

1 Introduction

The advancement of electronic devices went through the silent yet pivotal revolutions such as solid-state technology and integrated circuit miniaturization, among many others. However, through merging with the information and communication technologies (ICT), the increasing electronic integration was able to open new perspectives and markets in modern society. The ability to compute and to communicate is among the basic needs of the human community, and the progress of the ICT technologies showed unforeseen scenarios where in some cases – surprisingly – machines greatly surpassed our natural skills. However, in the enduring efforts to make artificial systems behave like humans, few areas are still lagging behind, and one of those is the ability to sense the environment as we do. In this framework, the pervasive implementation of sensing devices in consumer electronics is a highly pursued electronics industry paradigm. The number and variety of sensors implemented in consumer and industrial devices increased dramatically in the past decades: handheld devices, automotive, robotics, healthcare, and living assistance.

Unfortunately, artificial sensing requires an approach beyond the borders of information science with the necessity to cope with an incredible number of physical and chemical transduction processes. The interaction with the environment of synthetic systems still shows open issues both at the transduction and at the information processing levels.

Designing sensing devices is always an exciting and challenging task. Very often, the ultimate question is: “Will we be able to detect it?” The answer is hidden in both the technology capabilities and the environment status, and it is not clear, at first sight, where the limitation of the approach is and the reason for the failure. Therefore, the arguments should not be treated as a collection of individual cases but with a general approach using the appropriate tools of abstraction and formalization, with a constant look to the biomimetic inspiration of sensing. In doing so, adopting a strategic and theoretical perspective, we can foresee the integration of massive computation with the sensing capabilities as one of the next incoming revolutions of information engineering.

This first chapter aims to set up the framework on which the book will be shaped, and it is intentionally based on informal descriptions of concepts. This is a nonrigorous approach but is a fundamental step toward an abstraction process about artificial sensing: the ideas behind the general definition of sensors, their main performance-limiting processes, and essential tradeoffs. Using this inductive approach, we will first define concepts, leaving the formalization to the following chapters of the book. If the reader is facing this field for the first time, the arguments could appear vague and fuzzy; thus, this chapter should be eventually reread as the *last* one.

1.1 Sensing as a Cognitive Process

The concept of a sensor would not exist without life. To grow, reproduce, and survive, any organic entity should perceive external signals to evaluate them, either as an opportunity or a danger. *Sensing* is not a mathematical or physical abstraction derived by inorganic matter, but it is a biological process since any living being should perceive, measure, and evaluate external stimuli to take *actions*. As with many other engineering concepts, this *feedback* model is taken from nature, and sensing is the primary input of such a loop mechanism.

Focusing on human beings, the word *sensor* is derived from the verb *to sense*, referring to the capabilities of human beings to perceive reality by means of sight, hearing, taste, smell, and touch. Sensing is a fundamental part of what is referred to as cognitive sciences, an interdisciplinary field aimed at studying the human mind and knowledge processes.

It has been common practice to analyze the sensing process in interdependent stages such as *sensation*, *perception*, and *consciousness*. The definitions and the borders between these domains significantly differ among the scientific (in a broader sense) communities. However, there is a general agreement on this sensorial experience segmentation. This partition is also reflected in artificial sensing systems.

Sensation is the primary process of receiving, converting, and transmitting information resulting from the stimulation of sensory receptors. Sensory stimuli are taken from the environment by means of physical transduction processes such as those operated by the eyes, where the photons coming from the scene are focused onto the retina in the same manner as a photo camera. The cones and rods of the retina work as transducers, detecting external energy from every single photon and sending the information by means of electrical messages to the brain. On the other hand, perception is the process of selecting, identifying, organizing, and interpreting sensory information. It is not a passive reception of stimuli but early processing: the information is collected, organized, and transmitted by nerves to the brain. Edge detection of objects in seeing and touch is an example of perception. Finally, consciousness is the more elaborated knowledge process: it is the brain's deepest interpretation of neural responses to sensory stimuli. It involves the capacity to sense or perceive and active use of those abilities, depending on previous experience. Humans can experience conscious and unconscious perception. If we relate it to machines, definition, context setting, learning, and adaptation could be possible processes ascribing to a sort of artificial "consciousness." Here, the concept of consciousness is restricted to a functional/phenomenal process, distinguished from the problem of self-awareness where implications are highly speculated in philosophy, with open issues.

In the past centuries, when there were weak boundaries between scientific and philosophical studies, the sensing process was highly conjectured, especially when it was considered a fundamental step for human perception and knowledge. The connection of sensory stimuli with the brain was observed and studied since the time of the

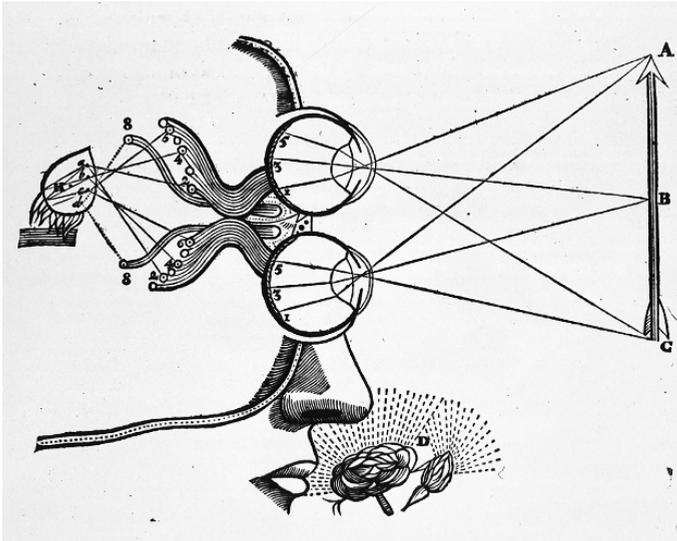


Figure 1.1 Picture taken from R. Descartes, *Tractatus de homine et de formatione foetus*, 1677 (edited posthumously), showing the process of seeing and smelling.

ancient Greeks and in Leonardo da Vinci's monumental work. Among others, it is interesting to note how Descartes remarkably analyzed in greater detail the sensing process in some of his writings, which were very useful in contributing to a general framework of cognitive sciences. As shown in Fig. 1.1, sensorial stimuli (sight and smell) are conveyed into an inner part of the brain, where they are interpreted. Even if some physiological aspects were not correct and Descartes's speculations were far beyond the pure phenomenological aspects of the matter (still unresolved and on debate today), the organization of the sensing process in several steps was profoundly analyzed, introducing modern concepts.

Helmholtz, maybe one of the last polymaths, gave another example of deep analysis of the sensing process in the nineteenth century. In his works (see excerpt illustrated in Fig. 1.2), he gave seminal contributions in the field of visual and auditory perceptions, envisioning a profound relationship between sensing and cognitive sciences. He claimed that human perception should be studied by the *process's physical, physiological, and psychological characters*. He even attempted to justify the perception of beauty related to the sensing process in some of his works.

It is no coincidence that the Greek word *aisthanesthai*, meaning “to perceive by senses and by the mind, to feel” is at the root of the word *aesthetics*: a branch of philosophy dealing with the perception and appreciation of art, taste, and beauty.

To summarize:

- Sensing is a biomimetic concept. Sensors engineering has frequently borrowed functionality models from life sciences psychophysiology and cognitive studies.
- A sensor should not be considered a pure transducer but part of an artificial cognitive process to grab as much information as possible from the environment.

THE RECENT PROGRESS OF THE THEORY OF VISION.



A COURSE OF LECTURES DELIVERED IN FRANKFORT AND HEIDELBERG, AND REPUBLISHED IN THE PREUSSISCHE JAHRBUCHER, 1868.

I. THE EYE AS AN OPTICAL INSTRUMENT.

THE physiology of the senses is a border land in which the two great divisions of human knowledge, natural and mental science, encroach on one another's domain; in which problems arise which are important for both, and which only the combined labour of both can solve.

Figure 1.2 Excerpt from the *Popular Lectures on Scientific Subjects*, by H. von Helmholtz, 1873 (translated from German) envisioning the need for a strong relationship between the sensing and the cognitive processes.

1.2 Aiming at a General Definition of Electronic Sensors

In electronic engineering, the word “sensor” embraces a broad class of systems designed for highly different applications. A “sensor” could be roughly referred to as “a system that transduces physical stimuli into data.” However, this definition is too vague and does not take the essence of artificial sensing; thus, a closer look should be taken to understand the standard framework better.

Figure 1.3 shows four examples of systems referred to as “sensors”: a weight scale, a microphone, a heart rate monitor, and a machine vision system. They all collect stimuli from the physical environment and convert them into data; however, they are dealing with increasing complexity to achieve the related tasks.

A scale operates a static force measurement. We do not care about the variation of weight within the measurement timeframe. On the other hand, the microphone needs to follow the pressure variation (sound) on a surface versus time, and its time-domain properties are a fundamental aspect of its design. Next, a heartbeat sensor uses patterns in ECG signals associated with heartbeat events. Finally, a machine vision system deals with many images to detect/count defective objects. The idea is that any application identifies specific conditions of the signal to be identified and measured by a custom sensing system. However, the idea of classifying sensors according to the kind of signal could be misleading.

We are looking for not the stimulus itself, but something more complex hidden in primary stimuli and is referred to as *information*. In simple words, the information content is the essence of what we are looking for in the sensing process. The concept of

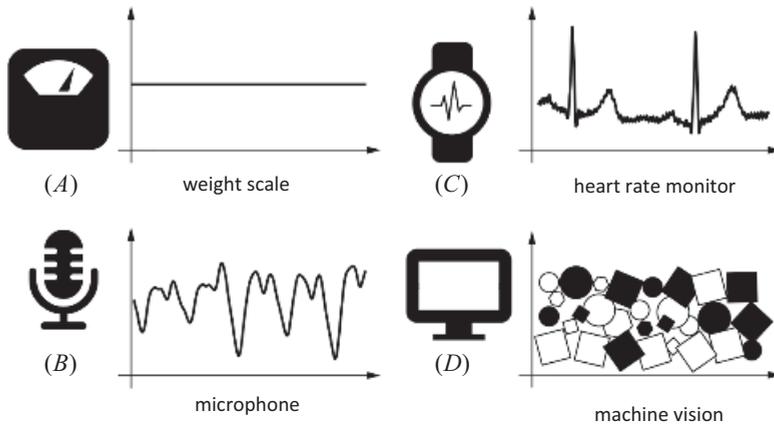


Figure 1.3 Input signals of systems commonly referred to as sensors. (A) Weight scale readout. (B) 1 ms of microphone recording from Etude no. 21 for piano by F. Chopin. (C) 2 s of heart rate biopotential recording. (D) A 2-D machine vision image processor.

information has been extensively treated and formalized in other disciplines; at the moment, it corresponds informally to the amount of knowledge that we gain during the sensing process aiming at the specific task application. We will use a more formal approach to the issue in Chapter 3.

1.2.1 Signals and Information

To illustrate the role of the information in the sensing process, we will use examples. In Fig. 1.4A is shown a heartbeat detector. The sensor's main task is to detect the number of beats in a given period of an ECG signal using a decision threshold. We informally link this to the "information" necessary for our application to understand the concept. The three signal examples of Fig. 1.4A are taken from a set of all possible ECG waveforms in the same time period, and we refer this to as *samples* in the *signal space*. In the first two cases, the system counts 8 beats, while in the last one, only 7. Therefore, we associate the result in a measurable space, referred to as *information space*. In other words, we say that samples in the signal space could be mapped in points in the information space. The acquisition process of a sensor is a correspondence between these two spaces. We will always refer to discrete information space.

In the second example of Fig. 1.4B, the sensing system should detect the number of circles/squares in images. Even in this case, the four sampled images belong to a very large signal space, for example, composed of all possible images of $N \times M$ black-and-white pixels. However, the "information" is relatively smaller than the signal space and could be organized in a two-dimensional space where the variables are the number of circles and the number of squares, respectively.

In these two examples, it is easy for human perception to identify the information in the signal space at first sight and check if the sensor system has correctly detected our task. However, there are other cases in which the information is more hidden than

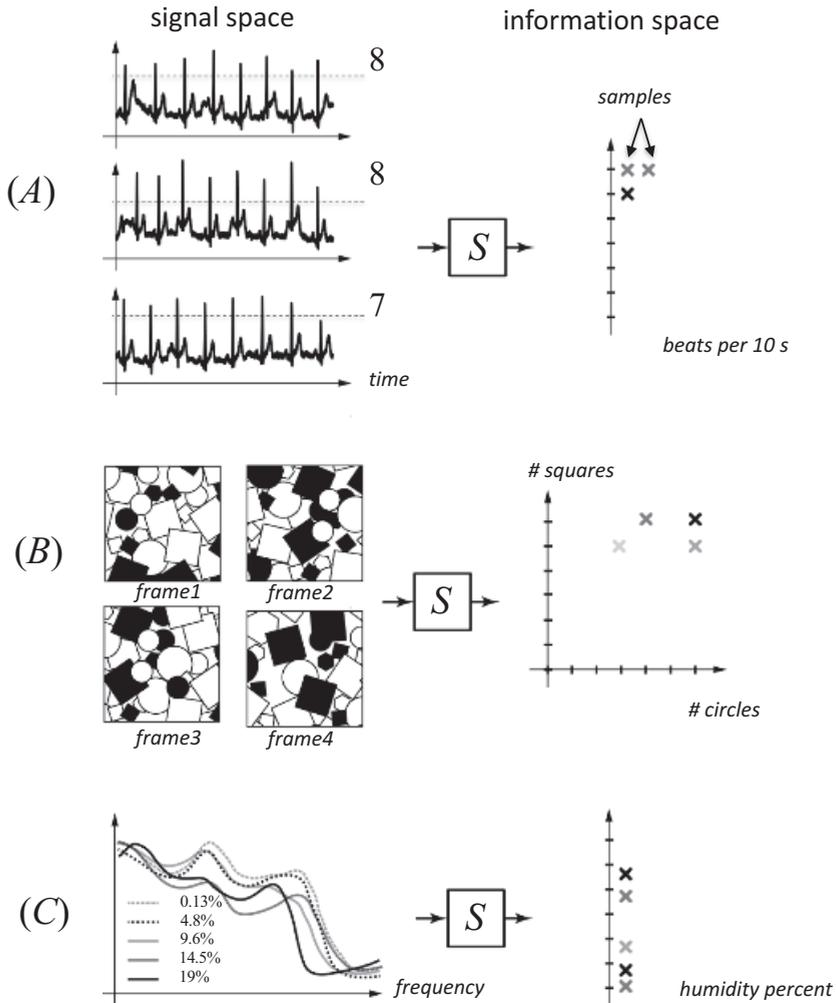


Figure 1.4 Signal and information spaces. (A) Biopotential heartbeat detector using a threshold. (B) Machine vision object counter. (C) Impedance spectrum humidity detector. Note that the first two examples have a discrete information space while the last is a continuous space.

previous examples, and machines could outperform human perception. For example, in the case of Fig. 1.4C, the signal is composed of five measured microwave impedance spectra related to a material having different water content (humidity). The idea is to use these spectra to implement a microwave humidity sensor, where the information is the percent humidity. It is hard to see any regular or monotonic behavior in spectra or in parts of them with respect to the stimulus (humidity). Our intuition concludes that there is no clear relationship between the humidity of the material and the spectra. In other words, it is not easy to see any significant information in the signal itself. However, suppose the signal is treated by suitable mathematical processing. In that case, we can set up a linear predictive model to

detect the humidity based on microwave spectra so that signals can be mapped into distinguishable and ordered levels in the information space. The latter example shows that the information could be very hidden in signals, even beyond human capabilities to distinguish them in raw data. For this reason, in these cases, the information to be extracted is often referred to as *latent variables*.

The preceding examples are related to cases of different *complexity* of the task and require different processing resources to extract the information.

To summarize:

- The sensing process should be defined by a *task*, which qualifies the kind of *information* that should be measured. Thus, the application (task) determines the characteristics of the information space.
- *Signals* are functions representing states of the sensed environment carrying information. All the possible configurations of signals define the signal space.
- The information space has smaller dimensions of the signal space, and it is discrete. This means that the multiple elements of the signal space may have the same element in the information space.
- The sensing process is a function implying that each *sample* of the signal space has a correspondence in the information space.

1.2.2 The Simplest Case of an Analog-to-Digital Interface

The previous section identified a distinction between signals and their information content, mapped into the information space. However, if we refer to the simple analog-to-digital (A/D) conversion of a signal in the “analog domain,” we can match more easily the two spaces since the analog value itself encodes the information. We can better understand this with the cases illustrated in Fig. 1.5. In Fig. 1.5A is shown a time-varying biopotential signal that is monitored by an A/D interface. Our task is to know the biopotential value evolution with respect to time, and thus it is precisely the information that we need. The A/D converter associates a specific analog value of the signal full scale with a binary-encoded discrete value. Therefore, the discrete values of the A/D converter are easily represented in the information space. The correspondence is made by associating an analog value with the converter’s closest discrete value. The case of Fig. 1.5B is even more straightforward: the information is the static analog value of a weight sensor. Therefore, each measure (sample) is directly mapped in the information space. As before, multiple analog values may be mapped into the same coded value by the converter.

In summary:

- In the simple case of an A/D interface, the association between *information* and *signal* is closer because the signal value itself *represents* the information that we need to detect.
- The correspondence is made by associating an analog value with the closest discrete level of the A/D converter.

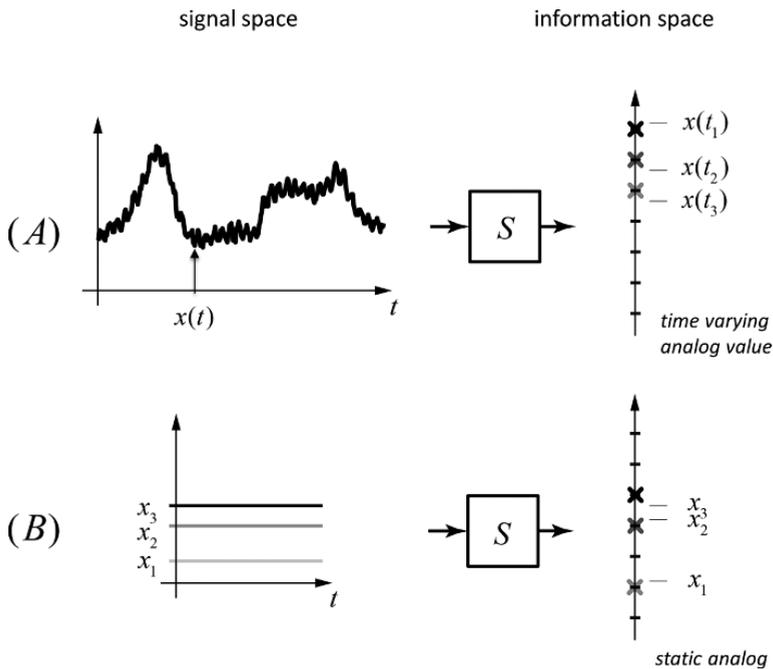


Figure 1.5 Sensing process in time-varying (A) and static (B) analog signals. The analog value of the signal space is associated with the closest discrete value of the information space.

1.2.3 The Role of Errors

Unfortunately, the sensing process's physical implementation is necessarily affected by *errors* due to the stochastic nature of random processes and nonidealities. Errors arise from either the environment or the sensing system itself and its nonperfect detection capabilities. Let us look at Fig. 1.6, where a biopotential is used to detect heartbeats utilizing a threshold as in Fig. 1.4A. In the absence of noise, during the time-lapse, we measure 8 beats, as shown in Fig. 1.6A. Now, assume that the sensing process is noisy and the same waveform as in Fig. 1.6A with added noise is shown in Fig. 1.6B. If we use the same detection approach, the count is no longer 8 but rather 10. The random process of noise changes the threshold crossing cases: there are some points (e.g., point M) that were not crossing the threshold before (without noise), while now they do, thanks to noise contribution. Conversely, there are other points (e.g., point N) that were crossing the threshold in the previous case, and now they do not pass the level because of the perturbation of noise. If we repeat the same procedure on a signal containing 8 beats in the presence of the noise, we may at one time count 7, another time 9, another time again 8, and so on. This means that we cannot say that the count is certain but, in the presence of noise, we can say that the “estimation of the count is given by 8 ± 2 .” Therefore, the presence of noise determines an *uncertainty* (Chapter 2) of the measure of ± 2 counts. The uncertainty due to noise could be visualized in the gray area across the tick equal to 8 in Fig. 1.6.

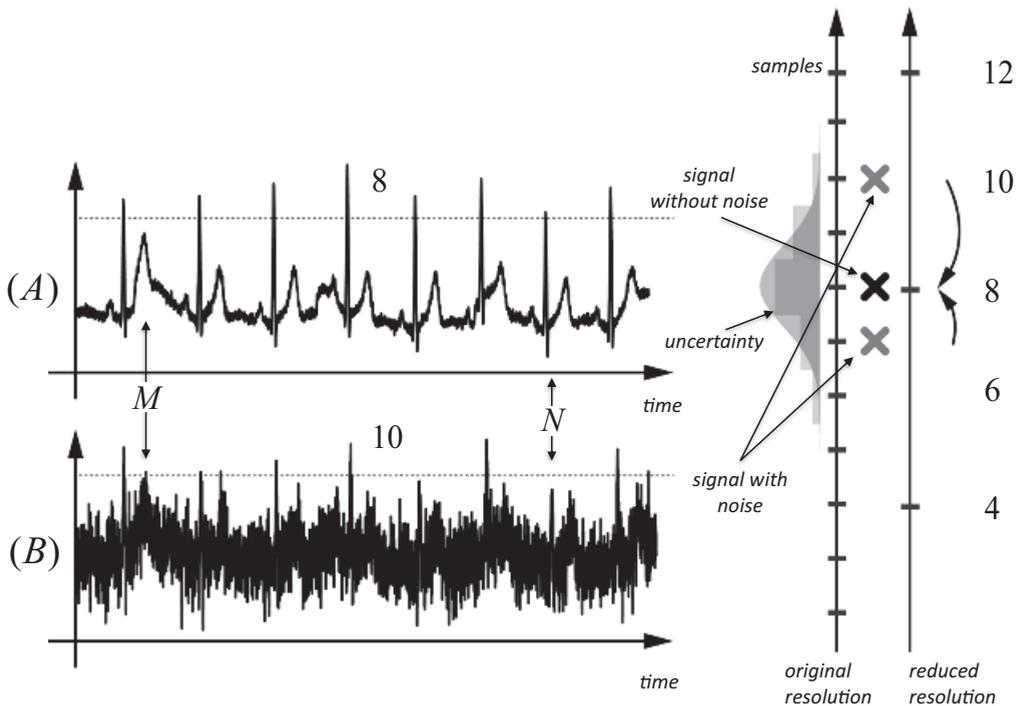


Figure 1.6 Effect of resolution reduction due to noise. Biopotential recording to calculate heartbeat (A). Same biopotential signal but with noise added (B). The addition of noise determines the increase of uncertainty of detection and thus a decrease of resolution.

In this example, the presence of errors (or noise) changes the information space situation. If we could have counted single beats without noise before, now we have an uncertainty of ± 2 counts. Therefore, previous levels are no longer truly distinguishable from each other because the *same* signal could give counts in the interval 8 ± 2 due to noise. This fact reveals that the information space subdivision is not appropriate because the sentence “count = 8” has information similar to that one of “count = 10.” This results in a *misclassification* because we classify 8 counts, although, in reality, they might be 10 or vice versa. This means that we have a high probability that the affirmations mentioned above reflect the same signal condition due to errors. This could be seen pictorially showing that the same uncertainty area covers samples. Therefore, it could be better to reduce the number of subdivisions (e.g., by grouping 4 levels) so that “count is between 6 and 10” and “count is between 2 and 6” has more significance from the information point of view because there is a lower probability that the two sentences correspond to the same signal. In this case, any sample giving a value in the uncertainty zone, identified by 8 ± 2 , will be associated with the center of the interval, whose value is 8. In other words, by enlarging the classification zones by considering uncertainty, we reduce possible misclassification errors.

We can thus refer to the subdivisions of information space as *resolution levels*. In the presence of noise, we might set the resolution level of the order of the uncertainty so

that it preserves significance from the information point of view, avoiding misclassification. As shown in Fig. 1.6, the higher the noise, the lower the resolution.

Turning back to the simplest case of the analog domain where information and signal are strictly related, we can observe that we could have infinitely small resolution levels in the absence of noise. It is thus the presence of noise that determines a finite resolution. This is congruent with the observation that any real sensing system (in which necessarily errors such as noise are present) does have a finite resolution and reinforces the already cited statement that the information space is discrete.

However, any application task always identifies a maximum in the needed information. For example, in the case of Fig. 1.4A, we do not need to detect more than 30 beats per 10 seconds, or in the one in Fig. 1.5A, we know that the biopotential will never surpass a value of hundreds of millivolts. The maximum achievable value in the information space is referred to as *full scale*. Therefore, the scale defines a *fixed number of resolution levels* under which the information space can be divided. The number of resolution levels expressed in terms of energy is referred to as *dynamic range* (Chapter 2). The number of information (resolution) levels achievable by a sensing system is a measure of the process's information.

To summarize:

- The physical sensing process is always affected by *errors* (e.g., noise) deriving from both the environment and from the sensing system itself.
- Errors define *uncertainty* in the measuring process, meaning that we cannot be sure of the result given by the system, but we can estimate the information with some degrees of confidence.
- The uncertainty sets the *resolution* of the sensing process to make the levels distinguishable from each other at the information space and reduce *misclassification*.
- Real applications always define a maximum level of the coded information or *full scale* in the information environment. This boundary identifies a limited number of *resolution levels* in the information space or *dynamic range* if expressed in energy terms.

One useful option is to express the number of discrete levels at the information space in terms of *bits*. This creates a significant link on the one hand within the context of A/D converters (Chapter 2) and, on the other hand, with information theory (Chapter 3). For example, referring to Fig. 1.6, if the number of resolution levels is coded in N bits and the noise reduces that number by grouping 4 adjacent levels, the resolution is degraded by 2 bits of information. We will return to this later.

Another important characteristic of the sensing process is the *minimum detectable signal* (MDS), which is the minimum variation of the signal to induce a significant variation in the information space (i.e., at least one resolution level). If we take the simpler example of Fig. 1.5A, where signal and information spaces overlap, the MDS is the amount of signal variation that is “equal” to that of noise. In other words, the signal to be detected should have a “strength” surpassing that one of noise. Thus the MDS is usually set to the noise “strength” and thus to the uncertainty of the process. Of

course, we should define a metric for making a comparison (Chapter 2). In any case, if there were no errors (e.g., noise) at all, the MDS could be infinitely small, resulting in an infinite detection capability.

Figure 1.7 shows a conceptual view of the sensing process as a mapping between signal space and information space. As shown in Fig. 1.7A, each cross identifies a sample in the signal domain corresponding to another point in the information space. Any point of a subset of the signal space is thus mapped into another point in the information space within the uncertainty area due to errors. Again, the number of elements in the signal space is, in general, higher than that of the information space (e.g., the number of all possible images containing the same number of objects is mapped in the same point in the information space.). Following the discussion of Fig. 1.6, we can set the resolution's size to encompass most of the uncertainty so that the discrete resolution levels have a higher degree of confidence to be distinguishable from each other. In doing so, most of the points within a discrete level are associated with one point of the information space, shown in the figure with the crossing point between lines. Another way to see this is that by enlarging the resolution levels, we reduce the overlap between the uncertainty zone and thus the misclassification. The higher the noise, the lower the resolution and the lower the resolution levels since the full scale is fixed.

Following the illustrations of Fig. 1.7, the transition of samples between two adjacent points of the information space corresponds to the transition between two

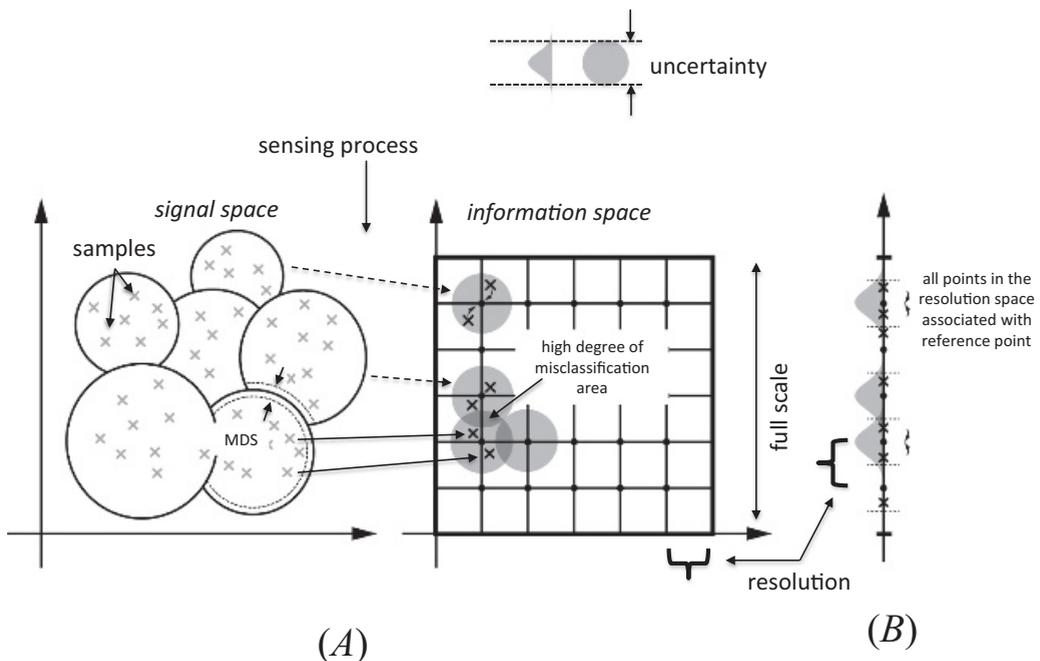


Figure 1.7 Sensing process as a function between signal and information space. Conceptual (A) multidimensional (B) one-dimensional representation.

signal space subsets. During the information space transition, the sample points cross overlapping zones of uncertainties. If it is in the middle of the overlapping, we are unsure whether the sample belongs to one discrete level or the adjacent one; thus, we have a high degree of misclassification. The overlapping misclassification area could be mapped in the signal space by an area across the subsets' boundaries. The thickness of this area identifies the MDS because it is the minimum variation of the signal required to have a distinguishable change in the information space.

The concept could be applied to the simplest case of a noisy analog to digital conversion shown in Fig. 1.7B. This is the simplest case where we can superpose the signal and the information space because the signal value itself encodes the information. Each sample in the analog signal space identifies a binary-encoded level in the information space; however, the noise implies the possibility of an error of assignment. To reduce that error, we can enlarge the resolution levels, but this implies that we reduce the conveyed information. As an example, let's take a noisy 8b converter. If the noise is so high that it covers 4 discrete levels (2 bits), this means that it is convenient to reduce the number of levels by adjacent groups of 4 levels to lower misclassification. Therefore, noise reduces the system's resolution from 8 bits into 6 equivalent bits ($8b - 2b = 6b$) to make information (resolution) levels distinguishable. In summary, the number of distinguishable resolution levels expressed in bits is a measure of the information gained by the sensing process.

The corruption of the information by errors is conceptually exemplified in Fig. 1.8. The signal coming from the environment is physically affected by noise. Therefore, the amount of information gained at the input is limited by this effect arising from the source's random physical processes. However, in sensor system design, we have to deal with other sources of errors that further restrict the amount of information conveyed. For example, analog interfaces have intrinsic noises due to thermal noise given by electronic devices. Furthermore, the transduction process might have nonoptimal characteristics (e.g., presenting nonlinear or saturation effects), or the detection algorithm could induce assignment errors due to poor characterization of the model. All these sources of errors reduce the amount of information conveyed to the output. From the resolution point of view, the sensor design is aimed at reducing to as low as possible the information corruption, leveraging on design constraints and tradeoffs.

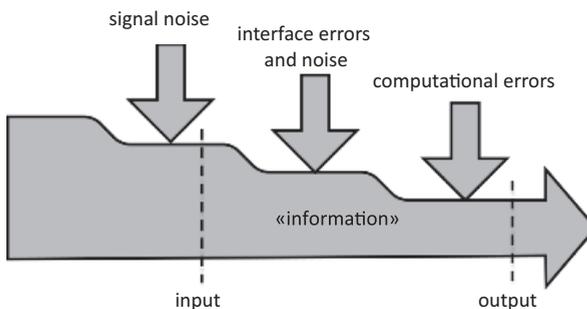


Figure 1.8 Degradation of the information in the sensing process due to errors/noise.

Another important point of the sensing process is the role of energy and time (Chapters 2 and 3). The corruption of information induced by errors could be accounted for by representing them in terms of energy/power whose effects set the detection limits. In other words, the amount of information conveyed by a sensing system could be referred to as the relationship between signal and errors (noise) energies. The simplest case of this tradeoff is the signal-to-noise-ratio.

As far as time is concerned, any measurement process requires some kind of energy/power to be taken from the signal itself to be classified. As in biological organisms, where sensing is useful to organize an action, the measurement process should be performed in a determined amount of time. As such, when the signal change versus time is limited, it is said that signals have a characteristic *bandwidth*. In the example of the microphone, we can observe that the change of states versus time is limited in the audio bandwidth. We have to implement an interface that follows signal changes in the whole audio bandwidth in an optimal design perspective. Therefore, a sensing system's capability to perform classification in a determined amount of time is a primary constraint.

Thus, we could make a broader definition of an "electronic sensor" as follows from previous arguments.

Electronic Sensor An electronic system that extracts the information required by the application from observed signals in a determined amount of time, namely an *information classifier*.

Turning back to our examples, we can see that this definition embraces the entirely different functions of the systems used as examples and gives some hints about the essence of the sensing process, based on the role of the information.

In summary:

- The number of distinguishable resolution levels expressed in equivalent bits at the readout is the measure of the information gained by the sensing process.
- One key aspect of sensor design is maximizing the number of resolution levels, which is equivalent to maximizing the information conveyed by the sensor system.
- Another key issue of sensor design is related to the time required by the sensor to resolve the task, which is linked to the bandwidth characterization of the electronic interface.
- The aforementioned characteristics allow us to define any kind of sensors as information classifiers performing the task in a limited amount of time.

1.3 Essential Building Blocks of Electronic Sensors

A preliminary and necessary step of the electronic sensing process is the transduction of the physical stimulus into electronic signals. This is performed by a *transducer*

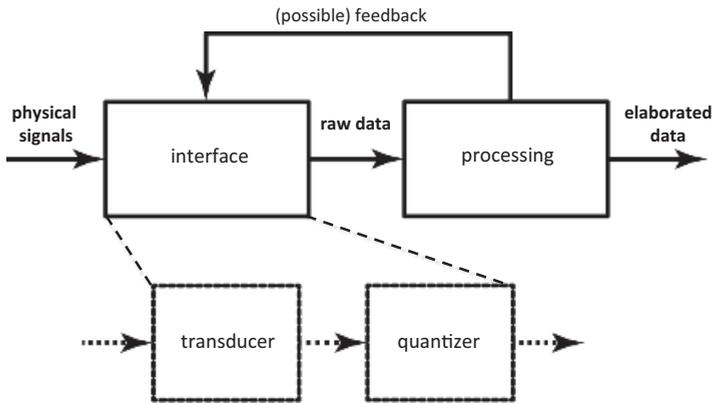


Figure 1.9 Building block structure of an electronic sensor.

interface, which converts the energy of physical signals into electronic states. Then, data should be organized and elaborated to extract the information depending on the specific task. This process could be performed by a system that follows the transducer and computes elaborations on the previous stage's raw data. The higher the complexity of information, the higher the “intelligence” (computational complexity) required by this block to achieve the result.

We can draw a general scheme of an electronic sensor structure illustrated in Fig. 1.9. Two parts could describe a generic sensor: the interface and the processing machine. The first one is devoted to the transduction and digitization of the signal, composed of two subblocks: the transducer and the quantizer. The transducer is the block that directly connects with the physical environment. It could be a simple amplifier or a more complex structure operating in the time or frequency domain. The quantizer is a block that is necessary to operate in the binary domain on which the elaboration block will work. A typical quantizer is an A/D converter, even if other digitizers might be used (Chapter 8).

A processing machine implements the second part to gain the maximum information from raw data. The main differences between the two parts can be summarized. The transducer converts the energy/power levels of physical signal states into raw data. Therefore, it acts as an energy detector, and the design should be mainly focused on optimization with respect to the energetic content of signals. One of the references of this optimization is the *signal-to-noise ratio*. Conversely, the processor searches the raw data's information to achieve the final task required by the sensor application. It makes complex elaborations of raw data that earlier stages cannot treat. For example, we can implement analog filters in the first blocks to optimize the signal-to-noise ratio, but we need data elaboration to implement complex algorithms, for example, Kalman's filters or machine learning classifiers.

From previous arguments, it is clear that the signal-to-noise-ratio (which is an energetic ratio) should *not* be considered as the reference for the ultimate limit of the sensor performance but just a first step to optimizing the overall chain. For example,

Table 1.1 Sensing approaches classified according to the information complexity (from lower to upper lines in the table).

Signal characteristics	Information extraction strategy
Unknown pattern/event classes	Unsupervised classification (learning)
Known pattern/event classes	Supervised classification (learning)
Pattern/event counting	Pattern matching
DC and AC measurement	Transducer and quantizer

there are radar or ECG signals whose energy is much lower than that of noise. However, it is possible to discriminate useful information using raw data processing techniques.

Looking back to the examples, we can see that the weight scale and the microphone are sensors that could be modeled by just the first block of the structure illustrated in Fig. 1.9. This is because the relationship between signal and information is closer to each other. On the other hand, the heartbeat detector needs to extract information at an upper level, for example, using pattern recognition or other advanced filtering techniques. These techniques can be implemented only at higher processing levels, on raw data. *Intelligent filtering*, *pattern recognition*, *compressive sensing* (Chapter 5 by M. Chiani), and *machine learning* are just a few examples of possible functionalities of this block.

Based on the role of the structure of Fig. 1.9, we can draw Table 1.1 showing different strategies of information detection according to the complexity of the information to be extracted. The bottom line is related to the transducer case, where the A/D conversion of the stimulus readily captures the information. This is the case of a direct measure where the transduction directly converts the intensity of a stimulus into information content. In the pattern/event line of the table, the information extraction is based on a priori knowledge of the pattern so that suitable algorithmic strategies could be conceived to get the right information content. The last upper two lines of the table represent a borderline case of sensing in which not only a measurement should be performed but also learning is involved. In supervised classification, the computational machine should learn the rule by way of predefined examples (training set). In the top line of the unsupervised learning case, the computational architecture should be able to identify classes of events/objects based on their inherent characteristics.

To summarize:

- Electronic sensors should be segmented into different processing stages, similar to the biological paradigm of the cognitive process aimed at extracting information from the environment.
- Depending on the information content's complexity, electronic sensors implement higher degrees of computation in the final operational blocks to increase the information classification.

A problematic and still debated argument is quantifying the complexity of information extraction. This is generally unknown; however, we can relate it to the minimum number of computational resources required to solve it.

1.4 At the Origin of Uncertainty: Thermal Agitation

Errors are differences from observed results from what we expect. One of the primary sources of error is *noise*, arising from random physical processes (Chapter 6), so that it is one of the main limiting factors of the sensing process. As discussed, noise limits the resolution of the electronic interfaces (Chapter 7); thus, the amount of information conveyed into the sensor.

If we look at Fig. 1.10, we can see a simple mechanical force transducer (Chapter 11). One end of a cantilever is anchored to a firm reference while a variable force is exerted on the other free end. On the cantilever's upper side, a laser beam is reflected toward a surface or a position-sensitive optical sensor (Chapter 9). Therefore, the position of the cantilever beam is proportional to the input force, realizing a force meter.

What is the limit of discrimination of this “sensor”? In principle, we can sense any slight variation of the input. If it is hard to distinguish variations on the screen, we can move the screen to a greater distance than the original one to look at variations clearer. In principle, this sensor has an “infinite” capability of discrimination (down to fundamental physical limits).

Unfortunately, nature made things a more complex way, and we know that any mechanical system is at a microscopic level subject to molecular agitation. In thermal equilibrium, any atom of the cantilever and any molecule of the surrounding gas is subject to a natural thermal agitation where the mean kinetic energy of any particle of the system is a microscopic expression of the temperature.

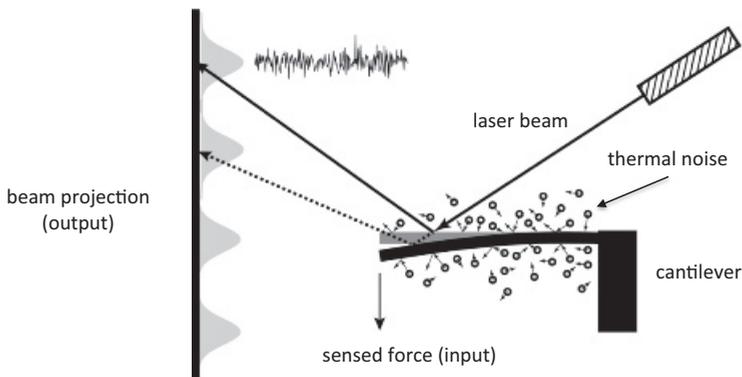


Figure 1.10 A mechanical sensor as a paradigm of electronic transducers. The discrimination of the sensor is limited by thermal noise.

Therefore, the cantilever is subject to natural and random displacements that determine an erratic movement of the beam projection; thus, the measure is affected by *uncertainty* so that we no longer can have an infinite capability to sense the input source in a limited amount of time. It is said that the discriminability of this sensor is thermal noise limited. We could average samples to increase the resolution of the system; however, this takes time and is limited by how fast the force signal is varying (signal bandwidth).

Understanding the balance between the signal (force) and the noise (thermal agitation) is necessary for sensor design. We will see that the above sensor mechanical paradigm is very much similar to electronic sensing systems. Even if fundamental physical limits are at the bottom of any sensor's resolution, thermal noise is one of the most common limitations in devices operating at room temperature and at the microscale.

1.5 Basic Constraints of Electronic Sensor Design

Looking back to all the examples of this chapter, we can devise several basic interdependences and constraints to be taken into account in sensor design.

- Resolution-bandwidth tradeoff. We can increase the resolution (i.e., reduce the uncertainty) by averaging the readouts by following the law of large numbers, as in the force sensor example. However, we must assume that the input force shall be stable during the averaging. This means that averaging is limited by signal bandwidth, and we cannot follow input signals that are faster than the averaging time. Therefore, the higher the resolution, the lower the bandwidth.
- Resolution–power consumption tradeoff. Depending on the complexity of the information to be extracted, we need to use higher amounts of power due to the larger computation requirements to gain higher information (resolution).
- Bandwidth–power consumption tradeoff. Since the computation requires energy consumption in real systems, the elaboration of higher information in a shorter time implies higher power consumption.

A block diagram of the foregoing relationships is shown in Fig. 1.11. The diagram shows the basic constraints of the sensor design divided into three main areas (Chapters 2 and 3). The first is related to the *information constraint* related to the amount of information conveyed by the sensing system, which is represented by the dynamic range. The *dynamic range* is in turn determined by the *operating range* and the input-referred *resolution* of the system. A second area is related to the *system's time constraint*, that is, the bandwidth. A third area is related to the *energy constraint*, which is represented by the power consumed by the sensing system.

A given electronic technology or architecture allows us to determine *figures of merit* (Chapter 3) relating and trading off the three areas mentioned above. Therefore, once we have two out of the three constraints set, we can determine the remaining. For example, once we have the required bandwidth and dynamic range for a given figure of

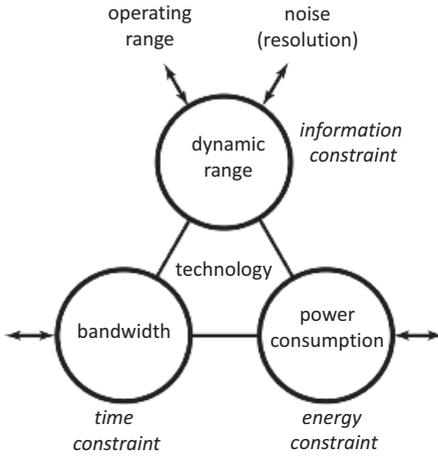


Figure 1.11 Basic design constraints of electronic sensor design.

merit, we can determine the minimum required power consumption for a given technology. Alternatively, we can determine the dynamic range for a given power budget and bandwidth, that is, its maximum achievable resolution.

Further Reading

- Brillouin, L., *Science and Information Theory*, 2nd ed. Mineola, NY: Dover Publications, 1962.
 Pierce, J. R., *An Introduction to Information Theory: Symbols, Signals & Noise*. Mineola, NY: Dover Publications, 1961.
 von Helmholtz, H., *Popular Lectures on Scientific Subjects*, Harvard Un. D. Appleton, 1873. L.