

1

Introduction

Most advanced biological and man-made physical systems require reliable sensing to function properly. The more sensors they integrate, the more complete and comprehensive is the information they can gather from their surroundings. For a long time, practical challenges have set limits to the number of sensors that can be embedded into physical systems, processes, or environments. Among these challenges are limited space, the difficulty and obtrusiveness of wiring, heat dissipation, and power supply. Miniaturisation of sensors has fundamentally changed the way we deal with these challenges. Furthermore, integrating processing and wireless communication capabilities into sensing systems has enabled not only dynamic programmability but also networking, so that data can be processed (aggregated, filtered, compressed) in a distributed manner or can be packed in packets and transferred to a different location where they can be processed by employing advanced signal-processing algorithms.

The past decade has witnessed an explosion of interest in wireless sensors and wireless sensor networks, for which there are a variety of applications. In civil engineering, these sensors and networks can be employed to monitor the integrity of infrastructure, such as pipelines, bridges, and buildings. In the medicine and healthcare domain, they have already proved to be indispensable, but they are also finding new applications in augmenting existing diagnosis and monitoring infrastructure and in enabling more independent and flexible lifestyles for patients. In agriculture and environmental science, wireless sensors and wireless sensor networks are useful for precision agriculture, for monitoring the quality, amount, and flow of water, and for studying wildlife without the need to interfere with it.

However, the usefulness of the applications that employ sensors depends on the depth of understanding pertaining to the design and operation of the sensors as well as the quality of the data-processing algorithms employed. The faith an application developer puts in a sensor should be based on a quantitative understanding of its reliability, accuracy, precision, sensing range, sensitivity, and lifetime as well as on the strength of the assumptions supporting the data-processing schemes. Otherwise, the relevance of the application will necessarily be limited to laboratory settings, or prototypes. On the other hand, a fundamental understanding of sensors and their design leads to innovative ideas and the identification of totally new application domains.

Interestingly, the basic concepts of sensors are straightforward to grasp and the electrical circuits required to realise a sensing system are relatively simple and comprehensible, for example, compared to the design of high-frequency communication systems. This is because, in most practical situations, sensors have to deal with low-frequency

signals that can be detected and processed by relatively simple electrical components. The purpose of this book is to acquaint the reader with:

- the fundamental principles of electrical, ultrasonic, optical, and magnetic sensing
- the broad range of issues pertaining to the design and manufacturing of microelectromechanical sensors (MEMS)
- the principles of energy harvesting and sensor integration
- the fundamental assumptions and methodologies pertaining to the processing of sensed data.

I have tried to make the book self-contained by discussing the necessary prerequisites within the book itself, so that the reader is not obliged to refer to other materials in order to understand the text.

1.1 System Overview

Figure 1.1 displays the most essential building blocks of a self-contained sensing system. These are the sensing system, the conditioning system, the processor, and the wireless transceiver. The figure is intended to give a complete picture, but we shall not be dealing with the wireless transceiver here. Whether or not these blocks are distinct from each other depends on many factors, such as the quality of the signal that can be sensed by the sensor, the targeted energy and space efficiency, and the ease with which the wireless sensor should be integrated into and interact with other systems. The processor and the wireless transceiver are usually connected with the rest of the system using standard buses that use standard or quasi-standard protocols. Hence, the main design issue is how to integrate the remaining building blocks. I shall briefly summarise the function of each building block and highlight the different trade-offs that influence its integration with the other blocks.

1.1.1 Sensing System

The process we wish to observe or monitor is called the measurand. It either releases some form of energy that describes its condition in some way, or external energy in the form of an electrical (radio-frequency), optical, or acoustic signal is applied to it, so that from the way it modifies some of the characteristics of the signal (magnitude, phase, frequency or a combination of these), its condition can be determined or inferred. The human body is a typical example of a measurand, because it is a remarkable signal generator. The human brain generates electromagnetic signals that can be sensed by electroencephalogram or magnetoencephalogram. Likewise, the human cardiovascular and muscular systems generate electric potentials that can be sensed by employing an electrocardiogram or electromyogram. In contrast, ultrasound systems release ultrasonic

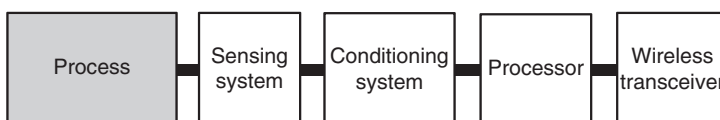


Figure 1.1 The main building blocks of a wireless sensor.

waves into a human organ and the spectrum of the reflected signal is analysed to determine the organ's condition. If a measurand's condition is determined from the signal it releases, then the sensing method is called "passive sensing". Otherwise, it is called active sensing. Passive sensing introduces less intrusion into the measurand compared to active sensing, but the amount of energy that can be collected through passive sensing is normally small.

1.1.2 Conditioning System

Regardless of the sensing mechanism employed, there are important conditions the sensor and the signal it produces should fulfil before useful features can be extracted from the signal. One of them is appropriate interfacing. When a sensor is employed to a measurand, apparently the measurand "perceives" the presence of the sensor, because the sensor draws some amount of power from it. This power must not affect the measurand's proper operation. A related issue to interfacing is impedance matching. If the impedance of the measurand as seen by the sensor is not matched by the sensor's own input impedance, maximum power does not transfer from the measurand to the sensor. Instead, power dissipation in the form of heat may be experienced at the interface, disturbing the measurand and reducing the efficiency of the sensor. Therefore, the impedance of the measurand (human body, water, air) should be taken into account when the sensor circuits are designed.

Even with the interfacing problem solved, the signal produced by the sensor may not accurately reflect the measurand's true condition for a number of reasons. Noise may be added to the sensed signal from the surrounding environment or from the internal circuits of the sensor itself. Likewise, some portion of the signal may be removed, suppressed, or distorted, because the sensor circuits act as filters. Therefore, the bandwidth of the desired signal and the bandwidth of the sensing circuits should be matched. It must be noted that in most real-world cases, the signal produced by a measurand contains a range of frequency components. The purpose of the conditioning system is to deal with all these issues. A conditioning component typically consists of a filter circuit and a differential amplifier, the order in which they appear usually depending on the nature of the measurand as well as the strength of the signal produced by the sensing system.

1.1.3 Analogue-to-digital Signal Conversion

This component is not directly shown in Figure 1.1 because it may be a part of the conditioning system or the processor or it may be a distinct entity. Regardless of its specific position, the analogue signal the sensor produces and the conditioning system pre-processes should first be converted to a digital bit stream before it can be further processed by a microcontroller or a digital signal processor (DSP). In some sensors, the analogue-to-digital converter (ADC) is an integral part of the conditioning system, while in others it is a separate block. Modern microcontrollers also integrate multiple general-purpose ADCs, to one of which the analogue signal coming from a conditioning circuit can directly be fed. Next to the transceiver and the processor, the ADC is the largest power consumer and hence care must be taken in choosing a suitable ADC. Several factors determine the choice of an ADC. For example, if the sensor signal is noisy, it is better not to use a powerful pre-amplifier lest the noise is amplified together with the useful signal. In this case it is better to use a high-resolution

ADC, so that an efficient DSP algorithm can eliminate the noise. But a high-resolution ADC consumes a large amount of power and generates a large amount of data, which require a sizeable resource to process, store, and communicate. If, on the other hand, there is a small noise component in the signal, then it is better first to amplify the signal and use a low-resolution ADC. If the ADC is not an integral component of the sensor or conditioning system, then it is possible to use the sensor for different applications which require different resolutions (accuracies), in which case separating the ADC from the conditioning stage enables the choice of suitable ADCs, independent of the sensing system.

1.1.4 Processor

The processor is a multi-purpose system, but as far as a wireless sensor is concerned, the level of data processing it can support is limited by factors such as available RAM, processor speed, battery capacity, the amount of heat dissipation that can be tolerated by the object or person, and the sensor's size. In wireless sensor networks and in wireless body-area networks, the processor is mainly responsible for low-level DSP (such as digital filtering and data compression) and for managing the various communication protocols which transfer the raw data to a nearby base station.

1.2 Example: A Wireless Electrocardiogram

A wireless electrocardiogram (ECG) measures the electric or action potentials that are generated in the heart and propagated through its electrical conduction system (a combination of nerve fibres, muscles, and tissues). These electrical potentials are responsible for creating and regulating the diastole–systole rhythms of the heart. Action potentials are produced at the sinoatrial (SA) node (located in the right atrium of the heart) and propagate through the atrial muscles to the atrioventricular (AV) node and further into the ventricular muscles of the heart through the His bundle, the left and the right bundle, and the Purkinje fibres (see Figure 1.2).

The propagating potential difference can be sensed by placing two or more electrodes on the skin at the right and the left sides of the heart (Figure 1.3). The magnitude of the pulses that can be picked up by the electrodes can reach up to 5 mV and their frequency varies between 0.05 Hz and 150 Hz. By analysing the shape, magnitude and the frequency of these pulses it is possible to determine several cardiac conditions.

Whilst the pulses themselves can be easily detected, the design of a safe and reliable electrocardiogram involves several components and DSP stages. Figure 1.4 displays the essential building blocks of an electrocardiogram. Between the electrodes and the rest of the sensing system there should be a protection mechanism to ensure that the system's operation does not interfere with the operation of the body. Both to prevent the ECG from overloading and interfering with the functions of the heart and to pick up as much voltage as possible, the ECG should have a high input impedance (because the body has a high output impedance, which has to be matched by the sensing system).

The electrodes capturing the action potentials of the heart also pick up electrical signals from their surroundings which have nothing to do with the action potentials of the heart and are therefore unwanted. The human skin itself produces a DC signal of up to

Figure 1.2 The generation and propagation of action potentials in the heart. (1) the sinoatrial (SA) node, (2) the atrioventricular (AV) node, (3) His bundle, (6 and 10) left and right bundle branches. Courtesy of J. Heuser (2007). Original image of the heart was by Patrick J. Lynch and C. Carl Jaffe, Yale University, Center for Advanced Instructional Media.

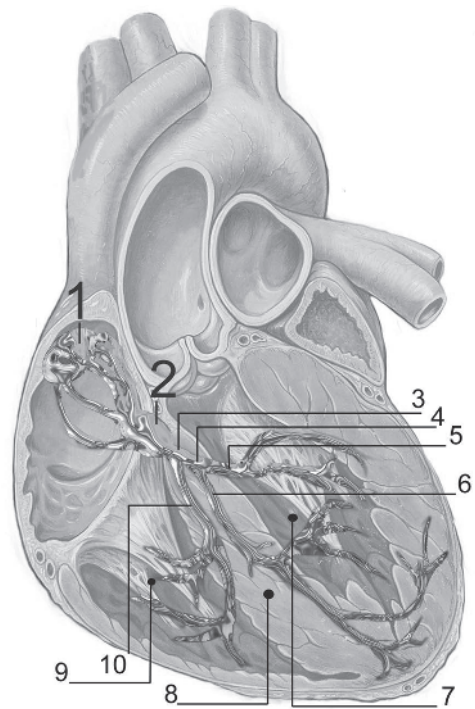


Figure 1.3 Two electrodes are used to measure cardiac action potentials.

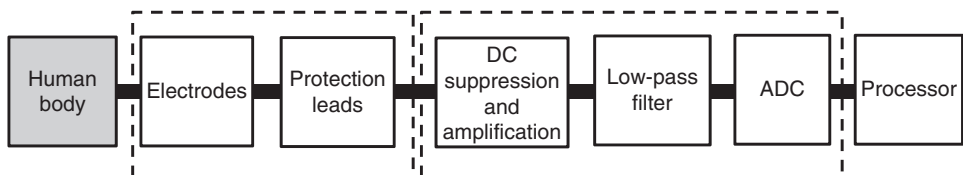


Figure 1.4 The essential building blocks of a wireless electrocardiogram.

300 mV, and other sources of noise include the power-line and the radio-frequency signals that are generated by nearby wireless and microwave devices as well as fluorescent lamps. Because of the small amplitude of the useful signal, it is not possible to separate all the unwanted signals from it right from the beginning. However, the pre-amplification and DC-suppression stages can remove the DC components by using relatively simple coupling capacitors. This same stage can amplify the rest to an appreciable level.

The pre-amplifier is typically a differential amplifier, the main purpose of which is to suppress all unwanted signals that have equal effects on all the electrodes. For example, noise that is generated by the power line will equally affect all the electrodes. Therefore, combining the outputs of the two electrodes in a differential mode (subtractive mode) suppresses the signals that are produced by the power line.

After the pre-amplification stage, an additional amplification is applied, followed by a low-pass filtering with a cut-off frequency of 150 Hz to remove all signals that have higher-frequency components. Then the analogue signal is converted to a digital signal and supplied to the processor. A DSP algorithm further processes the digital signal to improve the quality of the ECG measurements. For example, errors that can occur as a result of shaking or vibrating electrodes can be detected by a digital filter and corrected. In the case of a wireless ECG, the digital stream after the DSP stage can be packed into packets and transferred to a remote location where it can be further processed or displayed to a physician, who remotely monitors the patient.

Figure 1.5 displays a five-cord wireless ECG consisting of three distinct stages: the electrodes, the conditioning system, and the processor with a memory subsystem and a wireless transceiver. The memory subsystem enables data to be logged locally. Figure 1.6 highlights both the achievements and challenges of using a wireless ECG. Measurements were taken in our laboratory using a wireless ECG while a person was freely moving on a flat surface. Apparently, because of different movement-related artefacts, the measurements suffer from both long-term and short-term drift and signal distortions, which is why the various building blocks and signal processing algorithms are necessary. We return to this issue in subsequent chapters.

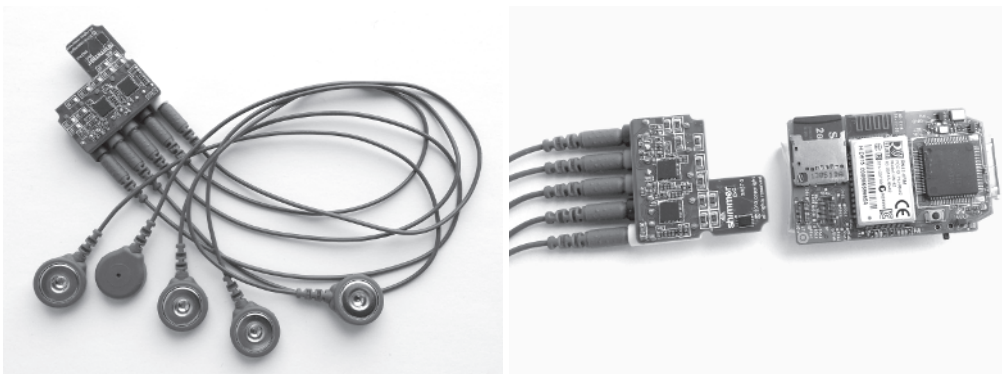


Figure 1.5 A wireless electrocardiogram with its processor and conditioning systems as well as electrodes as separate building blocks: Left, the electrodes and the conditioning system; right, the conditioning system and the processor.

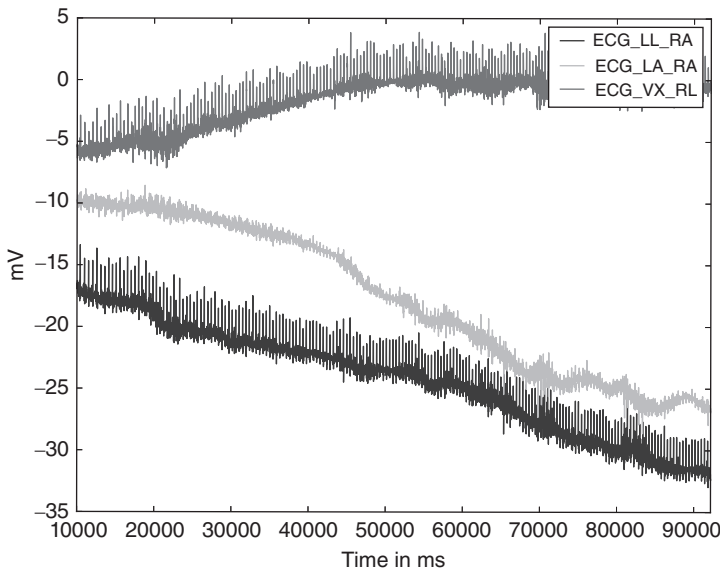


Figure 1.6 ECG measurements using the Shimmer wireless sensor platform while a person was freely moving on a flat surface. The measurements were taken from three different places: top trace, between the left leg and right arm; middle trace, left arm and right arm; bottom trace, neutral and right leg.

1.3 Organisation of the Book

The book is organised into several logical components. Chapter 2 provides an overview of emerging applications in the ubiquitous computing, wireless body-area network, and wireless sensor network communities. The typical features of these applications are the comprehensiveness of the sensing task, the intelligence and self-organising features embedded in the sensing systems, and the novelty of the applications themselves. The specific applications are selected to highlight the diversity of sensing techniques and the issues involved, or rather the challenges surrounding the design, deployment, and signal processing aspects of ubiquitous sensing.

Chapter 3 provides a brief introduction to signal conditioning and addresses the most essential aspects. Chapters 4–7 introduce the essential aspects of electrical, ultrasonic, optical, and magnetic sensing. These chapters cover essential aspects of most important regions of the wide spectrum of sensing.

Chapter 8 presents the most important aspects of medical sensing. I decided to give a separate treatment to this subject because of the growing number of medical applications in the communities listed at the beginning of this subsection.

Chapter 9 provides a comprehensive insight into the design and manufacturing of microelectromechanical sensors and demonstrates how the various sensing mechanisms (electrical, optical, magnetic, and so on) can be employed to develop practical sensors such as inertial, pressure, and fluid sensors.

Chapter 10 addresses an important issue in sensor deployment, namely energy harvesting. It describes the need for and the advantages of energy harvesting, discusses the

choice of suitable sensing mechanisms, proposes a conceptual architecture, and presents various prototypes, highlighting their merits and demerits.

Chapter 11 addresses practical issues surrounding sensor integration. In most practical cases, a sensor is a part of a more complex system, the operation of which, unfortunately, may produce undesirable effects on the quality of sensing, such as radiation and thermal noise. The chapter recommends several integration strategies.

Finally, the last chapter, Chapter 12, addresses the data processing aspects of sensing, the main objective of which is minimising uncertainty. The chapter describes how the outputs of sensors can be regarded as random variables and discusses the different evidence-combining techniques used to reduce sensing error. To make the subject both interesting and useful, I give several examples and endeavour to take the reader step by step into the different stages of estimation.