MINISTERUL EDUCAȚIEI, CERCETĂRII ȘI INOVĂRII
UNIVERSITATEA TEHNICĂ
DIN CLUJ-NAPOCA

Faculty of Electronics, Telecommunications, and the Information Technology
Bases of Electronics Department

METHODS FOR ANALYSIS AND CLASSIFICATION OF BIOLOGICAL SIGNALS

Ph.D. Thesis

Ph.D. student: Dipl. Eng. **Vasile Vlad MOCA**
Scientific supervisor: Prof. Dr. Eng. **Corneliu RUSU**

Cluj-Napoca
2010

Abstract

In a broad sense, *biological signals* or *bio-signals*, are those quantities (such as temperature, voice, or electroencephalogram) that can be continuously measured from living organisms. Regardless of their type, all biological signals reflect the physiological processes that produce them. Therefore, the study of bio-signals is essential for both medical applications and scientific studies.

The most interesting and fascinating structure of the human body is the brain. It allows us to process information, to store and remember memories, and to interact with the environment and our siblings. It is the home of our conscious thoughts and defines who we are. Despite rapid advances in neuroscience, due to its extremely high complexity (10^{12} neurons and $10^{14} – 5\times10^{14}$ synapses), the brain remains perhaps the least understood part of the human body. Nervous cells use electrical signals to encode, process, and transmit information, their activity being especially suited for investigations based on bio-electrical signals. Present recording techniques permit the acquisition of neuronal activity at microscopic (activity of single nervous cells), mesoscopic (local populations of cells), and macroscopic (large areas of the brain on the order of cm^2) scales. At all these spatial scales, dedicated signal processing techniques need constant additions and improvements to cope with demands imposed by the need to test new theories on how the brain functions. Consequently, we have conducted investigations of neuronal data at micro-, meso-, and macroscopic brain scales with emphasis on oscillatory (rhythmical) aspects that are intimately related to information processing. At each scale, signal-processing-related issues have been identified and solutions or alternatives have been proposed to address them.

At microscopic scales, single cell activity takes the form of point processes that are recorded as a succession of impulses called *spike trains*. For discrete data quantification of oscillatory behavior is problematic: Spectral analysis is not readily available and existing dedicated methods, based on autocorrelation function's ability to

reveal rhythmic (periodic) activity, suffer from stability issues and/or impose serious restrictions on the data in order to function properly (e.g. they require large amounts of data). To address this problem we have proposed a method based on the spectrum of the autocorrelation function, named the "*oscillation score*" that is able to reliably quantify the strength of oscillation. We show that, in contrast to existing methods, the oscillation score is simple and relatively fast, it is able to cope with low number of spikes (situation often found in data) and with multiple distinct oscillatory rhythms.

To calibrate the oscillation score and to compare it to other methods, we have designed a model of oscillatory spike trains able to reproduce statistical properties of recorded data. To the best of our knowledge, this is the first spike train model that allows for independent control over firing rates and oscillatory behavior. The properties of the model were explored in detail and were compared to those of data recorded from cat visual cortex. Results have shown that the model reproduces faithfully the global statistical and oscillatory aspects of neuronal data. The proposed model is useful for calibrating analysis methods (e.g. such as the oscillation score), for studies on artificial neural networks that need input with well-defined properties, and for comparative data-model studies (i.e. the benchmark of the oscillation score against other methods).

At mesoscopic scale (order of mm) the activity of many neurons sums up at the recording site to create a continuous signal termed the *local field potential*. Local field potentials are usually recorded with multichannel probes resulting in multivariate data. With the increase in the number of recorded channels many relations amongst variables remain hidden either because the number of possible relations is simply too large or because the existing relations are difficult to find (i.e. nonlinear relations). To address this issue we have proposed a method based on the concept of fractal dimension. The novelty comes from the fact that fractal dimension is assessed for data that is not expected to have fractal structure. We explored the method on both synthetic data with well-defined properties and on local field potentials recorded from cat visual cortex. As we have shown, changes in dimensionality at various geometric scales can reveal important aspects of the data, proving the usefulness of the method as a data mining technique. We have also shown that fractal dimension can be used to test a model's fit to the data. To this end, the dimensionality of the original data was compared with the dimensionality of

the data reconstructed based on the model's description of the data. Differences in dimensionality of the data and the model's reproduction reveal important information about the model. Our tests using the General Linear Model have proven that fractal dimension is of great help in estimating non-linear relationships within multivariate data.

At macroscopic scale (order of cm) we have investigated the problem of depth of anesthesia detection (DOA) based on single-channel electroencephalogram (EEG) recordings. To this end, we have created an artificial system employing Time Encoded Signal Processing And Recognition (TESPAR), a novel time-domain signal processing technique, and multi-layer perceptron artificial neural networks to asses DOA. The artificial system learned to discriminate between five DOA classes defined by imperfect human experts. Our evaluations have shown that the artificial system reproduced fairly well (within about 2% differences) the classification behavior of the human expert proving thus that TESPAR descriptors contain a high degree of DOA-related information. Moreover, investigation on the nature of information extracted by TESPAR has revealed that DOA related information is distributed across frequency bands. Importantly, our investigations have confirmed a novel view that acknowledges the usefulness of high frequency EEG (>80 Hz), which represents mostly muscle artifacts, for DOA assessment. The method we have introduced aims to offer an additional set of DOA relevant EEG descriptors that could be used to enhance already existing monitoring devices. After a brief exploration, we have found TESPAR to be useful also in the identification of epileptic activity. Thus, we concluded that TESPAR is a valuable feature extraction method able to condense signals of arbitrary length in compact and highly informative descriptors that reflect changes in the properties of the source producing the recorded signals.

These studies position the author at the confluence of signal processing, neurophysiology, and theoretical neuroscience, where, neuronal data are analyzed to gain insight into the governing principles of brain function and to produce concepts and predictions that are refined by going back to data. The author hopes that his studies will provide useful additions to the ultimate quest of neuroscience – understanding the brain.