Information Theoretical Measures for the Analysis of Physiological Signals

Martina Valente, IEEE Student Member, Giandomenico Nollo, EMBS Member, and Luca Faes, EMBS Member

Abstract—In the recent years, the healthcare system is undergoing a profound change shifting from a reactive care, in which people are treated when necessary, to proactive care, based on people’s involvement in their own health on a daily basis, by means of continuous monitoring and prevention. The continuous assessment of people’s health status results in many innovative hardware and software solutions entering the healthcare market. Among the last category, a field of great interest for the research community recently is the analysis of physiological signals within the paradigm of information theory. Physiological systems can be indeed considered as communication systems in which processing, storage and transfer of information can be analysed and studied. This approach has been particularly successful for many applications in neural, cerebral and cardiovascular domain offering valuable metrics and paradigms for describing underline physiological systems functioning and non-functioning. The purpose of this white paper is to provide the readers with a basic understanding of why information theory is a valuable tool for the study of the human being, to present some of the most common measures being used and to discuss a couple of paradigmatic applications.

I. INTRODUCTION

One of the main challenges smart cities need to cope with is the increase in world population. According to the United Nations 2015 World Population Prospects [1], the number of people is projected to increase by more than one billion within the next 15 years, reaching 8.5 billions in 2030, and to increase further to 9.7 billions in 2050 with 66 percent of people living in cities [2]. This rapid growth requires the development of new efficient and sustainable solutions in many different application scenarios, such as energy distribution, transportation, communications, healthcare, government, just to mention some examples.

Within the same context, the recent technological advances have fostered the widespread integration of ICTs in everyday activities. Smartphones, Cloud Computing and Internet of Things represent the fundamental constituents of nowadays’ smart systems, that drive the change for a smarter world based on instrumentation, interconnection, intelligence [3]. Instrumentation enables the collection of real-time data through embedded sensors that communicate over wired or wireless networks. As an example, one may think at video-surveillance cameras in public spaces, electricity meters in households, body area sensors for continuous monitoring of physiological signals. The interconnections favour new ways to share information. Finally, intelligence is needed in terms of new computing models, algorithms and advanced analytics to treat massive amounts of data provided by smart sensors and to enable better decisions for cities and citizens.

Fostered by demographic issues and technological trends, the healthcare system is experiencing a shift, going from an episodic care provided by hospitals when needed, to a proactive care, in which citizen are actively involved in the management of their own health, thus reducing the demand on healthcare structures and guaranteeing better prevention.

In order to empower people towards a more efficient and sustainable healthcare system, several solutions have been proposed to collect and analyse data, as well as to infer the cognitive and physical state of a subject, without the need for medical doctors to assess the condition themselves. Data collection is favoured by the massive use of mobile devices and wearable sensors having the capabilities to gather information from a wide variety of physiological districts non-invasively and most naturally. The huge amount of data being generated requires the development of new storage infrastructures as well as new analysis frameworks exploiting the diversified offerings that such a massive database can guarantee. Among the most recent developments in this field, it has been proposed to use information theory as an analysis framework. The idea comes from the observation that the human body intended as a network of physiological systems that continuously interact with each other is no different from a communication system, in which nodes connected by means of information channels produce, store and exchange information [4]. Many works have been published in this regards, including many different applications among which cardiovascular control [5], cardiorespiratory couplings [6], brain-heart interactions [7].

In order to better explain the relation between information theory and physiological signals and introduce some of the most commonly used measures, information theory is presented in the next section.

II. INFORMATION THEORY

The basis of information theory were first introduced by the electronics engineer and mathematician Claude Shannon in his seminal article, A Mathematical Theory of Communication [8]. Shannon’s work is mainly focused on coding and transmission of information in communication systems, an entity that he used to model as shown in Figure 1. An information source produces a message to be communicated to the receiving terminal. The transmitter encodes the message in order to make it suitable to be transmitted over the channel. The channel
is simply the medium over which the signal is transmitted, that can introduce noise. The receiver operates the inverse operation as the transmitter. Finally, the message is delivered to its destination.

In order to deal with general problems involving communications systems, Shannon proposed to represent each block constituting the communication system as a mathematical entity, which produces, codes, transmits or receives information. The main intuition Shannon had was to treat information, by then considered an abstract concept, as a measurable quantity, that can be quantified by means of its own unit of measure, the bit.

A new era of communication begun, characterized by techniques aimed at optimizing data processing and transmission in order to maximize the exchange of information while maintaining sufficient robustness against noise introduced by the channel. Concepts such as source coding and channel coding characterized communications after Shannon proposed its paradigm.

But how do these concepts relate to physiological signals? Recently, the study of the human body has shifted from a reductionist approach, where each organ system was studied independently in its own complex structural organization and regulatory mechanisms, to a broader perspective, in which the human organism is seen as a complex network of physiological interactions, where network nodes represent different physiological systems and network links indicate the dynamical interaction (coupling) between systems [9], happening through various feedback mechanisms and across different spatio-temporal scales [10].

The parallel with Shannon’s model of a communication system is straightforward: the role of source and destination is played by the different physiological systems, while the channel is represented by the network link that provides the connection.

In the attempt to model physiological systems, physiological signals are recorded from the main sites of interest. According to the model described before, a physiological signal is nothing more than the output of the source in Figure 1, that itself is described in terms of the amount of information that is produced per second. As any other signal, physiological signals are sequences of symbols, each of which possesses its own information content, that can be quantified in terms of bits. Given this fact, it is possible to obtain a number of different measures related to the physiological system of interest, by simply analysing the physiological signal that represents its output.

### III. Measures

#### A. Entropy

When a single symbol is considered, information content depends on its probability of occurrence. Intuitively, a symbol with a very high probability of occurrence has a low information content, while a symbol with a low probability of occurrence has a high information content. In mathematical terms, it has been proven [8] that the function that best models this behaviour is the logarithmic function. The Shannon information content of a symbol \( x \) is then defined as:

\[
h(x) = \log \frac{1}{p(x)},
\]

where \( p(x) \) is the probability of occurrence of the symbol \( x \). It is straightforward to notice from (1) that the Shannon information content \( h(x) \) ranges between zero and infinity. The first limit case happens when \( p(x) = 1 \). In this case, the symbol \( x \) brings no information, because there exist no alternative to it (\( p(x) = 1 \) means that one single symbol is allowed to happen, namely \( x \)). The second limit case happens when \( p(x) = 0 \). More realistically, let us consider a very low \( p(x) \). The information associated to symbol \( x \) turns out to be very high, because the symbol is infrequent and consequently very valuable. It is exactly the same concept behind collecting: the more an item is rare, the higher its value is. In case digital symbols are considered, value is expressed in terms of information content.

Up to here, one single symbol was considered. However, physiological signals are series of symbols. Thus, it is necessary to employ a measure that takes into account the average information yielded by the series. This kind of measure is called entropy and it is defined as the sum of the Shannon information content of each symbol weighted by the probability of occurrence of the same symbol. In mathematical terms:

\[
H(X) = \sum p(x)h(x) = \sum p(x) \log \frac{1}{p(x)},
\]

where \( X \) denotes the time series and \( H(X) \) the entropy.

It is common to consider entropy as a measure of information content. Another popular interpretation is to relate it to the concept of uncertainty. A simple example will help clarifying this relation.

Imagine to have a source that is able to output one single symbol, let us say "A", and nothing else. This source is 100% predictable, because no symbols other than "A" can be found at its output. Thus, a sequence of symbols produced by this kind of source will be constituted by a certain number of equal elements. What is the information content of this signal? It is zero because there is no uncertainty about the values the signal can take at each given time instant.
On the contrary, suppose there is a source whose output at each time instant can be one among a certain set of symbols, each of which has the same probability of occurrence. A typical example is a dice, where each of the six sides has the same probability as the others to happen. The information content of this kind of signal is maximum, being the uncertainty maximum as well (no information is available regarding the next symbol the source is going to issue).

B. Measures of complexity

Entropy itself is a measure that does not take into account the temporal evolution of signals. However, a wide range of signals, including physiological ones, is characterized by a certain degree of temporal correlation that can vary over time or across different states. Temporal correlation implies that the current value of the signal can be predicted up to a certain accuracy from the values that the signal took in its past.

Supposing the current time instant is \( n \), it is common to denote the current value of signal \( X \) as \( x_n \) and to represent the past of signal \( X \) by means of a vector of values \( x_n^- = \{x_{n-1}, x_{n-2}, x_{n-3}, \ldots \} \) that the signal took in its past.

In order to quantify the temporal correlation that a signal exhibits, or better its dynamic behaviour, two information theoretical measures are commonly used: conditional entropy and information storage. Before the mathematical description of these two quantities, let us introduce some basic probabilistic notation.

- The probability that a signal takes the value \( x_n \) at time instant \( n \) is simply \( p(x_n) \).
- The probability of observing a symbol \( x_n \) given a set of observed past symbols \( x_n^- \) is defined as \( p(x_n|x_n^-) \).
- The probability of observing simultaneously \( x_n \) and \( x_n^- \) is defined as \( p(x_n, x_n^-) \).

Using the above probabilities, conditional entropy is defined as [4]:

\[
CE(X) = H(X_n|X_n^-) = \sum p(x_n, x_n^-) \log \frac{1}{p(x_n|x_n^-)}
\]

and quantifies the average amount of information that is contained in the present of the signal and cannot be explained observing its past.

Information storage instead is defined as [4]:

\[
IS(X) = I(X_n; X_n^-) = \sum p(x_n, x_n^-) \log \frac{p(x_n|x_n^-)}{p(x_n)}
\]

and quantifies the average reduction in the uncertainty of the present following the observation of the past.

The sum of \( CE(X) \) and \( IS(X) \) gives the entropy of signal \( X \), namely \( H(X) \). Conditional entropy and information storage can be thus considered as decomposition factors of the entropy term.

When physiological signals are considered, conditional entropy is used as a measure of complexity while information storage yields the information about regularity. The more the conditional entropy is high, the more the signal is unpredictable given its past, the more its complexity is high. On the other hand, the more the information storage is high, the more a signal can be predicted from its past, the more it is regular.

C. Measures of coupling

Entropy, complexity and regularity are indexed that can be used to characterized single physiological signal (and consequently systems). However, information theoretical measures are also very helpful in the assessment of interdependencies between different systems. It is well accepted that different signals depend on one another: one may think at heart rate and blood pressure, respiration and heart rate, the response of different areas of the brain, etc. When more than one signal is considered, there exist a number of different measures of information theory that can serve the purpose of assessing the coupling between signals (transfer entropy [4] [11] [12], mutual information [11], interaction information [11], etc.). Different measures have different specific features, meanings, strengths and limitations. However, the general reasoning behind is the same: quantifying how much the past values of a signal (or multiple signals), called driver(s), are useful to predict the current value of another signal, called target (in case instantaneous effects are considered, the current value of the driver must be included in the computation as well). Further details will not be covered here because coupling between signals is not the objective of this white paper.

IV. Applications

A. Study of postural stress

Cardiovascular disease (CVD) is one of the main causes of disability and death in industrialized countries. An estimated 17.5 million people died from CVDs in 2012, representing 31 percent of all global deaths [13]. Thus, it is of great social and medical importance to study how cardiovascular dynamics change in response to different conditions, such as pharmaceutical drugs administration [14], pathological states [15], physically challenging situations [16].

When the last category is considered, the study of the physiological response to a physical stress is usually performed by means of the so called head-up tilt test. The objective is to recreate the situation leading to a postural stress in a controlled environment, such a laboratory or an equipped hospital room.

The test consists in having the patient lying on a flat table with foot support, secured to it by means of protective straps. The physiological parameters of the patient (heart rate, respiration, blood pressure, etc.) are recorded by means of suitable medical equipment during the whole test. At the beginning of the test, the patient is resting supine and horizontal on the table. Then, an orthostatic challenge (the physiological response to a change in posture) is caused tilting the table to a vertical position. Tilt angle and duration depend on the protocol.

When we change our posture, for example from supine to upright, our body activates a number of measures in order to keep our physiological state (heart rate, blood flow, respiration, etc.) as stable as possible (homeostasis). This is mainly
the task of the autonomic nervous system. Consequently, by monitoring physiological signals it is possible to gain information about the functioning/non-functioning of the autonomic nervous system.

Let us analyse the effect of the tilt on heart rate variability. Heart rate variability (HRV) is obtained from the ECG as a series of time intervals corresponding to the time differences between two consecutive R peaks (each R peak in the ECG signal is representative of a heart beat).

Figure 2a shows a 300 samples HRV time series of a subject at rest. The corresponding values of conditional entropy and information storage are also shown. As soon as the subject gets out of the rest condition because of tilt, the orthostatic stress causes a drop in the arterial blood pressure, which is sensed by the baroreceptors located on the vessel walls. Baroreceptors cause in turn the activation of the sympathetic limb of the autonomic nervous system that triggers an increase in the heart rate in order to keep the blood pressure at a stable level. The sympathetic activation is reflected in the HRV signal as a strong dominant oscillation, that results in an increase of the information storage, being the signal more regular and predictable than during rest condition. This effect is well visible in Figure 2b, as a low-frequency oscillation that was not in place in the signal at rest.

**B. Study of mental stress**

Stress is a huge problem in today’s society, both at the individual and at the societal level: it was estimated that in 2013 in the EU the cost of work-related depression was around 617 billion Euro annually [17].

Mental stress has been associated with the same physiological phenomena of upright posture [18]: when stress is experienced, the sympathetic branch of the autonomic nervous system is activated, while the parasympathetic one is suppressed, leading to what is commonly known as fight-or-flight response. In this condition, an increase in blood pressure, heart rate and respiratory frequency is experienced [18] [19] [20]. However, individual differences emerge in the mode of autonomic control to stress, characterized by various combinations of sympathetic and parasympathetic contributions [20].

Similarly to physical stress, the assessment of the physiological response to mental stress is performed by means of a range of different laboratory tests: among the most widely used it is possible to find mental arithmetic tasks, in which subjects have to perform arithmetic operations silently [6] or aloud [21], Stroop test [22], or speech presentations [21]. As a result, some recent studies [23], supported by previous findings [18], seem to suggest that blood pressure variations could provide interesting information regarding the physiological regulation during mental stress. In particular, it is shown that the complexity of the time series obtained as a sequence of maxima of blood pressure (i.e. systolic blood pressure) signal for each cardiac cycle (SBP time series) is found to increase when a mental stressor is applied.

Figure 3 shows the paradigmatic case of a subject whose SBP time series shows a certain degree of regularity at rest, that seems to be lost when a mental arithmetic task is performed. This kind of behaviour is reflected in the values of conditional entropy and information storage, that are found to increase and decrease. It is suggested that the physiological mechanism leading to this result could be the increased involvement of higher brain cortical structures in the control of blood pressure.

Not only cardiovascular signals are altered as a consequence of mental stress. There exist many studies in literature that take into account changes in the respiratory signal when mental tasks are performed [24] [6] [25]. When complexity information is considered, an increase is associated to mental stress, as an intuitive consequence of a less regular breathing characterized by pauses and erratic patterns.

Figure 4 shows the time series obtained sampling a respiratory signal at the onset of each cardiac R wave (i.e. every time the heart beats). When the subject is resting quietly, the respiratory pattern is regular and characterized by
a dominant oscillation, while during the mental task is much more complex, characterized by breaths of different duration and amplitude.

V. STRENGTHS AND LIMITATIONS

A lot of different techniques exist in literature for the analysis of physiological signals, including time domain measures, frequency domain measures, and non-linear approaches. Information domain techniques belong to the last category. In the following, a brief explanation about strengths and limitations of this kind of measures compared to other approaches is provided.

First of all, it is important to clarify what is the meaning of non-linear, also called model-free, approaches. As the name suggests, the peculiarity of this kind of technique is to avoid making strong assumptions about the nature of the investigated signal dynamics. As an advantage, it is robust against model misspecification, a situation happening when the underlying model does not fit the observed data. Model misspecification may be critical when complex parametric models are identified on short and noisy time series, as the ones extracted from most physiological signals.

When compared to linear methods, the main advantage of using model-free techniques is the possibility to explore non-linear effects [4]. Whereas linear methods quantify predominantly the magnitude of signals' fluctuations, they do not characterize the complex dynamics, that may provide additional information [26]. In this case, the benefit resides in the type of information that is possible to extract.

In addition, non-linear methods are robust against errors arising from the assumptions behind frequency domain methods. This last type of techniques assume that all the interactions among frequency components of a signal are linear. Even though the linear assumption can be sufficiently accurate to model a variety of physiological phenomena, most physiological signals are the result of complex non-linear interactions of different central and peripheral oscillators, whose frequencies and coupling strength change in time and across conditions [21]. As a result of coupling, the physiological rhythms exceed their own basic frequency bands because of the appearance of harmonics and sub-harmonics. In such situations, the information provided by frequency domain methods might be impaired. Thus, the use of non-linear methods such as information domain techniques is recommended.

It must be pointed out however that linear methods are often preferred in many practical applications.

First, this is due to the reduced computational load associated to linear model-based methods: once a model is defined, linear techniques are very fast and powerful in the computation of the parameters of interest.

Secondly, linear measures are often very intuitive when it comes to interpretation and have the ability to isolate phenomena that are instead obscure when other techniques are used. As an example, in the analysis of HRV signal, high frequency components (HF: 0.15-0.4 Hz) are representative of the involvement of respiration in the variation of heart rate. This fact appears to be very intuitive knowing that the respiratory frequencies are usually limited to the HF band and that inhaling causes an increment of the heart rate, while expiration causes a decrement. On the other hand, the information domain measures of complexity outlined above do not provide information related to specific frequency bands, thus combining the contributions coming from respiration with the ones coming from other sources of variability in a single index. It is possible however to isolate effects coming from different districts using measures of coupling similar to the ones briefly discussed in Section III-C.

VI. CONCLUSION

Health in smart cities is mostly relying on technologically advanced solutions in order to cope with the increasing demands of a continuous monitoring of patients and of more and more specialized services. According to the outcomes of recent studies, information theory was proven to be a valid tool for the assessment of a variety of physiological and pathological conditions, allowing to extract from signals very specific information that other commonly used measures failed to detect. The aim of this paper was to present the basics of information theory and its relation to physiological signals. For a better understanding, some of the fundamental measures have been introduced and discussed. In order to prove the applicability of such measures in practical contexts, paradigmatic applications have been presented, highlighting strengths and limitations of the employed techniques. Given the positive outcomes achieved, it is expected that more and more specialized measures of information theory will help understanding complex physiological phenomena, thus providing always better solutions for a smart and efficient healthcare system.

ACKNOWLEDGMENT

This project is supported by IEEE Smart Cities Student Grant Program.