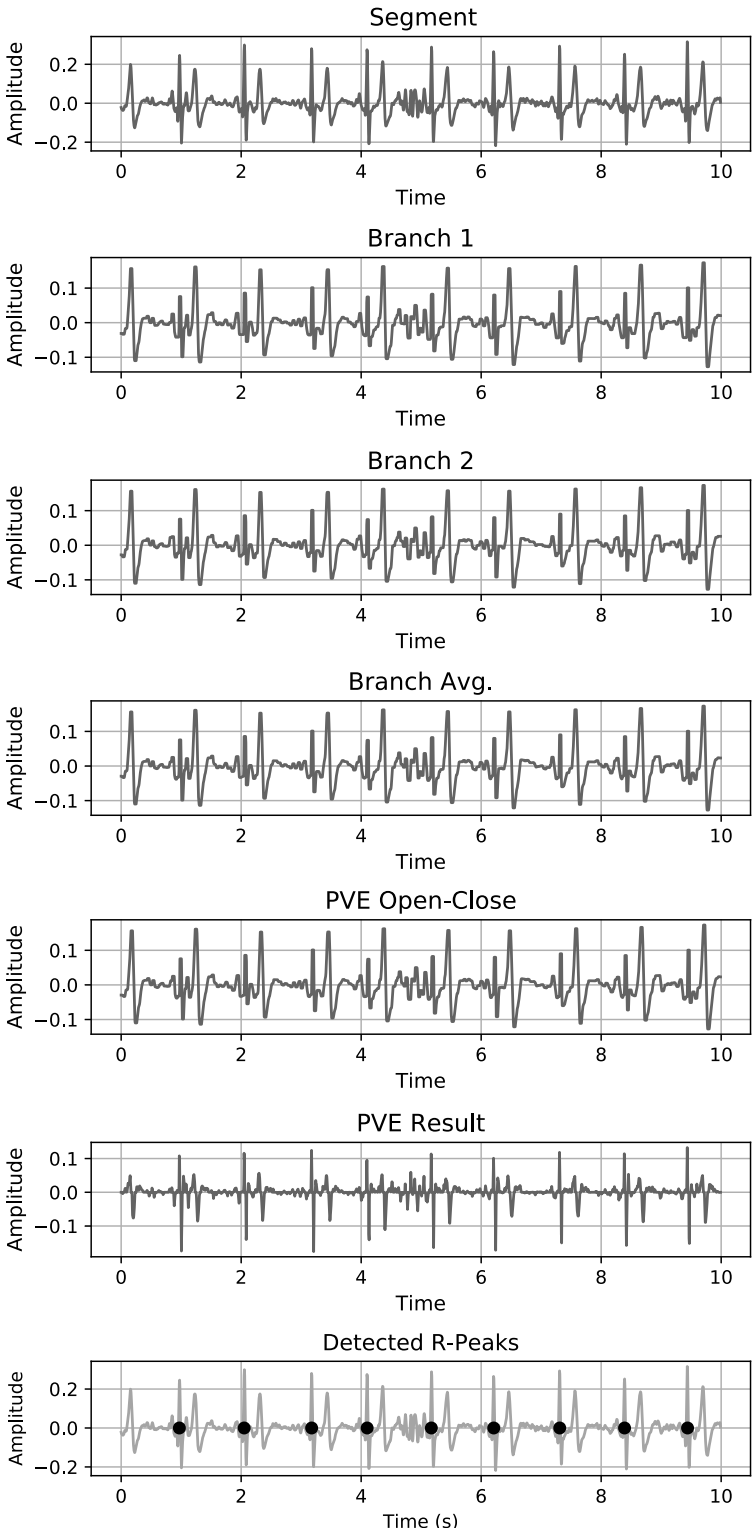

Figure 8.4: Example of the application and result of the Trahanias method for R-peak detection on an ECG signal (the signal used was an excerpt from the UofTDB collection).

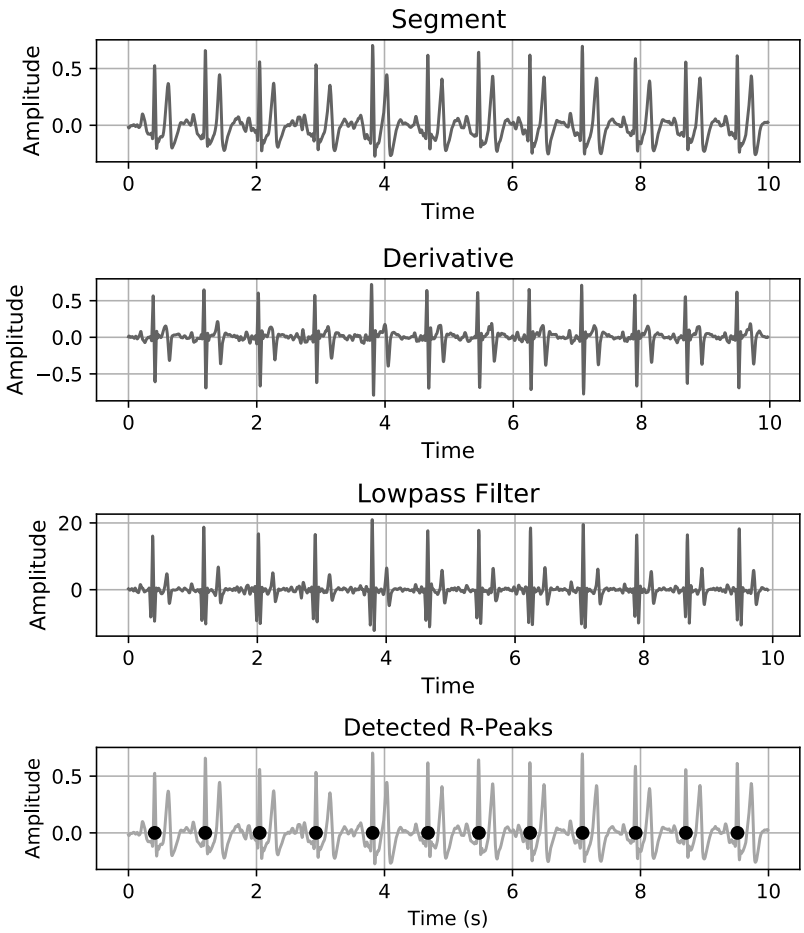

Figure 8.5: Example of the application and result of the Engelse-Zeelenberg method for R-peak detection on an ECG signal (the signal used was an excerpt from the UofTDB collection).

First, the algorithm scans the signal until it finds a local maximum where $y_2[n] > M$, where M is an adaptative threshold. Then, it scans the following signal samples to find a 160ms window, W, where a consecutive group of points verify the condition $y_2[n] < -M$. The original signal $x[n]$ inside the window W is then scanned to find the R-peak (that will correspond to the amplitude maximum inside the window).

Besides all this, the algorithm denies the detection of R-peaks for a period of 200 ms after the last detection. This process repeats throughout the signal, partial thresholds M_i are computed in 5 s sliding windows, according to the equation:

$$M_i = 0.6 \max(y_2[n]) \tag{8.8}$$

The threshold M is determined using a buffer composed by the five most recently computed partial thresholds, as shown by:

$$M = \begin{cases} M_i & \text{during the initial 5s of signal} \\ \frac{1}{5}\sum_{i=1}^{5} M_i & \text{for the rest of the signal} \end{cases} \tag{8.9}$$

8.2.3 Segmentation and Normalisation

The segmentation follows the detection of R-peaks on the signal. In this dissertation, like many successful prior art research works (Labati et al., 2013; Silva et al., 2013; Chun, 2016), a simple approach was followed, using fixed-width windows around the detected R-peaks to segment the heartbeats. The window frame limits were defined at 250 ms before the R-peak, and 400 ms after it, as an adaptation to those used by Silva et al. (2013), following empirical proof this was sufficient to include all waveforms in its entirety, considering normal heart rate variability, and without intrusion of anterior or posterior heartbeats.

Following the segmentation of the heartbeats, amplitude normalisation was conducted. Due to varying impedance in the driving settings used during acquisition, signals are prone to vary very widely in their amplitude range, even within short periods of time in the same acquisition. Because of this, it proves unreliable to apply normalisation to the signals as a whole. Hence, it was individually applied to each segmented heartbeat.

For the amplitude normalisation, three common methods were explored, namely, the *min-max* (Irvine et al., 2008; Fang and Chan, 2009; Li and Narayanan, 2010; Safie et al., 2011), the *z-score* (Odinaka et al., 2010), and the *max-div* (Tawfik et al., 2010; Lourenço et al., 2011a). Considering a signal or signal segment $x[n]$, its normalised version $y[n]$ can be obtained, through the *min-max* method, by subtracting its minimum amplitude value and then dividing it by the difference between the maximum and minimum amplitude values, following the equation:

$$y[n] = \frac{x[n] - \min(x[n])}{\max(x[n]) - \min(x[n])} \tag{8.10}$$

The *z-score* normalisation takes the signal or segment $x[n]$, subtracts the mean of the distribution of sample amplitudes, and divides it by the standard deviation, according to:

$$y[n] = \frac{x[n] - \mu(x[n])}{\sigma(x[n])} \tag{8.11}$$

The *max-div* technique simply takes the signal $x[n]$ and divides it by its maximum amplitude value, as in:

$$y[n] = \frac{x[n]}{\max(x[n])} \tag{8.12}$$

The *min-max* method limits the amplitude values of the signal to the range $[0;1]$. When applied to individual heartbeat segments, as in the case of this approach, the maximum amplitude value of 1 will most likely belong to the R-peak, while the minimum of 0 will be verified at either the Q or

S peaks. The *max-div* method also limits the amplitude with an upper boundary at 1, but without enforcing a lower amplitude limit. The *z-score* normalisation will, ideally, center the signals' baselines at 0, and set the amplitude distribution to, roughly, the same range.

Upon a quick but thorough exploration of these three methods, and faced with the results of each on the driving ECG signals, the best option was the *z-score* normalisation. Although it tends to allow for R-peak maximum amplitude variations between different heartbeats (that the other methods do not allow), the fact that the signals are centered around amplitude 0 makes the heartbeats, overall, more similar between themselves than with the other methods, which greatly enhances the performance of the posterior outlier detection and ensemble construction processes.

8.2.4 Outlier Detection and Removal

8.2.4.1 Proposed Approach: Normalised Cross-Correlation Clustering

Chan et al. (2008) proposed a method for the removal of noise-corrupted heartbeats based on the cross-correlation coefficients between pairs of PQRST segments obtained. The distribution of the coefficients was measured, and any segment with cross-correlation coefficient lower than one standard deviation from the mean coefficient was discarded, effectively removing heartbeats that strayed from the regular and almost perennial cyclic structure of ECG heartbeats. A similar approach was used by Labati et al. (2013), for QRS segments.

However, these methods, similarly to DMEAN, are based on average measurements between all segments extracted. In highly noise-ridden ECG signals, such as those obtained in off-the-person and driving settings, this aspect means that the frequent outliers will have an impact on the reference used for outlier detection and removal, and will, thus, lower the likelihood of being detected and removed.

Hence, to eliminate these vulnerabilities and allow for robust outlier removal in noisy settings, a reformulation of the method proposed by Chan et al. (2008) is proposed in this dissertation. First, cross-correlation was substituted by normalised cross-correlation (NCC), in order to obtain coefficient values in the range $[0;1]$. A matrix NCC, $N \times N$, of normalised cross-correlation coefficients between N heartbeat segments, was built following the equation:

$$NCC_{ij} = \max(c(x_i, x_j)), i, j = 1, ..., N \tag{8.13}$$

Where x_i and x_j denote two of the N heartbeat segments, and $c(a,b)$ is the vector of coefficients from the cross-correlation between a and b, of which only the maximum is considered (for $i = j$, this maximum coefficient will be equal to 1).

Root mean square error (RMSE), Dynamic Time Warping (DTW) and Fast DTW distances were also explored as metrics of dissimilarity between heartbeat segments for outlier detection. However, these were quickly discarded, as RMSE rendered much worse results than NCC, and DTW and Fast DTW, while slightly better, made the process far too lengthy to be effective in continuous settings.

The average of each row of *NCC* was then computed (vector A), returning an average value of cross-correlation of each heartbeat segment with all the other segments extracted (including itself). An average value of 1 means the segment considered is perfectly correlated with all the others, while an average value of 0 means the opposite.

$$A_i = \frac{1}{N}\sum_{j=1}^{N} NCC_{ij} \tag{8.14}$$

Knowing that ECG heartbeats, although presenting variability, are cyclical and always present the five characteristic waveforms (P, Q, R, S, and T), it stands to reason that, albeit contaminated by noise, these will present much higher average cross-correlation than saturation peaks, that are mostly random.

Thus, the N elements of vector A of average normalised cross-correlation coefficients were arranged in descending order. The highest elements of the vector were added to the sole cluster, and its mean value was computed, to initialise the clustering process (based on the k-means algorithm). The number of highest elements to be used should be defined weighing the total number of elements and the expected quality of the signal. If the number of elements is high, the initial cluster size should be larger, while if the signal quality is expected to be poor, the initial cluster should be smaller, to avoid the inclusion of outliers.

For each of the remaining elements of the vector, each iteration compared the element's value to the current mean of the cluster (m). If the cluster's mean value was within a distance of ε of the element's value, and the element's value of average normalised cross-correlation is above a empirically found threshold of 0.5, then the element was added to the cluster.

$$i = \begin{cases} template & \text{if } m - A_i \leqslant \varepsilon \text{ and } A_i \geqslant 0.5 \\ outlier & \text{otherwise} \end{cases} \tag{8.15}$$

To increase the difficulty of adding more elements to the cluster when its mean value approaches the threshold of 0.5, ε is dynamic and depends on the value of the cluster's mean m and a fixed ideal ε_0 (empirically defined as 0.1). It is computed, at each iteration, using the equation:

$$\varepsilon = \varepsilon_0 \left(m^2\right) \tag{8.16}$$

Finally, at the end, the segments that correspond to the elements added to the cluster are selected to be used in the construction of the ensemble segments, while those left out of the cluster are deemed as outliers and discarded. The results of this method are presented in the next section, along with a comparison with DMEAN, a common outlier detection algorithm described below.

8.2.4.2 DMEAN

DMEAN is the designation attributed to an algorithm for outlier detection and removal proposed by Lourenço et al. (2013), as a better alternative to the DBSCAN clustering algorithm, proposed by Ester et al. (1996), and specifically focused for outlier removal on biometric recognition using

ECG signals acquired in noisy off-the-person settings. It was used on ECG-based recognition by Lourenço et al. (2014) and Carreiras et al. (2016).

Briefly, DMEAN is a distance-based outlier detection algorithm, unlike the proposed method, that is based on clustering. DMEAN begins by computing a mean template $\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i$ of the set of N templates (x_i) extracted from a recording session.

The cosine distance between each template and the mean is then computed, following the equation:

$$D_{cos}(x_i, \mu_x) = 1 - \frac{\sum_{k=1}^{m} x_i[k]\mu_x[k]}{\sqrt{(\sum_{k=1}^{m} x_i[k]^2)(\sum_{k=1}^{m} \mu_x[k]^2)}} \tag{8.17}$$

Based on the computed distances, a set of measurements are performed: the minimum and maximum values of each template (x^{min} and x^{max}, respectively) and the template set's median minimum and maximum values ($\tilde{x}^{min}$ and $\tilde{x}^{max}$, respectively); and the first and second order statistical moments of the distances of all templates to the reference (μ_D, the mean, and σ_D, the standard deviation).

Finally, the decision step of the algorithm states that a template of the set is considered an outlier if it verifies at least one of the following conditions:

1. i_{max}, the index of the template's amplitude maximum, does not correspond to the index of the R-peak;

2. $x^{max} > 1.5\tilde{x}^{max}$ and $x^{min} < 1.5\tilde{x}^{min}$;

3. $D(x, \mu_x) > \mu_D + 0.5\sigma_D$.

The templates that do not verify any of these conditions are kept in the set. DMEAN's performance in outlier detection, on the signals used for this dissertation, is compared with the results obtained with the proposed approach. These results and their discussion are presented in section 8.3.

8.2.5 Ensemble Heartbeats

After the selection of the best heartbeat templates in the recording, through the aforementioned outlier detection process, an heartbeat ensemble is built resorting to a simple average heartbeat, following the examples of , Silva et al. (2013), Lourenço et al. (2014), and Chun (2016).

DTW-based approaches were also considered, namely Non-Linear Alignment and Averaging Filters (Gupta et al., 1996), Prioritised Shape Averaging (Niennattrakul and Ratanamahatana, 2009), and Resampling Shape Averaging (Ongwattanakul and Srisai, 2009). However, as the ensemble construction process must be able to work with a large amount of selected templates, and DTW is computationally expensive in these circumstances, these options were discarded. Nevertheless, for ensemble construction in non-continuous settings, the versatility of DTW should not be overlooked.

Table 8.1: R-peak detection results upon evaluation with segments of the UofTDB collection (the segments were obtained through the extraction of 25 second excerpts of recodings of different, randomly selected subjects).

Segment	**Trahanias**		**Pan-Tompkins**		**Engelse-Zeelenberg**	
Total Peaks	**Correctly Detected**	**False Peaks**	**Correctly Detected**	**False Peaks**	**Correctly Detected**	**False Peaks**
30	30	0	27	0	30	0
29	26	0	25	0	29	0
30	27	0	27	1	30	0
30	27	0	29	1	29	0
33	31	0	32	0	33	0
27	26	0	27	0	25	0
40	32	0	35	0	38	0
35	31	0	31	0	35	0
30	29	5	26	4	29	7
30	28	0	28	0	30	0
314	91.40%	**1.59%**	91.40%	1.91%	**98.09%**	2.23%

Table 8.2: R-peak detection results upon evaluation with segments of the CardioID coach bus driving database (segments were randomly chosen among those that presented length between 10 and 25 seconds of continuous contact, confirmed through the lead-on detector).

Segment		**Trahanias**		**Pan-Tompkins**		**Engelse-Zeelenberg**	
Length (s)	**Total Peaks**	**Correctly Detected**	**False Peaks**	**Correctly Detected**	**False Peaks**	**Correctly Detected**	**False Peaks**
17.58	16	15	9	0	10	14	10
13.90	10	6	1	1	2	10	4
13.72	13	12	3	9	5	12	4
18.49	23	23	5	0	5	21	3
14.39	18	17	2	12	4	17	2
11.06	8	7	5	2	6	2	6
16.61	10	3	8	3	11	6	8
20.68	21	21	6	0	5	19	5
18.66	18	18	4	15	7	18	6
16.04	13	12	11	0	9	11	7
	150	**89.33%**	**36.00%**	28.00%	42.67%	86.67%	36.67%

8.3 Results and Discussion

8.3.1 Reference Peak Detection

For the detection of R-peaks on the ECG signals, the three methods previously presented were explored: Pan-Tompkins, Trahanias, and Engelse-Zeelenberg. These were tested on randomly selected recordings from the collection provided by the University of Toronto (UofTDB) and the database of real coach bus acquisitions performed by CardioID Technologies.

First, the evaluation of the methods with UofTDB signals began with the random selection of ten subjects. The first recording of each of these was cut into excerpts of 25s, and the R-peaks were annotated manually. Then, the three algorithms were applied, and the results were analysed. These results are presented on Table 8.1.

The Engelse-Zeelenberg method showed an advantage over the Trahanias and the Pan-Tompkins algorithms on the detection of true R-peaks, correctly locating 98.09% of the total amount, versus 91.40% by the other methods. Overall, all methods seem quite robust against false detections. However, in the specific case of the ninth segment (cf. Table 8.1), that included an unusual moment of sensor saturation, the methods presented a great increase on the number of false R-peak detections.

Then, ten segments of bus driving acquisitions of CardioID were selected, randomly, among those that verified a length between 10 and 25 s of continuous contact with the steering wheel (proven through the lead-on detector signal). The results are presented on Table 8.2, and showed a significant performance deterioration. As forecasted by the results of the specific segment of UofTDB mentioned above, the frequency of false detections was substantially higher, due to the higher influence of noise and prevalence of sensor saturation events.

Overall, while the Engelse-Zeelenberg method was indubitably better for relatively cleaner signals, Trahanias showed the best results, both in correct R-peak localisation and avoidance of false peak detections, for highly noisy signals. Nevertheless, the unacceptably frequent false R-peak detections affect the heartbeat segmentation process that follows, and emphasise the need for robust outlier removal techniques.

8.3.2 Outlier Detection and Removal

Both DMEAN and the proposed approach for outlier detection, Normalised Cross-Correlation Clustering (NCCC), were evaluated on a set of thirty segments of continuous contact, of varying duration, from a six-day recording of coach bus driver ECG signals at the steering wheel, provided by CardioID Technologies.

The thirty segments were randomly chosen among the available, and the only selection was performed so as to have a well-distributed representation of the available segment lengths. Each of the segments was preprocessed using a combination of sixth-order Savitzky-Golay and 1s moving

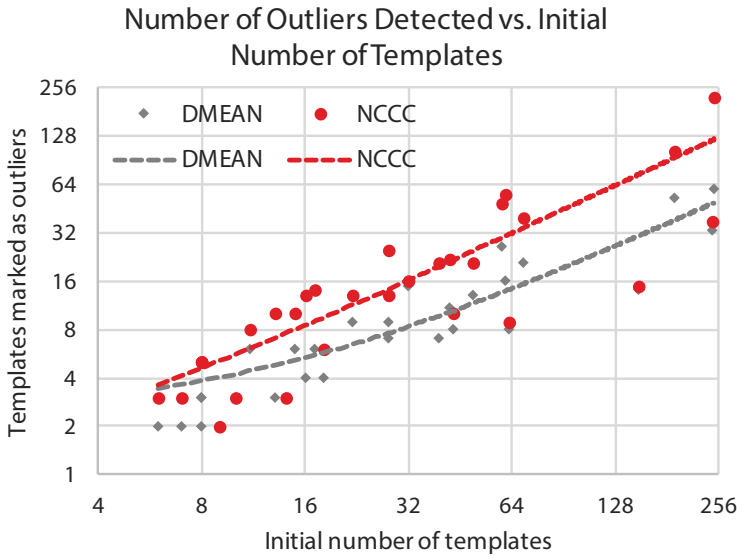

Figure 8.6: Outlier detection algorithm evaluation: number of templates marked as outliers versus the initial number of templates.

average filter (as described in Chapter 7), the R-peaks were detected using Trahanias (cf. subsection 8.2.2), and the heartbeats around those peaks were segmented and normalised using *z-score* (see subsection 8.2.3).

The algorithms of DMEAN and NCCC were then applied to the template sets extracted from each of the segments (see Fig. 8.7 for some representative examples). From the results, the initial number of templates and the number of outliers detected were computed, as well as the standard deviation of the template sets before and after the application of each method. As this dissertation also focus on continuous settings, and these methods may require large amounts of computations, the time required for the application of each algorithm in each segment was also measured.

Analysing the number of templates marked as outliers versus the total number of templates in the initial set (cf. Fig. 8.6), it is possible to conclude that, especially for larger initial sets, the NCCC method outperforms DMEAN in almost all situations.

The same happens when comparing the methods on the evolution of percentage of templates discarded with the increase of initial standard deviation (cf. Fig. 8.8). A higher initial standard deviation indicates a higher number of outliers on the initial set of templates. NCCC shows the ability to perform as well as DMEAN in better sets (with lower initial standard deviation), but also adapting and discarding more templates than DMEAN when applied in worse template sets.

For the comparison of NCCC and DMEAN using standard deviation, *z-score* was replaced by *min-max* normalisation, as the former sets the standard deviation of each template to 1.0 and, thus, it would always present this value independently of the outlier removal performance. It is important to notice that, although DMEAN appeared to perform better on *z-score* normalised heartbeat templates, NCCC outperformed it, generally, in either of the normalisation methods

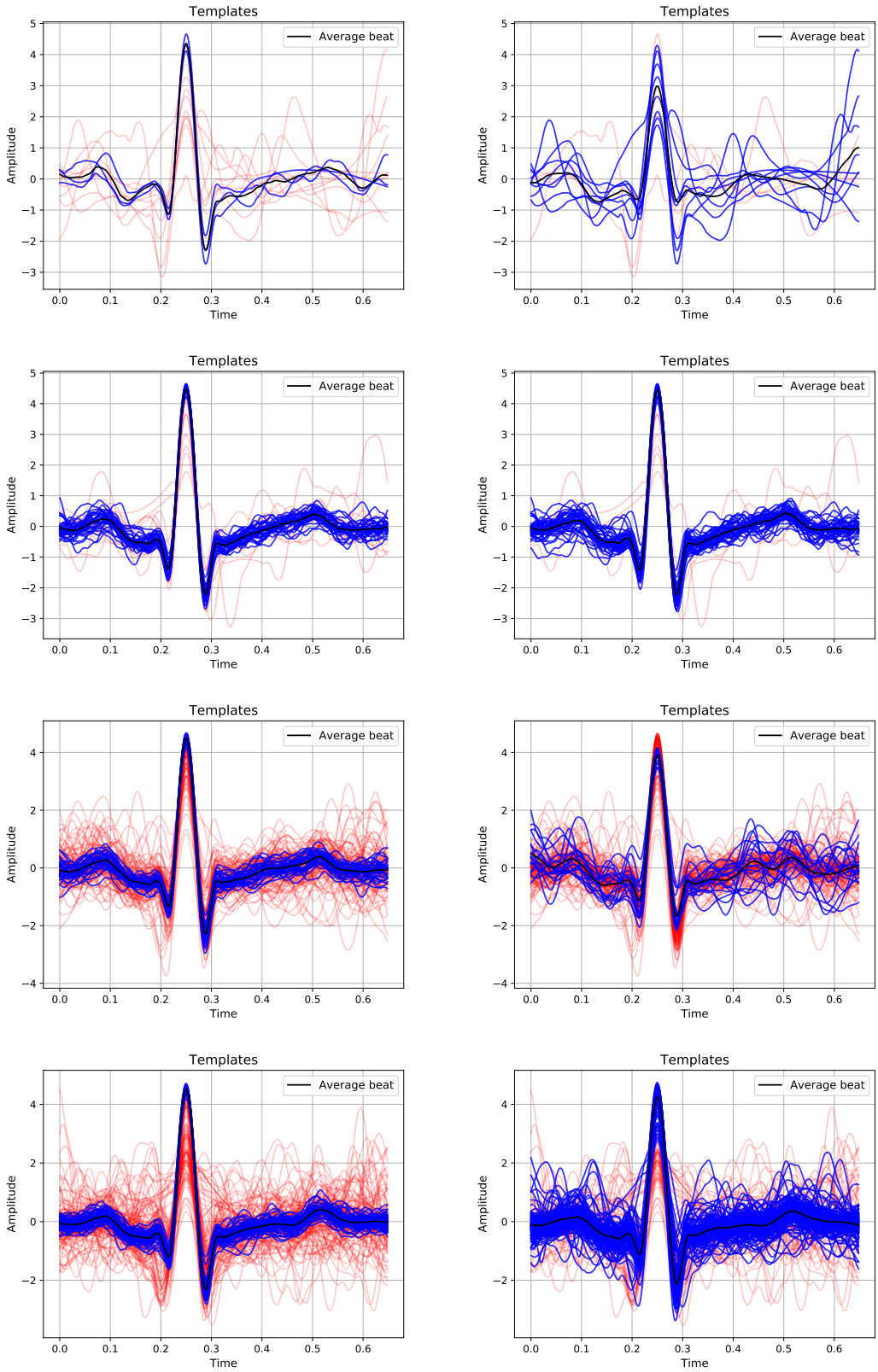

Figure 8.7: Outlier detection algorithm evaluation: comparison of the outlier detection algorithms' results, in sets with increasing number of templates (blue: normal templates; light red: outliers; on the left: using NCCC, the proposed approach; on the right: using DMEAN).

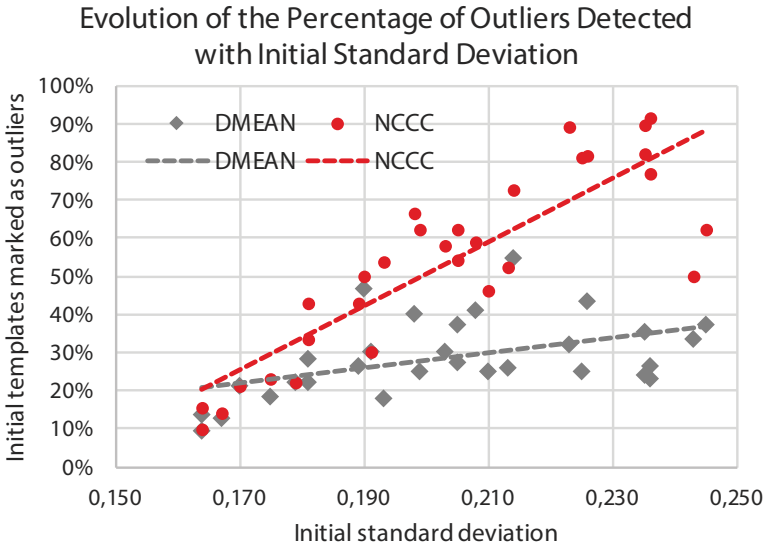

Figure 8.8: Outlier detection algorithm evaluation: percentage of outliers detected versus initial template set's standard deviation.

used.

However, by itself, the number or percentage of templates marked as outliers means little to nothing if it is not accompanied by improvement on the cohesion of the templates inside the set. By analysing the results of set's standard deviation before and after template selection (cf. Fig. 8.9), NCCC shows similar performance to DMEAN at lower initial standard deviation values, but its performance is, generally, much better than DMEAN for sets with higher initial standard deviation.

These results prove that the proposed approach is fit to be applied to highly noisy data, as well as template sets with few outliers. However, as NCCC requires the computation of cross-correlation between each possible pair of templates, while DMEAN uses a reference mean template, the only disadvantage is the time required that is much higher than with DMEAN, and increases quickly with the total number of templates (cf. Fig. 8.10).

The increased computational requirements of the proposed approach are a grave disadvantage, especially in the case of continuous systems working on hardware with less processing power. However, the performance benefits in highly noisy settings seem to surpass it. I hypothesise that, in the future, performance could be further improved, especially for smaller template sets, with the introduction of a supervised component, by using a general heartbeat template as atlas to reject outliers even when clean heartbeats are scarce.

8.3.3 Ensemble Heartbeats

Finally, the whole process of signal preparation culminates in the ensemble, where the average of each set of selected heartbeat templates is computed. It is important to carefully analyse the results

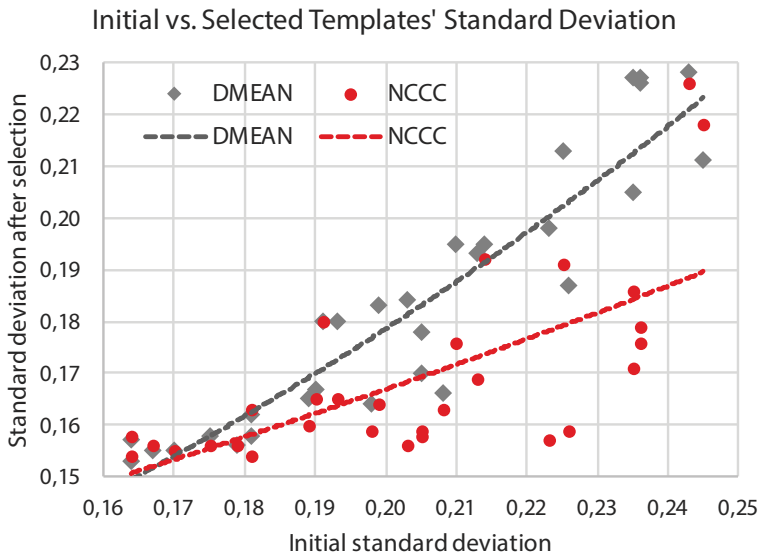

Figure 8.9: Outlier detection algorithm evaluation: template set's standard deviation before and after outlier removal.

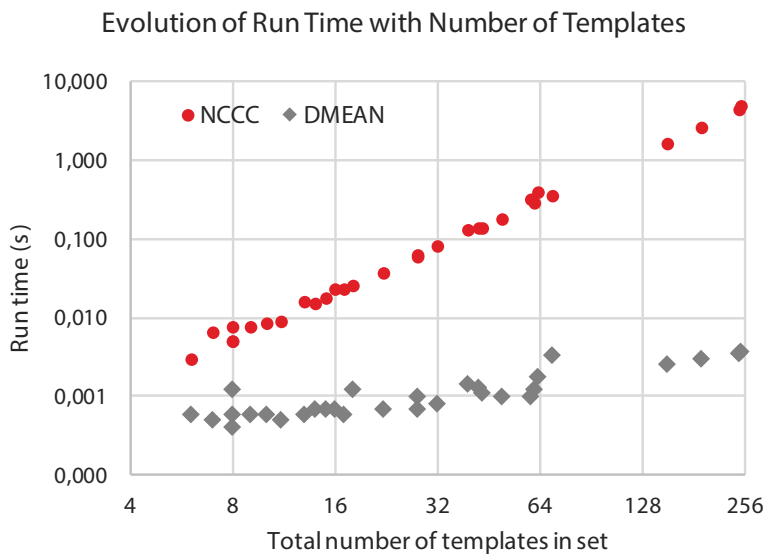

Figure 8.10: Outlier detection algorithm evaluation: evolution of time required with the total number of templates in the set.

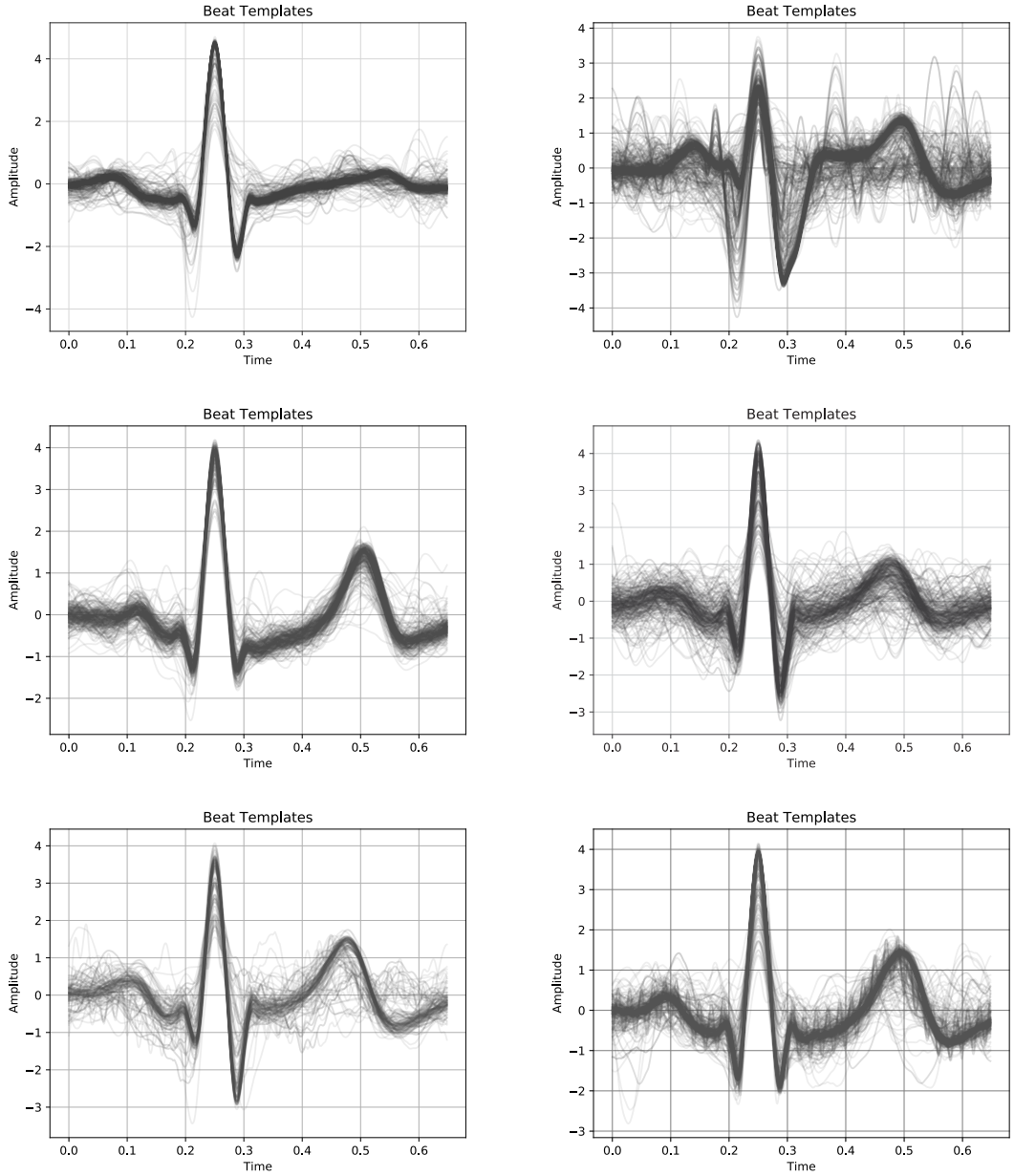

Figure 8.11: Ensemble heartbeats from each of the six subjects on the CardioID coach bus driving database.

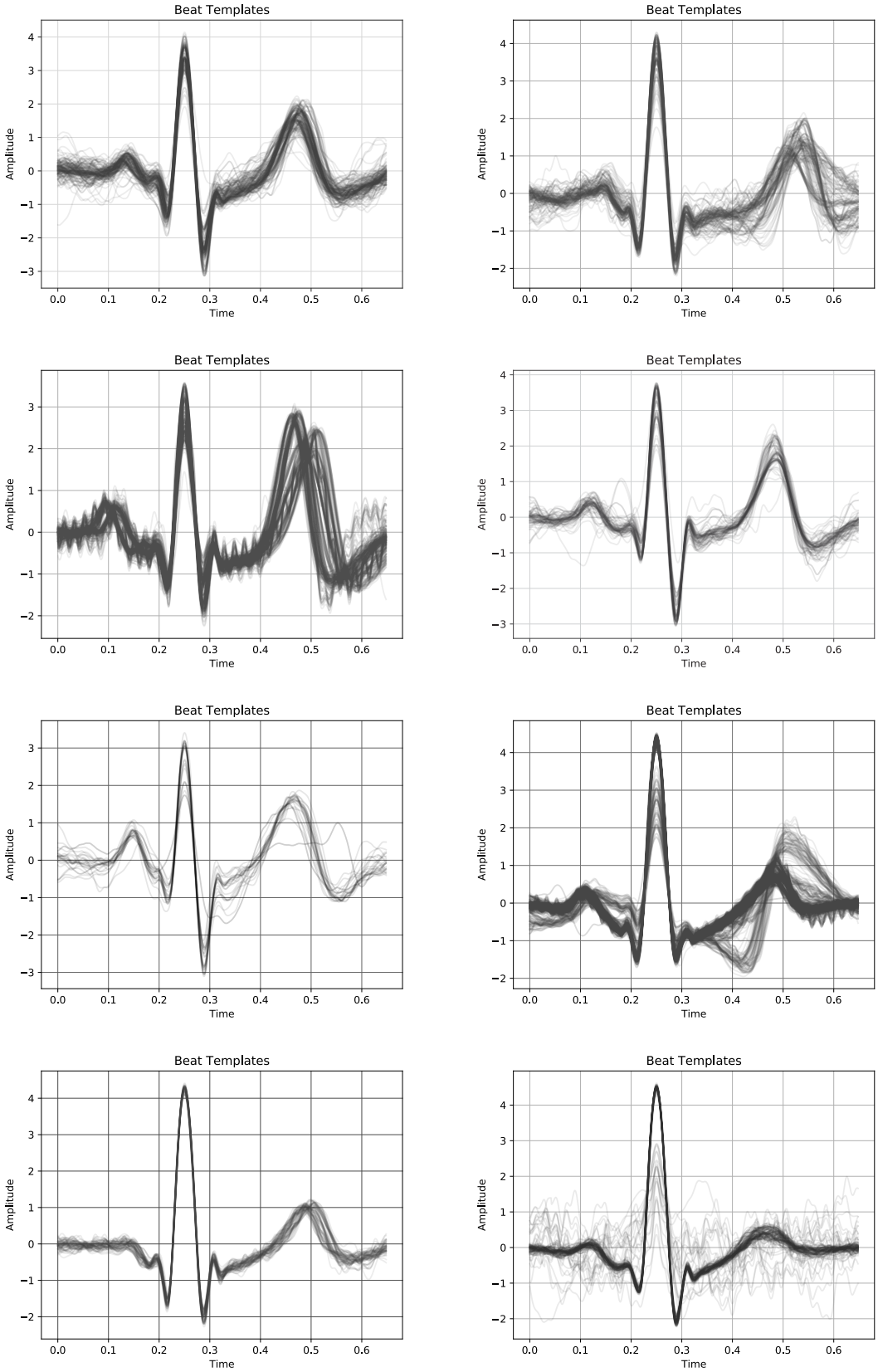

Figure 8.12: Ensemble heartbeats from each of the eight subjects on the Raspberry Pi simulated database.

of the ensembles, as they may reflect possible problems in the whole process, that must be dealt with before moving on to the posterior phases of feature extraction and recognition.

The off-the-person signals from the real coach bus driving collection and from the simulated acquisitions using the Raspberry Pi were applied the preprocessing step described in the previous chapter, and the preparation process here presented. The signals from both databases were segmented into five-seconds-long excerpts, to simulate the preparation process in expected continuous conditions.

Plots of the subjects from the real driving and simulation databases are presented, respectively, in Fig. 8.11 and Fig. 8.12. Two conclusions can immediately be drawn from their analysis. First, both preprocessing and preparation approaches fulfill their purpose to a very satisfying extent, as the ensemble segments seem clean and mostly unaffected by noise. Secondly, a high similarity can be observed between the majority of segments that belong to the same subject, meaning it can be reliably used to verify their identity.

Nevertheless, some variability is still found between ensemble segments of the same individuals. Some of it can be attributed to the normal intrasubject variability that each subject presents and, in the case of the driving signals, to their behaviour while driving, as some subjects present much better results than others.

The remaining variability reveals there is still some space to improve on the processes of denoising and signal preparation. However, the current quality of the ensemble segments appears to be sufficient to provide the needed conditions to achieve good biometric recognition performance.

8.4 Summary and Conclusions

This chapter focused on the preparation of the electrocardiographic signals for the processes of feature extraction and recognition, after their acquisition and preprocessing. The work presented addressed the existence of periods of sensor saturation in ECG acquired on the steering wheel, and the noise that was able to survive the preprocessing stage, that could influence negatively the process of recognition.

The combination of R-peak detection, heartbeat segmentation and amplitude normalisation, outlier detection and removal and, at last, heartbeat ensemble construction, proved to have a consistent and desirable performance in the cancellation of remaining noise and rejection of sensor saturation moments. Despite the difficulty associated with such noisy signals, the proposed approach in this imperative step of signal preparation proved able to enhance the quality of the extracted templates, and to significantly attenuate the variability caused by noise.

Future improvement seems very likely through the use of atlas heartbeat templates for enhanced outlier rejection in template sets with low amounts of relatively clean heartbeats. Despite this, the proposed method, together with the precedent step of preprocessing, provided the best conditions for the posterior tasks of feature extraction and recognition, discussed in the next chapters.

Chapter 9

Feature Extraction and Dimensionality Reduction

9.1 Problem Statement

In highly noisy signals, as those acquired in all off-the-person settings, holistic approaches are the most effective, as fiducial detection is difficult and unreliable. Extracting features through holistic approaches focuses, mainly, on the capture of important information that helps on the efficient discrimination between subjects (Li and Jain, 2009; Agrafioti et al., 2011).

To avoid losing important information, the feature sets extracted are usually large (composed by many features). However, some of these features may not contain information to discriminate between individuals, and can, thus, damage the recognition performance. Furthermore, if we consider continuous, time-sensitive systems, these may not be suitable for efficient biometric recognition. Thus, it is common for dimensionality reduction techniques to be applied, that reduce the number of features through selection or transformation, and ensures the smallest loss possible of discriminant information (Hejazi et al., 2016).

This chapter presents the process and results of the exploration of prior art feature extraction methods, accompanied by some own adjustments and improvements, and dimensionality reduction techniques, with the final goal of finding the most promising feature sets for the task of human recognition using ECG signals.

9.2 Feature Extraction

Five of the most promising feature sets from prior art research were selected to be implemented, adapted, and improved in this phase: autocorrelation coefficients, proposed by Plataniotis et al. (2006), Agrafioti and Hatzinakos (2008), Agrafioti et al. (2012), and Hejazi et al. (2016); cardioid plots, used by Sufi et al. (2011) and Iqbal et al. (2014); 1D Local Binary Patterns, recently proposed by Louis et al. (2016); Discrete Cosine Transform coefficients, explored by Tawfik et al. (2010); and Haar Wavelet Transform coefficients, as used by Dar et al. (2015a,b).

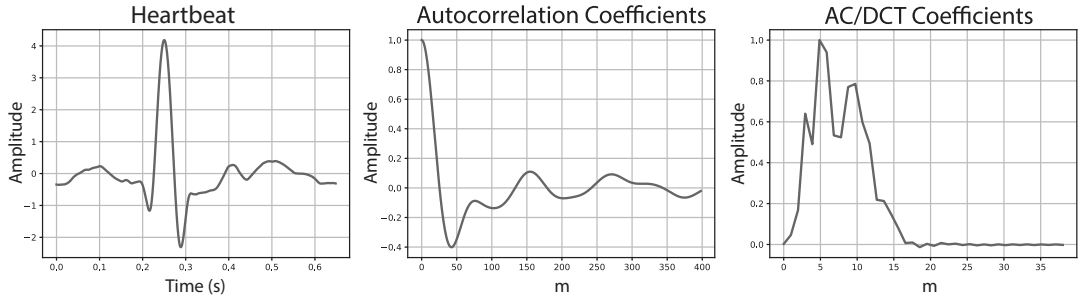

Figure 9.1: Autocorrelation and AC/DCT coefficients extracted from an example ensemble heartbeat (the example heartbeat was taken from the CardioID real driving database; m designates the index of the coefficient - cf. equation 9.1).

The ensemble heartbeats, and the QRS complexes cropped from those, without further feature extraction, were also explored as feature sets.

9.2.1 Autocorrelation Coefficients

The use of autocorrelation coefficients (AC) as features for ECG-based biometric recognition was first proposed by Plataniotis et al. (2006), and their work was afterwards complemented in Agrafioti and Hatzinakos (2008), Agrafioti et al. (2012), and, more recently, Hejazi et al. (2016). This approach has posed itself as one of the most reliable for individual recognition, and, in this dissertation, the aim is to assess if these capabilities are also aplicable to the highly noisy driving ECG signals.

Autocorrelation was initially proposed because, as it blends the temporal information into its coefficients, the generally difficult and dubious fiducial detection would not be required, and instead fixed-width windows could slide over the signal and render features just as able to discriminate between subjects.

As proposed by Plataniotis et al. (2006), considering a signal $x[n]$, the normalised autocorrelation coefficients can be computed through:

$$\hat{R}_{xx}[m] = \frac{\sum_{n=0}^{N-|m|-1} x[n]x[n+m]}{\hat{R}_{xx}[0]} \tag{9.1}$$

Essentially, a correlation is performed between the signal and $x[n+m]$, a shifted version of the signal, and the division by $\hat{R}_{xx}[0]$ (the expected maximum) causes the normalisation of the coefficients. From the coefficients obtained, only the k first were used as features ($k = 0.4f_s$, where f_s is the sampling frequency of the signal), due to their higher stability against noise (cf. Fig. 9.1).

Agrafioti and Hatzinakos (2008) proposed the reduction of this feature set using DCT, first proposed as dimensionality reduction approach by Ahmed et al. (1974), as well as LDA or, later by Hejazi et al. (2016), with PCA. The two latter methods are described later on.

DCT, also used for feature extraction and signal preprocessing, was already presented in subsection 7.2. Essentially, it represents a signal by a set of coefficients, that correspond to frequencies of cosine functions that compose the signal. The DCT reduction technique was applied to the autocorrelation coefficients and, among the DCT coefficients obtained, the first 10% were selected to constitute the new feature set (cf. Fig. 9.1).

9.2.2 Cardioid Plots

Cardioids are an alternative representation of signals that, as bidimensional plots, allow for the extraction of features commonly associated with image segmentation and recognition. Sufi et al. (2011) and Iqbal et al. (2014) proposed the use of this transformation of ECG heartbeats and QRS complexes as features for individual recognition.

Considering a heartbeat or QRS signal $x[n]$, the first step consists on its differentiation, to obtain:

$$y[n] = \{0, x[1] - x[0], ..., x[N] - x[N-1]\} \tag{9.2}$$

The cardioid can then be analysed by considering x and y as the abscissae and ordinates, respectively. As proposed by Iqbal et al. (2014), in this dissertation, the direct use of x and y as features was explored, both for heartbeats and QRS complexes.

Sufi et al. (2011) opted for the extraction of fiducial features from the obtained cardioid plots. A loop segregation procedure was used, after the computation of y, to ensure the cardioid plot begins and ends in the same location. From the loop, the centroid, top, bottom, left, and right extremas' coordinates were extracted, as well as its perimeter, and the area inside it.

An example of the cardioids obtained with ensemble heartbeats from the CardioID real driving database can be conferred in Fig. 9.2, as well as the locations of the fiducial features, also explored for biometric recognition in this dissertation.

9.2.3 DCT Coefficients

The Discrete Cosine Transform maps a time series into a set of cosine functions with different frequencies, thus representing the series by a set of coefficients (frequency bins) that describe its frequency content. DCT has been previously presented in this document, in subsection 7.2, as it has also been proposed for signal denoising by Choudhary and Manikandan (2015).

As proposed by Tawfik et al. (2010), the set of DCT coefficients (G_x) was obtained for each ensemble heartbeat on the datasets. While the researchers have only used the first twenty coefficients, for this dissertation all coefficients that correspond to the range $[0;40]$ Hz have been selected as features (cf. Fig 9.3). This allowed for the capture of the whole range of important frequencies, discarding the coefficients at higher frequencies that, generally, tend to zero.

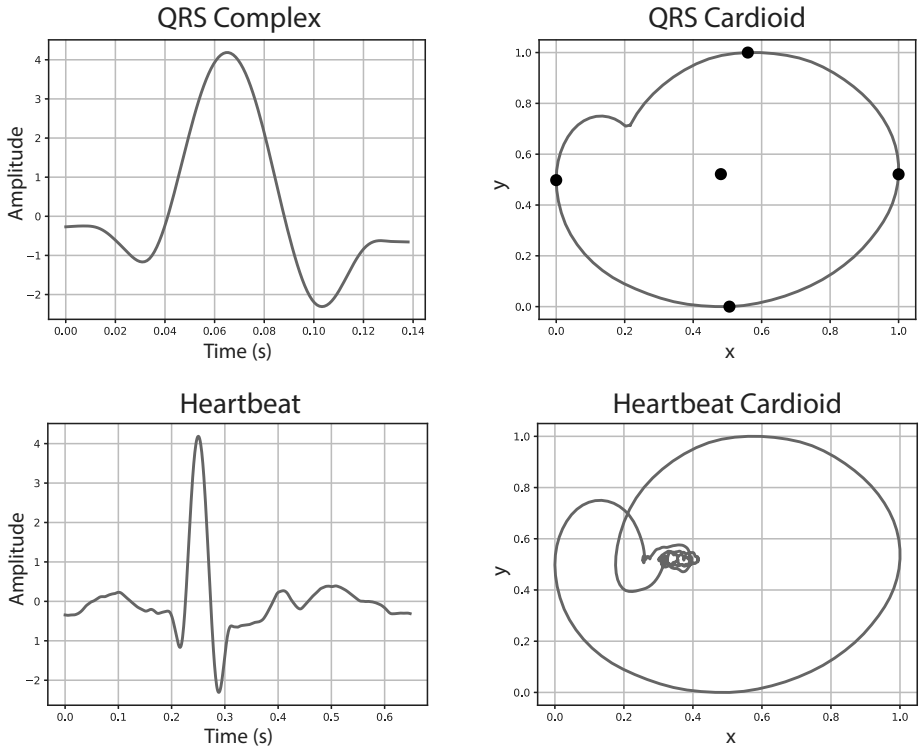

Figure 9.2: Cardioid plots and features from an example ensemble heartbeat (The example heartbeat was taken from the CardioID real driving database. Top row: QRS complex and cardioid plot, after loop segregation, along with the extremas and centroid (black dots); bottom row: heartbeat and cardioid).

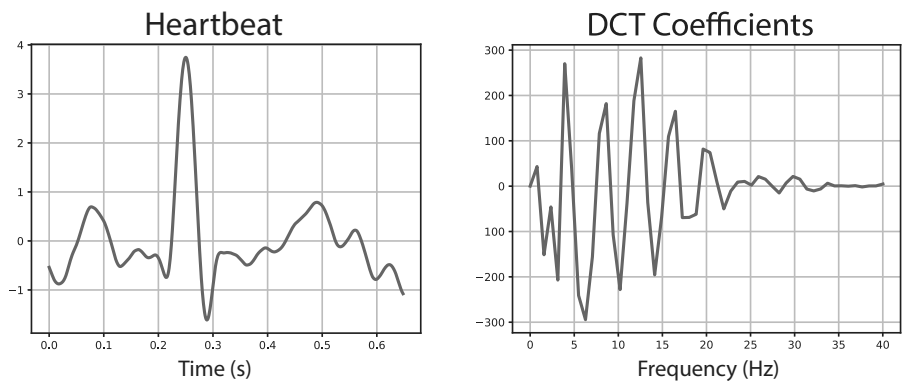

Figure 9.3: Selected DCT coefficients extracted from one example ensemble heartbeat (the heartbeat belongs to the CardioID coach bus driving database).

9.2.4 Fuzzy Tunable 1D Local Binary Patterns

Unidimensional LBP has been used for face recognition, by Benzaoui and Boukrouche (2013), and for arrhythmia detection, by Nikan et al. (2017). Louis et al. (2016) proposed an adaptation towards multi-resolution capabilities, for similar performance on signals acquired at different sampling frequencies, for the specific purpose of ECG-based biometric recognition.

Other researchers, such as Ren et al. (2015) and Gubbi et al. (2015) have proposed the use of Fuzzy Local Binary Patterns, where a tolerance range is defined around the reference value, and values inside that tolerance range were not attributed a score of 1 or 0, but instead a fuzzy value between these.

In this part of the dissertation, the goal was to merge the qualities of all these LBP approaches into a more flexible and tunable algorithm that allowed for the extraction of features at several scales, from signals acquired with different sampling frequencies, that could be as much resistant to signal noise as possible.

The proposed adaptation of 1D Local Binary Patterns focused on three main concerns:

1. *Equal results independently of the sampling frequency* - Following the prior art survey results, there is a high variety of sampling frequencies used for the acquisition of ECG signals and, minding the goal of wide applicability, this algorithm should be able to perform equally in any of these cases. Thus, independently of the original sampling frequency, the signal is resampled to a fixed sampling frequency, so that the results are equivalent;

2. *Noise resistance* - Because the signals acquired through off-the-person approaches are usually ridden with noise that prevails even after strong preprocessing procedures, this aims to avoid noise influence on the extracted features. While, generally, LBPs are based on a binary comparison of the value of a centre sample with other samples that surround it, in this approach, the comparison is between the mean value of a centre neighbourhood of samples and the mean values of other neighbourhoods of samples that surround it. Besides this, if the compared values are close enough, the use of fuzzy scores allows for the reduction of the influence of those small differences on the final pattern;

3. *High flexibility* - To adapt to various signals, and to accurately capture different characteristics on the extracted features, able to work towards diverse goals on pattern recognition, the method was built to be highly tunable, in what concerns the size of the centre neighbourhood, the neighbourhoods considered around it, the distances between them, and also the tolerance range for the fuzzy score attribution, and the number of levels for quantisation inside that range.

Hence, in each iteration, this LBP approach allows us to obtain a pattern vector P composed of $2p$ elements, where p is the number of neighbourhoods l considered on each side of the centre

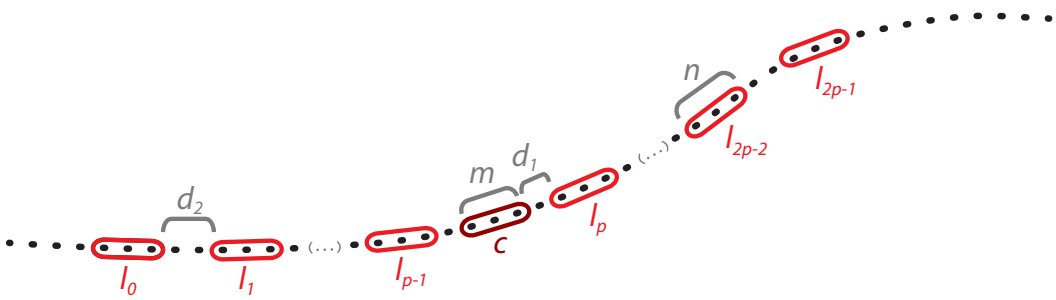

Figure 9.4: Illustration of one iteration of the application of the proposed Fuzzy Tunable 1D LBPs feature extraction algorithm (the signal samples are shown as black dots, the centre neighbourhood c is shown in dark red, and the neighbourhoods l in bright red).

neighbourhood c. An illustration of an iteration of this algorithm on an example signal can be conferred on Fig. 9.4. The first p elements of the P vector can be defined defined by:

$$P_i(x[c]) = f\left(\bar{l}_i - \bar{c}\right), i = 0, \ldots, p-1 \tag{9.3}$$

While the last p elements of P can be computed through:

$$P_{p+i}(x[c]) = f\left(\bar{l}_{p+i} - \bar{c}\right), i = 0, \ldots, p-1 \tag{9.4}$$

On these equations, $\bar{c}$, $\bar{l}_i$, and $\bar{l}_{p+i}$ denote, respectively, the mean values of the centre, anterior, and posterior neighbourhoods. $\bar{c_n}$ can be obtained with:

$$\bar{c} = \frac{1}{m}\sum_{k=0}^{m-1} x(b-(m-1)/2+k) \tag{9.5}$$

Where b denotes the centre sample index of the centre neighbourhood considered (note that m should be an odd integer). The anterior and posterior neighbourhoods' mean values (respectively, $\bar{l}_i$ and $\bar{l}_{p+i}$) can be computed through:

$$\bar{l}_i = \frac{1}{n}\sum_{k=0}^{n-1} x[s_{before}+k] \tag{9.6}$$

$$\bar{l}_{p+i} = \frac{1}{n}\sum_{k=0}^{n-1} x[s_{after}+k] \tag{9.7}$$

Where s_{before} and s_{after} denote the indexes of the first samples of each neighbourhood, that can be found with:

$$s_{before} = b + n(i-p) + d_2(i-p+1) - (m-1)/2 - d_1 \tag{9.8}$$

$$s_{after} = b + i(n+d_2) + (m-1)/2 + d_1 + 1 \tag{9.9}$$

Where d_1 is the distance (in samples) between the centre neighbourhood and the neighbourhoods closest to it, and d_2 is the distance between neighbourhoods. Still on the definition of P, the comparison between the centre neighbourhood and the others is performed by the function f, defined as:

$$f(a) = \begin{cases} 1 & \text{if } a \geqslant \varepsilon \\ 0.5\left(1.0 + \frac{round(x \times l)}{\varepsilon \times l}\right) & \text{if } \varepsilon > a > -\varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{9.10}$$

Where ε is the tolerance parameter for the fuzzy part of the algorithm, and l is the number of quantisation levels for the scores between 0 and 1, attributed when the difference falls on the range $]-\varepsilon;\varepsilon[$. It is simple to discard the fuzzy part of the algorithm, if desired, by setting $\varepsilon = 0$, thus getting binary pattern vectors.

Although it is possible to set desired locations (center sample indexes) for the patterns to be extracted, in this dissertation the algorithm worked by extracting LBPs from every sample of the signal that had enough samples before and after it. A step parameter s was also defined as a distance between center samples b between iterations.

To extract features from the available ECG signals, Fuzzy Tunable 1D LBP was used in two different modalities. The first contained a fuzzy range $\varepsilon = 2.0$ with $l = 10$, and the features consisted on a histogram h_s of the sums of all values of the LBP vectors. The second discarded the fuzzy capabilities, and returned a histogram h_p of the distribution of pattern vector's values.

9.2.5 Haar Wavelet Transform Coefficients

Like the Discrete Cosine Transform, the Discrete Wavelet Transform is also characterised by a wide applicability in signal and image processing and analysis. In the case of ECG-based biometrics, prior art approaches have used DWT for signal denoising, feature extraction, and even

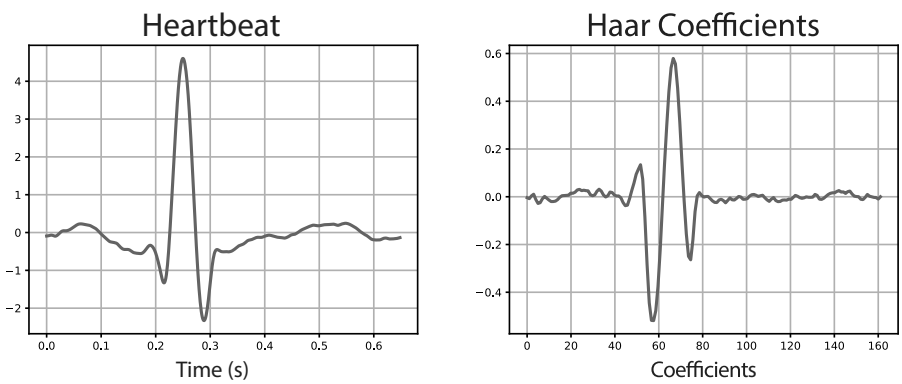

Figure 9.5: Illustration of the Haar Wavelet Transform coefficient features extracted from an example heartbeat (the ensemble heartbeat used was selected from the CardioID real driving database).

template matching for recognition.

As DWT was also explored, in this dissertation, for signal preprocessing, the method is described in detail in subsection 7.2. The single specificity of Haar Wavelet Transformed, used by Dar et al. (2015a,b), is that the wavelet function Ψ used on the filter bank of DWT is, specifically, the Haar wavelet.

Based on these settings proposed by Dar et al. (2015a,b), the detail coefficients output by the second level of decomposition were extracted to make up the feature set (cf. Fig. 9.5), as they have, empirically, proved to be the most promising for recognition.

9.3 Dimensionality Reduction

Two common dimensionality reduction algorithms were applied to the extracted feature sets: LDA (Agrafioti and Hatzinakos, 2008; Boumbarov et al., 2009; Matta et al., 2011; Agrafioti et al., 2012), and PCA (Hejazi et al., 2016), with the main goal of reducing the number of features on the sets, to also reduce the time required for recognition, and increase the performance of this posterior process.

9.3.1 Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) is a supervised algorithm used in pattern recognition, both for classification and dimensionality reduction (Hejazi et al., 2016). LDA is one of the most popular choices for dimensionality reduction in ECG-based biometrics.

Knowing a set of N samples x_n (with d features) with known class, LDA computes a linear tranformation w that maps the initial d-dimensional feature space into l dimensions (through $x_n^l = w^T x_n$, where $l < d$), and maximises discrimination between classes (Ye and Ji, 2009).

As detailed by Hejazi et al. (2016), w can be found through the maximisation of Fisher's ratio function:

$$J(w) = \underset{w}{\operatorname{argmax}} \frac{|w^T S_B w|}{|w^T S_W w|} \tag{9.11}$$

Here, S_B and S_W denote, respectively, interclass and intraclass variance matrices, computed through:

$$S_B = \frac{1}{N} \sum_{j=1}^{k} (\overline{x_j} - \overline{x})(\overline{x_j} - \overline{x})^T \tag{9.12}$$

$$S_W = \frac{1}{N} \sum_{j=1}^{k} \sum_{i=1}^{N_j} (x_{ij} - \overline{x_j})(x_{ij} - \overline{x_j})^T \tag{9.13}$$

Where x_{ij} denotes each of the N_j samples of class $j = 1, ..., k$, and $\overline{x_j}$ represents the mean sample of each class. Similarly, $\overline{x}$ represents the mean of the whole set of N training samples.

The problem can be simplified considering that w that maximises $J(w)$ verifies $(S_W^{-1})S_B w_i = \lambda_i w_i$. Hence, LDA consists in finding the l most significant eigenvectors of $(S_W^{-1})S_B$, that correspond to the largest eigenvalues.

Linear Discriminant Analysis offers the great advantage of being supervised, and thus project the data into a feature space that is specifically fitted for higher class discrimination (Martinez and Kak, 2001). However, this comes with the setbacks of requiring prior knowledge of the population that will use the biometric system, and considerably large sets of training data, to avoid overfitting while finding the transformation (Martinez and Kak, 2001; Matta et al., 2011).

9.3.2 Principal Component Analysis

Unlike LDA, Principal Component Analysis (PCA) is an unsupervised dimensionality reduction method and, although nowadays often thought of as inferior, they are less sensitive to the training set's size and do not require prior class knowledge (Martinez and Kak, 2001).

Similarly to LDA, PCA consists on an eigenvalue problem (Hejazi et al., 2016). The entire set of train samples $X = x_i, i = 1, ..., N$ is centered on the feature space by subtracting the mean $\bar{x}$. Then, a covariance matrix is computed, through:

$$S_{cov} = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})(x_i - \bar{x})^T \tag{9.14}$$

The vector w that allows the transformation of a new sample x^{new} to a lesser feature space through $y = w^T (x^{new} - \bar{x})$ can be found with:

$$\lambda_i w_i = S_{cov} w_i \tag{9.15}$$

The highest non-zero eigenvalues λ correspond to non-zero eigenvectors, that are used to define a linear transformation that projects the original feature space into one with fewer dimensions, that maximises the variance of the data.

9.4 Results and Discussion

The suitability of the feature set used as input for the recognition task at hand is paramount to achieve success. Average heartbeat ensembles, autocorrelation coefficients, cardioid plots, and 1D local binary patterns have proven, in prior art research works, their ability to help discriminate individuals. However, as circumstances change and we move towards less obtrusive off-the-person acquisition procedures, features may lose that ability.

Hence, the feature sets, extracted from the signals in the off-the-person databases of CardioID bus driving and UofTDB, had their class separability evaluated through a simple ten-fold cross-validation identification procedure. The use of the first database aims to evaluate the behaviour of the features on noisy settings, while the UofTDB dataset allows to assess their ability to discriminate individuals among large groups.

A Nearest Neighbour classifier and a Support Vector Machine were used for the decision, and the results are presented in Table 9.1. Quick parameter optimisations were performed in order to obtain approximately optimal results. The average time required for the extraction of each feature

Table 9.1: Feature sets separability evaluation: results of the conducted ten-fold cross-validation task (identification rate results are presented in percentage, NS - number of subjects, NF - number of features in the set).

		CardioID Db. (NS=6)			**UofTDB (NS=100)**		
Features	**Dim. Red.**	**NF**	**KNN**	**SVM**	**NF**	**KNN**	**SVM**
Ensemble Heartbeats	-	650	88.93	89.94	130	95.83	97.61
	LDA	5	92.02	92.70	15	96.99	97.98
	PCA	649	78.10	89.94	129	80.61	98.65
Ensemble Heartbeats	-	140	79.20	85.77	28	89.02	86.32
(Only QRS)	LDA	5	91.90	91.04	15	92.82	92.94
	PCA	139	85.03	90.74	27	90.43	94.05
Autocorrelation	-	400	84.36	84.17	80	89.57	93.68
Coefficients	DCT	40	80.06	83.37	8	89.69	89.75
	LDA	5	76.01	76.07	15	92.58	92.88
	PCA	399	59.45	76.75	79	85.15	92.45
Cardioid Plots	-	1300	86.56	91.78	1300	94.91	98.22
(using x and y,	LDA	5	88.59	88.34	15	97.79	98.53
entire beats)	PCA	1167	61.04	92.09	-	-	-
Cardioid Plots	-	280	83.01	87.36	280	70.42	76.65
(using x and y,	LDA	5	89.08	89.39	44	80.42	80.84
QRS only)	PCA	279	68.59	88.83	-	-	-
Cardioid Plots	-	8	67.55	73.99	8	48.07	52.99
(using area, perimeter,	LDA	5	68.71	71.10	8	46.02	52.66
centroid, and extremas)	PCA	7	70.49	72.09	7	42.20	48.81
DCT Coefficients	-	52	89.51	90.61	52	90.55	94.17
	LDA	5	90.12	89.75	52	89.84	94.20
	PCA	51	88.96	90.00	51	88.83	93.66
Fuzzy Tunable 1D LBP	-	40	77.30	77.85	40	60.99	69.16
(with fuzzy range)	LDA	5	79.26	78.83	38	60.69	70.81
	PCA	38	75.09	77.98	39	57.32	69.58
Fuzzy Tunable 1D LBP	-	256	75.64	76.13	256	63.29	76.24
(without fuzzy range)	LDA	5	75.28	73.37	99	66.01	77.80
	PCA	255	68.65	75.09	255	60.12	76.35
Haar Coefficients	-	163	90.74	90.86	33	88.10	90.99
	LDA	5	94.05	94.42	32	89.66	91.83
	PCA	162	81.94	92.45	32	88.43	91.62

set from one ensemble heartbeat was also determined, and such results can be observed in Fig. 9.6.

Independently of the dataset used, the four best feature sets were DCT coefficients, Haar coefficients, ensemble heartbeats, and cardioid plots (x and y, entire beats). While ensemble heartbeats and cardioids enabled the achievement of the best identification rates with the higher quality signals of UofTDB, the performance with driving signals proved to be much lower.

DCT and Haar coefficients, however, appeared to be much more resistant to the effects of noise, by revealing similarly high identification rates with both datasets. Overall, despite having far more subjects than the driving dataset, the UofTDB signals allowed for the achievement of

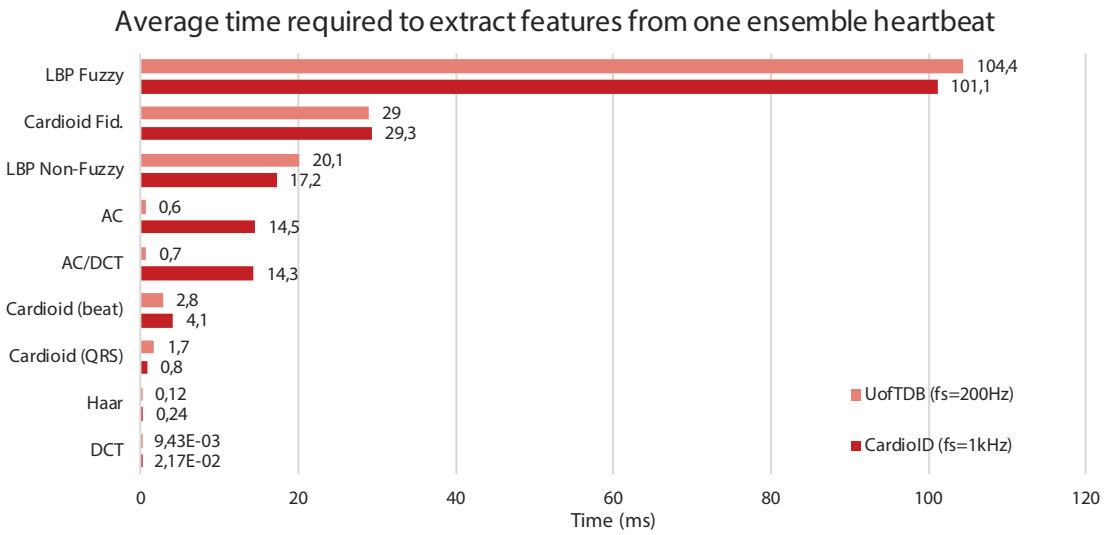

Figure 9.6: Average time required for the extraction of the explored feature sets from one ensemble heartbeat (results were obtained for CardioID signals, sampled at 1 kHz, and UofTDB signals, at 200 Hz, to assess the behaviour with different sampling frequencies).

higher identification rates. DCT and Haar coefficients have, also, shown much less time requirements for the feature extraction process.

Fuzzy Tunable 1D LBPs have proven to be the greatest underachievers, rendering identification rates more than ten percent below the best feature sets. The Autocorrelation coefficients, although widely used and praised in the prior art, and presenting good results with the UofTDB dataset, seemed to present a greater performance decline with the driving dataset, likely due to the greater influence of noise felt in those signals.

The dimensionality reduction methods, in general, presented little improvement to the identification rate results. Considering this and the extraordinary time requirements of such procedures, and minding the promising results of feature sets with relatively low number of features (DCT and Haar), the use of dimensionality reduction appears to be, in this situation, inefficient.

Hence, weighing the identification results with both datasets, the total number of features of each set, and the time required to extract such features from each ensemble heartbeat available, the most promising feature sets were DCT and Haar coefficients. Together with ensemble heartbeats, these were selected to be used in the final recognition process.

9.5 Summary and Conclusion

This chapter focused on the extraction of feature sets that could adequately capture the individual information contained in ECG signals, while resisting to the influence of noise remaining after the preprocessing and preparation stages. In the arduous task of continuously recognizing individuals from ECG signals acquired during travels, at the steering wheel, time, recognition performance, and resistance to noise, are equally important.

Following the results obtained with real driving signals and a large dataset of off-the-person signals from the UofTDB collection, the features extracted through the Discrete Cosine Transform and through Haar wavelet transform have proven to be the fastest and most promising.

Despite the adverse conditions created by the large number of subjects (in the UofTDB dataset) and the high noise influence (in the driving signals), these features enabled the achievement of consistent and promising identification results. Based on this, everything indicated that, with further exploration and optimisation of recogntion methods, detailed in the next chapter, even better results should be expected, and a complete and effective biometric recognition method should be obtained.

Chapter 10

Recognition

10.1 Problem Statement

Matching and decision compose the final step in the process of biometric recognition. After the enhancement of the signal's quality with the preprocessing stage, the preparation of the denoised trait's signal for use in the biometric system, and the collection of relevant information through the extraction of features, this step will, at last, allow the biometric system to output an identity (in the case of identification), to reject, or to accept an identity claim (in the case of authentication).

In the specific topic of continuous recognition of drivers, the main application resides on permanent remote fleet control and supervision. Knowing who is driving the vehicles at all times is a simple solution to ensure mandatory driver swaps take place (for long trips), or to establish where the responsibility lies in case of accident or reckless driving.

Authentication is especially fitted for fleets of buses, trucks, or taxis, where the claimed identities can be automatically retrieved from official timetables. Identification can be useful in fleets with no predefined schedules, or private automobiles, where the system will select among a set of enrolled identities, without the need for a claimed identity.

In this chapter, the authentication and identification performance results of several prior art and proposed algorithms is presented, using the most common metrics described in section 3.5. The methods were benchmarked, through validation with 70%-30% dataset split between train and test, and simulating a real, scarse data situation, where the train was performed only with the first thirty seconds of data from each user. For more information on the parameters of each optimised model, confer the Appendix A of this dissertation.

10.2 Selected Prior Art Approaches

10.2.1 Support Vector Machine

Support Vector Machines are binary classifiers, proposed by Cortes and Vapnik (1995), that operate by finding the optimal hyperplane, in the feature space, that maximises the margin that separates the two classes (Lin et al., 2014). The training samples closest to the boundary between

classes, designated as support vectors, are exclusively used for the computation of the hyperplane, undermining the influence of outliers.

This assumes the classes are linearly separable. However, this is generally not verified, even in ECG biometrics. To solve this problem, non-linear SVM classifiers use kernel functions that map the original feature space into an alternative space where the classes become linearly separable, and where an optimal hyperplane can be found (Lin et al., 2014).

SVM can also perform multiclass classification (as in identification), through a one-versus-one approach, by training $k(k-1)/2$ binary classifiers, where k is the number of classes. SVM has been used in many ECG biometrics prior art works, for identification and authentication tasks, such as Rezgui and Lachiri (2016), Hejazi et al. (2016), Lourenço et al. (2015), or Silva et al. (2013).

In this dissertation, SVM was used with Radial Basis Function (RBF) and polynomial kernels, and the parameters C (misclassification penalty parameter), γ (manages individual sample influence, for the RBF kernel), and degree (for the polynomial kernel) were optimised to obtain the best results possible.

10.2.2 Nearest Neighbours

The Nearest Neighbour Classifier (kNN for k-Nearest Neighbours), is one of the simplest methods, and also one of the most used in ECG-based recognition. Nearest Neighbours has been the option of several prior art works, such as those of Porée et al. (2016), Dar et al. (2015a), Brás and Pinho (2015), Matos et al. (2014), and Wang et al. (2013).

As described by Theodoridis and Koutroumbas (2009) and Murphy (2012), kNN is a non-parametric, non-linear classifier. Based on the location on the feature space of the object x to be classified, kNN will find the k nearest train samples, and analyse the labels of each. The object x will be predicted as having the class most frequently verified among the k labels found. Similarly, the class probabilities can be computed through the frequency of each class label among the k nearest train samples.

In this dissertation, Nearest Neighbour classifiers were explored, using Euclidean distance as the distance metric. The number of neighbours to be considered, k, was optimised based on the test set identification and authentication results.

10.2.3 Multilayer Perceptron

Multilayer Perceptron (MLP) is the common designation attributed to feed-forward neural networks (Bishop, 2006). MLP is one of the most commonly used neural networks, due to its simplicity, stability, and ability to work well with small training sets (Iqbal et al., 2014; Waili et al., 2016a).

Multilayer Perceptrons are composed of an input layer, an output layer, and a variable number of hidden layers between these. The layers are composed by neurons, that apply non-linear operations to their inputs (Bishop, 2006). Each neuron on the input layer receives a feature from the

feature set, while each output neuron emits a score for each class. Each hidden layer is composed by a variable and tunable number of neurons. A layer's neurons are fully interconnected with those of the immediately neighbouring layers.

The network training is performed through backpropagation, where the true scores are iteratively compared with the outputs, and the error is fed back through the network. The error is used to update the weights of the connections between neurons, until the predicted outputs converge to the true scores (Iqbal et al., 2014).

In this dissertation, Multilayer Perceptrons were explored for both authentication and identification problems. The number of hidden layers and the number of neurons in each were optimised in order to obtain the best results possible in each validation procedure.

10.2.4 Gaussian Mixture Model

Gaussian Mixture Models (GMM) model classes by fitting a set of K different, multi-dimensional normal distributions to the training samples. The probability density of a class will then be a linear combination of all those individual distributions (Bishop, 2006; Theodoridis and Koutroumbas, 2009):

$$p(x) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \Sigma_k) \tag{10.1}$$

Here, $k = 1, ..., K$ denotes each of the components, μ_k and Σ_k represent, respectively, the means and covariances of the components, and π_k are mixing coefficients used to define a weight for each component. The components are fitted to the training data through the Expectation Maximisation algorithm, that iteratively maximises the log-likelihood function estimates (Theodoridis and Koutroumbas, 2009).

Gaussian Mixture Models have been previously explored, for ECG-based biometrics, by Li and Narayanan (2010) and Matos et al. (2013). Likewise, GMM were used, along with GMM-UBM (described below), for identification and authentication procedures, in this dissertation. The optimal number of components varies with the data, so, this parameter was optimised for each validation procedure in order to obtain the best results possible.

10.2.5 Gaussian Mixture Model - Universal Background Model

Common approaches like the examples previously presented have been widely used in ECG biometrics, and generally present very good results. However, the problem with these approaches is that their performance is generally poor in situations with scarce training data.

In other scopes of the field of biometrics, especially when focusing on speaker and face detection, the problem of modelling the intrasubject variability of the traits while using reduced training data is of utmost importance. In these cases, an adaptation of Gaussian Mixture Models from Universal Background Models (designated as GMM-UBM) is generally highly regarded, and was successfuly used for speaker recognition by Reynolds et al. (2000) and Omar and Pelecanos (2010), and for gait-based biometrics by Neverova et al. (2016).

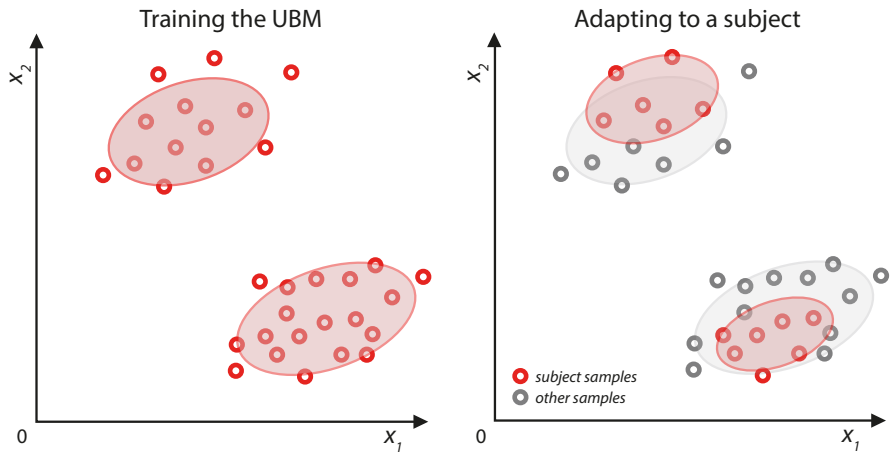

Figure 10.1: Illustration of the UBM training and adaptation to a subject of the GMM-UBM method, an example bidimensional feature space (actual results may differ, as they vary with the parameters used).

As proposed by Reynolds et al. (2000), GMM-UBM begins with the training of a GMM with all training data, regardless of the subjects. This general model is designated as the Universal Background Model, and is expected to better describe the intrasubject variability, due to the larger amount of data used.

The trained UBM is then adapted into a GMM for each subject (cf. Fig. 10.1). Each of the subject's train samples x_t will be evaluated in what concerns its alignment with the UBM's K mixture components (represented by k), following the equation:

$$Pr(k|x_t) = \frac{\pi_k p_k(x_t)}{\sum_{j=1}^{K} \pi_j p_j(x_t)} \tag{10.2}$$

Based on this and x_t, a set of statistics is computed:

$$n_k = \sum_{t=1}^{T} Pr(k|x_t) \tag{10.3}$$

$$E_k(x) = \frac{1}{n_k} \sum_{t=1}^{T} Pr(k|x_t) x_t \tag{10.4}$$

$$E_k(x^2) = \frac{1}{n_k} \sum_{t=1}^{T} Pr(k|x_t)(x_t)^2 \tag{10.5}$$

Here, x_t^2 represents the diagonal of the matrix resulting from $x_t x_t'$. The sufficient statistics of the UBM (π_k, μ_k, and σ_k^2: weights, means, and variances, respectively) are adapted to create a GMM for the subject, following the equations:

$$\hat{w}_k = [\alpha_k^w n_k / T + (1 - \alpha_k^w)\pi_k]\,\gamma \tag{10.6}$$

$$\hat{\mu}_k = \alpha_k^m E_k(x) + (1 - \alpha_k^m)\mu_k \tag{10.7}$$

$$\hat{\sigma}_k^2 = \alpha_k^v E_k(x^2) + (1 - \alpha_k^v)(\sigma_k^2 + \mu_k^2) - \hat{\mu}_k^2 \tag{10.8}$$

The scaling factor γ is used to ensure the sum of all mixture's weights is equal to one. A set parameter r^ρ (where $\rho \in \{w, m, v\}$) is used to compute α_k^ρ, a data-dependent adaptation coefficient for each mixture and parameter, through:

$$\alpha_k^\rho = \frac{n_k}{n_k + r^\rho} \tag{10.9}$$

The UBM (Θ_{UBM}) and the UBM-adapted GMM of each subject (Θ_s) are used for the identification and authentication tasks. The log-likelihood ratio for a test sample x, for each enrolled identity s is computed through:

$$\Lambda_s(x) = \log p(x|\Theta_s) - \log p(x|\Theta_{UBM}) \tag{10.10}$$

The log-likelihood ratios are thresholded to accept or reject an individual (in the case of authentication), or the maximum is selected to attribute an identity (in the case of identification).

On the specific case of ECG biometrics, this technique has not yet been explored in depth. Li and Narayanan (2010) were the first and, to date, the only researchers to use GMM-UBM for ECG-based recognition, using on-the-person MIT-BIH NSR signals, acquired on the chest of 18 subjects. These settings are, nowadays, unable to fit the expectations of current off-the-person settings (such as those on which this dissertation focuses). Hence, the goal is to evaluate its performance on these new, more adverse conditions.

10.3 Proposed Methods

10.3.1 Naïve Bayes

The aforementioned prior art method Gaussian Mixture Model shares many similarities to the Gaussian Naïve Bayes classifier. Both are generative models that fit Gaussian functions to a set of training samples to describe the distribution of each class: GMM uses a number of multidimensional Gaussian components, while Naïve Bayes describes each feature, separately, by a single, unidimensional function.

Naïve Bayes assumes independence between the features of each sample. Thus, given a set of train samples with D features, D independent normal distributions are fitted to the train samples for each class, through maximum likelihood estimation (Theodoridis and Koutroumbas, 2009; Murphy, 2012). Each class c conditional density can, then, be computed through the product of

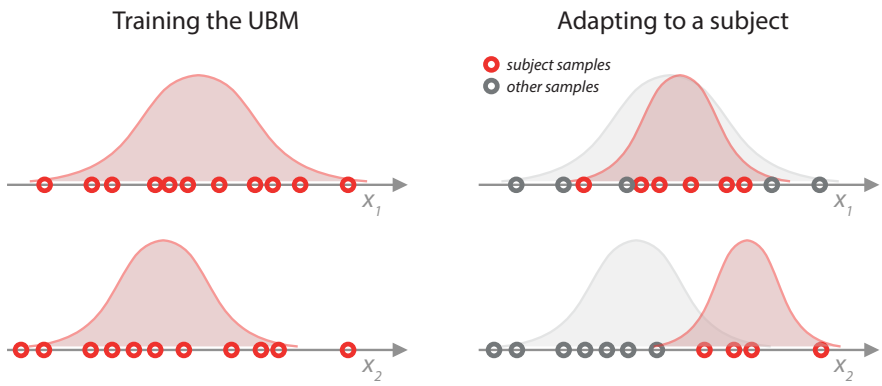

Figure 10.2: Illustration of the UBM training and adaptation to a subject of the Naïve Bayes UBM method, on two dimensions of an example feature space (actual results may differ, as they vary with the parameters used).

the D individual, one-dimensional densities:

$$p(x|y=c) = \prod_{j=1}^{D} \mathcal{N}(x_j|\mu_{jc}, \sigma_{jc}^2) \tag{10.11}$$

After training, class predictions can be performed by finding the class that maximises $p(y|x)$, through the Naïve Bayes formula:

$$p(y|x) = \frac{p(y)p(x|y)}{p(x)} \tag{10.12}$$

Considering that $p(x)$ is constant, the class of object x can be predicted through:

$$\hat{y} = \underset{y}{\operatorname{argmax}} p(y)p(x|y) \tag{10.13}$$

In this dissertation, to allow the use of thresholds in authentication procedures, the decisions were based on the probabilities $p(y|x)$ output by the trained models. To avoid the effect of imbalanced classes in the train sets, the priors $p(y)$ were set as equal for all identities.

10.3.2 Naïve Bayes - Universal Background Model

Given that only one UBM adaptation was already proposed in the prior art, GMM-UBM, It seemed reasonable to explore an UBM adaptation other than GMM-UBM, to allow a more diverse assessment of UBM capabilities in these settings. Given that Gaussian Naïve Bayes was already selected to be explored in this dissertation, the most profitable option would be to also use a Naïve Bayes Universal Background Model adaptation.

To the best of our knowledge, such an adaptation has never been proposed in the literature. Thus, the Naïve Bayes UBM classifier is here formulated and proposed, based on the GMM-UBM approach by Reynolds et al. (2000).

First, a Gaussian Naïve Bayes UBM is trained with all training samples, as if they belonged to the same subject (cf. Fig. 10.2). Then, based on that UBM, for each of the N subjects, using the set of their T training samples (x_t), the alignment of each with the UBM Gaussian distributions is computed through:

$$Pr(i|x_i^t) = e^{-\frac{(x_i^t-\mu_i)^2}{2\sigma_i^2}} \tag{10.14}$$

As Naïve Bayes uses an independent, unidimensional Gaussian function to model each feature, i denotes each feature, x_i^t represents the x_t sample's value of that feature, and μ_i and σ_i^2 represent, respectively, the mean and variance of the UBM's i-th feature Gaussian.

Afterwards, the process becomes very similar to GMM/UBM. $Pr(i|x_i^t)$ and x_t are used to compute the sufficient statistics:

$$n_i = \sum_{t=1}^{T} Pr(i|x_i^t) \tag{10.15}$$

$$E_i(x) = \frac{1}{n_i}\sum_{t=1}^{T} Pr(i|x_i^t)x_i^t \tag{10.16}$$

$$E_i(x^2) = \frac{1}{n_i}\sum_{t=1}^{T} Pr(i|x_i^t)(x_i^t)^2 \tag{10.17}$$

The new means and variances of the feature Gaussians ($\hat{\mu}_i$ and $\hat{\sigma}_i^2$, respectively) are determined according to the equations:

$$\hat{\mu}_i = \alpha_i^m E_i(x) + (1-\alpha_i^m)\mu_i \tag{10.18}$$

$$\hat{\sigma}_i^2 = \alpha_i^v E_i(x^2) + (1-\alpha_i^v)(\sigma_i^2+\mu_i^2) - \hat{\mu}_i^2 \tag{10.19}$$

As with GMM/UBM, α_i^ρ are adaptation coefficients, dependent on predefined r^ρ parameters, that control the weight of the subject's training samples over the UBM when adjusting the means and variances. They can be computed through:

$$\alpha_i^\rho = \frac{n_i}{n_i + r^\rho} \tag{10.20}$$

A prediction can be obtained by computing the posterior probability of the test sample belonging to each of the classes, using the Naïve Bayes models adapted from the UBM. For the purpose of authentication, the probability is compared with a threshold (following equation 10.12) or, for identification, the class that presents the highest posterior probability is chosen (as in equation 10.13).

10.4 Additional Recognition Procedures

10.4.1 User-Tuned Authentication

As previously detailed in subsection 3.4.2, biometric authentication systems work by measuring a trait of the person being recognised, and comparing it to stored templates of the claimed identity. If this comparison shows a high similarity or probability of the subject having the claimed identity, it is accepted, otherwise the claim is rejected.

The acceptance or rejection of an identity claim based on the comparison of acquisition and stored templates relies on a threshold. Generally, validation procedures are conducted for all subjects available, where false acceptance and false rejection rates are verified for several different thresholds, and a single, overall threshold that minimises both FAR and FRR is selected, to be used in all decisions made by the system.

However, following the concepts of biometric menagerie (cf. subsection 3.5.3), it is unwise to assume the fitness of similar settings for different individuals. Some subjects may be more similar to *chameleons*, and higher thresholds are more advisable to avoid false acceptances. Others, as *phantoms*, may be more prone to false rejections, and thus require lower thresholds.

Hence, in this dissertation, inspired on these concepts defined by the biometric menagerie, the idea of user-tuned authentication was explored. Based on the training data available, individual receiver operating characteristic (ROC) curves were analysed for each enrolled individual, and a bespoke threshold was defined. This approach is expected to not only improve overall performance, but also to ensure the equilibrium between security and convenience at all times.

10.4.2 Past Score Weighting

As emphasised throughout this dissertation, the existence of outliers is a very serious issue than can threaten the reliability of the recognition process. The very frequent noise contamination, losses of contact with the sensors, or impedance variations and sensor saturation are the most common origins of outliers in the case of ECG signals acquired on the steering wheel.

Although many of these are rejected by the outlier detection process on the signal preparation phase (cf. subsection 8.2.4), it is expected that noise contamination will never be completely eliminated and, in extreme cases, outlier detection and rejection could fail completely. Hence, outliers should be expected to occur, occasionally, in continuous acquisitions.

To avoid false decisions of acceptance, rejection, or identification, during continuous recognition processes, the probability scores output by the recognition methods can be postprocessed in order to account for recent past scores. This way, if the current sample is an outlier, giving past samples an influence in the current score is expected to avoid a false decision.

In this dissertation the idea of past score weighting was explored. Let $p_i(x_t)$ be the probability of the sample x_t (the most recently available feature vector) belonging to the class i (in the case of authentication, it denotes the claimed identity; for identification, it is each of the enrolled

Figure 10.3: Illustration of the past score weights relative to the samples' distance to the current sample (actual results may differ, as they vary with the parameters used).

identities), as output by the recognition methods. Moreover, x_{t-n} denotes each of the previous N samples considered.

Postprocessed scores $p_i^w(x_t)$ could, then, be obtained following the equation:

$$p_i^w(x_t) = \frac{1}{\sum_{n=0}^{N} w_{t-n}} \sum_{n=0}^{N} w_{t-n} p_i(x_{t-n}) \tag{10.21}$$

The probability scores of the current and past samples are, in 10.21, weighed by w_{t-n}, in order to control their influence on the final postprocessed scores. It is reasonable to guess that the current sample is more likely to belong to the same individual as the sample available one second ago, than the sample acquired one minute ago.

As illustrated by Fig. 10.3, the weights work towards the attenuation of past sample influence according to their distance to the current moment. A half-Gaussian function was used to compute those weights, having t_0 as the time of the current sample, $\mu = t_0$, and t_n as the time of each of the past samples, following:

$$w_{t-n} = e^{-\frac{(t_n - t_0)^2}{2\sigma^2}} \tag{10.22}$$

The parameter σ was set to 30s, as it was experimentally found that this value renders the best results. Likewise, the number of past samples to consider, N, was set to 15. However, the optimal values should be expected to vary for different settings.

10.5 Results and Discussion

10.5.1 Conventional Performance Assessment and Benchmarking

10.5.1.1 Authentication

To compare the performance of the devised approaches with the prior art authentication approaches, authentication validation was performed, splitting the datasets in 70% train and 30%

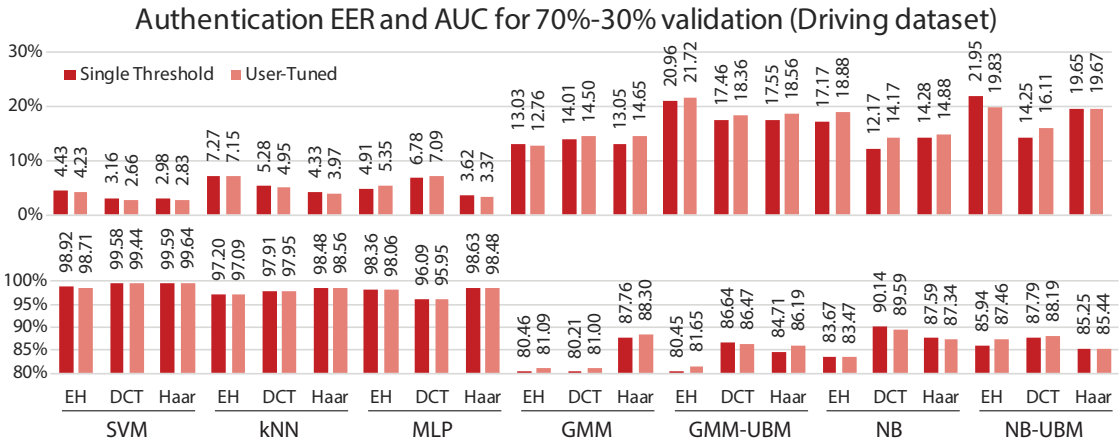

Figure 10.4: Authentication EER and AUC results, for 70%-30% split on the driving dataset.

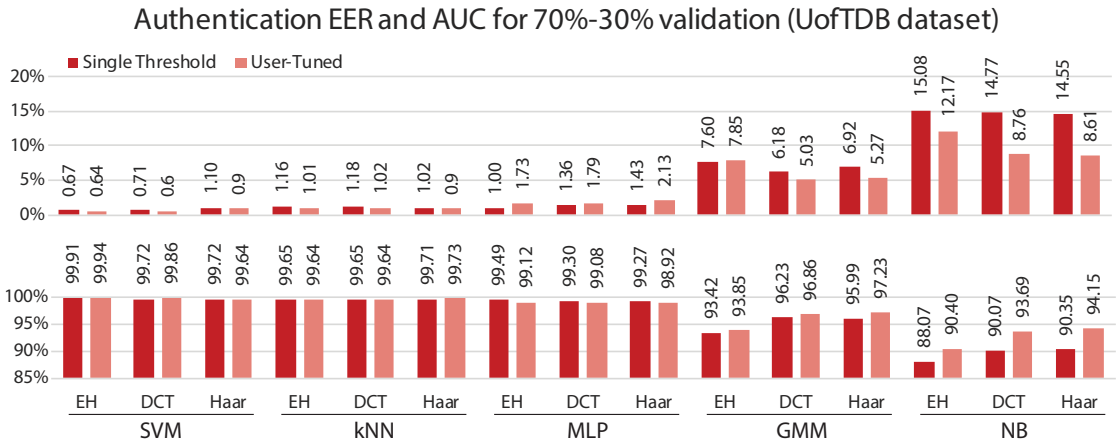

Figure 10.5: Authentication EER and AUC results, for 70%-30% split on the UofTDB dataset.

test. The datasets used were the driving dataset, described in detail in section 6.3, and a one-hundred-subject subset from the UofTDB collection.

The results, after parameter optimisation of the recognition methods, are presented in Fig. 10.4. According to both EER and AUC results, SVM, kNN, and MLP are the best options. The lowest equal error rate result, 2.66%, was obtained using SVM on DCT features, with user-tuned thresholds.

As for Gaussian Mixture and Naïve Bayes, the results show that they are not as fitting for these circumstances as the aforementioned best methods. Their combination with Universal Background Models was, as well, disappointing, likely because of the large size of the training set used.

Despite the different settings, the results obtained with the driving dataset are comparable with some of the best recent prior art results, such as those by Hejazi et al. (2016) (3.5% FAR, 4.83% FRR), Chun (2016) (2.4% EER), Lourenço et al. (2014) (2.5% EER), or (Labati et al., 2014) (5.36% EER).

Similarly, with the UofTDB dataset (cf. Fig. 10.5), the best results were obtained with SVM with DCT features, reaching an EER of 0.60% and AUC of 99.86%. Overall, SVM rendered the best results, followed closely by kNN, while GMM and Naïve Bayes greatly underperformed, hinting that these methods may possibly be inadequate for this task.

The results with the UofTDB dataset surpass those of many prior art approaches, and are comparable with the performance of the methods by Dar et al. (2015b) (0% EER with MIT NSR signals, and 4.1% FAR and 0.1% FRR with MIT Arrhythmia), Silva et al. (2013) (0.99% EER), Coutinho et al. (2013) (0.13% EER with PTB signals, and 0.29% EER with a private dataset), and Belgacem et al. (2013) (0.63% FAR and 0.66% FRR).

In most cases, the use of user-tuned thresholds, instead of a single threshold for all subjects, slightly improved the results on the driving dataset, reducing EER in 2.73%, in average. The same situation was verified with the UofTDB dataset, despite the higher quality of the signals and the larger population, with user tuned authentication reducing the EER in 1.09%.

10.5.1.2 Identification

The algorithms were tested in identification tasks, with the driving and UofTDB datasets, using a random split of data, 70% for train and 30% to be used as the test set. The results obtained can be analysed in Figures 10.6 and 10.7.

Identically to the authentication results previously presented, the best identification results were obtained with SVM and kNN, for both the driving and the UofTDB datasets. SVM with Haar transform features offered the best result, with 94.89% IDR. Although the settings are unique, this result is consistent with many recent prior art results, like those of Waili et al. (2016a) (96% IDR), Dar et al. (2015a) (95.85% IDR), or Dar et al. (2015b) (93.1% IDR).

With the data from the UofTDB collection, the tendency was, again, observed, with SVM with DCT features rendering the highest identification rate at 97.23%. Independently of the feature set, SVM presented the best results, followed by kNN and MLP, in all cases offering more than 95% IDR. Once again, these results are comparable to the best recent prior art approaches, such as Rezgui and Lachiri (2016) (98.8% IDR), Jahiruzzaman and Hossain (2015) (96.7% IDR), Pathoumvanh et al. (2014) (97% IDR), and Iqbal et al. (2014) (96.4% IDR).

Like authentication, GMM and Naïve Bayes, as well as their UBM adaptations, could not reach the high identification performance levels of the other models in validation with 70%-30% dataset division.

10.5.2 Continuous Performance Assessment

To evaluate the approaches in a way that would resemble more closely the settings of continuous use, only the first thirty seconds of each subject (26 samples, considering segments of 5 s with 1 s steps) were used, to simulate a very short enrollment process.

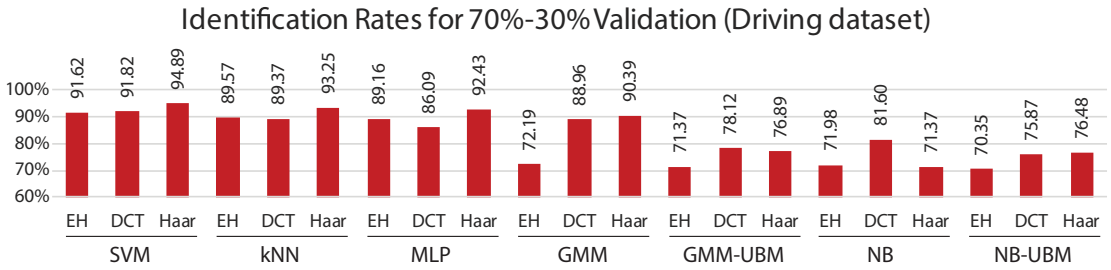

Figure 10.6: Identification rate results, for 70%-30% split on the driving dataset.

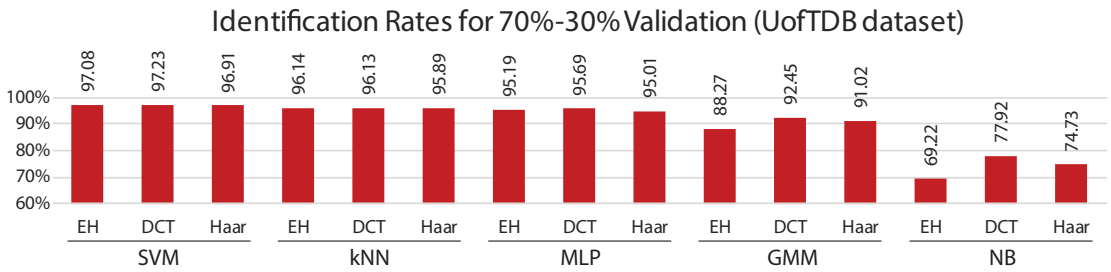

Figure 10.7: Identification rate results, for 70%-30% split on the UofTDB dataset.

The dataset extracted from the CardioID SteeringWheel v1 record (cf. section 6.3) was used, along with the timestamps of the samples extracted, to evaluate the continuous recognition performance of the explored methods.

Although template update is often used in the prior art, it was not used in this situation but is, nevertheless, encouraged for future developments and expected to bring significant performance improvements.

10.5.2.1 Authentication

The authentication results with and without past score weighting are presented, respectively, in Fig. 10.8 and Fig. 10.9. The best EER result was obtained using a Gaussian Mixture Model based on Ensemble Heartbeats, with 11.47% and an AUC of 91.05%. This was followed closely by an SVM with DCT features, that rendered an EER of 11.77% and AUC of 92.02%.

Overall, Universal Background Models managed to improve the results of GMM and Naïve Bayes without Past Score Weighting. However, when this technique was applied, GMM and Naïve Bayes manage to surpass the results of the respective combinations with UBM.

User-tuned thresholds led to modest improvements on most of the methods, especially those that already presented the best results. This technique, in average, reduced the EER by 2.73%, but only increased AUC by 0.12%. The application of past score weighting in authentication had a much better effect than user-tuned thresholds (cf. Fig. 10.10), improving the results in all methods

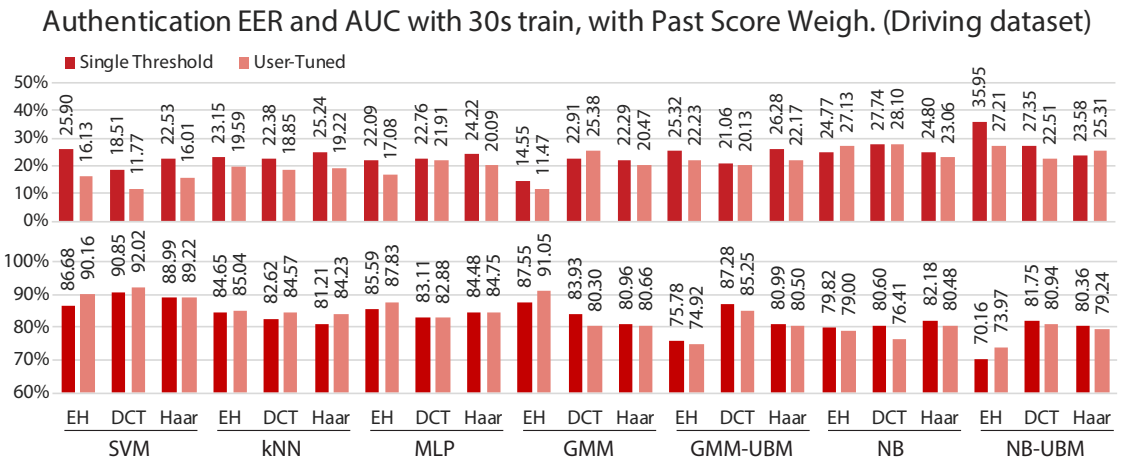

Figure 10.8: Authentication EER and AUC with 30s train sets, when using posterior past score weighting (with the driving dataset from CardioID).

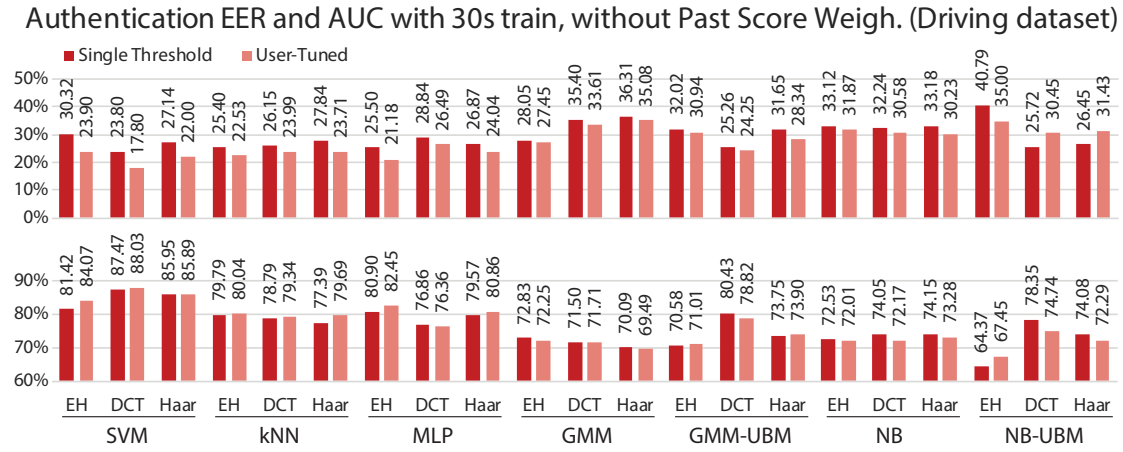

Figure 10.9: Authentication EER and AUC with 30s train sets, without posterior past score weighting (with the driving dataset from CardioID).

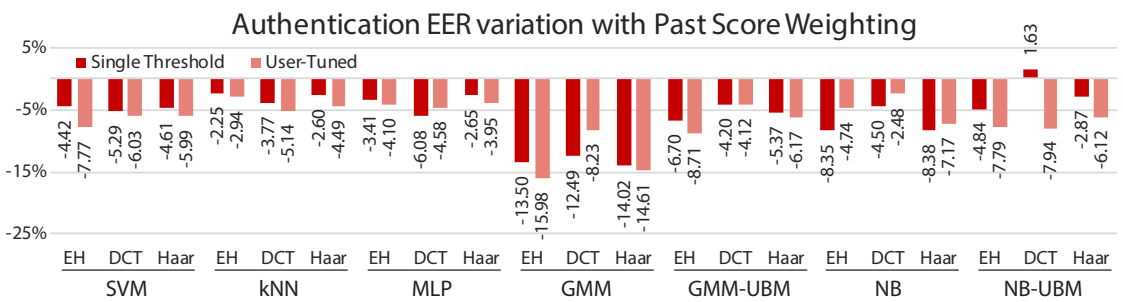

Figure 10.10: Improvement brought by past score weighting to the authentication EER results (using 30 s train sets, from the driving dataset from CardioID).

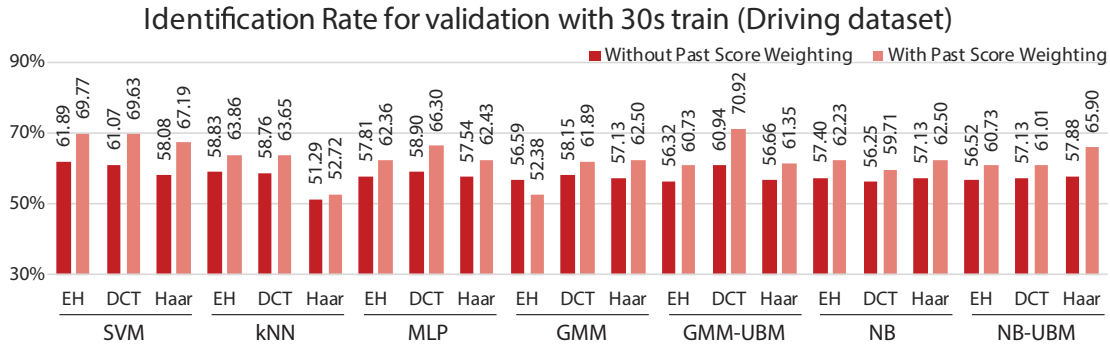

Figure 10.11: Identification rate results, with 30 s train sets, on the driving dataset.

but one. In average, past score weighting reduced the EER by 6.14% and increased the AUC by 6.48%.

The average combined improvement brought by the use of user-tuned thresholds and past score weighting was a reduction of 8.87% EER and an increase of 6.60% AUC, improving every method that has been explored.

Although these results correspond to a continuous recognition setting, not frequent in the prior art, the best result obtained (11.47% EER) was similar to state-of-the-art approaches focused on continuous biometrics. Louis et al. (2016) (7.89% EER), Matos et al. (2014) (14.0% EER), and Lourenço et al. (2012a) (9.39% EER) have reached similar results. With future improvements, such as the implementation of frequent template updates, the results are likely to surpass these and reach the performance level of conventional, non-continuous approaches.

10.5.2.2 Identification

The results of the limited train validation procedure, with the various methods described before, are presented in Fig. 10.11. These were, mostly, coherent with other results obtained and already discussed.

The results were, expectedly, much worse than the results of validation with larger train sets. The best result was obtained using GMM/UBM with DCT features and past score weighting, rendering an IDR of 70.92%. Despite the poor performance verified with GMM and GMM/UBM in the previously presented results, this proved that, in scarse train data settings, UBM can, indeed, be advantageous.

In fact, Naïve Bayes, its UBM adaptation, and GMM, all presented results surprisingly comparable to those of kNN and MLP. SVM was, still, one of the best methods, independently of the feature set, presenting IDR between 67% and 70%. Finally, past score weighting was, once again, a successful technique, improving the large majority of results, leading to the increase of IDR, in average, by 5.12%.

10.6 Summary and Conclusions

Throughout this chapter, the decision approaches used for recognition of individuals were described in detail, and their performance in the tasks of identification and authentication was measured.

Based on the results obtained, the method that presented the best and most consistent results was the Support Vector Machine with DCT features and user-tuned thresholds (for authentication). This approach rendered 2.66% EER and 91.82% IDR with the CardioID dataset, and 0.60% EER and 97.23% IDR with the UofTDB dataset, comparable with the recent successful prior art studies.

Identification and authentication performance seemed to decay more with lower quality of the signals (with the driving dataset) than with the size of the population (with the UofTDB dataset). Hence, should we consider a system working with driving signals and a large population, the results should be expected to be worse than those presented here, but not much.

In continuous settings, like those simulated here with the thirty-second train sets, the introduction of past score weighting and user-tuned thresholds brought significant improvements to the results. Should the system developers have large sets of data before deployment, GMM-UBM or NB-UBM could, likely, be an effective alternative to SVM, especially for identification tasks.

Furthermore, with the introduction of template update, proven successful by many prior art studies, the performance of a continuous recognition system is expected to be similar to some of the best prior art approaches, especially for authentication tasks.

Chapter 11

Conclusion

11.1 Summary and Final Remarks

This dissertation aimed to study the application of signal processing and machine learning algorithms to allow the biometrical recognition of individuals using ECG signals acquired on the steering wheel whilst driving. To achieve this goal, several prior art approaches, as well as novel techniques, were extensively explored for every module of a biometric system, and individually evaluated and compared, to reach the best results possible.

Beginning with preprocessing, this dissertation showed that, unlike medical-grade ECG signals, that can be properly denoised with simple bandpass filters, standard off-the-person signals, such as those from UofTDB, benefit more from DCT-based preprocessing. Driving signals acquired from the steering wheel, however, were better served with Savitzky-Golay combined with a Moving Average filter, likely due to requiring less assumptions on the noise frequencies.

As for signal preparation, Engelse-Zeelenberg proved to be better for R-peak detection in the cleaner signals of UofTDB, but the Trahanias method showed better results in both detection of true R-peaks and rejection of false peaks. Furthermore, due to the frequent sensor saturation periods that cause significant distortions on the signal, the need for outlier detection and removal in this kind of signals was evident. For this purpose, a novel clustering method based on Normalised Cross-Correlation was devised, and presented significantly better results than DMEAN, a prior art reference.

Feature extraction unveiled some surprises. Highly regarded feature sets such as the autocorrelation coefficients, cardioid plots and, recently, local binary patterns, appeared to be too reliant on the quality of the signals. Discrete Cosine and Haar transforms enabled the best and most consistent results in both the highly noisy driving signals, and the large population of the dataset from UofTDB. Dimensionality reduction with LDA or PCA brought no significant improvements to the results, especially considering the extra time required to process such transformations.

Finally, recognition brought promising results for authentication and identification. Using 70% of the datasets for train, the results for both recognition modes were comparable to those of some of the best prior art research works, even considering the more challenging settings. Using highly

limited train data, the results, while significantly worse, prove that the proposed approaches could, with some future improvements, be used to recognise individuals continuously, using steering wheel ECG signals, with performance comparable to the prior art.

11.2 Future Work

The promising results obtained with the approaches studied throughout this dissertation prove it is, indeed, possible to use highly noisy ECG signals, acquired on the steering wheel whilst driving, to accurately identify or authenticate an individual. However, based on these results and the surveyed prior art approaches, some paths towards improvement remain unexplored, and should be considered in future research or development initiatives.

One key future improvement pertains to the outlier detection and removal. Steering wheel signals are much more challenging than standard off-the-person signals used in the prior art, and the trends show that, with growing focus on the usability and unobtrusiveness of biometric systems, signals should become more and more challenging to work with. While the proposed Normalised Cross-Correlation Clustering approach manages to overcome the greater influence of outliers in the rejection process, future approaches should focus on the inclusion of clean reference heartbeat segments as atlas, to allow for better and more flexible outlier detection.

On the module of feature extraction, other prior art and novel features should be explored, especially those in the frequency domain, more similar to DCT and Haar. The combination of these and other feature sets should also be pondered, as it could enhance the performance of the system. For recognition, the idea of Universal Background Models could be further explored, especially based on other generative models. Likewise, Markov models and recurrent neural networks could, as sequential models, bring performance improvements, if explored in depth.

Finally, the system should be deployed to an embedded environment, to serve its purpose of recognising individuals, continuously, whilst driving. Here, based on the performance results, the implementation of template update techniques would be paramount to the acceptable performance of the system. Then, the performance should be evaluated, based on the continuous metrics presented in this dissertation, to assess the true continuous recognition capabilities of the system.

Appendix A

Decision Methods' Optimisation

Table A.1: Optimised parameters of the various decision algorithms, for authentication using the driving dataset.

Model	**Feat.**	**70%-30%**	**30s train**
SVM	EH	RBF, $C = 0.1$, $\gamma = 0.1$	RBF, $C = 0.1$, $\gamma = 0.1$
	DCT	RBF, $C = 1.0$, $\gamma = 1.0$	RBF, $C = 1.0\text{E}{-2}$, $\gamma = 0.1$
	Haar	RBF, $C = 1.0\text{E}2, \gamma = 1.0$	RBF, $C = 1.0$, $\gamma = 1.0$
kNN	EH	$k = 5$	$k = 10$
	DCT	$k = 5$	$k = 10$
	Haar	$k = 5$	$k = 15$
MLP[1]	EH	1 hidden layer	2 hidden layers
	DCT	3 hidden layers	3 hidden layers
	Haar	3 hidden layers	2 hidden layers
GMM[2]	EH	$K = 2$	$K = 2$
	DCT	$K = 2$	$K = 2$
	Haar	$K = 2$	$K = 2$
GMM-UBM[3]	EH	$K = 2$, $r = 0.1$	$K = 2$, $r = 0.1$
	DCT	$K = 2$, $r = 0.1$	$K = 2$, $r = 1.0\text{E}{-3}$
	Haar	$K = 2$, $r = 1.0$	$K = 2$, $r = 0.1$
NB-UBM[4]	EH	$r = 5.0$	$r = 2.0$
	DCT	$r = 20.0$	$r = 2.0$
	Haar	$r = 5.0$	$r = 2.0$

[1]With MLP, each hidden layer had 100 neurons, the activation function was the rectifier linear unit, and adaptive learing rates were used.

[2]With GMM, in all cases, diagonal covariance matrices were used for each component.

[3]With GMM-UBM, in all cases, $r = r^w = r^m = r^v$. Diagonal covariance matrices were used for each component.

[4]With NB-UBM, in all cases, $r = r^m = r^v$.

Table A.2: Optimised parameters of the various decision algorithms, for identification using the driving dataset.

Model	**Feat.**	**70%-30%**	**30s train**
SVM	EH	RBF, $C = 1.0$, $\gamma = 1.0\text{E}{-2}$	RBF, $C = 10.0$, $\gamma = 1.0\text{E}{-2}$
	DCT	RBF, $C = 1.0\text{E}3$, $\gamma = 1.0\text{E}{-5}$	RBF, $C = 1.0$, $\gamma = 1.0\text{E}{-5}$
	Haar	RBF, $C = 1.0\text{E}3$, $\gamma = 1.0$	RBF, $C = 1.0\text{E}3$, $\gamma = 1.0$
kNN	EH	$k = 1$	$k = 1$
	DCT	$k = 1$	$k = 1$
	Haar	$k = 1$	$k = 15$
MLP[1]	EH	2 hidden layers	3 hidden layers
	DCT	6 hidden layers	5 hidden layers
	Haar	4 hidden layers	5 hidden layers
GMM[2]	EH	$K = 10$	$K = 1$
	DCT	$K = 10$	$K = 7$
	Haar	$K = 10$	$K = 1$
GMM-UBM[3]	EH	$K = 10$, $r = 0.1$	$K = 10$, $r = 1.0$
	DCT	$K = 10$, $r = 0.1$	$K = 9$, $r = 1.0$
	Haar	$K = 10$, $r = 0.1$	$K = 7$, $r = 1.0$
NB-UBM[4]	EH	$r = 0.1$	$r = 0.1$
	DCT	$r = 0.1$	$r = 0.1$
	Haar	$r = 0.1$	$r = 0.1$

Table A.3: Optimised parameters of the various decision algorithms, for identification and authentication, using the UofTDB dataset.

Model	**Feat.**	**Authentication**	**Identification**
SVM	EH	RBF, $C = 10.0$, $\gamma = 0.1$	RBF, $C = 1.0\text{E}8$, $\gamma = 1.0\text{E}{-2}$
	DCT	RBF, $C = 100.0$, $\gamma = 1.0\text{E}{-4}$	RBF, $C = 1.0\text{E}8$, $\gamma = 1.0\text{E}{-4}$
	Haar	RBF, $C = 100.0$, $\gamma = 0.1$	RBF, $C = 1.0\text{E}8$, $\gamma = 1.0$
kNN	EH	$k = 30$	$k = 1$
	DCT	$k = 30$	$k = 1$
	Haar	$k = 30$	$k = 1$
MLP[1]	EH	3 hidden layers	4 hidden layers
	DCT	4 hidden layers	2 hidden layers
	Haar	4 hidden layers	3 hidden layers
GMM[2]	EH	$K = 2$	$K = 6$
	DCT	$K = 2$	$K = 6$
	Haar	$K = 2$	$K = 6$

Appendix B

Prior Art Overview

In this appendix chapter, a full summary of the techniques and methods applied for each step of each of the sixty-five research publications surveyed is presented (Table B.1). It intends to provide a full perspective on the prior art approaches for ECG-based biometric recognition. For more detailed information about each of the approaches and their preprocessing methods, feature extraction, recognition, and other processes, please refer to Chapter 5 of this dissertation.

Table B.1: Summary of the surveyed prior art ECG-based biometric recognition approaches (ordered ordered ascendingly by year. RP - reference point detection; SS - signal segmentation; AN - amplitude normalisation; TN - time normalisation; DR - dimensionality reduction; T - recognition type (I - identification; A - authentication; B - both were evaluated); C - continuous recognition; NS - number of subjects).

Researchers	Preprocessing	RP	SS	AN	TN	Features	DR	Matching	Decision	T	C	Dataset	NS	Results
Biel et al. (1999, 2001)	-	-	-	-	-	10 Lead I fiducial features	✓	PCA	SIMCA	I	×	Private	20	IDR 100%
Kyoso et al. (2000)	HPF + LPF + NF	✓	✓	×	×	PQ and QT times	×	Mahalanobis distance	LDA	I	×	Private	3	IDR 99.5%
Kyoso and Uchiyama (2001a,b)	HPF + LPF + NF	✓	✓	×	×	QRS duration and QT time	×	Mahalanobis distance	LDA	I	×	Private	9	IDR 94.2%
Shen et al. (2002)	-	✓	✓	✓	✓	RQ, RS, and ST amplitudes, QS and QT intervals, RS slope, and QRS triangle area	×	Correlation	DBNN	I	×	MIT NSR	20	IDR 100%
Palaniappan and Krishnan (2004)	LPF	✓	✓	×	×	Amplitudes of R, QR, RS, width of QRS, R-R interval, and beat form factor	×	Extracted features	MLP; SFA	I	×	MIT NSR	10	IDR[1]: MLP 96.2% SFA 83.6%
Israel et al. (2005)	BPF	✓	✓	×	×	RQ, RS, RP, RL, RP', RT, RS', RT', P width, T width, ST, PQ, PT, LQ, ST'	✓	LDA	Contingency matrix majority voting	I	×	Private	49	IDR[2]: Anxty. 97% Norm. 98%
Saechia et al. (2005)	-	✓	✓	×	✓	Fourier transform of PQRST (whole), P, QRS, and T	×	Extracted features	Neural Networks	A	×	-	-	FRR[3]: Whole 17.1% Apart 2.85%
Plataniotis et al. (2006)	BPF	×	×	×	×	Autocorrelation coefficients	✓	Normalised Euclidean distance	Normalised Gaussian Log-likelihood	B	×	PTB	14	IDR 100% FAR 0.02%
Zhang and Wei (2006)	-	✓	✓	×	×	Fiducial amplitudes, durations, intervals, levels, and areas	✓	Extracted features	Bayes-minimum-error-rate classification	I	×	Private	502	IDR[4]: L. I 85.3% L. II 92.0% L. V1 95.2% L. V2 97.4%
Molina et al. (2007)	Savitzky-Golay	✓	✓	×	×	R-R signal segments	×	DTW path	Nearest Neighbour	A	×	Private	10	EER 2%

[1]Results for the use of Multilayer Perceptron (MLP) or Simplified Fuzzy ARTMAP (SFA) on decision.
[2]Results for the use of normal signals and signals acquired from subjects with anxiety.
[3]Results when using the whole PQRST segment, or when using the P, QRS, and T segments apart.
[4]Results for each of the various leads used.

Researchers	Preprocessing	RP	SS	AN	TN	Features	DR	Matching	Decision	T	C	Dataset	NS	Results	
Wübbeler et al. (2007)	Moving Median + LPF	✓	✓	×	×	2D QRS segments (combination of leads I, II, and III)	×	Temporal derivatives distance	Nearest Neighbour	B	×	PTB	74	IDR EER	98.1% 2.8%
Agrafioti and Hatzinakos (2008)	BPF	×	✓	×	×	Normalised autocorrelation	✓	Correlation coefficient	Nearest Neighbour	I	×	PTB + NSR	27	IDR[5]: DCT LDA	 96.3% 100%
Chan et al. (2008)	NF	✓	✓	×	×	Signal-averaged ECG	×	Percent residual difference; correlation coeff.; and wavelet distance	Nearest Neighbour	I	×	Private	50	IDR[6]: PRD CC WD	 70% 80% 89%
Irvine et al. (2008)	BPF	×	✓	✓	✓	Covariance matrix eigenvectors	×	Extracted features	PCA	I	×	Private	39	IDR	100%
Boumbarov et al. (2009)	HPF	×	✓	×	×	Cardiac cycle vector matrix	✓	Extracted features	RBF NN	I	×	Private	9	IDR	83.3%
Fang and Chan (2009)	BPF	✓	✓	✓	×	Avg. beat phase space portrait	×	Spatial correlation; and Mutual Nearest Point Distance	Nearest Neighbour	I	×	Private	100	IDR[7]: 1L 3L	 93% 99%
Fatemian and Hatzinakos (2009)	DWT + Moving Average	✓	✓	×	✓	Heart-rate normalised single heartbeat template construction	×	Correlation coefficient	Nearest Neighbour	I	×	PTB + NSR	27	IDR	99.6%
Guennoun et al. (2009)	LPF	✓	×	×	×	Fiducial amplitude and time features	✓	Mahalanobis distance	Thresholding and Voting	A	✓	Private	16	FRR FAR	0.01% 0%
Coutinho et al. (2010)	BPF	✓	✓	×	×	Uniformly quantised avg. beats	×	Ziv-Merhav relative entropy	Nearest Neighbour	I	×	Private	19	IDR	99.5%
Fatemian et al. (2010)	DWT	×	✓	×	✓	Avg. ensemble heartbeat	×	Correlation coefficient	Nearest Neighbour	A	×	Private	21	IDR EER	95.4% 3.3%
Ghofrani and Bostani (2010)	BPF + NF	×	-	×	×	AR coeff.; PSD; Lyapunov exponents; Approximation Entropy; Higuchi Fractal Dimension; Shannon Entropy	×	Extracted features	kNN; MLP; PNN	I	×	PTB	12	IDR[8]: AR ApEn Hig. Lya. Sha.	 98.6% 94.3% 87.4% 96.7% 92.8%

[5]Results when using Autocorrelation with DCT and Autocorrelation with LDA.

[6]Results when using Percent residual difference , Correlation coefficient, and Wavelet distance.

[7]Results with the use of data from one lead, or from 3 leads.

[8]Results with Autoregressive coefficients (AR), Approximation Entropy (ApEn), Higuchi F. D. (Hig.), Lyapunov Exponents (Lya.), and Shannon Entropy (Sha.).

Researchers	Preprocessing	RP	SS	AN	TN	Features	DR	Matching	Decision	T	C	Dataset	NS	Results	
Li and Narayanan (2010)		-	✓	✓	✓	Hermite polynomial expansion; Cepstral features	✓	Extracted features	Fusion of SVM and GMM-UBM Supervector	B	×	MIT NSR	18	IDR EER	98.3% 0.5%
N and Jayaraman (2010)	HPF + BPF	✓	×	×	×	P, T, ST, PR, QRS and QT intervals	×	FLDA and DTW	Nearest Neighbour	B	×	MIT NSR	15	IDR	96%
Odinaka et al. (2010)	LPF + HPF + NF	✓	✓	✓	×	Log-STFT spectrogram Gaussian models	✓	Log-likelihood ratio	Nearest Neighbour	I	×	Private	269	IDR EER	99% 0.37%
Sasikala and Wahidabanu (2010)	Median Filters + DWT	✓	✓	×	×	Fiducial amplitudes and differences	×	Correlation coefficient	Maximisation and mean distance	I	×	MIT Arrh.	10	IDR	62.7%
Tawfik et al. (2010)	HPF + LPF	✓	✓	✓	✓	QRS DCT coefficients	×	Extracted features	Neural Networks	I	×	Private	22	IDR	99.1%
Ye et al. (2010)	BPF	✓	✓	×	×	Daubechies DWT and ICA	×	Extracted features	Gaussian RBF SVM	I	×	MIT Arrh.	47	IDR FPIR	99.6% 12.3%
												MIT NSR1[9]	18	IDR FPIR	99.3% 26.9%
												MIT LT	65	IDR FPIR	98.1% 28.7%
												MIT NSR2	18	IDR FPIR	97.5% 37.1%
Coutinho et al. (2011)	BPF	✓	✓	×	×	User-tuned Lloyd-Max quantised avg. ensemble beats	×	Ziv-Merhav cross parsing similarity	Nearest Neighbour	A	✓	Private	19	EER	0.36%
Lourenço et al. (2011a)	BPF	✓	✓	✓	✓	Avg. normalised beat	×	Euclidean distance	Nearest Neighbour	B	×	Private	16	IDR EER	94.3% 13%
Matta et al. (2011)	BPF	×	✓	×	×	Autocorrelation coefficients	✓	Euclidean distance	Nearest Neighbour	I	✓	Private	10	IDR TPIR3	75% 99%
Safie et al. (2011)	BPF	×	✓	✓	×	Pulse Active Ratio	×	Euclidean distance	Nearest Neighbour	A	×	PTB	112	EER[10]: Heal. Arrh. AUC: Heal. Arrh.	 9.98% 19.2% 94.5% 86.7%
Shen et al. (2011)	BPF	✓	✓	✓	✓	Amplitudes, durations, slopes, angles, and QRS area	×	Correlation coefficient and LDA	Nearest Neighbour	I	×	Private	168	IDR	98%

[9]MIT NSR1 - only short excerpts selected from the collection signals, and MIT NSR2 - long signals used.
[10]Results when using data only from healthy subjects, or from arrhythmia patients as well.

Researchers	Preprocessing	RP	SS	AN	TN	Features	DR	Matching	Decision	T	C	Dataset	NS	Results
Sufi et al. (2011)	-	✓	✓	×	×	Cardioid graph centroid, extremas, area, and perimeter	×	Straight line and percentage distances	Nearest Neighbour	B	×	MIT Arrh.	-	MIDR 1% FAR 0.5% FRR 0.5%
Agrafioti et al. (2012)	BPF	×	✓	✓	×	Autocorrelation coefficients	✓	Euclidean distance	Nearest Neighbour	A	✓	Private	42	EER 3.96%
Belgacem et al. (2012)	BPF	✓	✓	✓	×	Avg. beat Daubechies DWT	×	Extracted features	Random Forest Ensemble	B	×	Arrh. + ST-T + NSR + PTB + Private	80	Sens. 100% FAR 0.60% FRR 0.58%
Lourenço et al. (2012a)	BPF + LPF	✓	✓	✓	×	Segmented heartbeats	×	Extracted features	kNN and SVM	B	✓	Private	32	EER[11] 9.39% FAR 0% FRR 26.4%
Singh and Singh (2012)	-	✓	✓	✓	✓	Interval, angle, and amplitude fiducial features	×	Euclidean distance	Nearest Neighbour	A	×	ST-T + Arrh. + NSR + QT	73	EER 10.8%
Belgacem et al. (2013)	BPF	✓	✓	✓	×	Avg. beat Daubechies DWT	×	Extracted features	Random Forest Ensemble	B	×	Arrh. + ST-T + NSR + PTB + Private	80	IDR 100% FAR 0.63% FRR 0.66%
Coutinho et al. (2013)[12]	BFP	✓	✓	✓	×	Fid. latency and amplitude from mean waveform subsampling	×	Euclidean distance	Nearest Neighbour	B	✓	PTB Private	51 26	IDR 99.9% EER 0.01% IDR 99.6% EER 0.70%
Coutinho et al. (2013)[12]	BFP	✓	✓	✓	×	User-tuned Lloyd-Max quantised heartbeats	×	Ziv-Merhav cross parsing similarity	Nearest Neighbour	B	✓	PTB Private	51 26	IDR 99.4% EER 0.13% IDR 99.9% EER 0.29%
Labati et al. (2013)	HPF + NF	✓	✓	×	×	QRS segment set templates	×	Cross-correlation similarity matrix and score fusion	Nearest Neighbour	A	✓	E-HOL 24h	185	EER 5.36%
Matos et al. (2013)	HPF + NF	✓	✓	×	×	STFT spectrogram Gaussian models and Spectral zoom	✓	Log-likelihood ratio	Nearest Neighbour	A	×	Private	27	EER 10%
Silva et al. (2013)	BPF	✓	✓	×	×	Mean and median ensemble beats	×	Euclidean and cosine distances	kNN and SVM	A	×	Private	63	EER[13]: kNN 0.99% SVM 9.10%

[11] EER with kNN, FAR and FRR with SVM.

[12] More than one different approach was proposed in the same publication.

[13] Results with kNN and SVM.

Researchers	Preprocessing	RP	SS	AN	TN	Features	DR	Matching	Decision	T	C	Dataset	NS	Results
Wang et al. (2013)	-	×	✓	✓	×	Max-pooling representation elements	×	Extracted features	Nearest Neighbour	I	×	PTB	100	IDR 99.5%
Ergin et al. (2014)	-	✓	✓	×	×	Fusion of QRS fid., time domain, wavelet transform and PSD features	×	Extracted features	C4.5 Decision Tree and Bayesian Network	I	×	MIT NSR	18	F-score[14]: C4.5 0.97% Bayes 0.96%
Iqbal et al. (2014)	-	✓	✓	×	×	QRS cardioid graph coordinates	×	Extracted features	Multilayer Perceptron	I	×	Private	30	IDR 96.4%
Labati et al. (2014)	HPF + NF	✓	✓	×	×	QRS segments	×	Cross-correlation similarity matrix and score fusion	Nearest Neighbour	A	✓	E-HOL 24h	185	EER 5.36%
Lin et al. (2014)		×	×	×	×	Correlation dimension Lyapunov exponents, RMSE	×	Extracted features	SVM	I	×	Private	26	IDR 81.7%
Lourenço et al. (2014)	-	✓	✓	×	×	Mean ensemble beats	×	Extracted features	SVM	A	×	Private	63	EER 2.5%
Matos et al. (2014)	LPF	✓	✓	×	×	STFT window features	✓	Log-likelihood ratio	Nearest Neighbour	B	✓	Private	10	IDR 100% EER 14%
Pathoumvanh et al. (2014)	BPF	✓	✓	×	×	CWT	✓	Euclidean distance	Nearest Neighbour	I	×	Private	10	IDR[15]: Norm. 97% HRV 80%
Zhou et al. (2014)	BPF	✓	✓	×	×	Signal between 3 consec. R peaks	×	DTW path	Nearest Neighbour	A	×	Private	20	HTER 1.45%
Brás and Pinho (2015)	NF + Moving Average	×	×	✓	×	Kolmogorov-based normalised relative compression	×	Extracted features	Nearest Neighbour	I	×	PTB	52	IDR 99.9%
Choudhary and Manikandan (2015)	DCT	✓	✓	✓	✓	Avg. ensemble beat	×	RMSE; PRD; NCC; WWPRD; and WDIST	Nearest Neighbour	A	×	Arrh. + STC + QT + NSR + SLP	127	FAR 5.8% FRR 11.6%
Dar et al. (2015a)	Poly. line fitting	✓	✓	✓	×	Haar Transform	✓	Extracted features	Nearest Neighbour	B	×	MIT Arrh. MIT NSR ECG-ID	47 18 90	IDR 93.1% IDR 99.4% IDR 83.2%

[14]Results when using the C4.5 Decision Tree or the Bayesian Network

[15]Results for normal records, and acquisition under increased heart rate variability.

Researchers	Preprocessing	RP	SS	AN	TN	Features	DR	Matching	Decision	T	C	Dataset	NS	Results	
Dar et al. (2015b)	Poly. line fitting	✓	✓	✓	×	Haar Transform and HRV	✓	Extracted features	Random Forest Ensemble	B	×	MIT Arrh.	47	IDR	95.9%
														FAR	4.1%
														FRR	0.1%
												MIT NSR	18	IDR	100%
														FAR	0%
														FRR	0%
												ECG-ID	90	IDR	83.9%
														FAR	16.1%
														FRR	0.3%
Jahiruzzaman and Hossain (2015)	BPF	×	×	×	×	CWT and Chaotic Encryption	×	Extracted features	Unique CE sequences identification	I	×	MIT Arrh.	11	IDR	96.9%
Carreiras et al. (2016)	BPF	✓	✓	×	×	Segmented heartbeat	×	Extracted features	Nearest Neighbour	B	×	Private	618	EER	9.01%
														MIDR	15.6%
Chun (2016)[12]	DWT	✓	✓	×	×	Guided filtering avg. beat	×	DTW	Nearest Neighbour	A	×	ECG-ID	89	EER	5.2%
														AUC	0.987
Chun (2016)[12]	DWT	✓	✓	×	×	Guided filtering avg. beat	×	Euclidean distance	Nearest Neighbour	A	×	ECG-ID	89	EER	2.4%
														AUC	0.998
Chun (2016)[12]	DWT	✓	✓	×	×	Guided filtering avg. beat	×	Extracted features	PCA	A	×	ECG-ID	89	EER	2.4%
														AUC	0.998
Hejazi et al. (2016)	DWT	×	×	×	×	Autocorrelation coefficients	✓	Extracted features	SVM	I	×	Private	52	IDR	76.3%
														FAR	3.5%
														FRR	4.83%
Louis et al. (2016)	BPF	✓	✓	×	×	1D multi-res. LBP	×	Extracted features	Bagging	A	✓	UofTDB	1012	EER	7.89%
														FRR	0.39%
														FAR	1.57%
Porée et al. (2016)[12]	LPF	✓	✓	×	×	10 beat avg. ensemble	×	Extracted features	Discrimination coefficient	B	×	Private	14	IDR	100%
Porée et al. (2016)[12]	LPF	✓	✓	×	×	10 beat avg. ensemble	×	Extracted features	Nearest Neighbour	B	×	Private	14	IDR	100%
Rezgui and Lachiri (2016)	BPF	✓	✓	×	×	Amplitudes, areas, time intervals, and slopes of fiducials	×	Extracted features	SVM	I	×	NSR + Arrh.	48	IDR	98.8%
Waili et al. (2016a)	HPF + LPF	✓	✓	✓	×	12 QRS fid. amplitudes	×	Extracted features	Multilayer Perceptron	I	×	PTB	14	IDR	96%

References

M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas. Biometric authentication based on PCG and ECG signals: present status and future directions. *Signal, Image and Video Processing*, 8(4): 739–751, 2014. ISSN 1863-1711. doi: 10.1007/s11760-013-0593-4.

F. Agrafioti. *ECG in Biometric Recognition: Time Dependency and Application Challenges*. PhD thesis, University of Toronto, 2011.

F. Agrafioti and D. Hatzinakos. ECG Based Recognition Using Second Order Statistics. In *Communication Networks and Services Research Conference, 2008. CNSR 2008. 6th Annual*, pages 82–87, May 2008. doi: 10.1109/CNSR.2008.38.

F. Agrafioti, J. Gao, and D. Hatzinakos. Heart Biometrics: Theory, Methods and Applications. In J. Yang, editor, *Biometrics*, pages 199–216. InTech, 2011. doi: 10.5772/18113.

F. Agrafioti, F. M. Bui, and D. Hatzinakos. Secure Telemedicine: Biometrics for Remote and Continuous Patient Verification. *Journal of Computer Networks and Communications*, 2012: 11, 2012. doi: 10.1155/2012/924791.

N. Ahmed, T. Natarajan, and K. R. Rao. Discrete cosine transform. *IEEE Transactions on Computers*, C-23(1):90–93, Jan 1974. ISSN 0018-9340. doi: 10.1109/T-C.1974.223784.

N. Akhter, S. Tharewal, V. Kale, A. Bhalerao, and K. V. Kale. Heart-Based Biometrics and Possible Use of Heart Rate Variability in Biometric Recognition Systems. In R. Chaki, A. Cortesi, K. Saeed, and N. Chaki, editors, *Advanced Computing and Systems for Security: Volume 1*, pages 15–29. Springer India, New Delhi, 2016. ISBN 978-81-322-2650-5. doi: 10.1007/978-81-322-2650-5_2.

N. Belgacem, A. Nait-Ali, R. Fournier, and F. Bereksi-Reguig. ECG based human authentication using wavelets and random forests. *Int J Cryptography Inf Secur*, 2:1–11, 2012.

N. Belgacem, A. Nait-Ali, R. Fournier, and F. Bereksi-Reguig. ECG Based Human Identification Using Random Forests. In *The International Conference on E-Technologies and Business on the Web (EBW2013). Bangkok, Thailand*. 2013.

A. Benzaoui and A. Boukrouche. 1DLBP and PCA for face recognition. In *2013 11th International Symposium on Programming and Systems (ISPS)*, pages 7–11, April 2013. doi: 10.1109/ISPS.2013.6581486.

L. Biel, O. Pettersson, L. Philipson, and P. Wide. ECG analysis: a new approach in human identification. In *IMTC/99. Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference*, volume 1, pages 557–561, May 1999. doi: 10.1109/IMTC.1999.776813.

L. Biel, O. Pettersson, L. Philipson, and P. Wide. ECG analysis: a new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, 50(3):808–812, Jun 2001. ISSN 0018-9456. doi: 10.1109/19.930458.

C. M. Bishop. *Pattern Recognition and Machine Learning*. Springer, 2006. ISBN 978-0387-31073-2.

M. Blanco-Velasco, B. Weng, and K. E. Barner. ECG signal denoising and baseline wander correction based on the empirical mode decomposition. *Computers in Biology and Medicine*, 38:1–13, January 2008. doi: 10.1016/j.compbiomed.2007.06.003.

R. M. Bolle, J. H. Connell, S. Pankanti, N. Ratha, and A. W. Senior. *Guide to Biometrics*. Springer Science+Business Media, New York, NY, first edition, 2004. ISBN 978-1-4419-2305-9.

O. Boumbarov, Y. Velchev, and S. Sokolov. ECG personal identification in subspaces using radial basis neural networks. In *Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, 2009. IDAACS 2009. IEEE International Workshop on*, pages 446–451, Sept 2009. doi: 10.1109/IDAACS.2009.5342942.

R. Bousseljot, D. Kreiseler, and A. Schnabel. Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet. *Biomedizinische Technik*, 40(s1):317–318, 1995. doi: 10.1515/bmte.1995.40.s1.317.

S. Brás and A. J. Pinho. ECG biometric identification: A compression based approach. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5838–5841, Aug 2015. doi: 10.1109/EMBC.2015.7319719.

C. Carreiras, A. Lourenço, H. Silva, A. Fred, and R. Ferreira. Evaluating Template Uniqueness in ECG Biometrics. In J. Filipe, O. Gusikhin, K. Madani, and J. Sasiadek, editors, *Informatics in Control, Automation and Robotics: 11th International Conference, ICINCO 2014 Vienna, Austria, September 2-4, 2014 Revised Selected Papers*, pages 111–123. Springer International Publishing, Cham, 2016. ISBN 978-3-319-26453-0. doi: 10.1007/978-3-319-26453-0_7.

A. D. C. Chan, M. M. Hamdy, A. Badre, and V. Badee. Wavelet Distance Measure for Person Identification Using Electrocardiograms. *IEEE Transactions on Instrumentation and Measurement*, 57(2):248–253, Feb 2008. ISSN 0018-9456. doi: 10.1109/TIM.2007.909996.

Y. Choi, S. I. Han, S. H. Kong, and H. Ko. Driver Status Monitoring Systems for Smart Vehicles Using Physiological Sensors: A safety enhancement system from automobile manufacturers. *IEEE Signal Processing Magazine*, 33(6):22–34, Nov 2016. ISSN 1053-5888. doi: 10.1109/MSP.2016.2602095.

T. Choudhary and M. S. Manikandan. A novel unified framework for noise-robust ECG-based biometric authentication. In *Signal Processing and Integrated Networks (SPIN), 2015 2nd International Conference on*, pages 186–191, Feb 2015. doi: 10.1109/SPIN.2015.7095379.

I. I. Christov. Real time electrocardiogram QRS detection using combined adaptive threshold. *BioMedical Engineering OnLine*, 3(1):28, 2004. ISSN 1475-925X. doi: 10.1186/1475-925X-3-28.

S. Y. Chun. Single pulse ECG-based small scale user authentication using guided filtering. In *2016 International Conference on Biometrics (ICB)*, pages 1–7. June 2016. doi: 10.1109/ICB.2016.7550065.

R. Coppola and M. Morisio. Connected Car: Technologies, Issues, Future Trends. *ACM Computing Surveys*, 49(3):46:1–46:36, October 2016. ISSN 0360-0300. doi: 10.1145/2971482.

C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995. ISSN 1573-0565. doi: 10.1007/BF00994018.

D. P. Coutinho, A. L. N. Fred, and M. A. T. Figueiredo. One-Lead ECG-based Personal Identification Using Ziv-Merhav Cross Parsing. In *2010 20th International Conference on Pattern Recognition*, pages 3858–3861, Aug 2010. doi: 10.1109/ICPR.2010.940.

D. P. Coutinho, A. L. N. Fred, and M. A. T. Figueiredo. ECG-based continuous authentication system using adaptive string matching. *BIOSIGNALS 2011 - Proceedings of the International Conference on Bio-Inspired Systems and Signal Processing*, pages 354–359, 2011.

D. P. Coutinho, H. Silva, H. Gamboa, A. Fred, and M. Figueiredo. Novel fiducial and non-fiducial approaches to electrocardiogram-based biometric systems. *IET Biometrics*, 2(2):64–75, June 2013. ISSN 2047-4938. doi: 10.1049/iet-bmt.2012.0055.

L. T. D'Angelo, J. Parlow, W. Spiessl, S. Hoch, and T. C. Lüth. A system for unobtrusive in-car vital parameter acquisition and processing. In *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*, pages 1–7, March 2010. doi: 10.4108/ICST.PERVASIVEHEALTH2010.8877.

M. N. Dar, M. U. Akram, A. Shaukat, and M. A. Khan. ECG Based Biometric Identification for Population with Normal and Cardiac Anomalies Using Hybrid HRV and DWT Features. In *IT Convergence and Security (ICITCS), 2015 5th International Conference on*, pages 1–5, Aug 2015a. doi: 10.1109/ICITCS.2015.7292977.

M. N. Dar, M. U. Akram, A. Usman, and S. A. Khan. ECG biometric identification for general population using multiresolution analysis of DWT based features. In *2015 Second International Conference on Information Security and Cyber Forensics (InfoSec)*, pages 5–10, Nov 2015b. doi: 10.1109/InfoSec.2015.7435498.

S. R. Devasahayam. *Signals and Systems in Biomedical Engineering: Signal Processing and Physiological Systems Modeling*. Springer Science & Business Media, second edition, 2012. ISBN 978-1-4614-5331-4.

G. Doddington, W. Liggett, A. Martin, M. Przybocki, and D. Reynolds. Sheep, Goats, Lambs and Wolves: A Statistical Analysis of Speaker Performance in the NIST 1998 Speaker Recognition Evaluation. Technical report, National Institute of Standards and Technology, Gaithersburg, MD, 1998.

W. A. H. Engelse and C. Zeelenberg. A single scan algorithm for QRS-detection and feature extraction. *Proc IEEE Computers in Cardiology*, 6:37–42, 1979.

S. Ergin, A. K. Uysal, E. S. Gunal, S. Gunal, and M. B. Gulmezoglu. ECG based biometric authentication using ensemble of features. In *2014 9th Iberian Conference on Information Systems and Technologies (CISTI)*, pages 1–6, June 2014. doi: 10.1109/CISTI.2014.6877089.

M. Ester, H.-P. Kriegel, J. Sander, and X. Xu. A density-based algorithm for discovering clusters a density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, pages 226–231. AAAI Press, 1996.

S.-C. Fang and H.-L. Chan. Human identification by quantifying similarity and dissimilarity in electrocardiogram phase space. *Pattern Recognition*, 42(9):1824 – 1831, 2009. ISSN 0031-3203. doi: 10.1016/j.patcog.2008.11.020.

S. Z. Fatemian and D. Hatzinakos. A new ECG feature extractor for biometric recognition. In *2009 16th International Conference on Digital Signal Processing*, pages 1–6, July 2009. doi: 10.1109/ICDSP.2009.5201143.

S. Z. Fatemian, F. Agrafioti, and D. Hatzinakos. HeartID: Cardiac biometric recognition. In *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–5, Sept 2010. doi: 10.1109/BTAS.2010.5634493.

G. E. Forsen, M. R. Nelson, and R. J. Staron Jr. Personal Attributes Authentication Techniques. Technical report, Pattern Analysis and Recognition Corporation, Rome Air Development Center New York, October 1977.

A. Fratini, M. Sansone, P. Bifulco, and M. Cesarelli. Individual identification via electrocardiogram analysis. *BioMedical Engineering OnLine*, 14(1):78, 2015. ISSN 1475-925X. doi: 10.1186/s12938-015-0072-y.

N. Ghofrani and R. Bostani. Reliable features for an ECG-based biometric system. In *Biomedical Engineering (ICBME), 2010 17th Iranian Conference of*, pages 1–5, Nov 2010. doi: 10.1109/ICBME.2010.5704918.

A. L. Goldberger, L. A. N. 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. doi: 10.1161/01.CIR.101.23.e215.

J. Gomez-Clapers and R. Casanella. Wireless ECG and heart rate monitoring using dual ground dry electrodes. In *Proceedings of the 17th Symposium IMEKO TC 4 International Symposium*, 2010.

J. Gomez-Clapers and R. Casanella. A Fast and Easy-to-Use ECG Acquisition and Heart Rate Monitoring System Using a Wireless Steering Wheel. *IEEE Sensors Journal*, 12(3):610–616, March 2012. ISSN 1530-437X. doi: 10.1109/JSEN.2011.2118201.

P. J. Grother, G. W. Quinn, and P. J. Phillips. Report on the Evaluation of 2D Still-Image Face Recognition Algorithms. *NIST Interagency Report*, 7709, 2010.

A. Gubbi, M. F. Azeem, and Z. A. Ansari. New Fuzzy LBP Features for Face Recognition. *CoRR*, abs/1509.06853, 2015.

M. Guennoun, N. Abbad, J. Talom, S. M. M. Rahman, and K. El-Khatib. Continuous authentication by electrocardiogram data. In *Science and Technology for Humanity (TIC-STH), 2009 IEEE Toronto International Conference*, pages 40–42, Sept 2009. doi: 10.1109/TIC-STH.2009.5444466.

L. Gupta, D. L. Molfese, R. Tammana, and P. G. Simos. Nonlinear alignment and averaging for estimating the evoked potential. *IEEE Transactions on Biomedical Engineering*, 43(4):348–356, April 1996. doi: 10.1109/10.486255.

Z. Hassan, S. O. Gilani, and M. Jamil. Review of Fiducial and Non-Fiducial Techniques of Feature Extraction in ECG based Biometric Systems. *Indian Journal of Science and Technology*, 9(21), 2016. ISSN 0974 -5645.

J. A. Healey and R. W. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166, June 2005. ISSN 1524-9050. doi: 10.1109/TITS.2005.848368.

M. Hejazi, S. Al-Haddad, Y. P. Singh, S. J. Hashim, and A. F. A. Aziz. ECG biometric authentication based on non-fiducial approach using kernel methods. *Digital Signal Processing*, 52:72 – 86, 2016. ISSN 1051-2004. doi: 10.1016/j.dsp.2016.02.008.

S. Heuer, B. Chamadiya, A. Gharbi, C. Kunze, and M. Wagner. Unobtrusive in-vehicle biosignal instrumentation for advanced driver assistance and active safety. In *2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, pages 252–256, Nov 2010. doi: 10.1109/IECBES.2010.5742238.

R. Hoekema, G. J. H. Uijen, and A. van Oosterom. Geometrical aspects of the inter-individual variability of multilead ECG recordings. In *Computers in Cardiology 1999*, volume 26, pages 499–502, 1999. doi: 10.1109/CIC.1999.826017.

R. Hoekema, G. J. H. Uijen, and A. van Oosterom. Geometrical aspects of the interindividual variability of multilead ECG recordings. *IEEE Transactions on Biomedical Engineering*, 48(5): 551–559, May 2001. ISSN 0018-9294. doi: 10.1109/10.918594.

N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998. ISSN 1364-5021. doi: 10.1098/rspa.1998.0193.

S. C. Huang, B. H. Chen, S. K. Chou, J. N. Hwang, and K. H. Lee. Smart Car [Application Notes]. *IEEE Computational Intelligence Magazine*, 11(4):46–58, November 2016. ISSN 1556-603X. doi: 10.1109/MCI.2016.2601758.

F.-t.-Z. Iqbal, K. A. Sidek, N. A. Noah, and T. S. Gunawan. A comparative analysis of QRS and cardioid graph based ECG biometric recognition in different physiological conditions. In *IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA2014)*, November 2014.

J. M. Irvine, S. A. Israel, W. T. Scruggs, and W. J. Worek. eigenPulse: Robust human identification from cardiovascular function. *Pattern Recognition*, 41(11):3427 – 3435, 2008. ISSN 0031-3203. doi: 10.1016/j.patcog.2008.04.015.

S. A. Israel, W. T. Scruggs, W. J. Worek, and J. M. Irvine. Fusing face and ECG for personal identification. In *32nd Applied Imagery Pattern Recognition Workshop, 2003. Proceedings.*, pages 226–231, Oct 2003. doi: 10.1109/AIPR.2003.1284276.

S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold. ECG to identify individuals. *Pattern Recognition*, 38(1):133 – 142, 2005. ISSN 0031-3203. doi: 10.1016/j.patcog.2004.05.014.

F. Jager, A. Taddei, G. B. Moody, M. Emdin, G. Antolic, R. Dorn, A. Smrdel, C. Marchesi, and R. G. Mark. 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. *Medical & Biological Engineering & Computing*, 41(2):172–183, 2003.

M. Jahiruzzaman and A. B. M. A. Hossain. ECG based biometric human identification using chaotic encryption. In *Electrical Engineering and Information Communication Technology (ICEEICT), 2015 International Conference on*, pages 1–5, May 2015. doi: 10.1109/ICEEICT.2015.7307417.

A. K. Jain, R. Bolle, and S. Pankanti. *Biometrics: Personal Identification in Networked Society*. Kluwer Academic Publishers, 1999.

A. K. Jain, A. A. Ross, and K. Nandakumar. *Introduction to Biometrics*. Springer Publishing Company, Incorporated, 2011. ISBN 9780387773254.

G. Kaur, G. Singh, and V. Kumar. A Review on Biometric Recognition. *International Journal of Bio-Science and Bio-Technology*, 6:69–76, 2014.

D. Kreiseler and R. Bousseljot. Automatisierte EKG-Auswertung mit Hilfe der EKG-Signaldatenbank CARDIODAT der PTB. *Biomedizinische Technik*, 40(s1):319–320, 1995. doi: 10.1515/bmte.1995.40.s1.319.

M. Kyoso and A. Uchiyama. Development of an ECG identification system. In *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, volume 4, pages 3721–3723 vol.4, October 2001a. doi: 10.1109/IEMBS.2001.1019645.

M. Kyoso and A. Uchiyama. Development of ECG Identification System. *Japanese Journal of Medical Electronics and Biological Engineering*, 39(Supplement):559, May 2001b.

M. Kyoso, K. Ohishi, and A. Uchiyama. Development of ECG Identification System. *Japanese Journal of Medical Electronics and Biological Engineering*, 38(Supplement):392, May 2000.

R. D. Labati, R. Sassi, and F. Scotti. ECG biometric recognition: Permanence analysis of QRS signals for 24 hours continuous authentication. In *2013 IEEE International Workshop on Information Forensics and Security (WIFS)*, pages 31–36, Nov 2013. doi: 10.1109/WIFS.2013.6707790.

R. D. Labati, V. Piuri, R. Sassi, F. Scotti, and G. Sforza. Adaptive ECG biometric recognition: a study on re-enrollment methods for QRS signals. In *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, pages 30–37, Dec 2014. doi: 10.1109/CIBIM.2014.7015440.

P. Laguna, R. G. Mark, A. Goldberger, and G. B. Moody. A Database for Evaluation of Algorithms for Measurement of QT and Other Waveform Intervals in the ECG. *Computers in Cardiology*, 24:673–676, 1997.

M. Li and S. Narayanan. Robust ECG Biometrics by Fusing Temporal and Cepstral Information. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 1326–1329, Aug 2010. doi: 10.1109/ICPR.2010.330.

S. Z. Li and A. K. Jain. *Encyclopedia of Biometrics*. Springer Science & Business Media, second edition, 2009. ISBN 978-0-387-73003-5.

S. L. Lin, C. K. Chen, C. L. Lin, W. C. Yang, and C. T. Chiang. Individual identification based on chaotic electrocardiogram signals during muscular exercise. *IET Biometrics*, 3(4):257–266, 2014. ISSN 2047-4938. doi: 10.1049/iet-bmt.2013.0014.

W. Louis, M. Komeili, and D. Hatzinakos. Continuous Authentication Using One-Dimensional Multi-Resolution Local Binary Patterns (1DMRLBP) in ECG Biometrics. *IEEE Transactions on Information Forensics and Security*, 11(12):2818–2832, Dec 2016. ISSN 1556-6013. doi: 10.1109/TIFS.2016.2599270.

A. Lourenço, H. Silva, C. Carreiras, and A. Fred. *Outlier Detection in Non-intrusive ECG Biometric System*, pages 43–52. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-39094-4. doi: 10.1007/978-3-642-39094-4_6.

A. Lourenço, H. Silva, and A. Fred. Unveiling the Biometric Potential of Finger-based ECG Signals. *Intell. Neuroscience*, 2011:5:1–5:8, Jan 2011a. ISSN 1687-5265. doi: 10.1155/2011/720971.

A. Lourenço, H. Silva, D. P. Santos, and A. Fred. Towards a Finger Based ECG Biometric System. In *BIOSIGNALS 2011 - Proceedings of the International Conference on Bio-inspired Systems and Signal Processing, Rome, Italy*, January 2011b.

A. Lourenço, H. Silva, and A. Fred. ECG-based biometrics: A real time classification approach. In *2012 IEEE International Workshop on Machine Learning for Signal Processing*, pages 1–6, Sept 2012a. doi: 10.1109/MLSP.2012.6349735.

A. Lourenço, H. Silva, R. L. Lourenço, P. L. Leite, and A. Fred. Real Time Electrocardiogram Segmentation For Finger Based ECG Biometrics. In *Proc. International Conf. on Bio-inspired Systems and Signal Processing - Biosignals - INSTICC*, pages 49–54, February 2012b.

A. Lourenço, C. Carreiras, H. Silva, and A. Fred. ECG biometrics: A template selection approach. In *Medical Measurements and Applications (MeMeA), 2014 IEEE International Symposium on*, pages 1–6, June 2014. doi: 10.1109/MeMeA.2014.6860081.

A. Lourenço, A. P. Alves, C. Carreiras, R. P. Duarte, and A. Fred. CardioWheel: ECG Biometrics on the Steering Wheel. In A. Bifet, M. May, B. Zadrozny, R. Gavalda, D. Pedreschi, F. Bonchi, J. Cardoso, and M. Spiliopoulou, editors, *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2015, Porto, Portugal, September 7-11, 2015, Proceedings, Part III*, pages 267–270. Springer International Publishing, Cham, 2015. ISBN 978-3-319-23461-8. doi: 10.1007/978-3-319-23461-8_27.

T. S. Lugovaya. Biometric human identification based on electrocardiogram. Master's thesis, Faculty of Computing Technologies and Informatics, Electrotechnical University "LETI", Saint-Petersburg, Russian Federation, June 2005.

R. MacLeod and B. Birchler. *ECG Measurement and Analysis*. University of Utah, February 2014. URL `http://www.sci.utah.edu/~macleod/bioen/be6000/labnotes/ecg/l3-ecg.pdf`.

A. Mani and M. Nadeski. Processing solutions for biometric systems. Technical report, Texas Instruments Incorporated, 2015.

E. N. Marieb and K. Hoehn. *Human Anatomy & Physiology*. Pearson, Glenview, IL, ninth edition, 2013. ISBN 978-0-321-74326-8.

R. G. Mark, P. S. Schluter, G. B. Moody, P. H. Devlin, and D. Chernoff. An annotated ECG database for evaluating arrhythmia detectors. *IEEE Transactions on Biomedical Engineering*, 29(8):600, 1982.

A. M. Martinez and A. C. Kak. Pca versus lda. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(2):228–233, Feb 2001. ISSN 0162-8828. doi: 10.1109/34.908974.

A. C. Matos, A. Lourenço, and J. Nascimento. Biometric recognition system using low bandwidth ECG signals. In *e-Health Networking, Applications Services (Healthcom), 2013 IEEE 15th International Conference on*, pages 518–522, Oct 2013. doi: 10.1109/HealthCom.2013.6720731.

A. C. Matos, A. Lourenço, and J. Nascimento. Embedded System for Individual Recognition Based on ECG Biometrics. *Procedia Technology*, 17:265 – 272, 2014. ISSN 2212-0173. doi: 10.1016/j.protcy.2014.10.236.

R. Matta, J. K. H. Lau, F. Agrafioti, and D. Hatzinakos. Real-time continuous identification system using ECG signals. In *Electrical and Computer Engineering (CCECE), 2011 24th Canadian Conference on*, pages 001313–001316, May 2011. doi: 10.1109/CCECE.2011.6030676.

P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith. A dynamical model for generating synthetic electrocardiogram signals. *IEEE Transactions on Biomedical Engineering*, 50(3): 289–294, March 2003. ISSN 0018-9294. doi: 10.1109/TBME.2003.808805.

M. Merone, P. Soda, M. Sansone, and C. Sansone. ECG databases for biometric systems: A systematic review. *Expert Systems with Applications*, 67:189 – 202, 2017. ISSN 0957-4174. doi: 10.1016/j.eswa.2016.09.030.

G. G. Molina, F. Bruekers, C. Presura, M. Damstra, and M. van der Veen. Morphological synthesis of ECG signals for person authentication. In *Signal Processing Conference, 2007 15th European*, pages 738–742, Sept 2007.

G. B. Moody and R. G. Mark. The MIT-BIH Arrhythmia Database on CD-ROM and software for use with it. *Computers in Cardiology*, 17:185–188, 1990.

G. Moran and A. Fecowycz. *Cardiology Teaching Package: A Beginners Guide to Normal Heart Function, Sinus Rhythm & Common Cardiac Arrhythmias*. University of Nottingham. URL `http://www.nottingham.ac.uk/nursing/practice/resources/cardiology/`.

K. P. Murphy. *Machine Learning: A Probabilistic Perspective*. MIT Press, 2012. ISBN 978-0-262-01802-9.

V. N and S. Jayaraman. Human Electrocardiogram for Biometrics Using DTW and FLDA. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 3838–3841, Aug 2010. doi: 10.1109/ICPR.2010.935.

A. P. Nemirko and T. S. Lugovaya. Biometric human identification based on electrocardiogram. In *Proc. XII-th Russian Conference on Mathematical Methods of Pattern Recognition*, pages 387–390, 2005. ISBN 5-317-01445-X.

N. Neverova, C. Wolf, G. Lacey, L. Fridman, D. Chandra, B. Barbello, and G. Taylor. Learning human identity from motion patterns. *IEEE Access*, 4:1810–1820, 2016. ISSN 2169-3536. doi: 10.1109/ACCESS.2016.2557846.

V. Niennattrakul and C. A. Ratanamahatana. Shape averaging under time warping. In *2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*, volume 02, pages 626–629, May 2009. doi: 10.1109/ECTICON.2009.5137128.

K. Niinuma, U. Park, and A. K. Jain. Soft Biometric Traits for Continuous User Authentication. *IEEE Transactions on Information Forensics and Security*, 5(4):771–780, Dec 2010. ISSN 1556-6013. doi: 10.1109/TIFS.2010.2075927.

S. Nikan, F. Gwadry-Sridhar, and M. Bauer. Pattern recognition application in ecg arrhythmia classification. In *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 5: HEALTHINF, (BIOSTEC 2017)*, pages 48–56. 2017. ISBN 978-989-758-213-4.

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 *2010 IEEE International Workshop on Information Forensics and Security*, pages 1–6, Dec 2010. doi: 10.1109/WIFS.2010.5711466.

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 Transactions on Information Forensics and Security*, 7(6):1812–1824, Dec 2012. ISSN 1556-6013. doi: 10.1109/TIFS.2012.2215324.

M. K. Omar and J. Pelecanos. Training Universal Background Models for Speaker Recognition. In *Odyssey - The Speaker and Language Recognition Workshop*, pages 52–57, 2010.

S. Ongwattanakul and D. Srisai. Contrast Enhanced Dynamic Time Warping Distance for Time Series Shape Averaging Classification. In *Proceedings of the 2Nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, ICIS '09, pages 976–981, 2009. ISBN 978-1-60558-710-3. doi: 10.1145/1655925.1656102.

A. V. Oppenheim, R. W. Schafer, and J. R. Buck. *Discrete-Time Signal Processing*. Prentice-Hall, 1999. ISBN 0-13-754920-2.

R. Palaniappan and S. M. Krishnan. Identifying individuals using ECG beats. In *Proceedings of the International Conference on Signal Processing and Communications (SPCOM'04)*, pages 569–572, 2004.

J. Pan and W. J. Tompkins. A Real-Time QRS Detection Algorithm. *IEEE Transactions on Biomedical Engineering*, BME-32(3):230–236, March 1985.

S. Pathoumvanh, S. Airphaiboon, and K. Hamamoto. Robustness study of ECG biometric identification in heart rate variability conditions. *IEEJ Transactions on Electrical and Electronic Engineering*, 9(3):294 – 301, May 2014.

K. N. Plataniotis, D. Hatzinakos, and J. K. M. Lee. ECG Biometric Recognition Without Fiducial Detection. In *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6, Sept 2006. doi: 10.1109/BCC.2006.4341628.

F. Porée, G. Kervio, and G. Carrault. ECG biometric analysis in different physiological recording conditions. *Signal, Image and Video Processing*, 10(2):267–276, 2016. ISSN 1863-1711. doi: 10.1007/s11760-014-0737-1.

S. Pouryayevali, S. Wahabi, S. Hari, and D. Hatzinakos. On establishing evaluation standards for ECG biometrics. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3774–3778, May 2014. doi: 10.1109/ICASSP.2014.6854307.

S. Prabhakar, S. Pankanti, and A. K. Jain. Biometric recognition: security and privacy concerns. *IEEE Security Privacy*, 1(2):33–42, Mar 2003. ISSN 1540-7993. doi: 10.1109/MSECP.2003.1193209.

J. Ren, X. Jiang, and J. Yuan. Quantized fuzzy LBP for face recognition. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1503–1507, April 2015. doi: 10.1109/ICASSP.2015.7178221.

D. A. Reynolds, T. F. Quatieri, and R. B. Dunn. Speaker Verification Using Adapted Gaussian Mixture Models. *Digital Signal Processing*, 10(1):19 – 41, 2000. ISSN 1051-2004. doi: 10.1006/dspr.1999.0361.

D. Rezgui and Z. Lachiri. ECG biometric recognition using SVM-based approach. *IEEJ Transactions on Electrical and Electronic Engineering*, 11:S94–S100, 2016. ISSN 1931-4981. doi: 10.1002/tee.22241.

S. Saechia, J. Koseeyaporn, and P. Wardkein. Human Identification System Based ECG Signal. In *TENCON 2005 - 2005 IEEE Region 10 Conference*, pages 1–4, Nov 2005. doi: 10.1109/TENCON.2005.300986.

S. I. Safie, J. J. Soraghan, and L. Petropoulakis. Electrocardiogram (ECG) Biometric Authentication Using Pulse Active Ratio (PAR). *IEEE Transactions on Information Forensics and Security*, 6(4):1315–1322, Dec 2011. ISSN 1556-6013. doi: 10.1109/TIFS.2011.2162408.

K. Sakai, K. Yanai, S. Okada, and K. Nishii. Design of Seat Mounted ECG Sensor System for Vehicle Application. *SAE Int. J. Passeng. Cars – Electron. Electr. Syst.*, 6:342–348, 04 2013. doi: 10.4271/2013-01-1339.

P. Sasikala and R. Wahidabanu. Identification of Individuals using Electrocardiogram. *International Journal of Computer Science and Network Security*, 10(12):147 – 153, December 2010.

A. Savitzky and M. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36:1627–1639, 1964.

V. C. Scanlon and T. Sanders. *Essentials of Anatomy and Physiology*. F. A. Davis Company, Philadelphia, PA, fifth edition, 2007. ISBN 978-0-8036-1546-5.

R. W. Schafer. What Is a Savitzky-Golay Filter? [Lecture Notes]. *IEEE Signal Processing Magazine*, 28(4):111–117, July 2011. ISSN 1053-5888. doi: 10.1109/MSP.2011.941097.

R. J. A. Schijvenaars. *Intra-individual Variability of the Electrocardiogram: Assessment and exploitation in computerized ECG analysis*. PhD thesis, Erasmus University Rotterdam, 2000.

T.-W. D. Shen, W. J. Tompkins, and Y. H. Hu. One-lead ECG for identity verification. In *Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint*, volume 1, pages 62–63 vol.1, October 2002. doi: 10.1109/IEMBS.2002.1134388.

T.-W. D. Shen, W. J. Tompkins, and Y. H. Hu. Implementation of a one-lead ECG human identification system on a normal population. *Journal of Engineering and Computer Innovations*, 2 (1):12 – 21, 2011.

H. P. Silva, A. Fred, A. Lourenço, and A. K. Jain. Finger ECG signal for user authentication: Usability and performance. In *2013 IEEE Sixth International Conference on Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–8, Sept 2013. doi: 10.1109/BTAS.2013.6712689.

H. P. Silva, A. Lourenço, A. Fred, N. Raposo, and M. Aires-de-Sousa. Check Your Biosignals Here: A new dataset for off-the-person ECG biometrics. *Computer Methods and Programs in Biomedicine*, 113(2):503 – 514, 2014. ISSN 0169-2607. doi: 10.1016/j.cmpb.2013.11.017.

T. Sim, S. Zhang, R. Janakiraman, and S. Kumar. Continuous Verification Using Multimodal Biometrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):687–700, April 2007. ISSN 0162-8828. doi: 10.1109/TPAMI.2007.1010.

B. Singh, P. Singh, and S. Budhiraja. Various Approaches to Minimise Noises in ECG Signal: A Survey. In *2015 Fifth International Conference on Advanced Computing Communication Technologies*, pages 131–137, Feb 2015. doi: 10.1109/ACCT.2015.87.

Y. Singh and S. Singh. Evaluation of Electrocardiogram for Biometric Authentication. *Journal of Information Security*, 3(1):39–48, 2012. doi: 10.4236/jis.2012.31005.

F. Sufi, I. Khalil, and J. Hu. ECG-Based Authentication. In P. Stavroulakis and M. Stamp, editors, *Handbook of Information and Communication Security*, pages 309–331. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010. ISBN 978-3-642-04117-4. doi: 10.1007/978-3-642-04117-4_17.

F. Sufi, I. Khalil, and I. Habib. Cardioids-based faster authentication and diagnosis of remote cardiovascular patients. *Security and Communication Networks*, 4(11):1351–1368, 2011. ISSN 1939-0122. doi: 10.1002/sec.262.

Syris. Technical Document About FAR, FRR and EER. Technical report, Syris Technology Corporation, 2004.

A. Taddei, G. Distante, M. Emdin, P. Pisani, G. B. Moody, C. Zeelenberg, and C. Marchesi. The European ST-T Database: standard for evaluating systems for the analysis of ST-T changes in ambulatory electrocardiography. *European Heart Journal*, 13:1164–1172, 1992.

P. Tate. *Seeley's Principles of Anatomy and Physiology*. McGraw-Hill, New York, NY, second edition, 2009. ISBN 978–0–07–337819–0.

M. M. Tawfik, H. Selim, and T. Kamal. Human identification using time normalized QT signal and the QRS complex of the ECG. In *Communication Systems Networks and Digital Signal Processing (CSNDSP), 2010 7th International Symposium on*, pages 755–759, July 2010.

S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. Academic Press, fourth edition, 2009. ISBN 978-1-59749-272-0.

P. E. Trahanias. An approach to QRS complex detection using mathematical morphology. *IEEE Transactions on Biomedical Engineering*, 40(2):201–205, Feb 1993. ISSN 0018-9294. doi: 10.1109/10.212060.

A. van Oosterom, R. Hoekema, and G. Uijen. Geometrical factors affecting the interindividual variability of the ECG and the VCG. *Journal of Electrocardiology*, 33:219 – 227, 2000. ISSN 0022-0736. doi: 10.1054/jelc.2000.20356.

S. Wahabi, S. Pouryayevali, S. Hari, and D. Hatzinakos. On Evaluating ECG Biometric Systems: Session-Dependence and Body Posture. *IEEE Transactions on Information Forensics and Security*, 9(11):2002–2013, Nov 2014. ISSN 1556-6013.

T. Waili, R. M. Nor, A. W. B. A. Rahman, K. A. Sidek, and A. A. Ibrahim. Electrocardiogram Identification: Use a Simple Set of Features in QRS Complex to Identify Individuals. In P. Meesad, S. Boonkrong, and H. Unger, editors, *Recent Advances in Information and Communication Technology 2016: Proceedings of the 12th International Conference on Computing and Information Technology (IC2IT)*, pages 139–148. Springer International Publishing, Cham, 2016a. ISBN 978-3-319-40415-8. doi: 10.1007/978-3-319-40415-8_14.

T. Waili, R. M. Nor, H. Yaacob, K. Sidek, and A. W. A. Rahman. A hasty approach to ecg person identification. In *2016 International Conference on Computer and Communication Engineering (ICCCE)*, pages 267–271, July 2016b. doi: 10.1109/ICCCE.2016.65.

J. Wang, M. She, S. Nahavandi, and A. Kouzani. Human Identification From ECG Signals Via Sparse Representation of Local Segments. *IEEE Signal Processing Letters*, 20(10):937–940, Oct 2013. ISSN 1070-9908. doi: 10.1109/LSP.2013.2267593.

T. Wartzek, B. Eilebrecht, J. Lem, H. J. Lindner, S. Leonhardt, and M. Walter. ECG on the Road: Robust and Unobtrusive Estimation of Heart Rate. *IEEE Transactions on Biomedical Engineering*, 58(11):3112–3120, Nov 2011. ISSN 0018-9294. doi: 10.1109/TBME.2011.2163715.

G. Wübbeler, M. Stavridis, D. Kreiseler, R.-D. Bousseljot, and C. Elster. Verification of humans using the electrocardiogram. *Pattern Recognition Letters*, 28(10):1172 – 1175, 2007. ISSN 0167-8655. doi: 10.1016/j.patrec.2007.01.014.

N. Yager and T. Dunstone. The Biometric Menagerie. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(2):220–230, Feb 2010. ISSN 0162-8828. doi: 10.1109/TPAMI.2008.291.

C. Ye, M. T. Coimbra, and B. V. K. V. Kumar. Investigation of human identification using two-lead Electrocardiogram (ECG) signals. In *Biometrics: Theory Applications and Systems (BTAS), 2010 Fourth IEEE International Conference on*, pages 1–8, Sept 2010. doi: 10.1109/BTAS.2010.5634478.

J. Ye and S. Ji. Discriminant analysis for dimensionality reduction: An overview of recent developments. In *Biometrics*, pages 1–19. John Wiley and Sons, Inc., 2009. ISBN 9780470522356. doi: 10.1002/9780470522356.ch1.

Z. Zhang and D. Wei. A New ECG Identification Method Using Bayes' Theorem. In *TENCON 2006 - 2006 IEEE Region 10 Conference*, pages 1–4. Nov 2006. doi: 10.1109/TENCON.2006.344146.

Z. Zhao, L. Yang, D. Chen, and Y. Luo. A Human ECG Identification System Based on Ensemble Empirical Mode Decomposition. *Sensors*, 13(5):6832–6864, 2013. ISSN 1424-8220. doi: 10.3390/s130506832.

X. Zhou, Y. Lu, M. Chen, S. D. Bao, and F. Miao. A method of ECG template extraction for biometrics applications. In *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 602–605, Aug 2014. doi: 10.1109/EMBC.2014.6943663.