Biomedical Signal Processing and Control 43 (2018) 41–63

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Trends in biomedical signal feature extraction

Sridhar Krishnan[1], Yashodhan Athavale*,[2]

The Department of Electrical and Computer Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada

ARTICLE INFO

Article history:
Received 12 December 2016
Received in revised form 2 January 2018
Accepted 18 February 2018

Keywords:
Feature extraction
Biomedical signal processing
Pattern classification
Dimensionality reduction
Machine learning

ABSTRACT

Signal analysis involves identifying signal behaviour, extracting linear and non-linear properties, compression or expansion into higher or lower dimensions, and recognizing patterns. Over the last few decades, signal processing has taken notable evolutionary leaps in terms of measurement – from being simple techniques for analysing analog or digital signals in time, frequency or joint time–frequency (TF) domain, to being complex techniques for analysis and interpretation in a higher dimensional domain. The intention behind this is simple – robust and efficient feature extraction; i.e. to identify specific signal markers or properties exhibited in one event, and use them to distinguish from characteristics exhibited in another event. The objective of our study is to give the reader a bird's eye view of the biomedical signal processing world with a zoomed-in perspective of feature extraction methodologies which form the basis of machine learning and hence, artificial intelligence. We delve into the vast world of feature extraction going across the evolutionary chain starting with basic A-to-D conversion, to domain transformations, to sparse signal representations and compressive sensing. It should be noted that in this manuscript we have attempted to explain key biomedical signal feature extraction methods in simpler fashion without detailing over mathematical representations. Additionally we have briefly touched upon the aspects of curse and blessings of signal dimensionality which would finally help us in determining the best combination of signal processing methods which could yield an efficient feature extractor. In other words, similar to how the laws of science behind some common engineering techniques are explained, in this review study we have attempted to postulate an approach towards a meaningful explanation behind those methods in developing a convincing and explainable reason as to which feature extraction method is suitable for a given biomedical signal.

Contents

* Corresponding author.
E-mail addresses: krishnan@ryerson.ca (S. Krishnan), yashodhan.athavale@ryerson.ca (Y. Athavale).
[1] Senior Member, IEEE.
[2] Student Member, IEEE.

https://doi.org/10.1016/j.bspc.2018.02.008
1746-8094/

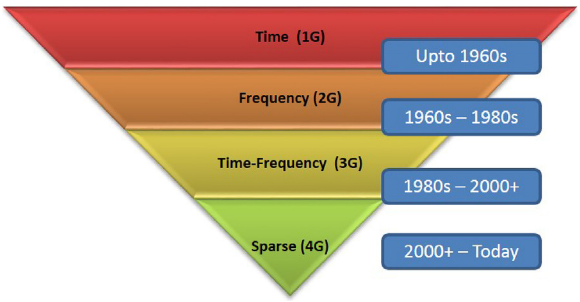

Fig. 1. Evolution of biomedical signal feature extraction.

1. Introduction

Signals are omnipresent. This statement certainly holds true when we are able to represent most stationary and non-stationary phenomenon in mathematical expressions. These representations are able to give us keen insights into those phenomena, and help us in identifying characteristic patterns of interest. Signal processing involves analysing analog/digital signals with the intention of measurement, reconstruction, quality improvement, compression, feature extraction and pattern recognition. Advancements in sensor technologies have come a long way by making signal data acquisition, storage and analysis easier, as well as opening doors for further improvisation considering unstructured big data (Fig. 1).

Someone might ask "If signal analysis should be easier to achieve with developments in signal processing algorithms, then how come we have to deal with increasingly complex mathematical representations and optimization problems?" The only suitable answer to this would be that – modern day information theory relies extensively on big data signals being churned out by sensors from our natural and digital environments, and treating these signals requires algorithms that are highly efficient in storage and computation. Although the underlying algorithms constitute complex mathematical operations, the flow of code is designed with the intention of processing maximum amount of signal data and discovering characteristic patterns in shortest time possible. This would be conducive only if we are able to seamlessly stream and process data. This in turn motivates us to design better tools for capturing useful information from signals at the source, and discarding unwanted signal artefacts, which would also lead to hardware optimization. As a quick remark, we would like to highlight that this concept has been successfully implemented in state-of-the-art compressive sensing techniques for signal acquisition, analysis and reconstruction.

1.1. Evolution of feature extraction methods

In simple terms, feature extraction is the process of unveiling hidden characteristic information about the input signal and its behavior of its sources. That is, we are able to represent a given input signal by a set of features which represent a specific behavior or pattern depicted by the signal; or a compact or useful representation of the signal [1–5]. Feature extraction is usually a dimensionality reduction or data compression/reduction process and helps in reducing the number of resources required to analyze an input signal. In other words, given a large input signal with multiple redundant components, performing feature extraction on it would yield a smaller set of representative data which could describe the original signal with sufficient accuracy and also help in building an efficient and robust pattern classifier system [1,6–10].

We suggest that the user derives application-dependent features rather than generic features, as they would better suit and depict signal behavior and underlying patterns. For example, when

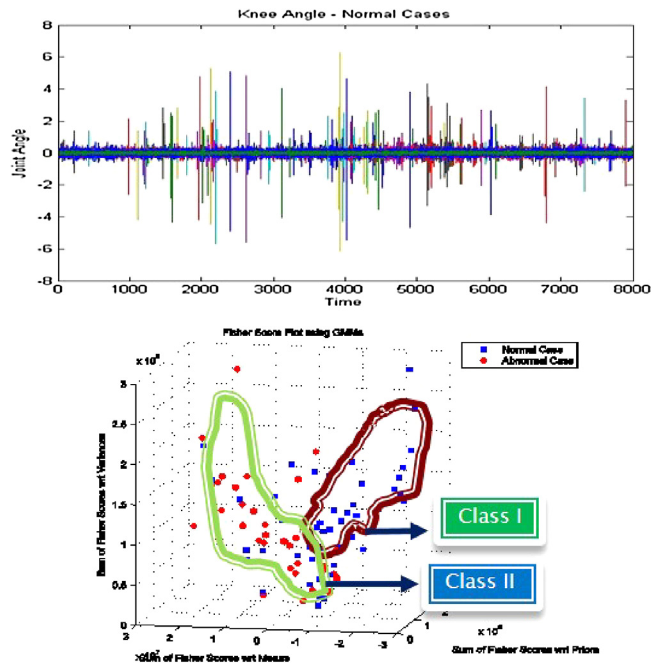

Fig. 2. Example of pattern classification.

we are attempting to analyze and classify music signals, we do not need mean or variance of the signals or even its root mean squares (RMS), since the room or environment settings will be tuned accordingly, but we might need to use them when audio signals are taken from non-modifiable sources. Before we proceed with extracting information from signals, we usually discretize the continuous analog signals into discrete digital signals using an A-to-D converter. This helps in identifying characteristic patterns over discrete time-intervals which otherwise cannot be observed if the signal is processed in analog form (Fig. 2).

At grass roots level, the easiest way to analyze time-domain signals is by filtering them, which helps in removing unwanted artefacts from the signals such as overlapping noise, third party/source components or values, and unwanted signal patterns. The most appropriate method for signal pre-processing method will be the one that can produce an output most suited to feature extraction. This method can be devised through two possible approaches: (a) if the artifact characteristics (such as noise patterns) are known, we can design appropriate signal filters, or (b) if the artifact properties are unknown, we need to pre-process the signal using trial and error approaches. Let us review some feature extraction methodologies applied to real-world biomedical signals in the past few decades. To make it simpler for the reader, let us group all the available signal processing and feature extraction techniques into the following four generations:

(1) Time domain
(2) Frequency domain
(3) Joint time–frequency domain
(4) Signal decomposition and sparse domains

The reader may note that the list of methods included in our review is by no means exhaustive, and that we have studied some key feature extraction methods in biomedical signal processing, and have attempted to find out the most efficient method from each generation. This study will further define the criteria to design an intelligent feature extractor specific for biomedical signals. In order to better explain and demonstrate our views on various feature extraction techniques, we have running examples