Electronic Devices and Systems for Monitoring of Diabetes and Cardiovascular Diseases

Gianmarco ANGIUS

Advisor: Prof. Luigi RAFFO
Curriculum: ING-INF/01 Elettronica
XXI Cycle
February 2009
Dedicated to -fra-
## Contents

1. Introduction ........................................................................................................................................... 14

2. Diabetes and Cardiovascular Diseases .............................................................................................. 16

   2.1 Diabetes ............................................................................................................................................. 17
       2.1.1 Related Complications .................................................................................................................. 17
       2.1.2 Risk Factors ................................................................................................................................. 18
       2.1.3 Interventions and Delivery Modes ............................................................................................... 19
           2.1.3.1 Preventing Type 1 Diabetes .................................................................................................... 19
           2.1.3.2 Preventing Type 2 Diabetes .................................................................................................... 19
           2.1.3.3 Detecting Diabetes or Prediabetes .......................................................................................... 19
           2.1.3.4 Managing Diabetes ................................................................................................................. 19
       2.1.4 Intensive Glucose Control ............................................................................................................. 19
       2.1.5 Diabetes Education ...................................................................................................................... 19
       2.1.6 Quality of Diabetes Care ............................................................................................................. 20

   2.2 Cardiovascular Diseases .................................................................................................................... 21
       2.2.1 Overview of Cardiovascular System ............................................................................................ 21
           2.2.1.1 The Heart ................................................................................................................................ 21
           2.2.1.2 The Vascular System .............................................................................................................. 22
           2.2.1.3 The Conduction System ........................................................................................................ 24
           2.2.1.4 Cardiovascular Control ....................................................................................................... 25
           2.2.1.5 Principles of Electrocardiography .......................................................................................... 26
           2.2.1.6 QRS Detection ....................................................................................................................... 31
               2.2.1.6.1 Algorithms Based on Amplitude and First Derivative ...................................................... 32
               2.2.1.6.2 Algorithms Based on First Derivative Only ................................................................. 34
               2.2.1.6.3 Algorithms Based on First and Second Derivative ......................................................... 35
       2.2.2 Sleep Apnoea .............................................................................................................................. 39
       2.2.3 Atrial Fibrillation ......................................................................................................................... 42
       2.2.4 Deep Vein Thrombosis and Pulmonary Embolism ...................................................................... 44
       2.2.5 Myocardial infarction or Heart Attack ......................................................................................... 47

3. Introduction to Signal Processing ......................................................................................................... 49

   3.1 Noise and Random Variables ........................................................................................................... 50
   3.2 Transforms ........................................................................................................................................ 51
   3.3 Convolution, Correlation and Covariance ......................................................................................... 54
   3.4 Fourier Analysis ............................................................................................................................... 57
       3.4.1 Discrete-Time Fourier Analysis .................................................................................................... 58
       3.4.2 Power Spectrum ........................................................................................................................ 60


3.4.3 Short-Term Fourier Transform: The Spectrogram ........................................ 61
3.5 Digital Filters ....................................................................................................... 62
   3.5.1 The Z-Transform ......................................................................................... 62
      3.5.1.1 Digital Transfer Function ...................................................................... 63
   3.5.2 Finite Impulse Response Filters ................................................................... 64
   3.5.3 Infinite Impulse Response Filters ................................................................. 65

4. **Wearable Devices and Systems for Cardiovascular Monitoring** .................. 67
   4.1 Overview of Wearable Devices ................................................................. 68
      4.1.2 Wearable Sensing of Physiological Parameters ............................... 68
   4.2 The Developed Wearable Device ................................................................. 70

5. **Continuous Detection and Screening of Sleep Apnea Disease** ................. 74
   5.1 Heart Rate Variability Analysis ................................................................. 75
   5.2 Detection of SA from frequency analysis of HRV ...................................... 79
   5.3 The Proposed System .................................................................................. 80

6. **Continuous Cuff-less Assessment of Blood Pressure from PPG signals** ...... 84
   6.1 Arterial Blood Pressure ............................................................................... 85
   6.2 Classical NIBP Measurements ................................................................... 85
      6.2.1 Auscultation ......................................................................................... 85
      6.2.2 Oscillometry ....................................................................................... 86
   6.3 Innovative NIBP Measurements ................................................................. 87
      6.3.1 Finapres ............................................................................................... 87
   6.4 Cuff-less NIBP Measurements .................................................................... 89
      6.4.1 Methods based on Pulse Transit Time .................................................. 89
      6.4.2 Methods based on PPG Signals ............................................................ 90
   6.5 Calibration of the PPG to BP for the Continuous Assessment in the Developed
   Wearable Device .............................................................................................. 91

7. **Continuous Monitoring of patients with HA, AF and DVT through Wavelet
   Transform of PPG signals** ............................................................................... 94
   7.1 Wavelet Analysis ........................................................................................ 95
   7.2 Wavelet Analysis in Biomedical Signal Processing ..................................... 102
   7.3 Wavelet Analysis of Acquired PPG signals ............................................... 103
      7.3.1 Analysis of Atrial Fibrillation PPG signals .......................................... 105
      7.3.2 Analysis of Deep Vein Thrombosis PPG signals ................................. 106
      7.3.3 Analysis of Heart Attack PPG signals .................................................. 107

8. **A DVB-T based System for Tele-Home Care Practice** ............................. 109
   8.1 Introduction to Interactive TV ....................................................................... 110
      8.1.1 Digital Video Broadcasting .................................................................. 110
      8.1.2 Digital Video Broadcasting Terrestrial ................................................. 111
      8.1.3 Multimedia Home Platform ................................................................. 113
         8.1.3.1 MHP Architecture ....................................................................... 113
List of Publications Related to the Thesis

List of Figures

**Figure 2.1:** Anterior view of the human heart showing the four chambers, the inlet and outlet valves, the inlet and outlet major blood vessels, the wall separating the right side from the left side, and the two cardiac pacing centers (the sinoatrial node and the atrioventricular node) [Dorf et al., 2006].

**Figure 2.2:** The cardiac microscopic capillary network.

**Figure 2.3:** The cardiac conduction system.

**Figure 2.4:** Electrical activity generated by the heart.

**Figure 2.5:** Einthoven leads.

**Figure 2.6:** Wilson’s central terminal and chest unipolar recordings.

**Figure 2.7:** Goldberger’s augmented leads.

**Figure 2.8:** ECG lead II typical nomenclature.

**Figure 2.9:** Common structure of the QRS detectors.

**Figure 2.10:** Obstructive apnoea.

**Figure 2.11:** Continuous Positive Airway Pressure for Obstructive Apnoea Treatment.

**Figure 2.12:** Overnight Polysomnography.

**Figure 2.13:** Comparison between a normal conduction and atrial fibrillation.

**Figure 2.14:** Atrial fibrillation events in a ECG recording.

**Figure 2.15:** Stroke caused by a clot due to atrial fibrillation disease.

**Figure 2.16:** A deep vein clot in femoral vein.

**Figure 2.17:** Automatic measurement of PT.

**Figure 2.18:** Pulmonary embolism.

**Figure 2.19:** Heart Attack caused by the occlusion of the coronary artery.

**Figure 3.1:** A waveform (upper plot) is multiplied by a window function (middle plot) to create a truncated version of the original waveform (lower plot).

**Figure 3.2:** The most popular windows function in signal processing.

**Figure 3.3:** Autocorrelation of two colored noise with different bandwidths.

**Figure 3.4:** Power Spectrum of a waveform composed of two sines (100Hz and 200Hz) sampled at 1Ksps.

**Figure 3.5:** Spectrogram of a 15Hz sine before $t = 1$ sec and 50Hz after $t = 1$ sec.

**Figure 3.6:** A unit delay process.

**Figure 3.7:** Butterwort,Chebyshev and Elliptic frequency response.

**Figure 4.1:** Examples of prototype of wearable medical devices/systems.

**Figure 4.2:** The h-shirt prototype.
**Figure 4.3:** The developed wearable device: block diagram (left), PCB top view (center) and PCB bottom view (right).

**Figure 4.4:** Standard SpO2 sensors: wrapper (left), fingertip (center) and ear lobe (right).

**Figure 4.5:** Bode diagram of the analog part of the wearable device developed.

**Figure 4.6:** The developed internet gateway.

**Figure 4.7:** Block diagram of the developed internet gateway.

**Figure 4.8:** The developed wireless USB dongle.

**Figure 4.9:** Block diagram of the developed wireless USB dongle.

**Figure 5.1:** Interval tachogram of 256 consecutive RR values in a normal subject at supine rest.

**Figure 5.2:** HRV spectra calculated by a FFT non-parametric algorithm (left plot) and by a parametric algorithm (right plot).

**Figure 5.3:** Automatic analysis and classification of SA events by the frequency analysis of HRV [Drinnan et al., 2000].

**Figure 5.4:** System overview for the continuous detection and screening of SA disease.

**Figure 5.5:** The glove embedding the wearable device.

**Figure 5.6:** NN intervals detection from lead-II ECG (A) and PPG (B).

**Figure 5.7:** The Graphic User Interface at the Remote Monitoring Station.

**Figure 6.1:** Using sphygmomanometer and stethoscope in auscultation method.

**Figure 6.2:** Cuff pressure trend in auscultatory method.

**Figure 6.3:** Cuff pressure trend in oscillometric method.

**Figure 6.4:** Finapres schema.

**Figure 6.5:** Finapres Finometer (left) and Finapres Portapres (right).

**Figure 6.6:** Definitions of PTT.

**Figure 6.7:** Arterial compliance curves based on parameters estimated from hydrostatic challenge.

**Figure 6.8:** Experimental calibration of the PPG signal to blood pressure in five different volunteers.

**Figure 6.9:** Normalized PPG versus blood pressure graph.

**Figure 7.1:** A Morlet mother wavelet (a = 1) with two dilations (a = 2 and 4) and one contraction (a = 0.5).

**Figure 7.2:** CWT of a 15Hz sine before t = 1 sec and 50Hz after t = 1 sec.

**Figure 7.3:** Haar, Morlet, Mexican Hat and Meyer wavelet [Misiti et al., 2008].

**Figure 7.4:** Coiflet wavelets [Misiti et al., 2008].

**Figure 7.5:** Symlet wavelets [Misiti et al., 2008].

**Figure 7.6:** Daubechies wavelets [Misiti et al., 2008].

**Figure 7.7:** Biorthogonal wavelets [Misiti et al., 2008].
Figure 7.8: Simple filter bank consisting of only two filters applied to the same waveform. The filters have low-pass and high-pass spectral characteristic.

Figure 7.9: A typical wavelet application using filter bank containing only two filters. The input waveform is first decomposed into sub-bands using the analysis filter bank. Some process is applied to the filtered signals before reconstruction. Reconstruction is performed by the synthesis filter bank.

Figure 7.10: A typical wavelet application using three filters where the downsampling and upsampling processes are shown.

Figure 7.11: Scatter plot of wavelet power at 1 Hz against entropy at 7 Hz showing differentiation of well children from unwell children.

Figure 7.12: PPG signal of a healthy subject acquired with the wearable device realized.

Figure 7.13: Wavelet decomposition used for the healthy and ill PPG signal analysis.

Figure 7.14: PPG signal of a subject with AF disease acquired with the wearable device realized.

Figure 7.15: Energy versus entropy of the DD2 level of the wavelet analysis of healthy PPGs and of PPGs from subjects with AF disease.

Figure 7.16: Normalized energy versus entropy of the DD2 level of the wavelet analysis of healthy PPGs and of PPGs from subjects with AF disease.

Figure 7.17: PPG signal of a subject with DVT disease acquired with the wearable device realized.

Figure 7.18: Energy versus entropy of the AD2 level of the wavelet analysis of healthy PPGs and of PPGs from subjects with DVT disease.

Figure 7.19: PPG signal of a subject with HA disease acquired with the wearable device realized.

Figure 7.20: Energy versus entropy of the AA2 level of the wavelet analysis of healthy PPGs and of PPGs from subjects with HA disease.

Figure 8.1: DVB standards over the world.

Figure 8.2: Scheme of a DVB-T transmission system.

Figure 8.3: MHP architecture.

Figure 8.4: MHP profiles.

Figure 8.5: Xlet lifecycle model.

Figure 8.6: ISIS architecture scheme.

Figure 8.7: Kyriacou’s architecture scheme.

Figure 8.8: U-R-Safe architecture scheme.

Figure 8.9: HEALTHWARE architecture scheme.

Figure 8.10: The main parts of the DVB-T based tele-home care realized system.

Figure 8.11: First prototype of the custom smart card.

Figure 8.12: The block diagram of our broadcast transmission system.

Figure 8.13: Picture of the Xlet sending two acquired glucometer measurements to the RDB.
Figure 8.14: Architecture of the first version of the system at patient’s home.

Figure 8.15: Primary FSF.

Figure 8.16: The ECG exam screen of the MHPHomecare application (ECG Exam FSF).

Figure 8.17: A schematic representation of the lead-I for the prototipal 3-lead ECG base station.

Figure 8.18: A screenshot of the prototipal 3-lead ECG base station (left) and RS-232 to IR adapter (right).

Figure 8.19: Results of the digital notch filter (left) and of the high-pass filter (right) on ECG signal acquired by the prototypal base station.
List of Tables

Table 5.1: Time domain measures of HRV analysis.
Table 5.2: Frequency domain measures of HRV analysis in a short-term recording.
Chapter 1

Introduction

Diabetes is a serious chronic disease which causes a high rate of morbidity and mortality all over the world. In 2007, more than 246 million people suffered from diabetes worldwide and unfortunately the incidence of diabetes is increasing at alarming rates. The number of people with diabetes is expected to double within the next 25 years due to a combination of population ageing, unhealthy diets, obesity and sedentary lifestyles. It can lead to blindness, heart disease, stroke, kidney failure, amputations and nerve damage. In women, diabetes can cause problems during pregnancy and make it more likely for the baby to be born with birth defects. Moreover, statistical analysis shows that 75% of diabetic patients die prematurely of cardiovascular disease (CVD). The absolute risk of cardiovascular disease in patients with type 1 (insulin-dependent) diabetes is lower than that in patients with type 2 (non-insulin-dependent) diabetes, in part because of their younger age and the lower prevalence of CVD risk factors, and in part because of the different pathophysiology of the two diseases. Unfortunately, about 9 out of 10 people with diabetes have type 2 diabetes. For these reasons, cardiopathes and diabetic patients need to be frequently monitored and in some cases they could easily perform at home the requested physiological measurements (i.e. glycemia, heart rate, blood pressure, body weight, and so on) sending the measured data to the care staff in the hospital. Several researches have been presented over the last years to address these issues by means of digital communication systems. The largest part of such works uses a PC or complex hardware/software systems for this purpose. Beyond the cost of such systems, it should be noted that they can be quite accessible by relatively young people but the same does not hold for elderly patients more accustomed to traditional equipments for personal entertainment such as TV sets.

Wearable devices can permit continuous cardiovascular monitoring both in clinical settings and at home. Benefits may be realized in the diagnosis and treatment of a number of major
diseases. In conjunction with appropriate alarm algorithms, they can increase surveillance capabilities for CVD catastrophe for high-risk subjects. Moreover, they could play an important role in the wireless surveillance of people during hazardous operations (military, fire-fighting, etc.) or during sport activities.

For patients with chronic cardiovascular disease, such as heart failure, home monitoring employing wearable device and tele-home care systems may detect exacerbations in very early stages or at dangerous levels that necessitate an emergency room visit and an immediate hospital admission.

Taking into account mains principles for the design of good wearable devices and friendly tele-home care systems, such as safety, compactness, motion and other disturbance rejection, data storage and transmission, low power consumption, no direct doctor supervision, it is imperative that these systems are easy to use and comfortable to wear for long periods of time.

The aim of this work is to develop an easy to use tele-home care system for diabetes and cardiovascular monitoring, well exploitable even by elderly people, which are the main target of a telemedicine system, and wearable devices for long term measuring of some parameters related to sleep apnoea, heart attack, atrial fibrillation and deep vein thrombosis.

Since set-top boxes for Digital Video Broadcast Terrestrial (DVB-T) are in simple computers with their Operating System, a Java Virtual Machine, a modem for the uplink connection and a set of standard ports for the interfacing with external devices, elderly, diabetics and cardiopathes could easily send their self-made exam to the care staff placed elsewhere.

The wearable devices developed are based on the well known photopletysmographic method which uses a led source/detector pair applied on the skin in order to obtain a biomedical signal related to the volume and percentage of oxygen in blood. Such devices investigate the possibility to obtain more information to those usually obtained by this technique (heart rate and percentage of oxygen saturation) in order to discover new algorithms for the continuous and remote or in ambulatory monitoring and screening of sleep apnoea, heart attack, atrial fibrillation and deep vein thrombosis.
Diabetes is a serious metabolic disease which affects people of all ages and reduces both person’s quality of life and life expectancy. By 2025, the worldwide prevalence is projected to be 6.3 percent, a number very very high. The 65 percent of all deaths in developed countries are caused by cardiovascular disease of people with diabetes.

Cardiovascular disease (CVD) is the number one cause of death worldwide [Mathers et al., 2006; Murray et al., 1996]. It covers a wide array of disorders, including diseases of the cardiac muscle and of the vascular system supplying the heart, brain, and other vital organs.

This chapter give a little review of diabetes and cardiovascular diseases, starting from some aspect of diabetes and ending with an overview of the cardiovascular system, principles of electrocardiography, algorithms for automatic QRS detection, and a little overview of some diseases of cardiovascular system which are the main field of application of the devices and system developed in this thesis.
2.1 Diabetes

Diabetes is a serious metabolic disease in which the body does not produce or properly use insulin [ADA, 2004]. Insulin is a naturally occurring hormone needed to transform sugar into energy required for daily life.

Diabetes takes three major forms.

- **Type 1 diabetes** results from destruction of the beta cells in the pancreas, leading to absolute insulin deficiency. It usually occurs in children and young adults and requires insulin treatment;
- **Type 2 diabetes**, which accounts for approximately 85 to 95 percent of all diagnosed cases, is usually characterized by insulin resistance in which target tissues do not use insulin properly;
- **Gestational diabetes**, is first recognized during pregnancy.

Other rare types of diabetes include those caused by genetic conditions, surgery, drug use, malnutrition, infections, and other illnesses.

The disease affects persons of all ages and reduces both a person’s quality of life and life expectancy and imposes a large economic burden on the health care system and on families. In 2003, the worldwide prevalence of diabetes was estimated at 5.1 percent among people age 20 to 79. By 2025, the worldwide prevalence is projected to be 6.3 percent, a 24 percent increase compared with 2003. The death rate of men with diabetes is 1.9 times the rate for men without diabetes, and the rate for women with diabetes is 2.6 times that for women without diabetes [Lee et al., 2000]. Premature mortality caused by diabetes results in an estimated 12 to 14 years of life lost [Manuel et al., 2004; Narayan et al., 2003]. Cardiovascular disease (CVD) causes up to 65 percent of all deaths in developed countries of people with diabetes [Geiss et al., 1995]. The World Health Organization (WHO) estimates that, in 2001, 959,000 deaths worldwide were caused by diabetes. More recent estimates by WHO suggest that the actual number may be triple this estimate [WHO, 2004].

2.1.1 Related Complications

Diabetes-related complications include:

- Microvascular diseases (e.g. retinopathy, blindness, nephropathy, and kidney failure);
- Macrovascular diseases (coronary heart disease, stroke, peripheral vascular disease and lower-extremity amputation);
- Influenza and pneumococcal infections;
- Eye complications (e.g. 39 percent of people with diabetes in India had eye-related complications [Rajala et al., 1998]).
Diabetes imposes large economic burdens on national health care systems and affects both national economies and individuals and their families.

- Direct medical costs include resources used to treat the disease. They range from 2.5 to 15.0 percent of annual health care budgets, depending on local prevalence and sophistication of the treatments available [IDF, 2003];
- Indirect costs include lost productivity caused by morbidity, disability, and premature mortality;
- Intangible costs refer to the reduced quality of life for people with diabetes brought about by stress, pain, and anxiety. With a value of 1 representing the health-related quality of life without illness and 0 representing death, people with type 2 diabetes had a value of 0.77 in the population of the United Kingdom prospective diabetes study [Clarke et al., 2002].

### 2.1.2 Risk Factors

Risk factors for diabetes vary by disease type:

- Type 1 diabetes is most likely a polygenic disease, and a number of potential environmental risk factors have been implicated, including dietary factors, infectious agents, chemicals and toxins [Akerblom et al., 1998];
- The risk for type 2 diabetes is higher in monozygotic twins and people with a family history of diabetes [Rich, 1990]. This finding strongly suggests that genetic determinants play a role, but so far few genes have been associated with type 2 diabetes.

Environmental factors include [Qiao et al., 2004]:

- Prenatal factors;
- Obesity;
- Physical inactivity;
- Dietary;
- Socioeconomic factors.

Exposure to diabetes in utero increases the risk of developing type 2 diabetes in early adulthood [Dabelea et al., 2000].

Sugar-sweetened beverages are associated with an increased risk of diabetes [Schulze et al., 2004]. High intakes of dietary fiber and of vegetables may reduce the risk of diabetes [Fung et al., 2002; Stevens et al., 2002].

The strongest and most consistent risk factors for diabetes and insulin resistance among different populations are obesity and weight gain [Haffner, 1998]: for each unit increase in body mass index, the risk of diabetes increases by 12 percent [Ford et al., 1997]. The distribution of fat around the trunk region, or central obesity, is also a strong risk factor for
diabetes [Yajnik, 2001]. Therefore, diabetes risk may be reduced by increasing physical activity. Conversely, a sedentary lifestyle and physical inactivity are associated with increased risks of developing diabetes [Hu et al., 2003].

2.1.3 Interventions and Delivery Modes

Interventions against diabetes include those for:

- preventing the disease;
- detecting the disease in its asymptomatic stage;
- managing the disease to reduce its complications.

2.1.3.1 Preventing Type 1 Diabetes

Not enough scientific evidence is available to indicate that type 1 diabetes can be prevented, although various interventions have been explored. Examples of tested interventions include eliminating or delaying exposure to bovine protein and using insulin or nicotinamide for people at high risk of developing the disease.

2.1.3.2 Preventing Type 2 Diabetes

Intensive lifestyle interventions involving a combination of diet and physical activity can delay or prevent diabetes among people at high risk [Eriksson et al., 1991; Knowler et al., 2002; Pan et al., 1997; Tuomilehto et al., 2001].

2.1.3.3 Detecting Diabetes or Prediabetes

The benefits of early detection of type 2 diabetes through screening are not clearly documented, nor is the choice of the appropriate screening test established. Questionnaires used alone tend to work poorly; biochemical tests alone or in combination with assessment of risk factors are a better alternative [Engelgau et al., 2000].

2.1.3.4 Managing Diabetes

Short follow-up of self-management training on knowledge, frequency, and accuracy of self-monitoring of blood glucose, self-reported dietary habits and glycemic control show high-quality evidence of positive effects.

2.1.4 Intensive Glucose Control

The aim of intensive glucose control is to lower the glucose level of a person with diabetes to a level close to that of a person without diabetes. The U.K. Prospective Diabetes Study
clearly demonstrates that lowering glucose levels can prevent or delay long-term diabetes complications [UKPDS, 1998].

2.1.5 Diabetes Education

People with diabetes play a central role in managing their disease. Thus, diabetes education is an integral part of diabetes care. The goal of diabetes education is to support the efforts of people with diabetes to understand the nature of their illness and its treatment, to identify emergency health problems at early, to adhere to self-care practices and to make necessary changes to their health habits [IDA, 2003]. Moreover, diabetes self-management reduces medical costs for diabetes care in the short term.

2.1.6 Quality of Diabetes Care

The quality of diabetes care generally remains suboptimal worldwide, regardless of a particular country’s level of development, health care system, or population [Garfield et al., 2003]. The Costs of Diabetes in Europe study, conducted in eight European countries, found suboptimal diabetes care in each country [Liebl et al., 2002]. In the United States, population-based surveys in the 1990s among adults age 18 to 75 with diabetes found that only 63 percent of them had had a dilated eye examination and only 55 percent had had a foot examination within the past year, 18 percent had poor glycemc control, 42 percent had good cholesterol control, and 66 percent had a blood pressure within the normal range [Saaddine et al., 2002].
2.2 Cardiovascular Diseases

2.2.1 Overview of Cardiovascular System

2.2.1.1 The Heart

The human heart is divided by a muscular wall into a right side and left side, each being one self-contained pumping station, but the two being connected in series (Fig. 2.1):

- the **left side** of the heart drives oxygen-rich blood through the aortic semilunar outlet valve into the **systemic circulation**, which carries the blood to each cell in the body from which it returns to the right side of the heart poor in oxygen and rich in carbon dioxide [Tortora et al., 1993];
- the **right side** of the heart then drives this oxygen-poor blood through the pulmonary semilunar (pulmonic) outlet valve into the pulmonary circulation, which carries the blood to the lungs where its oxygen supply is replenished and its carbon dioxide content is purged before it returns to the left side of the heart.

![Figure 2.1: Anterior view of the human heart showing the four chambers, the inlet and outlet valves, the inlet and outlet major blood vessels, the wall separating the right side from the left side, and the two cardiac pacing centers (the sinoatrial node and the atrioventricular node) [Dorf et al., 2006].](image)

Because of the anatomic proximity of the heart to the lungs, the right side of the heart does not have to work very hard to drive blood through the pulmonary circulation, so it functions
as a low-pressure ($P \leq 40$ mmHg) compared with the left side of the heart, which does most of its work at a high pressure (up to 140 mmHg) to drive blood through the entire systemic circulation to the furthest extremes of the organism [Caro et al., 1978].

Each cardiac side is divided into two chambers: a small upper receiving chamber, or atrium, separated by a one-way valve from a lower discharging chamber, or ventricle, which is about twice the size of its corresponding atrium. In order of size, the left atrium is the smallest chamber of the entire heart, holding at rest about 45 ml of blood and operating at pressures on the order of 0 to 25 mmHg. The right atrium is next (63 ml of blood and 0 to 10 mmHg of pressure), followed by the left ventricle (100 ml of blood and up to 140 mmHg of pressure) and the right ventricle (about 130 ml of blood and up to 40 mmHg of pressure).

During the filling phase (diastole), the inlet valves of the two ventricles (tricuspid valve from right atrium to right ventricle and bicuspid or mitral valve from left atrium to left ventricle) are open, and the outlet valves (pulmonary and aortic semilunar valve) are closed and the heart can expand to its end-diastolic-volume (EDV) [Kaley et al., 1977]. During the systole, electrically induced vigorous contraction of cardiac muscle drives the intraventricular pressure up, forcing the one-way inlet valves closed and the unidirectional outlet valves open as the heart contracts to its end-systolic-volume (ESV). Thus the ventricles normally empty about half their contained volume with each heart beat, the remainder being termed the cardiac reserve volume. The difference between the actual EDV and the actual ESV, called the stroke volume (SV), is the volume of blood expelled from the heart during each systolic interval. The ratio of SV to EDV is called the cardiac ejection ratio [Lentner, 1980]. If the stroke volume is multiplied by the number of systolic intervals per minute, or heart rate (HR), one obtains the total cardiac output (CO):

$$CO = HR \times (EDV - ESV)$$

(2.1)

### 2.2.1.2 The Vascular System

The vascular system is divided by a microscopic capillary (Fig. 2.2) network into an high-pressure upstream viscoelastic tubes that carry blood away from the heart, and a low-pressure downstream elastic conduits that return blood back to the heart. Except for their differences in thickness, the walls of the largest arteries and veins consist of the same three distinct, well-defined, and well-developed layers. These layers are:

- the thinnest tunica intima, a continuous lining (the vascular endothelium) consisting of a single layer of simple squamous endothelial cells;
- the thickest tunica media, composed of numerous circularly arranged elastic fibers, especially prevalent in the largest blood vessels on the arterial side (allowing them to expand during systole and to recoil passively during diastole);
- the medium-sized tunica adventitia, consisting entirely of connective tissue.
The largest blood vessels, such as the aorta, the pulmonary artery, the pulmonary veins, and others, have such thick walls that they require a separate network of tiny blood vessels just to service the vascular tissue itself [Chandran et al., 1992]. The thick-walled large arteries and main distributing branches are designed to withstand the pulsating 80 to 130 mmHg blood pressures that they must endure. The smaller elastic conducting vessels need only operate under steadier blood pressures in the range 70 to 90 mmHg, but they must be thin enough to penetrate and course through organs without unduly disturbing the anatomic integrity of the mass involved.

Figure 2.2: The cardiac microscopic capillary network.

It is not surprising that, at any given instant, about two-thirds of the total human blood volume residing in the venous system and the remaining one-third being divided among the heart (6.5%), the microcirculation (7% in systemic and pulmonary capillaries), and the arterial system (19.5 to 20%).

Using an electrical analogy, one can think of the human cardiovascular system as a voltage source (the heart), two capacitors (a large venous system and a smaller arterial system), and a resistor (the microcirculation taken as a whole). Blood flow and the dynamics of the system represent electrical inductance, and useful engineering approximations can be derived from such a simple model.
2.2.1.3 The Conduction System

The conductive system of the heart is composed of specialized fibers that can be subdivided into excitation and conduction fiber (Fig. 2.3). It consists of the sinoatrial (SA) node, intermodal tracks, Bachmann’s bundle, the atrioventricular (AV) node, the bundle of His, bundle branches and Purkinje fibers. Cardiac cells are able to depolarize at a rate specific for the cell type. The intrinsic rate of AV-nodal cells is about 50 beats per minute, whereas Purkinje fibers depolarize at a rate of no more than 40 bpm. During normal sinus rhythm, the heart is controlled by the SA node having the highest intrinsic rate of 60/100 bpm, depending on the hemodynamic demand. The right atrial internodal tracks and Bachmann’s bundle conduct the SA-nodal activation throughout the atria, initiating a coordinated contraction of the atrial walls. Meanwhile, the impulse reaches the AV node, which is the only electrical connection between atria and ventricles. The AV node introduces an effective delay, allowing the contraction of the atria to complete before ventricular contraction is initiated. Due to this delay, an optimal ventricular filling is achieved. Subsequently, the electrical impulse is conducted at a high velocity by the His-Purkinje system comprising the bundle of His, bundle branches, and Purkinje fibers. Once the bundle of His is activated, the impulse splits into the right bundle branch, which leads to the right ventricle and the left bundle branch serving the left ventricle. Both bundle branches terminate in Purkinje fibers [Guyton et al., 1996].

Figure 2.3: The cardiac conduction system.
2.2.1.4 Cardiovascular Control

Control of cardiovascular function is accomplished by mechanisms that are based either on the inherent physicochemical attributes of the tissues and organs themselves (intrinsic control) or on responses that can be attributed to the effects on cardiovascular tissues of other organ systems in the body (extrinsic control). For example, the accumulation of wastes and depletion of oxygen and nutrients that accompany the increased rate of metabolism in an active tissue both lead to an intrinsic relaxation of local capillary with a consequent reduction of the local resistance to flow and thereby allows more blood to perfuse the active region. On the other hand, the extrinsic innervation by the autonomic nervous system of smooth muscle tissues in the walls of arterioles allows the central nervous system to completely shut down the flow to entire vascular beds when this becomes necessary (such as during exposure to extremely cold environments). In addition to controlling the distribution of blood, physiologic regulation of cardiovascular function is directed mainly at four other variables:

- cardiac output;
- blood pressure;
- blood volume;

Cardiac output can be increased by increasing the heart rate, increasing the end-diastolic volume, decreasing the end-systolic volume, or doing all three things at once. Under the extrinsic influence of the sympathetic nervous system, HR can triple, EDV can increase by as much as 50% and ESV can decrease a comparable amount to about 30 to 35 ml or less.

The control of blood pressure is accomplished mainly by adjusting at the arteriolar level the downstream resistance to flow. This effect is conveniently quantified by a fluid-dynamic analogue to Ohm’s famous \( V = R \times I \) law in electrotechnics theory. Voltage drop \( V \) being equated to fluid pressure drop \( P \), electric current \( I \) corresponding to flow and electric resistance \( R \) being associated with an analogous vascular peripheral resistance.

The control of blood volume is accomplished mainly through the excretory function of the kidney. For example, antidiuretic hormone secreted by the pituitary gland acts to prevent renal fluid loss and thus increases plasma volume, whereas perceived extracellular fluid overloads such as those which result from the peripheral vasoconstriction response to cold stress lead to a sympathetic receptor-induced renal diuresis (urination) that tends to decrease plasma volume to sometimes dangerously low dehydration levels.
2.2.1.5 Principles of Electrocardiography

The electrocardiogram (ECG) is the recording on the body surface of the electrical activity generated by the heart (Fig. 2.4).

![Figure 2.4: Electrical activity generated by the heart.](image)

It was originally observed by Waller in 1889 [Waller, 1889] using his pet bulldog as the signal source and the capillary electrometer as the recording device. In 1903 Einthoven [Einthoven, 1903] enhanced the technology by using the string galvanometer as the recording device and using human subjects with a variety of cardiac abnormalities. Einthoven is chiefly responsible for introducing some concepts still in use today including the labeling of the various waves, defining some of the standard recording sites, and developing the first theoretical construct whereby the heart is modeled as a single time-varying dipole.

In order to record an ECG waveform, a differential recording between two points on the body are made. Traditionally each differential recording is referred to as a lead. Einthoven defined three leads numbered with the Roman numerals I, II, and III.
Figure 2.5: Einthoven leads.

They are defined as:

\[ I = V_{LA} - V_{RA} \]  \hspace{1cm} (2.2)
\[ II = V_{LL} - V_{RA} \]  \hspace{1cm} (2.3)
\[ III = V_{LL} - V_{LA} \]  \hspace{1cm} (2.4)
\[ III = I + III \]  \hspace{1cm} (2.5)

where RA = right arm, LA = left arm, LL = left leg. Because the body is assumed to be purely resistive at ECG frequencies, the four limbs can be thought of as wires attached to the torso. Hence, lead I could be recorded from the respective shoulders without a loss of cardiac information.

For 30 years the evolution of the ECG proceeded when Wilson \cite{Wilson et al., 1934} added concepts of a unipolar recordings (Fig. 2.6). He created a reference point by tying the three limbs together and averaging their potentials so that individual recording sites on the limbs or chest surface would be differentially recorded with the same reference point.
Wilson extended the biophysical models to include the concept of the cardiac source enclosed within the volume conductor of the body. From the mid-1930s until today the 12 leads composed of the 3 limb leads, 3 leads in which the limb potentials are referenced to a modified Wilson terminal [Goldberger, 1942] (the augmented leads (Fig. 2.7)), and 6 leads placed across the front of the chest and referenced to the Wilson terminal (Fig. 2.6) form the basis of the standard 12-lead ECG. The voltage difference from any two sites will record an ECG, but it is these standardized sites with the massive 90-year collection of empirical observations that has firmly established their role as the standard. A typical ECG recording from lead II is shown in Fig. 2.8.
Einthoven chose the letters of the alphabet from P–U to label the waves and to avoid conflict with other physiologic waves being studied at the turn of the century. The ECG signals are typically in the range of \(2\text{mV}\) and require a recording bandwidth of 0.05 to 100 Hz. Full technical specification for ECG equipment has been proposed by both the American Heart Association [Bailey et al., 1990] and the Association for the Advancement of Medical Instrumentation [AAMI, 1993].

Other subsets of the 12-lead ECG are used in limited mode recording situations such as the tape-recorded ambulatory ECG (usually 2 leads) or in intensive care monitoring at the bedside (usually 1 or 2 leads) or telemetered within regions of the hospital from patients who are not confined to bed (1 lead).

More recently the high resolution ECG has been developed whereby the digitized ECG is signal averaged to reduce random noise [Berbari et al., 1973; Berbari et al., 1977]. This approach, coupled with post-averaging high-pass filtering, is used to detect and quantify low-level signals (1.0 μV) not detectable with standard approaches.

There are hundreds of interpretive statements from which a specific diagnosis is made for each ECG, but there are only about five or six major classification groups for which the ECG is used. The first step in analyzing an ECG requires the determination of the rate and rhythm for the atria and ventricles. Included here would be any conduction disturbances either in the relationship between the various chambers or within the chambers themselves. Then
one would proceed to identify features that would relate to the presence or absence of scarring due to a myocardial infarction. There may also be evidence of acute events occurring that would occur with ischemia or an evolving myocardial infarction.

ECG signals may be corrupted by various kinds of noise.

- power line interference;
- electrode contact noise;
- motion artifacts;
- muscle contraction (electromyographic, EMG);
- baseline drift and ECG amplitude modulation with respiration;
- instrumentation noise generated by electronic devices used in signal processing;
- electrosurgical noise, and other, less significant noise sources [1].

**Power line interference** consists of 50/60 Hz pickup and harmonics which can be modeled as sinusoids and combination of sinusoids [2].

**Electrode contact noise** is transient interference caused by loss of contact between the electrode and skin, which effectively disconnects the measurement system from the subject. The loss of contact can be permanent, or can be intermittent, as would be the case when a loose electrode is brought in and out of contact with the skin as a result of movements and vibration. This switching action at the measurement system input can result in large artifacts since the ECG signal is usually coupled to the system. With the amplifier input disconnected, 50 Hz interference may be significant.

**Motion artifacts** are transient baseline changes caused by changes in the electrode-skin impedance with electrode motion. As this impedance changes, the ECG amplifier sees a different source impedance, which forms a voltage divider with the amplifier input impedance. Therefore, the amplifier input voltage depends on the source impedance, which changes as the electrode position changes. The usual cause of motion artifacts will be assumed to be vibrations or movement of the subject. The shape of the baseline disturbance caused by motion artifacts can be assumed to be a biphasic signal resembling one cycle of a sine wave. The peak amplitude and duration of the artifact are variables.

**Muscle contractions** cause artifact potentials to be generated. The baseline ECG is usually in the microvolt range and therefore is usually insignificant. The signals resulting from muscle contraction can be assumed to be transient bursts of zero-mean band-limited Gaussian noise. The variance of the distribution may be estimated from the variance and duration of the bursts.
The drift of the baseline with respiration can be represented as a sinusoidal component at the frequency of respiration added to the ECG signal. The amplitude and frequency of the sinusoidal component should be variables. The amplitude of the ECG signal also varies by about 15% with respiration.

When an ECG is corrupted by artifacts generated by electronic devices used in signal processing, the input amplifier has saturated and no information about the ECG can be reach.

Electrosurgical noise completely destroys the ECG and can be represented by a large amplitude sinusoid with frequencies approximately between 100 kHz and 1 MHz. Since the sampling rate of an ECG signal is 250 to 1000 Hz, an aliased version of this signal would be added to the ECG signal. The amplitude, duration, and possibly the aliased frequency should be variable.

2.2.1.6 QRS Detection
The QRS complex is the most striking waveform within the electrocardiogram. Since it reflects the electrical activity within the heart during the ventricular contraction, the time of its occurrence as well as its shape provide much information about the current state of the heart. Moreover, it serves as the basis for the automated determination of the heart rate, as an entry point for classification schemes of the cardiac cycle, and often it is also used in ECG data compression algorithms. In that sense, QRS detection provides the fundamentals for almost all automated ECG analysis algorithms.

Software QRS detection has been a research topic for more than 30 years. Within the last decade many new approaches to QRS detection have been proposed, such as algorithms from the field of artificial neural networks [Hu et al., 1993; Strintzis et al., 1992; Vijaya et al., 1998; Xue et al., 1992], genetic algorithms [Poli et al., 1995], wavelet transforms [Kadambe et al., 1999; Li et al., 1995; Bahoura et al., 1997; Cellar et al., 1997; Crowe et al., 1992] and filter banks [Afonso et al., 1999].

Typical frequency components of a QRS complex range from about 10 Hz to about 25 Hz. Therefore, almost all QRS detection algorithms use a filter stage prior to the actual detection stage in order to attenuate other signal components and artifacts, such as P-wave, T-wave, baseline drift, and coupling noise (Fig. 2.9).
Figure 2.9: Common structure of the QRS detectors.

Whereas the attenuation of the P- and T-wave as well as baseline drift requires high-pass filtering, the suppression of coupling noise is usually accomplished by a low-pass filter. The combination of low and high pass means effectively the application of a band-pass filter, in this case with cut-off frequencies at about 10 Hz and 25 Hz.

2.2.1.6.1 Algorithms Based on Amplitude and First Derivative

Let $x(0), x(1), \ldots, x(8191)$ represent a one-dimensional array of sample points of the synthesized digitized ECG.

**A1: Morlet-Mahoudeaux [Mahoudeaux et al., 1981].**

An amplitude threshold is calculated as a fraction of the largest positive valued element of that array:

$$\text{Amp. Th.} = 0.3 \times \max x(n) \quad \text{for} \quad 0 < n < 8191 \quad (2.6)$$

The first derivative $y(n)$ is calculated at each point of $x(n)$ such that

$$y(n) = x(n+1) - x(n-1) \quad \text{for} \quad 1 < n < 8190 \quad (2.7)$$

A QRS candidate occurs when three consecutive points in the first derivative array exceed a positive slope threshold and are followed within the next 100ms by two consecutive points which exceed the negative threshold. All data points in the ECG between the onset of the rising slope and before the end of the descending slope must meet or exceed the amplitude threshold.

$$y(i), \; y(i+1), \; y(i+2) > 0.5 \quad (2.8)$$

and

$$y(j), \; y(j+1) < -0.3 \quad \text{where} \quad (i+2) < j < (i+25) \quad (2.9)$$

and
\[ x(i), x(i + 1), ..., x(j + 1) \geq Amp. Th. \] \hspace{1cm} (2.10)

**A2:** Fraden-Neuman [Fraden et al., 1988].
A threshold is calculated as a fraction of the peak value of the ECG.

\[ Amp. Th. = 0.4 \times \max[x(n)] \text{ for } 0 < n < 8191 \] \hspace{1cm} (2.11)

\[ y_0(n) = x(n) \text{ if } x(n) \geq 0 \text{ for } 0 < n < 8191 \] \hspace{1cm} (2.12)

\[ y_0(n) = -x(n) \text{ if } x(n) < 0 \text{ for } 0 < n < 8191 \] \hspace{1cm} (2.13)

\[ y_1(n) = y_0(n) \text{ if } y_0(n) \geq Amp. Th. \] \hspace{1cm} (2.14)

\[ y_1(n) = Amp. Th. \text{ if } y_0(n) < Amp. Th. \] \hspace{1cm} (2.15)

The first derivative is calculated at each point of the clipped, rectified array:

\[ y_2(n) = y_1(n + 1) - y_1(n - 1) \text{ for } 1 < n < 8190 \] \hspace{1cm} (2.16)

A QRS candidate occurs when a point in \( Y_2(n) \) exceeds the fixed constant threshold:

\[ y_2(i) > 0.7 \] \hspace{1cm} (2.17)

**A3:** Gustafson [Gustafson et al., 1977].
The first derivative is calculated at each point of the ECG:

\[ y(n) = x(n + 1) - x(n - 1) \text{ for } 1 < n < 8190 \] \hspace{1cm} (2.18)

The first derivative array is then searched for points which exceed a constant threshold:

\[ y(i) \geq 0.15 \] \hspace{1cm} (2.19)

Then the next three derivative values \( y(i + 1), y(i + 2), y(i + 3) \) must also exceed 0.15. If the above conditions are met, point \( i \) can be classified as a QRS candidate if the next two sample points have positive slope amplitude products:

\[ y(i + 1)x(i + 1) \text{ and } y(i + 2)x(i + 2) > 0 \] \hspace{1cm} (2.20)
2.2.1.6.2 Algorithms Based on First Derivative Only

B1: Menard [Menard et al., 1981].
The first derivative is calculated for each point of the ECG, using the formula specified by:

\[ y(n) = -2x(n - 2) - x(n - 1) + x(n + 1) + 2x(n + 2) \quad \text{for} \quad 2 < n < 8189 \quad (2.21) \]

The slope threshold is calculated as a fraction of the maximum slope for the first derivative array.

\[ \text{SlopeTh.} = 0.7 \times \max(y(n)) \quad \text{for} \quad 2 < n < 8189 \quad (2.22) \]

The first derivative array was searched for points which exceed the slope threshold. The first point that exceeds the slope threshold is taken as the onset of a QRS candidate:

\[ y(i) > \text{SlopeTh.} \quad (2.23) \]

B2: Holsinger [Holsinger et al., 1971].
The first derivative is calculated for the ECG:

\[ y(n) = x(n + 1) - x(n - 1) \quad \text{for} \quad 1 < n < 8190 \quad (2.24) \]

This array is searched until a point is found that exceeds the slope threshold:

\[ y(i) > 0.45 \quad (2.25) \]

A QRS candidate occurs if another point in the next three sample points also exceeds the threshold:

\[ y(i + 1) > 0.45 \quad (2.26) \]

or

\[ y(i + 2) > 0.45 \quad (2.27) \]

or

\[ y(i + 3) > 0.45 \quad (2.28) \]
2.2.1.6.3 Algorithms Based on First and Second Derivative

**C1: Balda** [Balda et al., 1977].
The absolute values of the first and second derivative are calculated from the ECG:

\[ y_0(n) = abs[x(n + 1) - x(n - 1)] \quad \text{for} \quad 2 < n < 8189 \]  
\[ y_1(n) = abs[x(n + 2) - 2x(n) + x(n + 2)] \quad \text{for} \quad 2 < n < 8189 \]  

These two arrays are scaled and then summed:

\[ y_2(n) = 1.3y_0(n) + 1.1y_1(n) \quad \text{for} \quad 2 < n < 8189 \]  

This array is scanned until a threshold is met or exceeded:

\[ y_2(t) \geq 1.0 \]  

Once this occurs, the next eight points are compared to the threshold. If six or more of these eight points meet or exceed the threshold, the criteria for identification of a QRS candidate is met.

**C2: Ahlstrom-Tompkins** [Ahlstrom et al., 1983].
The rectified first derivative is calculated from the ECG:

\[ y_0(n) = abs[x(n + 1) - x(n - 1)] \quad \text{for} \quad 3 < n < 8188 \]  

The rectified first derivative is then smoothed:

\[ y_1(n) = \frac{y_0(n-1)+2y_0(n)+y_0(n+1)}{4} \quad \text{for} \quad 3 < n < 8188 \]  

The rectified second derivative is calculated:

\[ y_2(n) = abs[x(n + 2) - 2x(n) + x(n - 2)] \quad \text{for} \quad 3 < n < 8188 \]  

The rectified, smoothed first derivative is added to the rectified second derivative:

\[ y_3(n) = y_1(n) + y_2(n) \quad \text{for} \quad 3 < n < 8188 \]  

The maximum value of this array is determined and scaled to serve as primary and secondary thresholds:
\[ PrimaryTh. = 0.8 \times \max[y_3(n)] \quad for \quad 3 < n < 8188 \quad (2.37) \]

\[ SecondaryTh. = 0.1 \times \max[y_3(n)] \quad for \quad 3 < n < 8188 \quad (2.38) \]

The array of the summed first and second derivatives is scanned until a point exceeds the primary threshold. In order to be classified as a QRS candidate, the next six consecutive points must all meet or exceed the secondary threshold:

\[ y_3(i) \geq PrimaryTh. \quad (2.39) \]

and

\[ y_3(i + 1), y_3(i + 2), \ldots, y_3(i + 6) > SecondaryTh. \quad (2.40) \]

2.2.1.6.4 Algorithms Based on Digital Filters

**D1: Engelese-Zeelenberg** [Engelese et al., 1979].

The ECG is passed through a differentiator with a 62.5 Hz notch filter.

\[ y_0(n) = x(n) - x(n - 4) \quad for \quad 4 < n < 8191 \quad (2.41) \]

The differentiated, filtered data is then passed through a digital low-pass filter.

\[ y_1(n) = y_0(n) + 4y_0(n - 1) + 6y_0(n - 2) + 4y_0(n - 3) + y_0(n - 4) \quad (2.42) \]

Two thresholds are used, equal in magnitude but opposite in polarity. The output of the low-pass filter is scanned until a point with amplitude greater than the positive threshold is reached. This point is the onset of a 160ms search region. The number of alternate threshold crossings is used to classify the initial crossing as either a baseline shift, a QRS candidate, or as noise:

If \( y_1(i) > 21.0 \), then search region onset = \( i \) \quad (2.43)

If no other threshold crossings occur within the 160ms search region, the occurrence is classified as a baseline shift. Otherwise, the following three conditions are tested:

1) If \( y_1(i + j) < -21.0 \) \quad for \quad 0 < j < 40 \quad (2.44)

2) If \( y_1(i + j) < -21.0 \) \quad for \quad 0 < j < 40 \quad (2.45)
and

\[ y_1(1 + k) > 21.0 \quad \text{for} \quad j < k < 40 \quad (2.46) \]

3) If \( y_1(i + j) < -21.0 \quad \text{for} \quad 0 < j < 40 \)

and

\[ y_1(i + k) > 21.0 \quad \text{for} \quad j < k < 40 \quad (2.48) \]

and

\[ y_1(i + 1) < -21.0 \quad \text{for} \quad k < 1 < 40 \quad (2.49) \]

If any of the above conditions apply, the occurrence is classified as a QRS candidate. If additional threshold crossings occur, the occurrence is classified as noise.

**D2: Okada [Okada, 1979].**

The first stage smooths the ECG using a three-point moving average filter:

\[ y_0(n) = \frac{x(n-1) + 2x(n) + x(n+1)}{4} \quad \text{for} \quad 1 < n < 8191 \quad (2.50) \]

The output of the moving averaging filter is passed through a low-pass filter.

\[ y_1(n) = \frac{1}{2m+1} \sum_{k=n-m}^{n+m} y_0(k) \quad \text{for} \quad m < n < (8191 - m) \quad (2.51) \]

The difference between the input and output of the lowpass filter is squared:

\[ y_2(n) = (y_0(n) - y_1(n))^2 \quad \text{for} \quad m < n < (8191 - m) \quad (2.52) \]

The squared difference is filtered:

\[ y_3(n) = y_2(n) (\sum_{k=n-m}^{n+m} y_2(k))^2 \quad \text{for} \quad m < n < (8191 - m) \quad (2.53) \]

A fourth array is formed using the following formula:

\[ y_4(n) = y_3(n) \quad \text{if} \quad [y_0(n) - y_0(n - m)][y_0(n) - y_0(n + m)] > 0 \quad (2.54) \]

\[ y_4(n) = 0 \quad \text{otherwise} \]
The maximum value of this array is determined and scaled to form the threshold:

\[
Th. = 0.125 \times \max[y_4(n)] \quad \text{for} \quad m < n < (8191 - m)
\]  

(2.55)

A QRS candidate occurs when a point in \( Y_4(n) \) exceeds the threshold:

\[
\text{If } y_4(n) > \text{Th.} \quad \text{then} \quad \text{QRS candidate}
\]  

(2.56)

Okada suggests setting \( m \) equal to three for best results. However, it was determined using the tuning procedure that larger values of \( m \) result in improved performance for several types of noise.
2.2.2 Sleep Apnoea

Sleep Apnoea (SA) is a common sleep disorder characterized by brief interruptions of breathing during sleep, with an estimated prevalence of about 4% in men and 2% in women [Young et al., 1993]. It can cause a total blockage of throat’s airway (i.e. apnoea) or a partial blockage of the airway which reduces the amount of oxygen that is taken into the body by 50% (i.e. hypopnoea). The amount of time in which a person stops breathing during the sleep is usually 10 seconds or longer and sometimes for as long as a minute [Shahar et al., 1991].

There are two major forms of SA:
- Obstructive Sleep Apnoea (OSA);
- Central Sleep Apnoea (CSA).

Obstructive apnoeas and hypopnoeas result from complete or partial collapse of a narrowed pharynx during sleep [Bradley et al., 1986] (Fig. 2.10).

![Figure 2.10: Obstructive apnoea.](image)

In subjects with normal pharyngeal anatomy, the partial withdrawal of pharyngeal dilator muscle tone that accompanies the onset of sleep is insufficient to cause pharyngeal collapse. However, the pharynx of patients with OSA is anatomically narrowed and highly compliant. In this setting, the superimposition of the normal withdrawal of pharyngeal dilator muscle tone at sleep onset causes the pharynx to occlude, triggering apnoea or hypopnoea [Remmers et al., 1978].

Treatments for obstructive sleep apnoea include Continuous Positive Airway Pressure (CPAP), a machine that delivers air pressure through a mask placed over the nose while sleeping (Fig. 2.11). With CPAP, the air pressure is somewhat greater than that of the surrounding air, and is just enough to keep upper airways passages open, preventing apnoea and snoring.
Central Sleep Apnoea, commonly referred to as Cheyne-Stokes respiration, is a form of periodic breathing in which apnoeas and hypopnoeas alternate with periods of hyperventilation. It is usually initiated during sleep by a further acute increase in ventilation and reduction in arterial carbon dioxide pressure (PaCO2) because the brain fails to send the appropriate signals to the breathing muscles to initiate respiration [Naughton et al., 1993]. When PaCO2 falls below the threshold level required to stimulate breathing, the central drive to respiratory muscles and airflow stop, and central apnoea ensues [Bradley et al., 1992].

SA is associated with a wide range of health implications and increased cardiovascular diseases and mortality. It has been linked to depression, irritability, sexual dysfunction, learning and memory difficulties, high blood pressure (hypertension), heart attack and stroke. At least 50% of patients with heart failure have OSA or CSA, responsible to the adverse consequences of the cardiovascular system [Sin et al., 1999]. It is important to notice that the effects of sleep apnoea on the cardiovascular system are not confined to sleep. Daytime sympathetic nervous activity and blood pressure, in fact, are increased in patients with OSA [Somers et al., 1995]. A continuous screening of sleep apnoea disorder is needed in order to have a major benefit of the treatment on cardiovascular outcomes.

The gold standard in diagnosing of SA is overnight polysomnography (PSG) (Fig. 2.12). It uses non-invasive methods for the recording of various physiological parameters including EEG, ECG, EMG, nasal airflow, abdominal and thoracic movements and blood oxygen saturation (SpO2) [Haponik et al., 1984].

Figure 2.11: Continuous Positive Airway Pressure for Obstructive Apnoea Treatment.
Fig. 2.12 shows that as a result of sleep apnoea, the oxygen level in the bloodstream falls and carbon dioxide which is a waste product of the body's chemical processes, accumulates. After a while the changes in these two chemicals in the bloodstream stimulate breathing again and the patient starts to make respiratory efforts.
2.2.3 Atrial Fibrillation

Atrial Fibrillation (AF) is a supraventricular arrhythmia characterized by uncoordinated atrial activation with consequent deterioration of atrial mechanical function [Fuster et al., 2001]. An arrhythmia is a problem with the speed or rhythm of the heartbeat. AF occurs when rapid, disorganized electrical signals in the heart’s two upper chambers cause them to contract very fast and irregularly (Fig. 2.13). As a result, blood pools in the atria and isn’t pumped completely into the heart’s two lower chambers (ventricles). When this happens, the heart’s upper and lower chambers don’t work together as they should [Wyse et al., 2002].

![Comparison between normal conduction and atrial fibrillation.](image)

There are three types of atrial classification:

- **paroxysmal atrial fibrillation**: the abnormal electrical signals and rapid heart rate begin suddenly and then stop on their own;
- **persistent atrial fibrillation**: the abnormal heart rhythm continues until it is stopped with treatment;
- **permanent atrial fibrillation**: the normal heart rhythm cannot be restored with the usual treatments;

Both paroxysmal and persistent atrial fibrillation may become more frequent and eventually result in permanent AF.

On the electrocardiogram, AF is described by the replacement of consistent P waves by rapid oscillations that vary in size, shape, and timing, associated with an irregular, frequently rapid ventricular response [Fuster et al., 2001] (Fig. 2.14).
AF can be isolated or associated with other arrhythmias, often atrial flutter or atrial tachycardia. Atrial flutter can arise during treatment with antiarrhythmic agents prescribed to prevent recurrent AF. Untreated, the atrial rate typically ranges from 240 to 320 beats per minute (bpm).

The clinical risk factors for AF include advancing age, diabetes, hypertension, congestive heart failure, rheumatic and nonrheumatic valve disease, myocardial infarction and stroke [Capucci et al., 1995]. In fact, during AF, the atria don’t pump all of their blood to the ventricles and some blood may pools in the atria. When this happens, a blood clot (a thrombus) can form. If the clot breaks off and travels to the brain, it can cause a stroke (Fig. 2.15).

Figure 2.14: *Atrial fibrillation events in a ECG recording.*

Figure 2.15: *Stroke caused by a clot due to atrial fibrillation disease.*
2.2.4 Deep Vein Thrombosis and Pulmonary Embolism

Thrombosis is the formation of blood clots within blood vessels. Deep Vein Thrombosis (DVT) is a blood clot that forms in a vein deep in the body [Abelsethe et al., 1996]. Most deep vein blood clots occur in the lower leg but they also can occur in other parts of the body (brachial vein, radial vein, ulnar vein, femoral vein, popliteal vein, peroneal vein, anterior tibial vein, posterior tibial vein).

A DVT can occur without symptoms, but in many cases the affected extremity will be painful, swollen, red, warm and the superficial veins may be engorged [Levenson et al., 1990].

The gold standard in treatment of DVT is Oral Anticoagulant Therapy (OAT), also called blood thinner. OAT decrease the blood's ability to clot or also stop existing blood clots from getting bigger. However, blood thinners can't break up blood clots that have already formed (the body dissolves most blood clots with time). Warfarin and heparin are two blood thinners used to treat DVT. Warfarin is given in pill form. Heparin is given as an injection [Kazmier, 1992].

Figure 2.16: A deep vein clot in femoral vein.
The most common side effect of blood thinners is bleeding. This happens if the medicine thins the blood too much. People treated with blood thinners usually use regular blood tests to measure their blood's ability to clot. These blood tests are called PT (Prothrombin Time) tests (Fig. 2.17).

![Figure 2.17: Automatic measurement of PT.](image)

These tests help the doctor make sure the patient is taking the right amount of medicine.

Blood clots can form in deep veins when:

- damage occurs to a vein's inner lining. This damage may result from injuries caused by physical, chemical, and biological factors. Such factors include surgery, serious injury, inflammation, or an immune response;
- blood flow is sluggish or slow. Lack of motion can cause sluggish or slowed blood flow. This may be caused after surgery, by illnesses and in bed or travelling for a long time;
- the blood is thicker or more likely to clot than usual. Certain inherited conditions (such as the genetic factor V Leiden), treatment with hormone replacement therapy or using birth control pills, increase blood's tendency to clot.

A blood clot in a deep vein can break off and travel through the bloodstream. The loose clot is called an embolus. When the clot travels to the lungs and blocks blood flow, the condition is called Pulmonary Embolism (PE) (Fig. 2.18) [Tsai et al., 2002].
Figure 2.18: *Pulmonary embolism.*

PE is a serious condition that can cause:
- permanent damage to part of lungs from lack of blood flow to lung tissue;
- low oxygen levels in blood;
- damage to other organs from not getting enough oxygen;
- death.
2.2.5 Myocardial infarction or Heart Attack

Myocardial infarction (MI), commonly known as a heart attack (HA), occurs when the blood supply to part of the heart is interrupted [AHA, 2006]. This is most commonly due to occlusion (blockage) of a coronary artery (Fig. 2.19).

Figure 2.19: Heart Attack caused by the occlusion of the coronary artery.

The resulting restriction in blood supply (ischemia) and oxygen shortage, if left untreated for a sufficient period, can cause damage and/or death (infarction) of heart muscle tissue. Heart attacks are the leading cause of death for both men and women all over the world [WHO, 2004].

Classical symptoms of acute myocardial infarction include sudden chest pain (typically radiating to the left arm or left side of the neck), shortness of breath, nausea, vomiting, palpitations, sweating, and anxiety. Women may experience fewer typical symptoms than men, most commonly shortness of breath, weakness, a feeling of indigestion, and fatigue [Kosuge et al., 2006]. Approximately one quarter of all myocardial infarctions are silent, without chest pain or other symptoms.
Important risk factors are previous cardiovascular disease (such as angina, a previous heart attack or stroke), older age (especially men over 40 and women over 50), tobacco smoking, high blood levels of certain lipids (triglycerides, low-density lipoprotein or "bad cholesterol") and low high density lipoprotein (HDL, "good cholesterol"), diabetes, high blood pressure, obesity, chronic kidney disease, heart failure, excessive alcohol consumption, the abuse of certain drugs (such as cocaine), and chronic high stress levels [D’Aiuto et al., 2006].

Immediate treatment for suspected acute myocardial infarction includes oxygen, aspirin, and sublingual glyceryl trinitrate (nitroglycerin) [Pearson et al., 2003]. The patient will receive a number of diagnostic tests, such as an electrocardiogram, a chest X-ray and blood tests to detect heart muscle damage. In patients who have multiple blockages and who are relatively stable bypass surgery of the blocked coronary artery is an option.
Biomedical signals, like all the other electronic signals related to a phenomenon of interest, need to be processed in some ways in order to remove noise and in order to extract useful information carried by electronic waveforms.

This chapter gives an introduction to digital signal processing, starting from basic concept of noise, Fourier transform, and ending with a brief overview of digital filters. Such techniques are used in deep in all hardware and software developed in this thesis.
3.1 Noise and Random Variables

Noise is an inherent component of most measurements. In addition to physiological and environmental noise, electronic noise arises from the transducer and associated electronics and is intermixed with the signal being measured. It is usually represented as a random variable, \( x(n) \). Since the variable is random, describing it as a function of time is not very useful. It is more common to discuss other properties of noise such as its probability distribution, range of variability, or frequency characteristics. While noise can take on a variety of different probability distributions, the central limit theorem implies that most noise will have a Gaussian or normal distribution. The central limit theorem states that when noise is generated by a large number of independent sources it will have a Gaussian probability distribution regardless of the probability distribution characteristics of the individual sources. The probability of a Gaussianly distributed variable, is specified in the well-known normal or Gaussian distribution equation:

\[
p(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \tag{3.1}
\]

Two important properties of a random variable are its mean, or average value, and its variance. The mean value of a discrete vector is evaluated as:

\[
\overline{x} = \frac{1}{N} \sum_{k=1}^{N} x_k \tag{3.2}
\]

The sample variance \( \sigma^2 \), is calculated as shown below, and the standard deviation \( \sigma \), is just the square root of the variance.

\[
\sigma^2 = \frac{1}{N-1} \sum_{k=1}^{N} (x^k - \overline{x})^2 \tag{3.3}
\]

\[
\sigma = \sqrt{\sigma^2} \tag{3.4}
\]

When multiple measurements are made, multiple random variables can be generated. If these variables are combined or added together, the means add so that the resultant random variable is simply the mean, or average, of the individual means. The same is true for the variance.

\[
\overline{\sigma^2} = \frac{1}{N} \sum_{k=1}^{N} \sigma_k^2 \tag{3.5}
\]
However, the standard deviation is the square root of the variance and the standard deviations add as the $\sqrt{N}$ times the average standard deviation. Accordingly, the mean standard deviation is the average of the individual standard deviations divided by $\sqrt{N}$.

$$\sum_{k=1}^{N} \sigma_k = \sqrt{N \bar{\sigma}^2} = \sqrt{N} \bar{\sigma}$$ \hspace{1cm} (3.6)

In other words, averaging noise from different sensors, or multiple observations from the same source, will reduce the standard deviation of the noise by the square root of the number of averages.

$$\text{(mean standard deviation)} \quad \text{MSD} = \frac{1}{N} \sum_{k=1}^{N} \sigma_k = \frac{1}{\sqrt{N}} \bar{\sigma}$$ \hspace{1cm} (3.7)

MSD indicates that averaging can be a simple, yet powerful signal processing technique for reducing noise when multiple observations of the signal are possible. Such multiple observations could come from multiple sensors, but in many biomedical applications, the multiple observations come from repeated responses to the same stimulus. In ensemble averaging, a group of time responses are averaged together on a point-by-point basis and an average signal is constructed by taking the average, for each point in time, over all signals in the ensemble. Two essential requirements for the application of ensemble averaging for noise reduction are needed: multiple observations, and a reference signal closely to the response.

### 3.2 Transforms

In signal processing there are waveforms, images and entities that operate on waveforms or images [Hubbard, 1998]. The entities can be functions that modify the data and functions used to analyze or probe the data. Functions that modify data are also termed operations or transformations. A transform can be thought of as a re-mapping of the original data into a function that provides more information than the original.

Many of the common transforms are achieved by comparing the signal of interest with some sort of probing function. This comparison takes the form of a correlation (produced by multiplication) that is averaged (or integrated) over the duration of the waveform, or some portion of the waveform:

$$X(m) = \int_{-\infty}^{+\infty} x(t) f_m(t) dt$$ \hspace{1cm} (3.8)

where $x(t)$ is the waveform being analyzed, $f_m(t)$ is the probing function (the basis) and $m$ is some variable of the probing function, often specifying a particular member in a family of
similar functions. For example, in the Fourier Transform \( fm(t) \) is a family of harmonically related sinusoids and \( m \) specifies the frequency of an individual sinusoid in that family.

Since most signal processing operations are implemented using digital electronics, functions are represented in discrete form as a sequence of numbers:

\[
x(n) = [x(1), x(2), ..., x(N)]
\]  

and the integral becomes summation over a finite range:

\[
X(m) = \sum_{n=1}^{N} x(n)f_{m}(n)
\]

where \( f_m(n) \) is a discrete version of the family of probing functions.

When either \( x(t) \) or \( fm(t) \) are of infinite length, they must be truncated to fit within the confines of limited memory storage. The length of either function can be shortened by simple truncation or by multiplying the function by yet another function that has zero value beyond the desired length. A function used to shorten another function is termed a window function, and its action is shown in Fig. 3.1.

**Figure 3.1:** A waveform (upper plot) is multiplied by a window function (middle plot) to create a truncated version of the original waveform (lower plot).
Simple truncation can be viewed as multiplying the function by a rectangular window, a function whose value is one for the portion of the function that is retained, and zero elsewhere. The consequences of this artificial shortening will depend on the specific window function used.

If a window function is used, Eq. (3.10) becomes:

\[ X(m) = \sum_{n=1}^{N} x(n)W(n)f_m(n) \]  

where \( W(n) \) is the window function.

If the probing function is of finite length (have finite support) and this length is shorter than the waveform, then it might be appropriate to translate or slide it over the signal and perform the comparison (correlation, or multiplication) at various relative positions between the waveform and probing function. This sliding comparison is given in discrete form by the equation:

\[ X(m, k) = \sum_{n=1}^{N} x(n)f_m(n - k) \]  

where the variable \( k \) indicates the relative position between the two functions. This approach is adopted in the filters used in the Continuous Wavelet Transform.

A variation of this approach can be used for long probing functions, where the probing function is shortened by windowing to a length that is less than the waveform. Then the shortened probing function can be translated across the waveform in the same manner as a probing function that is naturally short. The equation for this condition becomes:
\[ X(m, k) = \sum_{n=1}^{N} x(n)f_m(n)W(n - k) \]  
(3.13)

where \( f_m(n) \) is a longer function that is shortened by the sliding window function \( W(n - k) \). This is the approach taken in the *Short-Term Fourier Transform*.

### 3.3 Convolution, Correlation and Covariance

**Convolution** is an important concept in linear systems theory, solving the need for a time domain operation equivalent to the Transfer Function. Transfer Function is a frequency domain concept that is used to calculate the output of a linear system to any input. Convolution can be used to define a general input–output relationship in the time domain analogous to the Transfer Function in the frequency domain.

The input \( x(t) \), the output \( y(t) \), and the function linking the two through convolution \( h(t) \), are all functions of time. Hence, convolution is a time domain operation.

The basic concept behind convolution is superposition. The first step is to determine a time function \( h(t) \) that tells how the system responds to an infinitely short segment of the input waveform. If superposition holds, then the output can be determined by summing (integrating) all the response contributions calculated from the short segments. The way in which a linear system responds to an infinitely short segment of data can be determined simply by noting the system’s response to an infinitely short input \( \delta(t) \).

Given that the impulse response describes the response of the system to an infinitely short segment of data, and any input can be viewed as an infinite string of such infinitesimal segments, the impulse response can be used to determine the output of the system to any input. The response produced by an infinitely small data segment is simply this impulse response scaled by the magnitude of that data segment. The contribution of each infinitely small segment can be summed, or integrated, to find the response created by all the segments.

\[ y(t) = \int_{-\infty}^{+\infty} h(\tau)x(t - \tau)d\tau = \int_{-\infty}^{+\infty} h(t - \tau)x(\tau)d\tau \]  
(3.14)

To determine the impulse of each infinitely small data segment, the impulse response is shifted a time \( \tau \) with respect to the input, then scaled (i.e., multiplied) by the magnitude of the input at that point in time.
For most systems $h(\tau)$ is finite, so the limit of integration is finite. Moreover, a real system can only respond to past inputs, so $h(\tau)$ must be 0 for $\tau < 0$ (negative $\tau$ implies future times in Eq. (3.14), although for computer-based operations, where future data may be available in memory, $\tau$ can be negative.

For discrete signals, the integration becomes a summation and the convolution equation becomes:

$$y(n) = \sum_{k=1}^{N} h(n-k)x(k) = \sum_{k=1}^{N} h(n)x(k-n) \tag{3.15}$$

Correlation and Covariance connote similarity: how one thing is like another. Mathematically, correlations are obtained by multiplying and normalizing. Both covariance and correlation use multiplication to compare the linear relationship between two variables, but in correlation the coefficients are normalized to fall between zero and one. This makes the correlation coefficients insensitive to variations in the gain of the data acquisition process or the scaling of the variables. The operations of correlation and covariance can be applied to two or more waveforms, to multiple observations of the same source, or to multiple segments of the same waveform.

Correlation and covariance operations can not only be used to compare different waveforms at specific points in time, they can also make comparisons over a range of times by shifting one signal with respect to the other. The crosscorrelation function is an example of this process. The correlation function is the lagged product of two waveforms, and the defining equation, given here in both continuous and discrete form, is quite similar to the convolution equation:

$$r_{xx}(t) = \int_{0}^{T} y(t)x(t+\tau)d\tau \tag{3.16}$$

$$r_{xx}(n) = \sum_{k=1}^{N} y(k+n)x(k) \tag{3.17}$$

These equations show that the only difference in the computation of the crosscorrelation versus convolution is the direction of the shift. In convolution the waveforms are shifted in opposite directions. This produces a causal output: the output function is the creation of past values of the input function.

Crosscorrelation shows the similarity between two waveforms at all possible relative positions of one waveform with respect to the other, and it is useful in identifying segments of similarity. A special case of the correlation function occurs when the comparison is between two waveforms that are one in the same. This is termed the autocorrelation function and it provides a description of how similar a waveform is to itself at various time lags. The autocorrelation function will naturally be maximum for zero lag ($n = 0$) because at zero lag the comparison is between identical waveforms. Usually the autocorrelation is
scaled so that the correlation at zero lag is 1. The function must be symmetric about \( n = 0 \), since shifting one version of the same waveform in the negative direction is the same as shifting the other version in the positive direction.

The autocorrelation function is related to the bandwidth of the waveform. The sharper the peak of the autocorrelation function the broader the bandwidth. For example, in white noise, which has infinite bandwidth, adjacent points are uncorrelated, and the autocorrelation function will be nonzero only for zero lag.

![Figure 3.3: Autocorrelation of two colored noise with different bandwidths.](image)

The crosscovariance function is the same as crosscorrelation function except that the means have been removed from the data before calculation. Accordingly, the equation is a slight modification of Eq. (3.17), as shown below:

\[
\text{(crosscovariance)} \quad Cov(n) = \sum_{k=1}^{N} [y(k + n - \bar{y})][x(k) - \bar{x}]
\]

(3.18)

The terms correlation and covariance, when used alone (i.e., without the term function), imply operations similar to those described in Eqs. (16) and (17), but without the lag operation. The result will be a single number.

Of particular interest is the covariance and correlation matrices. These analysis tools can be applied to multivariate data where multiple responses, or observations, are obtained from a single process. The covariance and correlation matrices assume that the multivariate data are arranged in a matrix where the columns are different variables and the rows are
different observations of those variables. In signal processing, the rows are the waveform time samples, and the columns are the different signal channels or observations of the signal. Since the diagonals in the correlation matrix give the correlation of a given variable or waveform with itself, they will all equal to 1, and the off-diagonals will vary between ± 1. In its usual signal processing definition, the correlation matrix is a normalized version of the covariance matrix.

### 3.4 Fourier Analysis

Sometimes the frequency content of the waveform provides more useful information than the time domain representation. The accurate determination of the waveform’s spectrum requires that the signal be periodic, or of infinite length, and noise-free. The problem is that in many biological applications only a portion of the waveform is available for analysis. Moreover, biosignals are often corrupted by substantial amounts of noise or artifact. If only a portion of the actual signal can be analyzed, and/or if the waveform contains noise along with the signal, then all spectral analysis techniques must necessarily be approximate. The various spectral analysis approaches attempt to improve the estimation accuracy of specific spectral features.

The Fourier transform approach takes advantage of the fact that sinusoids contain energy at only one frequency. If a waveform can be broken down into a series of sines or cosines of different frequencies, the amplitude of these sinusoids must be proportional to the frequency component contained in the waveform at those frequencies [Boashash, 1992]. From Fourier series analysis, we know that any periodic waveform can be represented by a series of sinusoids that are at the same frequency as, or multiples of, the waveform frequency. This family of sinusoids can be expressed either as sines and cosines, each of appropriate amplitude, or as a single sine wave of appropriate amplitude and phase angle. Expressed as an equation, this procedure becomes:

\[
a(m) = \frac{1}{T} \int_0^T x(t) \cos\left(\frac{2\pi mt}{T}\right) dt \tag{3.19}
\]

\[
b(m) = \frac{1}{T} \int_0^T x(t) \sin\left(\frac{2\pi mt}{T}\right) dt \tag{3.20}
\]

where \( T \) is the period of the waveform and \( m \) is set of integers, possibly infinite. This gives rise to a family of sines and cosines having harmonically related frequencies \( \frac{m}{T} \).

Since this approach represents waveforms by harmonically related sinusoids, the approach is sometimes referred to as harmonic decomposition. For periodic functions, the Fourier
transform and Fourier series constitute a bilateral transform: the series of sine and cosine components can be summed to reconstruct the original waveform:

\[ x(t) = \frac{a(0)}{2} + \sum_{m=0}^{\infty} a(k) \cos\left(\frac{2\pi mt}{T}\right) + \sum_{m=0}^{\infty} b(k) \sin\left(\frac{2\pi mt}{T}\right) \]  

(3.21)

Spectral information is usually presented as a frequency plot, a plot of sine and cosine amplitude versus component number, or the equivalent frequency. Rather than plot sine and cosine amplitudes, it is more intuitive to plot the amplitude and phase angle of a sinusoidal wave using the rectangular-to-polar transformation.

Some waveforms are symmetrical or anti-symmetrical about \( t = 0 \), so that one or the other of the components \( a(k) \) or \( b(k) \), will be zero. Specifically, if the waveform has mirror symmetry about \( t = 0 \), that is \( x(t) = x(-t) \), than multiplications by a sine functions will be zero irrespective of the frequency, and this will cause all \( b(k) \) terms to be zeros. Such mirror symmetry functions are termed even functions. Similarly, if the function has anti-symmetry, \( x(t) = -x(t) \), a so-called odd function, then all multiplications with cosines of any frequency will be zero, causing all \( a(k) \) coefficients to be zero.

<table>
<thead>
<tr>
<th>Function type</th>
<th>Symmetry</th>
<th>Coefficient values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Even</td>
<td>( x(t) = x(-t) )</td>
<td>( b(k) = 0 )</td>
</tr>
<tr>
<td>Odd</td>
<td>( x(t) = -x(t) )</td>
<td>( a(k) = 0 )</td>
</tr>
</tbody>
</table>

### 3.4.1 Discrete-Time Fourier Analysis

The discrete-time Fourier series analysis is an extension of the continuous analysis procedure described above, but modified by two operations: sampling and windowing. The sampling process makes the spectra repetitive at frequencies \( m/T \). Hence the discrete Fourier series of any waveform is theoretically infinite, but since it is periodic and symmetric about \( fs/2 \), all of the information is contained in the frequency range of 0 to \( fs/2 \) (\( fs /2 \) is the Nyquist frequency). This follows from the sampling theorem and the fact that the original analog waveform must be bandlimited so that its highest frequency \( f_{\text{max}} \) is less than \( fs/2 \) if the digitized data is to be an accurate representation of the analog waveform.

The equations for computing Fourier series analysis of digitized data are the same as for continuous data except the integration is replaced by summation. Usually these equations are presented using complex variables notation so that both the sine and cosine terms can be represented by a single exponential term using Euler’s identity:

\[ X(m) = \sum_{n=0}^{N-1} x(n)e^{-j2\pi mn/N} \]  

(3.22)
where \( N \) is the total number of points and \( m \) indicates the family member, i.e. the harmonic number. This number must now be allowed to be both positive and negative when used in complex notation: \( m = -(N/2), \ldots , (N/2)-1 \).

The inverse Fourier transform can be calculated as:

\[
x(n) = \frac{1}{N}\sum_{m=0}^{N-1} X(m)e^{(-j2\pi fnmT_s)} \tag{3.23}
\]

Applying the rectangular-to-polar transformation is also apparent that \(|X(m)|\) gives the magnitude for the sinusoidal representation of the Fourier series while the angle of \(X(m)\) gives the phase angle for this representation, since \(X(m)\) can also be written as:

\[
X(m) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{2\pi mn}{N}\right) - j \sum_{n=0}^{N-1} x(n) \sin\left(\frac{2\pi mn}{N}\right) \tag{3.24}
\]

As mentioned above, for computational reasons, \(X(m)\) must be allowed to have both positive and negative values for \(m\); negative values imply negative frequencies, but these are only a computational necessity and have no physical meaning.

The discrete Fourier transform produces a function of \(m\). To convert this to frequency note that:

\[
f_m = mf_1 = \frac{m}{T_p} = \frac{m}{NT_s} = \frac{mf_s}{N} \tag{3.25}
\]

where \(f_1\) is the fundamental frequency, \(T_s\) is the sample interval, \(f_s\) is the sample frequency, \(N\) is the number of points in the waveform and \(T_p = NT_s\) is the period of the waveform. Substituting \(m = f_mT_s\) the equation for the discrete Fourier transform can also be written as:

\[
X(f) = \sum_{n=0}^{N-1} x(n)e^{(-j2\pi fnmT_s)} \tag{3.26}
\]

which may be more useful in manual calculations.

If the waveform of interest is truly periodic, then the approach described above produces an accurate spectrum of the waveform. In this case, such analysis should properly be termed Fourier series analysis, but is usually termed Fourier transform analysis. This latter term more appropriately applies to aperiodic or truncated waveforms. The algorithms used in all cases are the same, so the term Fourier transform is commonly applied to all spectral analyses based on decomposing a waveform into sinusoids.
From the discrete Fourier series equation above, the number of points produced by the operation is $N$, the number of points in the data set. However, since the spectrum produced is symmetrical about the midpoint, $N/2$ (or $f_s/2$ in frequency), only half the points contain unique information. If the sampling time is $T_s$, then each point in the spectra represents a frequency increment of $1/(NT_s)$. As a rough approximation, the frequency resolution of the spectra will be the same as the frequency spacing $1/(NT_s)$ [Boudreaux et al., 1995]. Increasing the sample interval $T_s$, should improve the frequency resolution but since that means a decrease in $f_s$, the maximum frequency in the spectra, $f_s/2$ is reduced limiting the spectral range. One simple way of increasing $N$ even after the waveform has been sampled is to use zero padding. Zero padding is legitimate because the undigitized portion of the waveform is always assumed to be zero. Under this assumption, zero padding simply adds more of the unsampled waveform. The zero-padded waveform appears to have improved resolution because the frequency interval is smaller. In fact, zero padding does not enhance the underlying resolution of the transform since the number of points that actually provide information remains the same. However, zero padding does provide an interpolated transform with a smoother appearance.

### 3.4.2 Power Spectrum

The power spectrum is commonly defined as the Fourier transform of the autocorrelation function. In continuous and discrete notation, the power spectrum equation becomes:

$$PS(f) = \int_{0}^{T} r_{xx}(\tau) e^{-2\pi f \tau} d\tau \quad (3.27)$$

$$PS(f) = \sum_{n=0}^{N-1} r_{xx}(n) e^{-2\pi n f T_s} \quad (3.28)$$

![Figure 3.4: Power Spectrum of a waveform composed of two sines (100Hz and 200Hz) sampled at 1Ksps.](image)
3.4.3 Short-Term Fourier Transform: The Spectrogram

Classical spectral analysis techniques represent powerful signal processing tools if one is not especially concerned with signal timing. They provide a complete and appropriate solution for waveforms that are stationary (e.g. waveforms that do not change their properties over the length of the analysis). Waveforms of biological origin are not stationary, and change substantially in their properties over time. Moreover, it is these changes with time that are often of primary interest. Fourier analysis provides a good description of the frequencies in a waveform, but not their timing. However, timing is encoded in the phase portion of the transform, and this encoding is difficult to interpret and recover.

A time–frequency method is usually based on the slicing of the waveform of interest into a number of short segments and performing the analysis on each of these segments, usually using the standard Fourier transform. A window function is applied to a segment of data, effectively isolating that segment from the overall waveform, and the Fourier transform is applied to that segment. This is termed the spectrogram or Short-Term Fourier Transform since the Fourier Transform is applied to a segment of data that is shorter, often much shorter, than the overall waveform.

The basic equation for the spectrogram in the continuous domain is:

\[ X(t, f) = \int_{-\infty}^{+\infty} x(\tau)w(t - \tau)e^{-j\pi f \tau}d\tau \]  

(3.29)

where \( w(t-\tau) \) is the window function and \( \tau \) is the variable that slides the window across the waveform \( x(t) \).

The discrete version of the spectrogram is:

\[ X(m, k) = \sum_{n=1}^{N} x(n)[W(n - k)e^{j\pi nk/N}] \]  

(3.30)

Figure 3.5: Spectrogram of a 15Hz sine before \( t = 1 \) sec and 50Hz after \( t = 1 \) sec.
There are two main problems with the spectrogram:

- selecting an optimal window length for data segments that contain several different features may not be possible;
- the time–frequency tradeoff: shortening the data length $N$, to improve time resolution will reduce frequency resolution which is approximately $1/(NT_s)$.

Shortening the data segment could also result in the loss of low frequencies that are no longer fully included in the data segment. Hence, if the window is made smaller to improve the time resolution, then the frequency resolution is degraded and vice versa. This time–frequency tradeoff has been equated to an uncertainty principle where the product of frequency resolution $B$ and time $T$ must be greater than some minimum:

$$BT \geq \frac{1}{4\pi}$$  \hspace{1cm} (3.31)

Despite these limitations, the spectrogram has been used successfully in a wide variety of problems, particularly those where only high frequency components are of interest and frequency resolution is not critical.

### 3.5 Digital Filters

A filter can be viewed as a linear process in which the input signal’s spectrum is reshaped in some well-defined manner. Filters differ in the way they achieve this spectral reshaping, and can be classified into two groups based on their approach. These two groups are termed finite impulse response (FIR) filters and infinite impulse response (IIR) filters.

#### 3.5.1 The Z-Transform

The frequency-based analysis is the most useful tool for analyzing systems or responses in which the waveforms are periodic or aperiodic, but cannot be applied to transient responses of infinite length, such as step functions, or systems with nonzero initial conditions. These shortcomings motivated the development of the Laplace transform in the analog domain. Laplace analysis uses the complex variable $s$ ($s = \sigma + j\omega$) as a representation of complex frequency in place of $j\omega$ in the Fourier transform.

The Z-transform is a digital operation analogous to the Laplace transform in the analog domain, and it is used in a similar manner. The Z-transform is based around the complex variable $z$, where $z$ is an arbitrary complex number, $|z|e^{j\omega}$. This variable is also termed the complex frequency, and as with its time domain counterpart, the Laplace variable $s$, it is possible to substitute $e^{j\omega}$ for $z$ to perform a strictly sinusoidal analysis.
The Z-transform follows the format of the general transform equation and is also similar to the Fourier transform equation:

\[ X(z) = \sum_{n=-\infty}^{\infty} x(n) z^{-n} \quad (3.32) \]

where \( z \) is an arbitrary complex variable. The probing function for this transform is simply \( z^{-n} \). In any real application, the limit of the summation will be finite, usually the length of \( x(n) \). When identified with a data sequence, \( z^{-n} \) represents an interval shift of \( n \) samples, or an associated time shift of \( nT_s \) seconds. This time shifting property of \( z^{-n} \) can be formally stated as:

\[ Z(x(n - k)) = z^{-k}Z(x(n)) \quad (3.33) \]

For example, the time shifting characteristic of the Z-transform can be used to define a unit delay process \( z^{-1} \). For such a process, the output is the same as the input, but shifted (or delayed) by one data sample (Fig. 3.6).

![Figure 3.6: A unit delay process.](image)

### 3.5.1.1 Digital Transfer Function

As in Laplace transform analysis, one of the most useful applications of the Z-transform lies in its ability to define the digital equivalent of a transfer function. By analogy to linear system analysis, the digital transfer function is defined as:

\[ H(z) = \frac{Y(z)}{X(z)} \quad (3.34) \]

Most transfer functions can be complicated, including polynomials of \( z \) in both the numerator and denominator, just as analog transfer functions contain polynomials of \( s \):

\[ H(z) = \frac{b_0 + b_1z^{-1} + b_2z^{-2} + \cdots + b_Nz^{-N}}{1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_Dz^{-D}} \quad (3.35) \]

While \( H(z) \) has a structure similar to the Laplace domain transfer function \( H(s) \), there is no simple relationship between them. For example, unlike analog systems, the order of the numerator \( N \), need not be less than, or equal to, the order of the denominator \( D \), for stability. In fact, systems that have a denominator order of 1 are more stable that those having higher order denominators.
From the digital transfer function $H(z)$, it is possible to determine the output given any input. In the Z-transform domain this relationship is simply:

$$Y(z) = H(z)X(z) = X(z)\sum_{k=0}^{N-1} b(k)z^{-n}$$

The input/output or difference equation analogous to the time domain equation can be obtained by applying the time shift interpretation to the term $z^{-n}$:

$$y(n) = \sum_{k=0}^{K} b(k)x(n-k) - \sum_{l=0}^{L} a(l)y(n-l)$$

All basic digital filters can be interpreted as linear digital processes, and, in fact, the term digital filter is often used interchangeably with digital systems. Filter design, then, is simply the determination of the appropriate filter coefficients, $a(n)$ and $b(n)$, that provide the desired spectral shaping. If the frequency spectrum of $H(z)$ is desired, it can be obtained substituting $z = e^{j\omega}$:

$$H(m) = \frac{Y(m)}{X(m)} = \frac{\sum_{n=0}^{N-1} b(n)e^{-j2\pi mn}}{\sum_{n=1}^{D-1} a(n)e^{-j2\pi mn}} = \frac{\text{fft}(b(n))}{\text{fft}(a(n))}$$

where FFT indicates the Fourier transform.

### 3.5.2 Finite Impulse Response Filters

Finite Impulse Response (FIR) filters have transfer functions that have only numerator coefficients, i.e. $H(z) = B(z)$. This leads to an impulse response that is finite, hence the name. They have the advantage of always being stable and having linear phase shifts. In addition, they have initial transients that are of finite durations. The downside of FIR filters is that they are less efficient in terms of computer time and memory than IIR filters. FIR filters are also referred to as non-recursive because only the input (not the output) is used in the filter algorithm. FIR filters are also referred to as a moving average process. This term is used for any process that uses a moving set of multiplier weights, even if the operation does not really produce an average.

The general equation for an FIR filter is a simplification of Eq. (3.37) that becomes:

$$y(k) = \sum_{n=1}^{L} b(n)x(k-n)$$

where $b(n)$ is the coefficient function (also referred to as the weighting function) of length $L$, $x(n)$ is the input, and $y(n)$ is the output. This is identical to the convolution equation with the impulse response $h(n)$ replaced by the filter coefficients $b(n)$. Eq. (3.39) indicates that the
filter coefficients (or weights) of an FIR filter are the same as the impulse response of the filter. Since the frequency response of a process having an impulse response $h(n)$ is simply the Fourier transform of $h(n)$, the frequency response of an FIR filter having coefficients $b(n)$ is just the Fourier transform of $b(n)$:

$$X(m) = \sum_{n=0}^{N-1} b(n)e^{-j2\pi mn/N} \quad (3.40)$$

The inverse operation, going from a desired frequency response to the coefficient function $b(n)$, is known as filter design. Since the frequency response is the Fourier transform of the filter coefficients, the coefficients can be found from the inverse Fourier transform of the desired frequency response.

3.5.3 Infinite Impulse Response Filters

The primary advantage of Infinite Impulse Response (IIR) filters over FIR filters is that they can usually meet a specific frequency criterion, such as a cutoff sharpness or slope, with a much lower filter order (i.e., a lower number of filter coefficients). The transfer function of IIR filters includes both numerator and denominator terms unlike FIR filters which have only a numerator. The major disadvantage of IIR filters is that they have nonlinear phase characteristics. Since IIR filters have transfer functions that are the same as a general linear process having both poles and zeros, many of the concepts of analog filter design can be used with these filters. One of the most basic of these is the relationship between the number of poles and the slope, or rolloff of the filter beyond the cutoff frequency. The asymptotic downward slope of a filter increases by 20 db/decade for each filter pole, or filter order. Determining the number of poles required in an IIR filter given the desired attenuation characteristic is a straightforward process.

Another similarity between analog and IIR digital filters is that all of the well-known analog filter types can be duplicated as IIR filters. Specifically the Butterworth, Chebyshev Type I and II, and elliptic designs can be implemented as IIR digital filters. Butterworth filters provide a frequency response that is maximally flat in the pass-band and monotonic overall. To achieve this characteristic, Butterworth filters sacrifice rolloff steepness; hence, the Butterworth filter will have a less sharp initial attenuation characteristic than other filters. The Chebyshev Type I filters feature faster rolloff than Butterworth filters, but have ripple in the pass-band. Chebyshev Type II filters have ripple only in the stop-band and a monotonic pass-band, but they do not rolloff as sharply as Type I. The ripple produced by Chebyshev filters is termed equiripple since it is of constant amplitude across all frequencies. Finally, elliptic filters have steeper rolloff than any of the above, but have equiripple in both the pass-band and stop-band. In general, elliptic filters meet a given performance specification with the lowest required filter order (Fig. 3.7).
Figure 3.7: Butterwort, Chebyshev and Elliptic frequency response.
Wearable devices allow clinicians, physiologists, professional and people themselves to have access to more and more information from the human body through miniature sensors placed in some places of interest.

This chapter gives an overview of wearable devices used for biomedical signals monitoring and show the wearable device developed and realized for monitoring of certain cardiovascular disease, such as sleep apnoea, continuous blood pressure measurement and wavelet analysis of PPG signals measured in patients with heart attack, deep vein thrombosis and atrial fibrillation.
4.1 Overview of Wearable Devices

Wearable systems are taking more and more importance in our life. They present to us in numerous forms, from standalone devices like smart phones, MP3 players, Personal Data Assistant (PDA), mobile phones to more advanced systems offering sensing and interaction with the user or remote professional. Wearable sensing of human physiological signal is for instance performed to address health application needs as well as wellness, sports activities and clinicians objectives. These systems allow clinicians, physiologists, professional and people themselves to have access to more and more information from the human body and the environment. They typically rely on wireless, miniature sensors enclosed in patches or bandages, or in items that can be worn, such as a ring or a shirt. They take advantage of hand-held units to temporarily store physiological data and then periodically upload that data to a database server via a wireless LAN or a cradle that allow Internet connection. The data sets recorded using these systems are then processed to detect events predictive of possible worsening of the patient’s clinical situation or they are explored to assess the impact of clinical interventions.

4.1.1 Wearable Sensing of Physiological Parameters

The early prototypes of wearable medical systems are mainly based on body-worn devices, which are designed to be miniaturized, integrated, networked, digitalized, and smart for e-Health application [Poon et al., 2008]. Body-worn medical systems can typically gather data from sensors located only in a specific and usually relatively small body area and cannot fulfill, alone, all the needs for sensing, actuating, displaying, and interacting with the user. An alternative approach to achieve wearable monitoring is to integrate sensors in clothing items. Intelligent biomedical clothing [Lyemberis et al., 2003] refers usually to clothes with sensors that are close to or in contact with the skin [Lyemberis et al., 2007]. The sensors are either embedded in the fabric of the garment, or it is the fabric itself that is used as a sensor or a sensor suite [Marculescu et al., 2003; Carpi et al., 2005]. Several prototypes of wearable medical devices or systems are shown in Fig. 4.1.

Figure 4.1: Examples of prototype of wearable medical devices/systems.
As an example of wearable systems, LifeShirt is a noninvasive, continuous ambulatory monitoring system that can collect data during a person’s daily routine [LifeShirt, 2008; Grossman, 2004]. However, it is unsuitable for lengthy continuous monitoring due to the cumbersome recorder and peripheral attachment to be carried around. To address the limitations of the first prototype developed, researchers shifted their focus on smart textiles. Smart textiles have become one of the thriving areas of application for materials and nanotechnology research for the purpose of health monitoring. Smartshirt, collects analog signals through conductive fiber sensors and transfers them through a conductive fiber grid knitted in the T-shirt [Sensatex, 2007].

Up-to-date, sensing body movements or physiological signals, particularly the ECG and respiration, remains the main target of a wearable device. The measurement of blood pressure, for example, is still a challenge for wearable device because of the conventional cuff-based principle utilized for blood pressure measurement. In a recent work [Zhang et al., 2006; Chan et al., 2008] a Health-Shirt (h-Shirt) that can monitor blood pressure cufflessly and continuously has been proposed. The h-Shirt is designed to simultaneously record ECG and PPG and uses them to estimate blood pressure with the Pulse Transit Time (PTT) method. ECG is captured from the two wrists, with a reference electrode placed on the forearm to avoid respiration induced noise. The electrodes were made of e-textile materials. A photoreflective sensor capture PPG from the fingertip. Fig. 4.2 shows the shirt when the user is measuring blood pressure and ECG continuously.

![Figure 4.2: The h-shirt prototype.](image)

Other studies performed in the area of wearable devices include:

- a wearable physiological monitoring system for space and terrestrial applications named Life Guard [Mundt et al., 2005];
• the Armband SenseWear wearable body monitor that has been used to study body movement and energy expenditure in healthy as well as subjects with chronic obstructive pulmonary disease [Pitta et al., 2006];
• the WEALTHY system that is made up of a sensorized cotton shirt that integrates carbon-loaded elastometer strain sensors and fabric bio-electrodes, enabling the monitoring of respiration, ECG, electromyogram, body posture and movement [WEALTHY, 2002];
• a textile based wearable system, MagIC, designed for unobtrusively recording cardiorespiratory and motion signals during daily life and in a clinical environment on patients with different cardiovascular conditions [Rienzo et al., 2005];
• the MyHeart wearable monitoring system that focuses on integration of unobtrusive sensors into everyday garments and miniaturized on-body electronic modules for data processing and storage with dedicated software for data analysis like ECG preprocessing and motion artifact detection, computation of heart rate and HRV [Pacelli et al., 2006; Lauter, 2004];

4.2 The Developed Wearable Device

The developed wearable device is depicted in Fig. 4.3.

Figure 4.3: The developed wearable device: block diagram (left), PCB top view (center) and PCB bottom view (right).

Battery powered for patient’s safety, the device can drive both the red and infrared emitters of a standard SpO2 sensor (Fig. 4.4).
The analog signal conditioning block is composed of a single-supply two-stage inverting amplifier with variable gain able to converts the photocurrent coming from the SpO2 sensor in a voltage with a dynamic range appropriate for the A/D converter input (Fig. 4.5).

Figure 4.5: Bode diagram of the analog part of the wearable device developed.

The digital block is composed of a high performance and low power 16-bit digital signal controller (i.e. Microchip dsPIC33FJ12MC202). It controls all the blocks of the device, filters the PPG signal in order to remove motion and environmental artifacts, and continuously analyzes the digitalized version of the PPG signals in order to execute application algorithm and needs.
Through a radio frequency link in the ISM band based on a IEEE 802.15.4 transceiver at 2.4GHz (i.e. Microchip MRF24J40), the wearable device can communicates with an internet gateway or with a wireless USB dongle for remote and continuous data analysis, clinical monitoring and alarm catch.

The picture of the realized internet gateway with its block diagram are shown in Fig. 4.6 and Fig. 4.7 respectively.

![Figure 4.6: The developed internet gateway.](image)

![Figure 4.7: Block diagram of the developed internet gateway.](image)

The internet gateway is mainly composed of a Microchip transceiver at 2.4GHz connected to an IEEE 802.3 10baseT Ethernet controller embedded in a low power microcontroller (i.e. Microchip PIC18F67J60).

The picture of the realized wireless USB dongle with its block diagram are shown in Fig. 4.8 and Fig. 4.9 respectively.
The wireless USB gateway is mainly composed of a Microchip transceiver at 2.4GHz connected to a USB controller embedded in a low power microcontroller (i.e. Microchip PIC18F4550).
Chapter 5

Continuous Detection and Screening of Sleep Apnea Disease

As depicted in section 2.2.2, sleep apnoea is a common sleep disorder characterized by brief interruptions of breathing during sleep, which causes the major part of unknown death in healthy subjects. The 90 percent of patient with sleep apnoea don’t know to have this disease, because of the small number of symptoms and because the gold standard in diagnosing of SA is overnight polysomnography.

This chapter show the biomedical system and application in which the wearable device developed is used for the continuous detection and screening of sleep apnoea disease at home thanks to a real-time algorithm based on the heart rate variability analysis.
5.1 Heart Rate Variability Analysis

Heart Rate Variability (HRV) analysis is the standard to describe variations of instantaneous heart rate and RR intervals. It is mainly used for the non-invasive evaluation of the synergetic action of the two branches of the Autonomous Nervous System to maintain cardiovascular parameters in their optimal ranges [Levy et al., 1994]. In response of physical exercise, thoughts and emotions, the parasympathetic system slow HR, while sympathetic system accelerate it.

The clinical importance of HRV became appreciated in the late 1980s, when it was confirmed that HRV was a strong and independent predictor of mortality after an acute myocardial infarction [Kleiger et al., 1987; Malik et al., 1989; Bigger et al., 1992]. With the availability of new, digital, high-frequency, 24-hour multichannel ECG recorders, HRV has the potential to provide additional valuable insight into physiological and pathological conditions and to enhance risk stratification.

There are two main methods to perform an HRV analysis: one in the time domain and one in the frequency domain.

In the **time domain method**, starting from short-term 5-minute recordings and nominal 24-hour long-term ECG recordings, the normal-to-normal (NN) intervals (all intervals between adjacent QRS complexes) are determined (the so called tachogram) (Fig. 5.1).

![Figure 5.1: Interval tachogram of 256 consecutive RR values in a normal subject at supine rest.](image)
Simple time domain variables that can be calculated include the mean NN interval, the mean heart rate, the difference between the longest and shortest NN interval, the difference between night and day heart rate, and so forth. From a series of instantaneous heart rates or cycle intervals, particularly those recorded over longer periods, more complex statistical time domain measures can be calculated (Table 5.1). These variables may be derived from analysis of the total ECG recording or may be calculated using smaller segments of the recording period. The latter method allows comparison of HRV to be made during varying activities, for example, rest, sleep, and so on.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDNN</td>
<td>ms</td>
<td>Standard deviation of all NN intervals.</td>
</tr>
<tr>
<td>SDANN</td>
<td>ms</td>
<td>Standard deviation of the average of NN intervals in all 5-min segments of the entire recording.</td>
</tr>
<tr>
<td>RMSSD</td>
<td>ms</td>
<td>The square root of the mean of the sum of the squares of differences between adjacent NN intervals.</td>
</tr>
<tr>
<td>SDNN index</td>
<td>ms</td>
<td>Mean of the standard deviations of all NN intervals for all 5-min segments of the entire recording.</td>
</tr>
<tr>
<td>SDSD</td>
<td>ms</td>
<td>Standard deviation of differences between adjacent NN intervals.</td>
</tr>
<tr>
<td>NN50 count</td>
<td></td>
<td>Number of pairs of adjacent NN intervals differing by more than 50 ms in the entire recording.</td>
</tr>
<tr>
<td>pNN50</td>
<td>%</td>
<td>NN50 count divided by the total number of NN intervals.</td>
</tr>
</tbody>
</table>

Table 5.1: Time domain measures of HRV analysis.

The simplest variable to calculate is the standard deviation of the NN intervals (SDNN), that is, the square root of variance. Since variance is mathematically equal to total power of spectral analysis, SDNN reflects all the cyclic components responsible for variability in the period of recording. In many studies SDNN is calculated over a 24-hour period and thus encompasses short-term heart rate variations as well as the lowest-frequency components seen in a 24-hour period. As the period of monitoring decreases, SDNN estimates shorter and shorter cycle lengths. It also should be noted that the total variance of HRV increases with the length of analyzed recording [Saul et al., 1988]. Thus, on arbitrarily selected ECGs, SDNN is not a well-defined statistical quantity because of its dependence on the length of recording period. In practice, it is inappropriate to compare SDNN measures obtained from recordings of different durations. On the contrary, durations of the recordings used to determine SDNN
values should be standardized. Short-term 5-minute recordings appear to be appropriate options.

Other commonly used statistical variables calculated from segments of the total monitoring period include SDANN, the standard deviation of the average NN intervals calculated over short periods, usually 5 minutes, which is an estimate of the changes in heart rate due to cycles longer than 5 minutes, and the SDNN index, the mean of the 5-minute standard deviations of NN intervals calculated over 24 hours, which measures the variability due to cycles shorter than 5 minutes. The most commonly used measures derived from interval differences include RMSSD, the square root of the mean squared differences of successive NN intervals, NN50, the number of interval differences of successive NN intervals greater than 50 ms, and pNN50, the proportion derived by dividing NN50 by the total number of NN intervals. All of these measurements of short-term variation estimate high-frequency variations in heart rate and thus are highly correlated.

In the frequency domain, the Power Spectral Density (PSD) analysis of the tachogram provides the basic information of how power distributes as a function of frequency.

In short-term recordings, three main spectral components are distinguished [Drinnan et al., 2000]:

- VLF (Very Low Frequencies);
- LF (Low Frequencies);
- HF (High Frequencies).

The variety of frequency domain measures of HRV analysis for short-term recordings are depicted in Table 5.2. The measurement of VLF, LF, and HF power components is usually made in absolute values of power but LF and HF may also be normalized relatively to the total power minus the VLF component.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Units</th>
<th>Description</th>
<th>Freq. range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total power</td>
<td>ms²</td>
<td>The variance of NN intervals over the temporal segment.</td>
<td>&lt;0.4Hz</td>
</tr>
<tr>
<td>VLF</td>
<td>ms²</td>
<td>Power in VLF range.</td>
<td>&lt;0.04 Hz</td>
</tr>
<tr>
<td>LF</td>
<td>ms²</td>
<td>Power in LF range.</td>
<td>0.04-0.15 Hz</td>
</tr>
<tr>
<td>LF norm</td>
<td>ms²</td>
<td>LF power in normalized units.</td>
<td>0.04-0.15 Hz</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LF/(total power-VLF)*100</td>
<td></td>
</tr>
</tbody>
</table>
Table 5.2: Frequency domain measures of HRV analysis in a short-term recording.

The distribution of the power and the central frequency of LF and HF are not fixed but may vary in relation to changes in autonomic modulations of heart period [Pagani et al., 1986; Furlan et al., 1990; Berger et al., 1986]. The representation of LF and HF in normalized units emphasizes the controlled and balanced behavior of the two branches of the autonomic nervous system. Moreover, the normalization tends to minimize the effect of the changes in total power on the values of LF and HF components. Nevertheless, normalized units should always be quoted with absolute values of the LF and HF power in order to describe completely the distribution of power in spectral components.

Spectral analysis also may be used to analyze the sequence of NN intervals of the entire 24-hour period. The result then includes a ULF component, in addition to VLF, LF, and HF components. Because of the important differences in the interpretation of the results, the spectral analyses of short-term and long-term ECGs should always be strictly distinguished.

Methods for the calculation of PSD may be generally classified as nonparametric and parametric. In most instances, both methods provide comparable results (Fig. 5.2). The advantages of the nonparametric methods are the simplicity of the algorithm used (FFT) and the high processing speed, while the advantages of parametric methods are the smoother spectral components that can be distinguished independent of preselected frequency bands and easy post-processing of the spectrum with an automatic calculation of low and high-frequency power components.
Increased HF power represent too quicker changes in heart rate due to excessive sympathetic activity while increased LF power represent slower changes in heart rate due to parasympathetic activity.

5.2 Detection of SA from frequency analysis of HRV

Previous scientific researches have demonstrated that HRV analysis in the time and frequency domain [Drinnan et al., 2000] can give reasonably high detection accuracy of the sleep apnoea occurrence for manual and automatic screening. These results also show that HRV spectrum analysis give the most informative feature than the statistical method of the time domain method [Xu et al., 2008].

Patients with sleep apnoea tend to have a spectral peak lying between 0.01 and 0.05 Hz, with the width of the peak indicating variability in the recurrence rate of the apnoea [Drinnan et al., 2000]. In this way, by a fixed threshold of 3.15 used to discriminate the ratio of the content of two spectral regions obtained by dividing the area under the spectral curve between 0.01 and 0.05 Hz by the area between 0.005 and 0.01 Hz, an automated classification of SA events can give an overall classification score of 90%. The block diagram of the algorithm of this automatic classification procedure is depicted in Fig. 5.3.
5.3 The Proposed System

The aim of this work is to develop a wearable biomedical system capable of performing real-time short-term HRV analysis for the continuous and real-time monitoring of the sleep apnoea disease at home.

The proposed system is composed of two main parts hosted respectively in the Remote Monitoring Station (RMS) and at the patient’s home (Fig. 5.4).

Figure 5.3: Automatic analysis and classification of SA events by the frequency analysis of HRV [Drinnan et al., 2000].

Figure 5.4: System overview for the continuous detection and screening of SA disease.
On the patient’s side, there are only the internet gateway for the patient’s data transmission to the RMS and the wearable device embedded in a comfortable glove in order to allow the continuous detection and screening during daily life (Fig. 5.5).

Figure 5.5: The glove embedding the wearable device.

The wearable device performs in real-time the short-term (5-min) HRV analysis from NN intervals measured in the PPG signal coming from a standard SpO2 wrapped sensor placed in one of the fingers. Fig. 5.6 shows that the detection of NN intervals from the PPG signal is equivalent to the traditional method based on a QRS detection algorithm from an ECG channel. Moreover, the use of a standard SpO2 wrapped sensor placed in one of the fingers instead of the use of electrocardiogram electrodes makes the realized device patient safe, inexpensive and easy to use.

Figure 5.6: NN intervals detection from lead-II ECG (A) and PPG (B).
Applying a fixed threshold at the frequency analysis of the heart rate variability [Drinnan et al., 2000], the device can continuously detect how many sleep apnoea events has been occurred in the previous 5-min and can activate an alarm if the number of sleep apnoea events per hour cross a guard level.

Through the radio frequency link in the ISM band, the smart glove communicates with the internet gateway connected with a remote monitoring station for the continuous data analysis, clinical monitoring and alarm catch.

On the RMS side there is only a simple PC with a TCP/IP server and a Graphic User Interface (GUI) (Fig. 5.7). The server receives the data stream of the patients connected and the GUI performs the real-time HRV analysis in both frequency and time domain in order to provide to a physician the best framework for the automatic and manual detection of sleep apnoea occurrence.

Figure 5.7: The Graphic User Interface at the Remote Monitoring Station.

In this GUI, there are:

- a listview that lists the name and the battery level of the wearable device of all patients connected with the RMS;
- a little text area that shows selected patient info, such as patient’s address, cell number and medicine usually taken;
- a plot area (PPG) that plots the real time PPG signal acquired;
- a plot area (Tachogram) that plot the short term tachogram;
• a plot area (PSD) that plots the power spectrum density of the short term tachogram acquired for the selected patient;
• a info area with some time and frequency domain index for the last PSD for the selected patient.
Chapter 6

Continuous Cuff-less Assessment of Blood Pressure from PPG signals

The importance of accurate blood pressure measurement in the diagnosis of arterial hypertension, and in the precise titration of antihypertensive treatment, is largely acknowledged in clinical practice. The relevance of this issue has recently been further emphasized by hypertension management guidelines [ESH, 2003].

Standard Non-Invasive Blood Pressure (NIBP) monitoring in a clinical or research setting is often performed using techniques that evolved in the 19th century. However, traditional sphygmomanometry is unable to monitor the short-term dynamic variability that occurs with BP and cannot be used for continuous monitoring of BP.

A noninvasive beat to beat measurement of BP would be extremely valuable. A number of approaches have been developed, including most notably finger blood-volume clamping [Penaz, 1969] and arterial tonometry. Although some studies have suggested that there is usually reasonable accuracy with these systems, the technology itself is generally expensive, cumbersome, and prone to movement artifact.

An alternative technique involves measuring the transit time of the pulse pressure wave through the arterial tree, but this is an unreliable marker of beat-to-beat blood pressure [Payne et al., 2006].

This chapter give an overview of traditional an innovative method for non-invasive blood pressure measurement, and illustrate the calibration of the under research cuff-less method of the assessment of BP from PPG signals into the wearable device developed in this thesis.
6.1 Arterial Blood Pressure

The maximum pressure is called the systolic pressure while the minimum pressure is called the diastolic pressure. The integral of the pressure over one cardiac cycle is called the mean pressure, or mean arterial pressure (MAP). The difference between the systolic and diastolic pressure is known as the pulse pressure. Blood pressure measurements are usually acquired from the arterial system and may be measured invasively and non-invasively.

6.2 Classical NIBP Measurements

6.2.1 Auscultation

Auscultation is the blood pressure reference method for non-invasive blood pressure measurements [AAMI, 1992]. It requires a sphygmomanometer and a stethoscope (Fig. 6.1).

![Image of sphygmomanometer and stethoscope in auscultation method.]

Figure 6.1: Using sphygmomanometer and stethoscope in auscultation method.

A sphygmomanometer consists of an inflatable bladder with a cuff, rubber bulb for cuff inflation and a manometer for pressure reading.

Using this method, a clinician inflates the external cuff surrounding the tissue over the brachial artery until the inflated pressure is above systole. The clinician then slowly releases the pressure while listening for heart sounds below the level of the cuff. When the systolic peaks are higher than the cuff pressure, audible sounds called Korotkoff sounds are
generated by the flow of blood and vibrations of the vessel under the cuff. The pressure reading at the first detection of the Korotkoff sounds indicates systolic pressure. As the cuff pressure is further decreased, the Korotkoff sounds become muffled and disappear. The pressure reading during the transition from muffling to silence is the diastolic pressure [Geddes, 1970].

![Figure 6.2: Cuff pressure trend in auscultatory method.](image)

### 6.2.2 Oscillometry

The most widely used automatic indirect blood pressure method is oscillometry [Ramsey, 1979]. With this method, an external cuff is inflated above the systolic pressure over the brachial artery. The external pressure is then slowly decreased, while measuring the amplitude of pressure oscillations that are created by expansion of the arterial wall during each beat as blood is forced through the artery. The systolic pressure is determined when the cuff pressure significantly increases in amplitude. The mean pressure is determined when the cuff pressure reaches the maximum amplitude. Because there is no clear transition in amplitude during diastole pressure, proprietary algorithms are used to measure diastolic pressure (Fig. 6.3). However, usually the diastolic pressure is determined by the last major oscillation. Although the mechanism behind this phenomenon of pressure amplitudes is unknown, the MAP measurement is very robust. However, measurement of systolic pressure is less robust, and measurement of diastolic pressure may be questionable, depending on the proprietary algorithms utilized.

![Figure 6.3: Cuff pressure trend in oscillometric method.](image)
Cuff size relative to arm circumference affects the accuracy of noninvasive measurements [Van Montfrans et al., 1987]. The original Riva Rocci method is still the most reliable but its precision depend on the expertise of the operator and do not allow a continuous measurement.

### 6.3 Innovative NIBP Measurements

#### 6.3.1 Finapres

In 1969, Jan Penaz of the University J.E. Purkyne in Czechoslovakia patented a continuous non-invasive blood pressure method that is commonly referred to as the volume clamp method or method of Penaz [Penaz, 1969]. This method was subsequently developed by Karel Wesseling [Cejnar at al., 1988; Yamakoshi et al., 1983] in the 1980s. Wesseling’s implementation is the basis of the Finapres (FINger Artery PRESSure) monitor manufactured by Ohmeda [Molhoek et al., 1984; Wesseling et al., 1990]. Using photoplethysmography, light from a light emitting diode is passed through a finger and detected. Simultaneously, a cuff is securely placed over the same finger with the cuff pressure continuously adjusted to maintain a constant transmural pressure across the digital arteries (Fig. 6.4).

![Finapres schema](image)

During self calibration, the cuff pressure at which the pulsation amplitudes and therefore the finger blood volume are maximized is assumed to be an indication of the pressure during which the artery is in a relaxed state and the arterial diameter is maximal. After calibration, the monitor servo system attempts to control the pressure to this set point. At the set point, it is assumed that the transmural pressure equals zero and an approximately linear relationship exists between the intraarterial pressure and detected light. The systolic, mean, and diastolic finger pressures detected at this set point are assumed to equal the intraarterial radial systolic, mean and diastolic pressures. The last version of the Finapres is
available in ambulatory version and in portable version. The ambulatory version is called Finometer instead the portable version is called the Portapres (Fig. 6.5).

![Finapres Finometer (left) and Finapres Portapres (right).](image)

**Figure 6.5:** Finapres Finometer (left) and Finapres Portapres (right).

The Finapres devices allow continuous blood pressure monitoring but it has a annoying cuff that not allow a long-term measurement and sometimes it not respect the AAMI specifications (±8 mmHg) [AAMI, 1992].
6.4 Cuff-less NIBP Measurements

6.4.1 Methods based on Pulse Transit Time

Pulse transit time (PTT), defined as the time interval for a pressure pulse to travel from one arterial site to another [McDonald, 1974], is one of the proposed parameters for the non-invasive beat to beat blood pressure estimation. PTT is usually measured as the time interval from the characteristic points of electrocardiogram and photoplethysmographic signal in the same heart cycle (Fig. 6.6).

**Figure 6.6: Definitions of PTT.**

Previous literature reported that the PTT is inversely related to blood pressure:

\[
SBP = \frac{1}{b} \ln \left( \frac{L^2 \rho_b}{PTT^2} - 1 \right) \tag{6.1}
\]

where \( L \) is the distance traveled by the pulse, \( b \) is a subject-dependent coefficient characterizing the artery properties, and \( \rho \) is the blood density.

\[
SBP = A \frac{L}{PTT} + B \tag{6.2}
\]

where \( L \) is a biometric parameter, typically related to the subject’s body size, \( A \) is a sensitivity factor and \( B \) is a constant.
6.4.2 Methods based on PPG Signals

PTT techniques require two independent channels to measure the time interval, which are less convenient to users in practical application. Moreover, relationships based on PTT are not reliable enough to make the PTT-based approach as a surrogate marker of SBP, and DBP and MBP cannot be predicted from PTT [Payne et al., 2006].

For these reasons, some works were made to estimate blood pressure by using photoplethysmographic signal. Several features (pulse amplitude, area, width, half width, systolic and diastolic time (DT), etc.) extracted from PPG signals in time domain were claimed to correlate with BP [Awad et al., 2001; Teng et al., 2003]. Experimental findings of this method demonstrate the promise of the method for the real-time characterization of a patient’s vascular compliance curve using a standard sigmoidal model. Once parameterized, the curve can be utilized for the real-time measurement of arterial blood pressure using non-invasive photoplethysmographic sensor.

It is well known that the arterial wall demonstrates a highly nonlinear relationship between vascular volume $V$ and transmural pressure $P_{tm}$ [Ando et al., 1991], which is defined as the difference between the internal pressure $P_{hem}$, external hydrostatic pressure $P_{hydro}$, and any additional externally applied pressure $P_{ext}$:

$$P_{tm} = P_{hem} + P_{hydro} - P_{ext}$$  \hspace{1cm} (6.3)

Typically, this nonlinear relationship can be characterized by fitting a sigmoid function to experimentally derived pressure-volume data,

$$PPG \propto V = \frac{b_1}{1 + e^{-b_2P_{tm}}}$$  \hspace{1cm} (6.4)

where $b_1$ and $b_2$ are fitting parameters with units of volume and inverse pressure, respectively.

$$b_1 = \lim_{P_{tm} \to \infty}(V)$$  \hspace{1cm} (6.5)

The remaining model parameter, $b_2$, requires additional knowledge about the pressure-volume relationship. One method for estimating this parameter is to assume that the middle section of the sigmoid is approximately linear, with slope $b_2$. If valid, the value of $b_2$ can then be estimated from any two points (pressures) within this linear region (Fig. 6.7). The main difficulty with this approach is that it appears to require knowledge of the compliance curve. In particular, it requires knowledge of the pressures containing the central region of the compliance curve.

90
6.5 Calibration of the PPG to BP for the Continuous Assessment in the Developed Wearable Device

The aim of this work is the development of a wearable biomedical system for the continuous and real-time monitoring of blood pressure through the validation of the PPG calibration method. Developing such a device has the advantage to be easy to wear thanks to the use of a standard SpO2 sensor. Such wearable device can easily be embedded, for example, in a comfortable glove using a fingertip sensor or in a shoe using a standard wrapped sensor placed in one of the foot fingers.

The calibration procedure was made on five healthy volunteers wearing the device embedded in a comfortable cotton glove with the SpO2 standard sensor wrapped in the index finger. In order to calibrate the PPG signal to blood pressure, the amplitude of the PPG signal is measured in eight different heights for each subject and the diastolic and systolic value are measured with the aid of a standard sphygmomanometer.

Fig. 6.8 shows the resultant plots obtained plotting the amplitude of PPG signal versus the measured blood pressure for every height and for every volunteer.
These experimental data confirm the standard sigmoidal model defined in Eq. (6.4). Once parametrized through the calibration process illustrated above, the sigmoidal curve can be used to compute a real-time cuffless estimation of MBP using only a non-invasive PPG sensor. However, this method is subjective and the calibrations curve obtained are different for each volunteer.
Hence, is necessary a normalization of the amplitude of the PPG signals in order to obtain a unique calibration curve for all subjects. Fig. 6.9 shows a graph obtained normalizing the amplitude of the PPG signals respect to the PPG signal with the less amplitude.

**Figure 6.9:** Normalized PPG versus blood pressure graph.

The graph obtained shows that this normalization process don’t permit to have a unique objective calibration curve useful for every patient. Hence this method permit to calibrate a sigmoidal curve that correlate the PPG amplitude of a patient but the standardization of a calibration protocol is needed in order to successfully utilize such as method. Moreover, the calibrated sigmoidal curve is more sensible to motion artifacts, so a little shift of the SpO2 sensor strongly influence the PPG amplitude and hence the sigmoidal curve.
Chapter 7

Continuous Monitoring of patients with HA, AF and DVT through Wavelet Transform of PPG signals

Due to the wide variety of signals and problems encountered in medicine and biology, the spectrum of applications of the wavelet transform in biomedical signal processing is extremely large [Aldroubi et al., 1996]. The main difficulty in dealing with biomedical objects is the extreme variability of the signals and the necessity to operate on a case by case basis.

Wavelet analysis in biomedical signal processing ranges from the analysis of the more traditional physiological signals such as the electrocardiogram (ECG), to the very recent imaging modalities including positron emission tomography (PET) and magnetic resonance imaging (MRI) [Wright, 2000]. It is used usually to remove noise, artefacts, or transient features.

This chapter show how it is possible to correlate PPG signal to heart attack, deep vein thrombosis and atrial fibrillation through a wavelet analysis in order to implement in the developed wearable device a clinical aid in the treatment of these diseases by anticoagulant oral therapy.
7.1 Wavelet Analysis

The wavelet transform can be used as another way to describe the properties of a waveform that changes over time, but in this case the waveform is divided not into sections of time, but segments of scale [Akansu et al., 1992].

In wavelet analysis, a variety of different probing functions may be used, but the family always consists of enlarged or compressed versions of the basic function, as well as translations. This concept leads to the defining equation for the continuous wavelet transform (CWT):

\[
CWT_x = W(a, b) = \int_{-\infty}^{+\infty} x(t) \frac{1}{\sqrt{|a|}} \psi^*(\frac{t-b}{a})dt
\]  

(7.1)

where \( b \) acts to translate the function across \( x(t) \) and the variable \( a \) acts to vary the time scale of the probing function \( \Psi \). If \( a \) is greater than one, the wavelet function \( \Psi \), is stretched along the time axis, and if it is less than one (but still positive) it contacts the function. Negative values of \( a \) simply flip the probing function on the time axis. While the probing function \( \Psi \) could be any of a number of different functions, it always takes on an oscillatory form, hence the term wavelet. The * indicates the operation of complex conjugation, and the normalizing factor \( \frac{1}{\sqrt{a}} \) ensures that the energy is the same for all values of \( a \). If \( b = 0 \), and \( a = 1 \), then the wavelet is in its natural form, which is termed the mother wavelet [Rao et al., 1998].

![Figure 7.1: A Morlet mother wavelet (a = 1) with two dilations (a = 2 and 4) and one contraction (a = 0.5).](image-url)
The wavelet coefficients $W(a, b)$, describe the correlation between the waveform and the wavelet at various translations and scales: the similarity between the waveform and the wavelet at a given combination of scale and position $a, b$. In other words, the coefficients provide the amplitudes of a series of wavelets, over a range of scales and translations, that would need to be added together to reconstruct the original signal. From this perspective, wavelet analysis can be thought of as a search over the waveform of interest for activity that most clearly approximates the shape of the wavelet. This search is carried out over a range of wavelet sizes: the time span of the wavelet varies although its shape remains the same.

If the wavelet function $\Psi$ is appropriately chosen, then it is possible to reconstruct the original waveform from the wavelet coefficients just as in the Fourier transform [Strang et al., 1997]. Since the CWT decomposes the waveform into coefficients of two variables, $a$ and $b$, a double summation (or integration) is required to recover the original signal from the coefficients:

$$x(t) = \frac{1}{C} \int_{a=-\infty}^{+\infty} \int_{b=-\infty}^{+\infty} W(a, b)\psi_{a,b}(t) \, da \, db$$  \hspace{1cm} (7.2)$$

$$C = \int_{-\infty}^{+\infty} |\psi(\omega)|^2 \frac{d\omega}{|\omega|}$$  \hspace{1cm} (7.3)$$

Figure 7.2: CWT of a 15Hz sine before $t = 1$ sec and 50Hz after $t = 1$ sec.

Figure 7.3: Haar, Morlet, Mexican Hat and Meyer wavelet [Misiti et al., 2008].
Figure 7.4: Coiflet wavelets [Misiti et al., 2008].

Figure 7.5: Symlet wavelets [Misiti et al., 2008].

Figure 7.6: Daubechies wavelets [Misiti et al., 2008].
Figure 7.7: Biorthogonal wavelets [Misiti et al., 2008].
The CWT has one serious problem: it is highly redundant. The CWT provides an oversampling of the original waveform: many more coefficients are generated than are actually needed to uniquely specify the signal. This redundancy is usually not a problem in analysis applications but will be costly if the application calls for recovery of the original signal. For recovery, all of the coefficients will be required and the computational effort could be excessive. In applications that require bilateral transformations, we would prefer a transform that produces the minimum number of coefficients required to recover accurately the original signal.

The discrete wavelet transform (DWT) achieves this requirement by restricting the variation in translation and scale, usually to powers of 2. When the scale is changed in powers of 2, the discrete wavelet transform is sometimes termed the dyadic wavelet transform. The DWT may still require redundancy to produce a bilateral transform unless the wavelet is carefully chosen such that it leads to an orthogonal family. In this case, the DWT will produce a non-redundant, bilateral transform [Wickerhauser, 1994].

In the DWT, a new concept is introduced termed the scaling function, a function that facilitates computation of the DWT. To implement the DWT efficiently, the finest resolution is computed first. The computation then proceeds to coarser resolutions, but rather than start over on the original waveform, the computation uses a smoothed version of the fine resolution waveform. This smoothed version is obtained with the help of the scaling function. In fact, the scaling function is sometimes referred to as the smoothing function. The definition of the scaling function uses a dilation or a two-scale difference equation:

$$\phi(t) = \sum_{n=-\infty}^{\infty} \sqrt{2} c(n) \phi(2t - n)$$  \hspace{1cm} (7.4)

where $c(n)$ is a series of scalars that defines the specific scaling function. This equation involves two time scales ($t$ and $2t$) and can be quite difficult to solve.

In the DWT, the wavelet itself can be defined from the scaling function:

$$\psi(t) = \sum_{n=-\infty}^{\infty} \sqrt{2} d(n) \phi(2t - n)$$  \hspace{1cm} (7.5)

where $d(n)$ is a series of scalars that are related to the waveform $x(t)$ and that define the discrete wavelet in terms of the scaling function.

While the DWT can be implemented using the above equations, it is usually implemented using filter bank techniques. The use of a group of filters to divide up a signal into various spectral components is termed sub-band coding. The most basic implementation of the DWT uses only two filters as in the filter bank shown in Fig. 7.8.
The waveform under analysis is divided into two components, \( y_{lp}(n) \) and \( y_{hp}(n) \), by the digital filters \( H_0(\omega) \) and \( H_1(\omega) \). The spectral characteristics of the two filters must be carefully chosen with \( H_0(\omega) \) having a low-pass spectral characteristic and \( H_1(\omega) \) a high-pass spectral characteristic. The high-pass filter is analogous to the application of the wavelet to the original signal, while the low-pass filter is analogous to the application of the scaling or smoothing function. If the filters are invertible filters, then it is possible, at least in theory, to construct complementary filters that will recover the original waveform from either of the sub-band signals, \( y_{lp}(n) \) (approximation) and \( y_{hp}(n) \) (detail). The original signal can often be recovered even if the filters are not invertible, but both sub-band signals will need to be used. Signal recovery is illustrated in Fig. 7.9 where a second pair of filters, \( G_0(\omega) \) and \( G_1(\omega) \) operate on the high and low-pass sub-band signals and their sum is used to reconstruct a close approximation of the original signal.

The Filter Bank that decomposes the original signal is usually termed the analysis filters while the filter bank that reconstructs the signal is termed the syntheses filters. In some analysis applications only the sub-band signals are of interest and reconstruction is not needed, but in many wavelet applications, some operation is performed on the sub-band signals, \( y_{lp}(n) \) and \( y_{hp}(n) \), before reconstruction of the output signal. In such cases, the output will no longer be exactly the same as the input.
There is one major concern with the general approach schematized in Fig. 7.9: it requires the generation of twice as many points as are in the original waveform. This problem will only get worse if more filters are added to the filter bank. If the analysis filters are correctly chosen, then it is possible to reduce the length of \( y_{lp}(n) \) and \( y_{hp}(n) \) by one half and still be able to recover the original waveform. To reduce the signal samples by one half and still represent the same overall time period, we need to eliminate every odd point. This operation is known as downsampling. The downsampled version of \( y(n) \) would then include only the samples with even indices \([y(2), y(4), y(6), \ldots]\) of the filtered signal.

If downsampling is used, then there must be some method for recovering the missing data samples (those with odd indices) in order to reconstruct the original signal. An operation termed upsampling accomplishes this operation by replacing the missing points with zeros.

**Analysis Filters**

\[
x(n) \xrightarrow{H_0(\omega)} y_{lp}(n) \xrightarrow{\downarrow 2} y_{hp1}(n) \xrightarrow{H_1(\omega)} y_{hp2}(n) \xrightarrow{\downarrow 2} y_{lp}(n)
\]

**Synthesis Filters**

\[
y_{lp}(n) \xrightarrow{\uparrow 2} G_0(\omega) \xrightarrow{\Sigma} y_{hp2}(n) \xrightarrow{\downarrow 2} G_1(\omega) \xrightarrow{\Sigma} y_{hp1}(n) \xrightarrow{\downarrow 2} H_0(\omega) \xrightarrow{\Sigma} x'(n) \xrightarrow{H_2(\omega)}
\]

**Figure 7.10:** A typical wavelet application using three filters where the downsampling and upsampling processes are shown.

Designing the filters in a wavelet filter bank can be quite challenging because the filters must meet a number of criteria. The most important concern is the ability to recover the original signal after passing through the analysis and synthesis filter banks.

A way to choose an optimal mother wavelet function is the algorithm developed in [Castro et al., 2000]. From the wavelet family set, the optimal mother wavelet is chosen by mean of best reconstruction quality and minimum coefficients obtained from the 4\(^{th}\) decomposition depth.
\[ R = 100 \frac{||X||^2}{||X_c||^2} \]  \hspace{1cm} (7.6)

where \( R \) is the reconstruct value, \( X \) is the signal and \( X_c \) is the reconstructed signal. The minimum reconstruction value is fixed to be 99%, which guarantees the preservation of important signal components.

7.2 Wavelet Analysis in Biomedical Signal Processing

Due to the wide variety of signals and problems encountered in medicine and biology, the spectrum of applications of the wavelet transform in biomedical signal processing is extremely large [Aldroubi et al., 1996]. The main difficulty in dealing with biomedical objects is the extreme variability of the signals and the necessity to operate on a case by case basis. Often, one does not know a priori what is the pertinent information and at which scale it is located. Moreover, the information of interest in biomedical signals is often a combination of features that are well localized temporally or spatially and others that are more diffuse. This requires the use of analysis methods sufficiently versatile to handle events that can be at opposite extremes in terms of their time-frequency localization.

Wavelet analysis in biomedical signal processing ranges from the analysis of the more traditional physiological signals such as the electrocardiogram (ECG), to the very recent imaging modalities including positron emission tomography (PET) and magnetic resonance imaging (MRI) [Wright, 2000]. It is used usually to remove noise, artefacts, or transient features.

In ECG processing, the wavelet analysis is used for QRS detection [Li et al., 1995]. In heart sound analysis, the WT is a useful tool for the time-frequency analysis and characterization of the murmurs, which are primarily caused by blood flow turbulence and characteristic of cardiac disease such as aortic stenosis or valve defects [Khadra et al., 1991]. In Computer-Assisted Tomography (CAT) and in Magnetic Resonance Imaging (MRI), the wavelet analysis is used for image enhancement and for detection of microcalcifications in mammograms [Clarke et al., 1994; Qian et al., 1994; Strickland et al., 1994; Yoshida et al., 2001]. In functional image analysis such as PET and fMRI, the wavelet analysis is utilized for investigating the neuronal activity of the brain in vivo [Ruttimann et al., 1993; Rio et al., 1994; Unser et al., 1995].

In PPG wavelet analysis, Leonard et al. discovered a wavelet-based algorithm for identification of unwell children from pulse oximeter waveform [Leonard et al., 2004]. In the past the photoplethysmogram has traditionally been thought to be of no value except to determine if a pulse oximeter is picking up a good signal and thus giving an accurate
reading [Moyle, 1994]. However, other works have shown that other cardiorespiratory variables influence the waveform [Vegfors et al., 1994; Middleton et al., 2000].

Photoplethysmogram signals were obtained from 20 children fulfilling the study entry criteria and from 12 children of a group control. All traces were analyzed using wavelet techniques, first to remove the effects of noise and drift, and then to determine the wavelet based features that permit separation of the control group from the unwell children. Maximal separation was achieved when the wavelet power across the 1 Hz frequency band was plotted against the entropy at the 7 Hz band (Fig. 7.11). The probability that the two samples come from populations that are identical is $p = 0.00002$.

![Figure 7.11: Scatter plot of wavelet power at 1 Hz against entropy at 7 Hz showing differentiation of well children from unwell children.](image)

### 7.3 Wavelet Analysis of Acquired PPG signals

Using the wearable device depicted in section 4.3, the Photoplethysmogram signals were obtained from 27 patients with atrial fibrillation disease, 30 patients with deep vein thrombosis disease, 15 patient with heart attack and from 20 healthy subjects of a group control (mean age 40 years). The aim of this work was to discover a wavelet analysis able to discriminate between the PPGs of the control group and the PPGs of the patients with
disease. In order to exclude results in the wavelet analysis imputable to different therapies from the patients sets, all of these patients follow the same anticoagulant oral therapy.

Starting from the PPG signals of the healthy subjects and in concordance with the Castro algorithm for choosing the optimal mother wavelet function (Eq. (7.6)), the optimal wavelet function used was the Sym7.

![Figure 7.12: PPG signal of a healthy subject acquired with the wearable device realized.](image)

The wavelet decomposition used is shown in Fig. 7.13. Using the filter coefficients for the Sym7 wavelet, all of PPG signals are decomposed in detail and approximation until the second level of filter banks.

![Figure 7.13: Wavelet decomposition used for the healthy and ill PPG signal analysis.](image)

In order to discriminate between healthy and ill PPG signals, three different parameters was used after the wavelet transform:

\[
\text{(energy)} \quad E_x = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2 \quad (7.7)
\]

where \( x \) is the waveform being analyzed, \( N \) is the number of samples of \( x \), and \( \bar{x} \) is the mean value of \( x \).

\[
\text{(normalized energy)} \quad E_{\text{norm},x} = \frac{E_x}{\sqrt{\sum_{i=1}^{N} (E_{x_i})^2}} \quad (7.8)
\]
where $E_x$ is the energy of $x$, $E_{x_i}$ is the energy of any other signal in the same set, and $N$ is the number of signals in the set.

\[
S_x = \sum_{i=1}^{N} \ln x_i^2
\]  

(entropy) \hspace{1cm} (7.9)

where $S_x$ is the entropy of $x$ and $N$ is the number of samples of $x$.

### 7.3.1 Analysis of Atrial Fibrillation PPG signals

In spite of the PPG signal in patients with AF disease is clearly different from a PPG signal of a healthy subject, plotting the energy (Fig. 7.15) or the normalized energy (Fig. 7.16) of the wavelet analysis at various scale of approximation versus the entropy, don’t permit to discriminate PPGs of patients with AF disease from PPGs of healthy subjects.

**Figure 7.14:** PPG signal of a subject with AF disease acquired with the wearable device realized.

**Figure 7.15:** Energy versus entropy of the DD2 level of the wavelet analysis of healthy PPGs and of PPGs from subjects with AF disease.
7.3.2 Analysis of Deep Vein Thrombosis PPG signals

The PPGs acquired from patients suffering from DVT are very similar to those acquired from healthy subjects. The wavelet analysis don’t permit to distinguish from healthy PPGs and ill PPGs, although the greater part of ill PPGs are clustered (Fig. 7.18).
Figure 7.18: Energy versus entropy of the AD2 level of the wavelet analysis of healthy PPGs and of PPGs from subjects with DVT disease.

7.3.3 Analysis of Heart Attack PPG signals

Figure 7.19: PPG signal of a subject with HA disease acquired with the wearable device realized.

In spite of the PPG signals in patients with HA disease are very similar to PPG signals of healthy subjects, the wavelet analysis permit to clearly distinguish between healthy PPGs and ill PPGs, simply plotting energy versus entropy of the wavelet transform of the second level of approximation (AA2).
Figure 7.20: Energy versus entropy of the AA2 level of the wavelet analysis of healthy PPGs and of PPGs from subjects with HA disease.
In order to reduce the costs for both the patients and the public administration, tele-home care systems are today often used for clinical practice. They permit to avoid the overcrowding of ambulatories for simple routine examinations and, in the case of chronic patients that must be frequently monitored, to avoid the costs and the waste of time of an ambulatory examination for simple physiological measurements that they could easily perform at home, sending the recorded data to the care staff in the appropriate department. However, these telemedicine systems are usually PC based or use complex hardware and software products. Even if such technologies are quite popular today, they are frequently used only by relatively young people, which are not the primary tele-home care target. Cardiopathic and diabetic patients need to be frequently monitored and in some cases they could easily perform at home the requested physiological measurements (i.e. glycemia, heart rate, blood pressure and body weight) sending the measured data to the care staff in the hospital.

This chapter shows a low-cost tele-home care system based on the Digital Video Broadcast Terrestrial (DVB-T) technology for the remote monitoring of patients with diabetes and suffering from cardiovascular diseases, even exploitable by elderly people.
8.1 Introduction to Interactive TV

8.1.1 Digital Video Broadcasting

The Digital Video Broadcasting (DVB) Project is an Alliance of about 300 companies aimed to define the Digital Television standardization. There are different standards based on the transmission mean [DVB, 1999]:

- DVB-S for Satellite;
- DVB-C for Cable;
- DVB-T for Terrestrial;
- DVB-H for handheld devices.

These standards define the physical layer and data link layer of the distribution system. Devices interact with the physical layer via a synchronous parallel interface (SPI), synchronous serial interface (SSI), or asynchronous serial interface (ASI). All data is transmitted in MPEG-2 transport streams with some additional constraints (DVB-MPEG).

These distribution systems differ mainly in the modulation schemes used and error correcting codes used, due to the different technical constraints. DVB-S uses QPSK, 8PSK or 16-QAM. DVB-C uses QAM: 16-QAM, 32-QAM, 64-QAM, 128-QAM or 256-QAM. Lastly, DVB-T uses 16-QAM, 64-QAM or QPSK in combination with COFDM and can support hierarchical modulation.

Figure 8.1: DVB standards over the world.
8.1.2 Digital Video Broadcasting Terrestrial

Digital Video Broadcasting Terrestrial (DVB-T) is the DVB European standard for the broadcast transmission of digital terrestrial television. It is designed to transmit a compressed digital audio/video/data MPEG-2 stream, using Orthogonal Frequency Divisional Multiplexing and a modulation technique compatible with the traditional 8MHz bandwidth of the analog transmission channel (i.e. QPSK, 16QAM and 64QAM). The compressed video, audio, and data streams of one TV channels are multiplexed together to form a Programme Elementary Stream (PES). The basic digital stream that a set-top box can receive is an MPEG-2 Transport Stream (TS), which is formed joining together one or more PES. From the receiver’s point of view, all the TV channel and services are received at once but it only demultiplexes and then decodes the selected content, one at a time, from the received TS. Fig. 8.2 shows the scheme of a DVB-T transmission system [Morris et al., 2005].

![Scheme of a DVB-T transmission system](image)

**Figure 8.2: Scheme of a DVB-T transmission system.**

- **Source coding and MPEG-2 multiplexing (MUX):** compressed video, compressed audio, and data streams are multiplexed into PSs (Programme Streams). One or more PSs are joined together into an MPEG-2 TS (MPEG-2 Transport Stream); this is the basic digital stream which is being transmitted and received by home Set-Top Boxes.
Allowed bitrates for the transported data depend on a number of coding and modulation parameters: it can range from about 5 to about 32 Mbps.

- **Splitter:** two different TSs can be transmitted at the same time, using a technique called Hierarchical Transmission (HT). It may be used to transmit, for example, a standard definition SDTV signal and a high definition HDTV signal on the same carrier. Generally, the SDTV signal is more robust than the HDTV one. At the receiver, depending on the quality of the received signal, the STB may be able to decode the HDTV stream or, if signal strength lacks, it can switch to the SDTV one.

- **MUX adaptation and energy dispersal:** the MPEG-2 TS is identified as a sequence of data packets, of fixed length (188 bytes). With a technique called energy dispersal, the byte sequence is decorrelated.

- **External encoder:** a first level of error correction is applied to the transmitted data, using a non-binary block code, a Reed-Solomon RS (204, 188) code, allowing the correction of up to a maximum of 8 wrong bytes for each 188-byte packet.

- **External interleaver:** convolutional interleaving is used to rearrange the transmitted data sequence, in such a way that it becomes more rugged to long sequences of errors.

- **Internal encoder:** a second level of error correction is given by a punctured convolutional code, which is often denoted in STBs menus as FEC (Forward error correction). There are five valid coding rates: 1/2, 2/3, 3/4, 5/6, and 7/8.

- **Internal interleaver:** data sequence is rearranged again, aiming to reduce the influence of burst errors. This time, a block interleaving technique is adopted, with a pseudo-random assignment scheme.

- **Mapper:** the digital bit sequence is mapped into a base band modulated sequence of complex symbols. There are three valid modulation schemes: QPSK, 16-QAM, 64-QAM.

- **Frame adaptation:** the complex symbols are grouped in blocks of constant length (1512, 3024, or 6048 symbols per block). A frame is generated, 68 blocks long, and a superframe is built by 4 frames.

- **Pilot and TPS signals:** in order to simplify the reception of the signal being transmitted on the terrestrial radio channel, additional signals are inserted in each block. Pilot signals are used during the synchronization and equalization phase, while TPS signals (Transmission Parameters Signalling) send the parameters of the transmitted signal and to unequivocally identify the transmission cell. The receiver must be able to synchronize, equalize, and decode the signal to gain access to the information held by the TPS pilots. Thus, the receiver must know this information beforehand, and the TPS data is only used in special cases, such as changes in the parameters, resynchronizations, etc.

- **OFDM Modulation:** the sequence of blocks is modulated according to the OFDM technique, using 2048, 4096, or 8192 carriers (2k, 4k, 8k mode, respectively).
Increasing the number of carriers does not modify the payload bit rate, which remains constant.

- **Guard interval insertion**: to decrease receiver complexity, every OFDM block is extended, copying in front of it its own end (cyclic prefix). The width of such guard interval can be 1/32, 1/16, 1/8, or 1/4 that of the original block length. Cyclic prefix is required to operate single frequency networks, where there may exist a not eliminable interference coming from several sites transmitting the same program on the same carrier frequency.

- **DAC and front-end**: the digital signal is transformed into an analog signal, with a digital-to-analog converter (DAC), and then modulated to radio frequency (VHF, UHF) by the RF front-end. The occupied bandwidth is designed to accommodate each single DVB-T signal into 5, 6, 7, or 8 MHz wide channels.

### 8.1.3 Multimedia Home Platform

The Multimedia Home Platform (MHP) defines a generic interface between interactive digital applications and the terminals on which those applications execute. It enables digital content providers to address all types of terminals ranging from low-end to high-end set-top boxes, integrated digital TV sets and multimedia PCs. The MHP extends the existing, successful DVB open standards for broadcast and interactive services in all transmission networks [MHP, 2008].

#### 8.1.3.1 MHP Architecture

The architecture of the MHP is defined in terms of three layers:

- resources;
- system software
- applications.

![MHP Architecture Diagram](image)

*Figure 8.3: MHP architecture.*
Typical MHP resources are MPEG processing, I/O devices, CPU, memory and a graphics system. The system software uses the available resources in order to provide an abstract view of the platform to the applications.

Applications include an application manager (also known as a "navigator") to control the MHP and the applications running on it.

### 8.1.3.2 MHP System Core

MHP is based around a platform known as DVB-J, which includes a virtual machine as defined in the Java Virtual Machine specification from Sun Microsystems. A number of software packages provide generic application program interfaces (APIs) to a wide range of features of the platform. MHP applications access the platform only via these specified APIs. MHP implementations are required to perform a mapping between these specified APIs and the underlying resources and system software.

### 8.1.3.2 MHP Profiles

The MHP specification provides a consistent set of features and functions required for the Enhanced Broadcast, Interactive Broadcast and Internet Access profiles.

![MHP profiles](image)

**Figure 8.4: MHP profiles.**

The Enhanced Broadcast profile is intended for broadcast (one way) services, while the Interactive Broadcast profile supports additional interactive services. The Internet Access profile allows MHP to use the worldwide communication network provided by the Internet.
8.1.4 JavaTV Xlet

JavaTV API are an extension of the Java platform designed for developers who are producing Java-based interactive television content. The Java TV API gives programs written in the Java programming language control of broadcast television receivers and set-top boxes.

An Xlet is a JavaTV application that follows a lifecycle model defined by a specific finite state machine (Fig. 8.5) [Calder et al., 2000].

![Xlet Lifecycle Diagram](image)

*Figure 8.5: Xlet lifecycle model.*

In the **Loaded** state the Xlet has been loaded by the broadcaster in the set-top box but it is yet uninitialized. From here, it can be initialized or **Destroyed**, if an exception occurs. In the **Paused** state the Xlet is initialized and waits to become **Active**, then it starts performing all the required operations.

8.1.5 DSM-CC Object Carousel

Interactive TV generally consists of small applications, typically java applications, which are transmitted together with the A/V content in a transport stream. The application files are carried in the transport stream in a Digital Storage Media Command and Control (DSM-CC) Object Carousel [Borko et al., 1995]. The DSM-CC Object Carousel is a protocol that enables the transmission of objects from a server to a client. The objects are transmitted in cycles, such that a client can connect to the transmission and receive all the objects without the need for a hack channel from the client to the server. In interactive TV a user can tune to a channel in the middle of a program, so the files for any interactive application have to be repeated frequently in the transport stream. Besides transmitting files the Object Carousel is also used to transmit three other types of object: Directory, Stream and Stream event objects. The directory objects convey the file system structure information, so the names of the files and the directory structure. The Stream objects allow the broadcaster to send A/V related commands like pause, play and stop. Stream event objects are used to carry the
information of the Stream event descriptors in the stream. Stream event descriptors are used to provide a mechanism whereby the interactive application can respond to “events” in the A/V content.

8.2 Others Tele-Home Care Projects in Literature

The primary activity of telemedicine systems was the transmission of diagnostic medical images using television for medical consultation from physician to physician in remote places. Recently, telemedicine systems through the Internet via satellite have become a reality by means of high throughput mixed satellite-web communication channels. This section presents some home-oriented telemedicine projects making use of DVB and Internet, without the claim to be exhaustive.

The Interactive Satellite Multimedia Information System (ISIS) project [Pierucci et al., 2000], realized a telemedicine system based on satellite communication for furnishing interactive services for residential users together with the traditional TV distribution, porting typical applications developed for terrestrial network (Internet) to the satellite digital video broadcasting technology (DVB-S). By a dual-band terminal, connected in uplink via Eutelsat satellite and in downlink via the Italsat satellite, the ISIS system provides a small and low-cost transceiver for sites not connected to the Internet or connected only in dial-up low-speed mode (Fig. 8.6). Through his Java application for the DVB-S transceiver, called Medical Environment for Diagnostic Images (MEDI), the project can manage a medical image database for remote expert consultation and so it demonstrated the feasibility of satellite-based interactive multimedia services for telemedicine purposes.

Figure 8.6: ISIS architecture scheme.
In 2001, a multipurpose health care telemedicine system with a base unit and a mobile unit was developed in Greece too [Kyriacou et al., 2001]. At patient’s home, the mobile unit allows the transmission of vital bio-signals and static images of the patient to the base unit at the physician site (office or hospital). The mobile unit device is compliant with some of the commercial main vital signs monitor and it is able to transmit ECGs, non-invasive blood pressure (NIBP), body temperature, percentage of arterial oxygen saturation (SpO2) and heart rate. Based on the TCP/IP protocol, the communication between the two parts ensures safe data transmission and the possibility to use different telecommunication means (GSM or satellite links) (Fig. 8.7).

Figure 8.7: Kyriacou’s architecture scheme.

Launched in January 2001, the Universal Remote Signal Acquisition For hEalth (U-R-Safe) tele-home care project [Mailhes et al., 2003], created a mobile telemedicine system for home monitoring to be used by elderly people and disabled patients in Europe. Via short range Wireless Personal Area Network (WPAN), wearable ECG and SpO2 sensors are connected to a portable electronic device, able to send the recorded data to a remote central through the TCP/IP protocol and the wireless public network (GPRS, UMTS and GEO satellites). The portable unit is also capable of sending an alarm when patient fells sick, falls or pushes a button (Fig. 8.8).
The Standard and Interoperable Satellite Solution to Deploy Health Care Services Over Wide Area (HEALTHWARE) project [Loghelongue, 2007], is an integrated project of the Aeronautics and Space thematic priority of the 6th Framework Program (FP6) for satellite telecommunications systems and telemedicine applications. Thanks to the Digital Video Broadcasting - Return Channel by Satellite (DVB-RCS) technology, that offers satellite reception and transmission capabilities from anywhere, the project aims at developing and validating DVB-RCS based telemedicine solutions. It will focus on the areas of chronic respiratory disease, cardiology and oncology, through four main applications: medical training, tele-consultation, second opinions and monitoring and remote assistance at home.

Figure 8.8: U-R-Safe architecture scheme.

Figure 8.9: HEALTHWARE architecture scheme.
Even if these systems are very interesting, none of them owns the possibility to be used by untrained people to perform single exams rather then continuous monitoring, with an immediate visual confirmation of the quality of the signal measurement. None of them could be used by the patient to control his/her health state through simple measurements even without sending the exam to a remote care center.

### 8.3 System Overview

This system is conceived to keep as low as possible the costs of the extra hardware with respect to a minimum setup composed of a TV and a DVB-T set-top box connected to the telephone line. This way the usability by untrained people not accustomed to digital equipments but the TV set one is also preserved. A global representation of the entire system is depicted in Fig. 8.10.

![Figure 8.10: The main parts of the DVB-T based tele-home care realized system.](image)

At the patient’s home, the TV set is equipped with a standard interactive set-top box. Standard Multimedia Home Platform (MHP) set-top boxes are simple computers with their Operating System, a Java Virtual Machine and a set of standard ports for I/O interfacing, such as a smart card reader and an internal 56Kbps modem for pay-per-view services. From the patient’s point of view, the system exploits the remote control of the set-top box as
control input for the user, the TV screen as output for him, and the telephone line to send the data to a Remote Data Base (RDB). For this work, has been used a Telesystem TS7.4DT set-top box, with the version 21p1 of the real-time Operating System (i.e. Osmosys) and implementing the MHP 1.02 profile with some enhancements, primarily the addition of java packages for both the smart card and the internal modem management.

The interactive application (Java Xlet) is loaded on the patient’s set-top box through an ether broadcast DVB-T transmission, while the user is watching the broadcaster’s channels. It can run in the set-top box thanks to the Java Virtual Machine. The Xlet mainly manages the acquisition and transmission to the Remote Data Base of the patient’s data, but also enables the visualization on the TV screen of the physician’s feedbacks from the Remote Hospital. To allow a secure one-to-one sensitive data transmission in a one-to-many communication channel, like in DVB-T broadcasting, a double key encrypting technique is implemented. In such a way, patients can view and read just their physician’s messages.

A microcontroller-based custom Smart Card (SC), which is programmed by the Remote Hospital, allows an easy use of the system, providing all the necessary data to the Xlet to configure the framework for a specific patient. The SC stores the patient’s name and their personal identification code, the treating physician’s name and their identification number, a permission exam code, some information about the Internet Service Provider for the Internet access (user ID, password, telephone number), the Remote Data Base IP address and the private key needed to read the physician’s feedback. This way the patient does not have to remember any code, number, password, and different patients can share the same hardware platform simply inserting their personal smart card. According to the international recommendations about the patient’s safety, the SC also provides the Xlet with the data exams collected by the acquisition unit (the Base Station) through a wireless link at 2.4GHz in the free ISM band. Fig. 8.11 shows the picture of the first prototype of the realized custom smart card. It is build around a low-cost microcontroller connected through the SPI protocol with a Mi-Wi P2P compliant wireless transceiver.

![Figure 8.11: First prototype of the custom smart card.](image)
The Base Station (BS) enables the acquisition of some physiological parameters needed for the monitoring of diabetic and cardiopathic patients (i.e. blood pressure, heart rate, body temperature, body weight and glycemia) through wired serial RS232 compliant connections to three commercial and certificate biomedical devices (i.e. Menarini Glucocard G+ glucometer, OMRON 705IT sphygmomanometer and A&D UC-321P personal precision electronic body scale). It is important to notice that the BS could communicate with others biomedical devices, depending on the application scenario. To improve patient’s safety, the BS is battery-powered and communicates with the Xlet running on the set-top box through a wireless link with the SC. In this way, the DVB-T equipment and the acquisition unit are wireless coupled, improving patient’s safety. From the patient’s point of view, the BS is only a black box with a very simple user interface: the patient has only to connect the commercial biomedical device to the BS and press a big button. From this point, the Xlet will do all the necessary to send the acquired data to the Remote DB. The SC and BS units are conceived taking into account the low-cost target, then resorting to a very simple and low-cost microcontroller, i.e. the Microchip PIC18F4620. Based on a 8-bit 10 MIPS RISC CPU, it has 64KB of Flash memory, 4KB of RAM, 1KB of EEPROM, a hardware multiplier, 13 channel 10-bit ADC, an USART module and a SPI/I2C controller. The wireless module used in the SC and in the BS is based on the Microchip MRF24J40 transceiver at 2.4GHz with a built-in ZigBee and Mi-Wi protocol stack.

Exploiting the full potential of the DVB-T technology and of the Java Xlet application loaded on the patient’s set-top box, after the acquisition of the physiological parameters of interest, the uplink connection supported by the set-top box is used to send the exam results to a Remote DB for physician analysis.

At the Remote Hospital, only a simple PC acting as a TCP/IP client is needed. Through the Internet, the physician can browse into the patient’s data for clinical examination and can send feedbacks directly to a specific patient using the DVB-T channel through a secure Internet connection with the broadcasting station.

As a matter of fact, we are testing the system on the field in cooperation with a local DVB-T broadcaster (i.e. VIDEOLINA [VIDEOLINA]) but for our preliminary tests we also developed the broadcasting part. Our DVB-T broadcast system is based on the CreaTV MiniLab by Media Solution (Fig. 8.12). Through the JustDvb-It 2.0 software, video and audio MPEG-2 encoded contents are multiplexed together with the Xlet application to build the Single Program Transport Stream (SPTS) to broadcast. A PCI card DekTec DTA-110, a multistandard modulator with UHF Upconverter, is coupled with a 20 dBm RF amplifier (RFbay LPA-4-14) for the ether transmission of the SPTS. To establish the bit rate needed to encode audio, video and multimedia contents in the MPEG-2 format, it is necessary to know the coding and modulation parameters used to transmit the SPTS. We have chosen 1/2 for the
convolutional rate, 16QAM for the modulation type and 1/4 of the original block length for the guard interval. The resulting bit rate is 9953 Kbps.

**Figure 8.12:** The block diagram of our broadcast transmission system.

Once the Xlet has been loaded on the set-top box, it starts to read the data stored in the SC. Without any valid smart card identification the application cannot proceed. The permission exam code stored in the SC is used to allow the patient to perform only some specific exams among those available, hence disabling all the other ones. For example, Figure 3 shows a screenshot of the Xlet sending two acquired glucometer measurements to the Remote DB. A small area on the top-right corner of the Xlet is devoted to the background TV program streaming. A pop-up window alerts the patient if he/she is exiting the Xlet without sending the results to the RDB (in this case the exam results will be lost).

**Figure 8.13:** Picture of the Xlet sending two acquired glucometer measurements to the RDB.
8.4 Overview of the First Version of the System

The first architecture of the system at patient’s home is depicted in Fig. 8.14.

Figure 8.14: Architecture of the first version of the system at patient’s home.

In this version of the system, the smart card is used only to identify the patient providing all the information needed to carry out the proper exams, whereas the biomedical parameters feed the Xlet (called MHPHomecare) through the serial RS-232 port (present only in set-top boxes upgradable).

The microcontroller-based Base Station is controlled by the Xlet and by the user through the set-top box remote control.

8.4.1 Patient’s User Interface

In this first version, the graphic user interface (GUI) of the Xlet (hereafter called MHPHomecare) allow the patient to easily control the acquisition process by means of the remote control. The user can explore the application moving through the different full-screen frames (hereafter called FSFs), choosing among the possible options and functions by means of the remote control keys, as if it would be a PC keyboard.

Three of the full-screen frames are generic whilst there is another FSF for each exam. The three generic FSFs are:

- the primary FSF, the first FSF shown when the application is started. It waits for the smart card insertion in the set-top box for user authentication and then it shows the patient’s name (storing the other authentication data only for the transmission purposes), waiting for patient’s commands (Fig. 8.15);
- the about FSF, that shows the application credits;
• the reduced FSF, that corresponds to the window minimization to allow the patient to watch a TV program while the application is running in background.

![Figure 8.15: Primary FSF.](image)

### 8.4.2 The First Prototypal Base Station

A simple 3-lead ECG is used to test the whole system, so there is only another FSF (a screenshot of this FSF is shown in Fig. 8.16), called ECG exam. The ECG Exam FSF enables the ECG acquisition and shows in real-time the samples acquired from the Base Station and all the other information sent by it about the current exam, such as the hearth rate (in bpm) and a warning message in case of poor quality signal. Once the exam has been saved, from the same FSF it is possible to send it to the server of the RCC, to review it on the TV (with zoom and scroll functions), to save a new exam or to return to the primary FSF.

![Figure 8.16: The ECG exam screen of the MHPHomecare application (ECG Exam FSF).](image)
The three channel electrocardiogram consists in a classical ECG amplifier coupled with a very simple and low cost Digital Signal Controller (DSC), i.e. the Microchip™ dsPIC30F4013 (Fig. 8.17). This is a 16-bit 30MIPS DSC enhanced with DSP hardware, such as a 17-bit x 17-bit multiplier, a 40-bit ALU, two 40-bit saturating accumulators and a 40-bit bidirectional barrel shifter. It also provides several embedded peripherals such as a 12-bit successive approximation analog to digital converter, and an USART, both of them employed in this prototype.

![Figure 8.17: A schematic representation of the lead-I for the prototypical 3-lead ECG base station.](image)

The fundamental characteristics of the Base Station (frequency bandwidth, sampling frequency and the number of bits of the A/D convertor) are chosen according to the recommendations of American Heart Association for standardization and specifications in automated electrocardiography [Bailey et al., 1990]. The ECG amplifier is a low-pass system with a variable gain up to 2000 and a cutoff frequency of 100 Hz. It consists of an optically isolated instrumentation amplifier with high CMRR connected to the wrists of the patient and a simple inverting operational amplifier connected to the right ankle for electromagnetic coupling noise rejection purpose, so that only 3 disposable (or reusable) electrodes are needed. The DSC is programmed to sample the ECG signal at 250 Hz, and it is interfaced through its internal USART to the DVB-T set-top box. According to the international recommendations about the patient’s safety, the physical connection between the Base Station and the set-top box is represented by an infrared (IR) link. The Base Station embeds a
module for the IR transmission of the serial stream, whereas the set-top box exploits an external IR coupler connected to the RS-232 serial port (Fig. 8.18).

Figure 8.18: A screenshot of the prototypal 3-lead ECG base station (left) and RS-232 to IR adapter (right).

In order to reduce both the residual 50/60 Hz environmental electrical noise the base station implemented a digital notch filter, whereas a high-pass digital FIR filter (cutoff frequency 0.05 Hz) was added for the baseline drift elimination (Fig. 8.19). All these computations are performed in real-time on the Base Station, since the set-top box is quite slow in processing due to the Java VM and to the applications running on it.

![Notch Filter](image1)
![High-Pass FIR Filter](image2)

Figure 8.19: Results of the digital notch filter (left) and of the high-pass filter (right) on ECG signal acquired by the prototypal base station.
On the Base Station has been also implemented a simple QRS detector algorithm based on amplitude and first derivative (section 2.2.1.6.1) for heart rate calculation and visualization on the TV screen. The RR interval measurement is performed sample by sample, and if it is too far from the standards a warning code is sent to the set-top box to ask the user to control the correct connection of the electrodes. The average heart rate is sent when the acquisition stops and is referred to the average RR during the acquisition window.
Most of the interventions to prevent and treat diabetes and cardiovascular diseases significantly affect the use of health services. Therefore, more effective and cheaper ways to prevent the major complications of diabetes and CVD are needed.

Diabetes and CVD care through blood pressure control is cost saving mainly because of its large health benefits and relatively low intervention costs. The National Health Service argues strongly that newer systems of care and newer ways of thinking are needed to tackle complex diseases such as diabetes and CVD.

The need to know the long-term effects of diabetes diseases on CVD has inclined to develop new system and biomedical devices for non-invasive and remote monitoring/screening of blood glucose and other parameters related to prediabetes, diabetes, sleep apnoea, heart attack, atrial fibrillation and deep vein thrombosis.

The developed wearable device, battery-powered for patient’s safety, allows the continuous, real-time monitoring of some cardiovascular disorders, and can trigger an alarm if the state of the disorder can be considered dangerous. It is easy to wear and easy to use, thanks to its reduced dimensions and the long term battery autonomy. Moreover, through a wireless link with an internet module, the wearable device allows a long time continuous monitoring by a remote monitoring station, which can alert the physicists in case of dangerous situations. Thanks to the comfortable standard SpO2 sensor applied in one of the fingers, the device can be applied to the continuous monitoring of sleep apnoea through HRV analysis, continuous monitoring of blood pressure through a hydrostatic calibration of PPG signal, and
can aid physicians to the treatment of anticoagulant oral therapy of patients with heart attack through a wavelet analysis of the PPG waveforms.

The presented tele-home care system, based on a framework built around the well-established DVB-T platform, allows remote monitoring of the patients’ conditions directly from their home, without any installation procedure. After 2012, DVB-T will be the only terrestrial television system in Europe, so its diffusion will be capillary, at a very low cost for the final user. Exploiting low-cost technologies for the acquisition unit will then lead to the development of a user-friendly cheap solution for tele-home care. Some strong points of the proposed system are:

- **low-cost**, since beyond the TV set only a low cost microcontroller-based acquisition unit and a custom wireless smart card are required;
- **user-friendly interface**, also for elderly people accustomed to TV but not to computers;
- **flexibility**, since the functionality of the system can change exploiting the same platform;
- **easy maintainability**, since Xlet broadcasting enables upgrade, and the smart card improves usability and completely avoids any installation activity at the patient’s home.

The overall system has been implemented, tested and is currently used for some trials on the field by twenty 65 years old in average selected diabetic and cardiopathes patients. At the end of the trial period, each patient will fill a standard questionnaire (i.e. QUEST, Quebec User Evaluation of Satisfaction with assistive Technology) in order to evaluate patient’s satisfaction and the related services he/she experienced.
References


[VIDEOLINA] http://www.videolina.it


[Wilson et al., 1934] Wilson F.N., Johnston F.S., Hill I.G.W., “The interpretation of the galvanometric curves obtained when one electrode is distant from the heart and the other near or in contact with the ventricular surface.” Am. Heart J. 10: 176; 1934.


[Xu et al., 2008] Xu W., Liu X., Pan Z., “Definition of Sleep Apnea event by one minute HRV spectrum analysis.” ICBBE; 2292-2294; 2008.


