UNIVERSITE SAAD DAHLAB BLIDA1

Faculté de Technologie

Département d'Electronique

THESE DE DOCTORAT

Spécialité : Electronique

NEURAL PROSTHETIC FOR FUNCTIONAL REPAIR

Par

Abdessalam KIFOUCHE

Devant jury composé de :

CHERFA Yazid	Professeur	U. Blida1	Président
FERDJOUNI Abdelaziz	MCA	U. Blida1	Examinateur
BENADDA Belkacem	Professeur	U. Tlemcen	Examinateur
RABEHI Abdelaziz	MCA	U. Tissemsilt	Examinateur
BENAMMAR Abdessalem	DR	CRTI-Cheraga	Examinateur
GUESSOUM Abderrezak	Professeur	U. Blida1	Directeur

Soutenue à Blida, le 15/07/2021

بسم الله الرحمن الرحيم

ABSTRACT

Brain machine interfaces capture brain signals in order to restore communication and movement to disabled people suffering from motor disorders or even brain palsy. Disorders come from the absence of brain signal transmission or total communication with the external world of the brain. In brain regions, the ensemble firing of a population of neurons represents Spatio-temporal patterns that are transformed into outgoing Spatio-temporal patterns which encode complex cognitive tasks. This transformation is dynamic, non-stationary (time varying) highly nonlinear. Hence, modeling such complex biological patterns requires specific model structures to uncover the underlying physiological mechanisms and their influences on system behavior.

In this study, we have investigated motor cortex responses recorded during movement in freely moving rats to provide evidence for the relationship between these patterns and special behavioral tasks. Recent multi-electrode technology allows the record of the simultaneous neuron activities in behaving animals. Intra-cortical data are processed in successive stages: spike detection, spike sorting, and intention extraction from the firing rate signal. We focus on the following important questions: is it realistic to link time events from the primary motor cortex with some time-delay mapping tool? Are some inputs more suitable than others for this mapping? Shall we consider separated channels or a special representation based on multidimensional statistics?

In this thesis, we concentrate mostly on the analysis of parallel spike trains when certain critical assumptions for the method to work are violated by the data. We have put efforts in defining explicitly if the underlying assumptions are indeed fulfilled. We propose different algorithms for the hand trajectory decoding process using two principal stages where the first is the online pattern detector based on template matching for spikes sorting and data preparation. The second stage is the decoder using the output of the previous stage. In the decoder, we use a linear system, based on the linear regressor and the Kalman filter, and nonlinear models

using the extended Kalman filter, time delay neural network and neural network aided the Kalman filter. We also show that these algorithms can demonstrate improved performance on inference tasks.

الملخص

تلتقط الواجهات آلة-دماغ إشارات الدماغ من أجل استعادة التواصل والحركة للأشخاص المعاقين الذين يعانون من اضطرابات حركية أو حتى شلل في الدماغ. تأتي الاضطرابات من غياب إرسال إشارات الدماغ أو الاتصال الكامل بالعالم الخارجي للدماغ. في مناطق الدماغ ، يمثل إطلاق مجموعة من مجموعات الخلايا العصبية أنماطًا مكانية وزمانية يتم تحويلها إلى أنماط مكانية وزمانية صادرة تشفر مهمة معرفية معقدة. هذا التحول ديناميكي وغير ثابت (متغير بمرور الوقت) وغير خطي بدرجة كبيرة. ومن ثم ، فإن نمذجة مثل هذه الأنماط البيولوجية المعقدة تتطلب هياكل نموذجية محددة للكشف عن الآليات الفسيولوجية الأساسية وتأثيراتها على سلوك النظام.

في هذه الدراسة، قمنا بالتحقيق في استجابات القشرة الحركية المسجلة أثناء الحركة في الفئران التي تتحرك بحرية لتقديم دليل على العلاقة بين هذه الأنماط والمهمة السلوكية الخاصة. تسمح تقنية الأقطاب المتعددة الحديثة بتسجيل الأنشطة العصبية المتزامنة في الحيوانات. تتم معالجة البيانات داخل القشرة في مراحل متتالية: اكتشاف السبايكات، وفرز السبايكات، واستخراج النية من إشارة معدل الإطلاق. نركز على الأسئلة المهمة التالية: هل من الواقعي ربط الأحداث الزمنية من القشرة الحركية الأساسية ببعض أدوات رسم الخرائط ذات التأخير الزمني؟ هل بعض المدخلات أكثر ملاءمة من غيرها لهذا التعيين؟ هل نعتبر القنوات المنفصلة أم التمثيل الخاص المبني على إحصاءات متعددة الأبعاد؟

في هذه الأطروحة، نركز في الغالب على تحليل قطار السبايكات المتوازي عندما تنتهك البيانات افتراضات حرجة معينة لطريقة العمل. لقد بذلنا جهودًا لتحديد ما إذا كانت الافتراضات الأساسية قد تحققت بالفعل. نقترح خوارزميات مختلفة لعملية فك ترميز مسار اليد باستخدام مرحلتين رئيسيتين حيث تكون الأولى هي كاشف الأنماط على الخط بناءً على مطابقة القالب لفرز السبايكات وإعداد البيانات. المرحلة الثانية هي وحدة فك التشفير باستخدام مخرج المرحلة السابقة. في مفكك الشفرة ، نستخدم نظامًا خطيًا ، يعتمد على الانحدار الخطي ومرشح كالمان ، والنماذج غير الخطية باستخدام مرشح كالمان الممتد ، والشبكة العصبية المتأخرة ، والشبكة العصبية لمساعدة مرشح كالمان. نظهر أيضًا أن هذه الخوارزميات يمكنها إظهار أداء محسن في مهام الاستدلال.

RESUME

Les interfaces cerveau-machine captent les signaux cérébraux afin de restaurer la communication et le mouvement des personnes handicapées souffrant de troubles moteurs ou même de paralysie cérébrale. Les troubles proviennent de l'absence de transmission de signal cérébral ou de communication totale avec le monde externe du cerveau. Dans les régions du cerveau, le déclenchement d'ensemble de population de neurones représente des modèles spatio-temporels qui sont transformés en modèles spatio-temporels sortants qui codent une tâche cognitive complexe. Cette transformation est dynamique, non stationnaire (variable dans le temps) et fortement non linéaire. Par conséquent, la modélisation de ces modèles biologiques complexes nécessite des structures de modèle spécifiques pour découvrir les mécanismes physiologiques sous-jacents et leurs influences sur le comportement du système.

Dans cette étude, nous avons étudié les réponses du cortex moteur enregistrées pendant le mouvement des rats en mouvement libre pour fournir des preuves de la relation entre ces modèles et une tâche comportementale particulière. Une technologie multi-électrodes récente permet l'enregistrement des activités neuronales simultanées chez les animaux qui se comportent. Les données intra-corticales sont traitées par étapes successives: détection des pics (spikes), tri des pics et extraction d'intention du signal de cadence de déclenchement. Nous nous concentrons sur les questions importantes suivantes: est-il réaliste de relier les événements temporels du cortex moteur primaire à un outil de plan des délais? Certaines entrées sont-elles plus adaptées que d'autres pour ce mappage? Devons-nous considérer des canaux séparés ou une représentation spéciale basée sur des statistiques multidimensionnelles?

Dans cette thèse, nous nous concentrons principalement sur l'analyse des trains de pointes parallèle lorsque certaines hypothèses critiques pour que la méthode fonctionne sont violées par les données. Nous nous sommes efforcés de définir explicitement si les hypothèses sous-jacentes sont effectivement remplies. Nous proposons différents algorithmes pour le processus de décodage de la trajectoire de la main en utilisant deux étapes principales où la première est le détecteur en ligne de formes basé sur la correspondance de forme pour le tri des pointes et la préparation des données. Le deuxième étage est le décodeur utilisant la sortie de l'étage précédent. Dans le décodeur, nous utilisons un système linéaire, basé sur un modèle de régression linéaire et le filtre de Kalman, et des modèles non linéaires utilisant le filtre de Kalman étendu, le réseau de neurones à retard temporel et le réseau de neurones dans l’aide du filtre de Kalman. Nous montrons également que ces algorithmes peuvent démontrer des performances améliorées sur les tâches d'inférence.

ACKNOWLEDGMENT

First of all, I would like to thank my supervisor of this thesis, Mr. Abderrezak Guessoum, professor at the University of Blida1, and express my deep gratitude to him for agreeing to lead and direct this work. I thank him warmly for his advice, his guidance, and the support that I found with him, and that he never ceased to testify to me for many years.

I would like to address my deep thanks to Mr. Yazid CHERFA, professor at the University of Blida1 for accepting to chair the committee in charge of examining my work.

I would like also to address my deep thanks to Mr. Abdelazize FERDJOUNI, MCA at the University of Blida1, Belkacem BENADDA, Professor at the University of Tlemcen, Mr. Abdelazize RABEHI, MCA at the University of Tissemsilt and Mr. Abdessalem BENAMMAR, Research Director at the CRTI Cheraga for accepting the examination of this work.

Works presented in this thesis was firstly in joint thesis between the University of Blida1 and the University of Evry (France) under the direction of Mr. Abderrezak Guessoum in jointly led with Mr. Vincent Vigneron, a professor at the University of Evry who I would like to address my deep gratitude and my warm thanks for his advice, guidance, and support that I found with him, and that he never ceased to testify to me for many years.

DEDICATION

To my parents, my wife and my kids M. Abderrahmane and Zineb Israa.

To all who know me.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

In the world, more than one billion people (about 15% of the world population) are suffering from disabilities or handicaps, and the number continues to increase rapidly. Disability takes many forms. While some conditions associated with disability lead to poor health and significant health care needs, this is not the case for all. Nonetheless, all people with disabilities have the same general health care needs as the rest of the population and therefore should have access to mainstream health care services. Article 25 of the Convention on the Rights of Persons with Disabilities reaffirms the right of such persons to obtain quality health care without discrimination. However, in reality, few countries offer satisfactory quality services to people with disabilities. Disabled people are all times dependent parsons. However, contemporary prosthetic devices are criticized for the lack of aestheticisms and the difficulty in control. For these reasons, we intend to improve the functionality of the myoelectric prostheses by embedding machine learning using brain signals for repairing at least the minimal activity of the disabled part of the body. The developed prosthetic device aims to restore partially the motor function of handicapped people by transforming the recorded brain signal from the motor cortex into a control signal in real time for the limb or the arm for instance. This is called brain machine interface.

The Brain Machine Interface (BMI), called also a Brain Computer Interface (BCI) and sometimes called Mind-Machine Interface (MMI), or sometimes called Direct Neural Interface (DNI), Synthetic Telepathy Interface (STI), is a system for reading information from the brain to be translated by a computer into a motor command or simple acts of communication for a down BMI and also a system for bringing information into the brain in a way that enables the equivalent of sensation or perception for the up BMI. Or more simply the Brain Computer Interface is the way for making a direct communication pathway between the brain with external devices in order to assisting, augmenting, and repairing human cognitive or sensory-motor functions. Neural prostheses offer the possibility to translate electrical neural activity from the brain into control signals for guiding paralyzed upper limbs, prosthetic arms, or computer cursors. Several research groups have already demonstrated that monkeys, as well as human, are capable to learn how to drive a robot arm, or more generally communicate with the outside world, simply by activating neurons ensembles that participate in natural arm movements. Brain computer interface requires four main stages: brain signal recording, spikes sorting, mapping between the desired actions and the brain signal, and finally the application on any device.

The different activities of the brain produce electrical and magnetic signals over and inner the head. Therefore, sensors can be used to detect diverse forms of variations in electrical or magnetic activities, at different times over different areas of the brain, to study brain activity. Measuring brain activity can be realized in two manners depending on invasiveness where the noninvasive is used for measuring electrical and magnetic activities from the outside of the head, this makes these techniques with low spatial precision except for some magnetic measure. Invasive method for recording brain signal based on electrical activity collection near neurons using sensors implanted by surgery, these techniques offer high Spatio-temporal precision. Multiple electrode arrays allow neurophysiologists to record the spiking activities of an increasing number of neurons. For instance, they have made feasible the recording of a large number of hippocampal cells along with the rat's position in its environment and hence the quantitative analysis of how rat's brain encodes spatial information in short term memory and use it for voluntary or non-voluntary action. Investigating these questions requires a collection of statistical tools to analyze how the animal's position is represented in terms of firing patterns of place cells. In this work, we have investigated motor cortex responses recorded ECoG during movement in freely moving rats to provide evidence for the relationship between these patterns and special behavioral tasks.

Recorded action potentials resulting from different cells are essentially distinguished by their shape. The difference of shape is due to the dynamics of generation of action potentials which is different for different cells, as well as to the filtering operated by the extracellular medium located between the electrode and the recorded neuron: according to the conductivity of the medium and the distance from the electrode to the neuron, the resulting shape of the recorded action potential can be very different. It should be noted that the action potentials from the same neuron are in theory very similar, but see their shapes evolve over time if the position of the electrode relative to the neuron is not stable. The spike sorting is a process; of three stages (detection, feature, and classification); makes it possible to classify the action potentials recorded by the same electrode into several groups corresponding to the different cells encountered near an electrode. Spikes sorting provides for each channel where each one corresponds to a neuron. Channels are represented by ones "1" separated by long suites of zeroes "0" or silences, the ones that signify the presence of spikes. We process spike trains in order to extract a 'firing rate'; the emphasis is on getting things to work robustly, with minimal efforts and with minimal delays, since the decoding must be real-time. Estimation algorithms can be designed to decode the desired trajectory from the neural activity patterns. We proposed for the spikes sorting an online detection using nonlinear energy operator and classification based on template matching where this was a dynamic algorithm able to detect new shapes (channels or neurons). A representation to get the continuous-time spiking rate is used by applying a moving Gaussian window [1, 2].

A control system could then generate appropriate signals for continuously guiding a paralyzed or prosthetic arm through space. Such control is indeed a daunting ultimate goal but would provide a 3 presumably natural control suitable for clinically viable systems. The core part is the spike train generated from various neurons. Sorting correctly the spikes with respect to their source improve significantly the decoding performance. Hence, before dealing with the decoding problem, the spikes should be sorted and the firing rate of each neuron extracted. In literature, various methods have been introduced to decode brain activities. This can be formulated systematically in a state-space framework; the hand representation (position on the two axes, velocity, and acceleration) can be modeled by a probabilistic relationship between this motion and the firing rates with a Kalman filter (KF). State-space methods provide a coherent framework for modeling stochastic dynamical systems [3]. A co-adaptive KF to train a rat for cortical control tasks can also be used. This method is able to estimate hidden states despite the lack of an accurate model of the system [4]. Particle filters can estimate the hidden states from a sequence of measurements using the Monte Carlo algorithm, without any assumption on the distribution of the observations [5]. In the decoding problem, we proposed the use of linear and nonlinear technics based on the linear regressor and the Kalman filter for linear modeling. The recurrent Kalman filter with delayed input is used to decode the hand trajectory from the primary motor cortex ElectroCorticoGraphy and also was used to aid the Kalman filter. Nonlinear decoding algorithms were better than linear because of the nonlinearity mapping between the hand position and brain signal [6].

The remainder of this thesis is organized as follows: In chapter 1 is described the background literature of brain computer interfaces based systems presenting its ideas, classifications, and applications. Chapter 2 presents the origin of brain signals and different techniques for recording these signals. It describes also the used database used in the rest of this work. Chapter 3 expands the spikes sorting with detail of the detection and isolation using nonlinear energy operator two technique of classification based on supervised support vector machine and templates matching and the representation of spiking rates. Chapter 4 concentrates on feature selection and data organization before decoding algorithms presented in chapter 5 where linear and nonlinear decoders were used to improve their ability to make a direct control from the brain to an external device (lever for bringing water). And finally, we summarized and distinguish results and future works in the conclusion.

1. BACKGBROUND

1.1. What is BMI?

Brain Machine Interface (**BMI**), can be found also under the terms Brain Computer Interface (**BCI**), Mind Machine Interface (**MMI**), Direct Neural Interface (**DNI**), or Synthetic Telepathy Interface (**STI**) is a system for:

- reading information from the brain to be translated by a computer into a motor command or simple acts of communication for a down BCI,
- bringing information into the brain in a way that enables the equivalent of sensation or perception in an up BCI,

Or more simply, the BCI is the way for making a direct communication pathway between the brain with external devices (world) in order to assisting, augmenting, and repairing human cognitive or sensory-motor functions.

In a down BMI, neural prostheses offer the possibility to translate electrical neural activity from the brain into control signals for guiding paralyzed upper limbs, prosthetic arms, or computer cursors. Several research groups have already demonstrated that monkeys as well as human are capable to learn how to drive a robot arm, or more generally communicate with the outside world, simply by activating neurons ensemble that participate in natural arm movements [7].

In general, BCIs have four principal components: (1) a record of the brain's invasive or non-invasive activities is necessary to know the user desired action. (2) the user must have a feedback of his control (artificial arm, cursor ... etc) to be able to adjust his mental commend, (3) the BMI must also be in real-time, and finally, (4) the system must rely on intentional control. That is, the user must choose to perform a mental task whenever s/he wants to accomplish a goal with the BCI.

Brain signals were discovered and recorded first time by Hans Berger as an electrical activity known as ElectroEncephaloGraphy (EEG) on July 6, 1924 [8, 9]. By analyzing EEG traces, he also was able to identify oscillatory activity, such as alpha (or Berger's) wave (8-13 Hz). The first recording device of Berger was very rudimentary based on silver wires implanted under the scalps of his patients. These were later replaced by silver foils attached to the patients' heads by rubber bandages. Berger connected these sensors to a Lippmann capillary electrometer, with disappointing results. However, more sophisticated measuring devices, such as the Siemens double-coil recording galvanometer, which displayed electric voltages as small as one ten thousandths of a volt, led to success. Berger analyzed the interrelation of alternations in his EEG wave diagrams with brain diseases. EEGs permitted completely new possibilities for the research of human brain activities.

Multiple electrode arrays allow neurophysiologists to record the spiking activities of an increasing number of neurons. For instance, they have made feasible the recording of a large number of hippocampal cells along with the rat's position in its environment and hence the quantitative analysis of how rat's brain encodes spatial information in short term memory and use it for voluntary or non-voluntary action. Investigating these questions requires a collection of statistical tools to analyze how the animal's position is represented in terms of firing patterns of place cells. This is the *decoding problem*. We process spike trains in order to extract a 'firing rate'; the emphasis is on getting things to work robustly, with minimal efforts and with minimal delays, since the decoding must be real-time. Estimation algorithms can be designed to decode the desired trajectory from the neural activity patterns. A control system could then generate appropriate signals for continuously guiding a paralyzed or prosthetic arm through space. Such control is indeed a daunting ultimate goal but would provide a presumably natural control suitable for clinically viable systems. The core part is the spike train generated from various neurons. Sorting correctly the spikes with respect to their source improve significantly the decoding performance. Hence, before dealing with the decoding problem, the spikes should be sorted and the firing rate of each neuron extracted.

To obtain and comprehend the signal processed in the biological network of neurons; spikes should be assigned to specific neural sources. This classification procedure is termed spike sorting. A proper sorting of spikes with respect to their origin ameliorates considerably the decoding operation. Then, before representing the position of the animal in terms of firing patterns, sorting spikes must be accomplished and the spike's rate of each neuron is calculated [10, 11].

Nevertheless, signal instabilities could arise from physical factors like the implants movement, reaction of the tissue or material degradation. Whatever the causes, signal changes can be significant, reaching 60% of the waveforms as reported by Dickey et al. [12]. In prior studies, we had used a multi-scale seriation approach for clustering spike trains [13]. Various methods were proposed for spike sorting. Some of these methods implemented pattern matching algorithms for identifying similar spikes, the main idea being the extraction of discriminative features from the shape of spikes and classification of them with supervised [14, 15] or unsupervised methods [16]. Some authors [10, 17], used independent component analysis (ICA) for distinguishing the spikes according to their sources; the independence assumption of the firing neurons helps to identify spikes from the same source. In [18] the occurrence time information of spikes and features related to the shape simultaneously is applied to estimate the Inter Spike Interval (ISI) for each neuron and sort the spikes using a Monte Carlo algorithm.

The decoding function that links neural activity to behavior may be relatively unstable as well as degrading decoding performance. The decoding problem starts using the firing rate extracted from the spike sorting stage; one can deal with the decoding problem. This can be formulated systematically in a state-space framework: In [2, 6], we used the time-delay neural network to decode the hand trajectory from the primary motor cortex signal of a rat, Wu et al. in [3] modeled the hand representation (position on the two axes, velocity, and acceleration) and the probabilistic relationship between this motion and the firing rates with a Kalman filter (KF). State-space methods provide a coherent framework for modeling stochastic dynamical systems. Kostov et al. [19] improved the possibility of real-time control with a presence of a partial spinal cord injury by using an adaptive logic network. Gage et al. in [4] implemented a co-adaptive KF to train a rat for cortical control tasks. This method is able to estimate hidden states despite the lack of an accurate model of the system. Recently Brockwell [5] used particle filters to estimate the hidden states from a sequence of measurements using the Monte Carlo algorithm, without any assumption on the distribution of the observations.

The state-space method offers a coherent outline for modeling stochastic dynamical systems. Note that despite the man-made neuroprosthesis will control a complex-living system, it shall not necessarily obey biological control principles:

Estimation procedures can be used to translate the brain signal into desired motor actions that can be used as a suitable command of a robot arm or other prosthetic interface. The final goal is an undeniable control, but it undoubtedly provides natural control suitable for clinically viable systems.

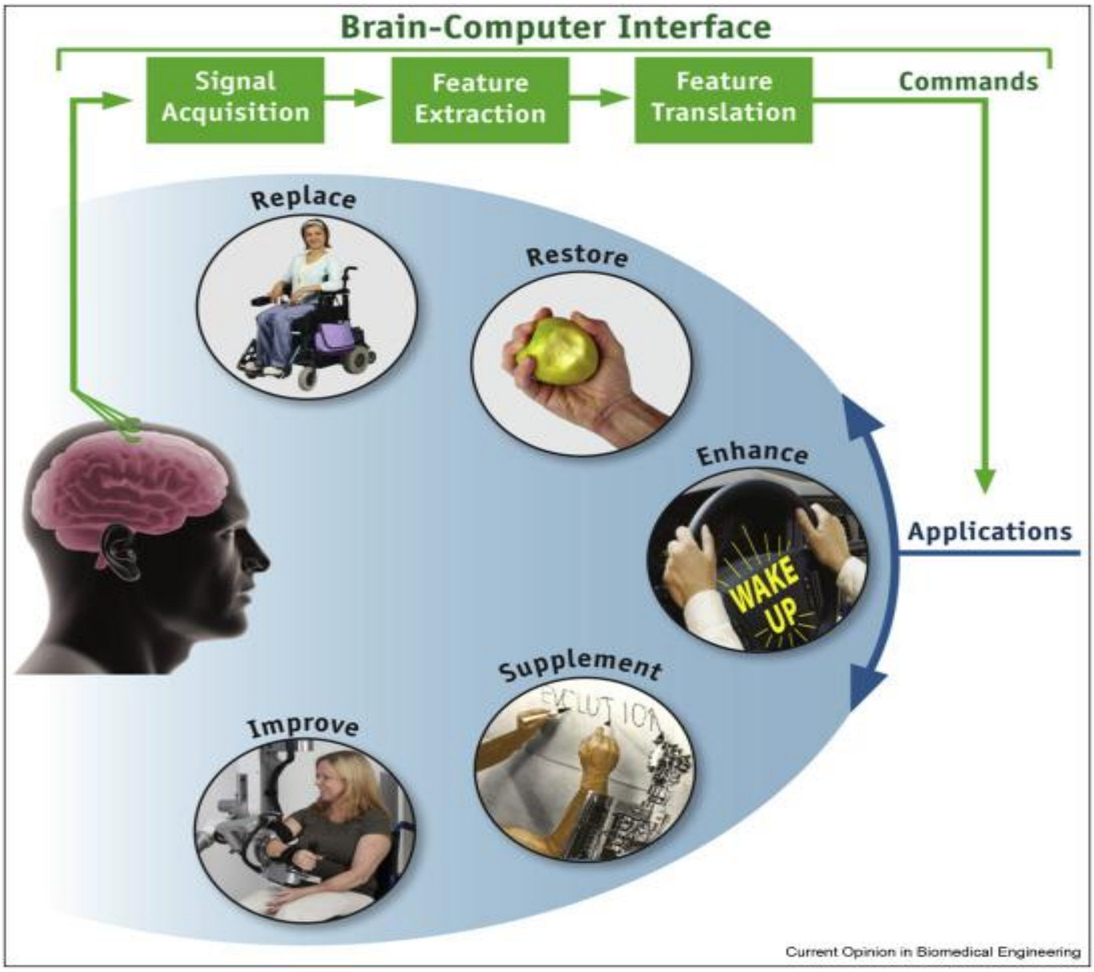


Figure 1.1: BMIs based system map and applications.

1.2. BMI Types and Classification

BMIs based systems are different if we talk about the type of the used source of signal or the realization and time of responses and other parameters. So the classification is necessary for designers and users to know how they can trust in their study and the possibility of realization with the ability of the users to adapt themselves with new commands added to brains. BCIs based systems can be classified depending to:

(1) The invasiveness: in which we talk about the source of the used signals where we focus our discussion on the recording technics with (and/ or without) surgery on the surface of the brain.

(2) The direction of the signal from where the signal is recorded and to where will be exploited or easily the up and down BMIs (the direction will be discussed in realization part).

(3) Synchronicity in where we explain the cause of brain signal with the realization time of the desired actions.

(4) The dependence of the BMIs talk about the used signal comes from secondary signals of the brain or principal signals originated from the nervous system.

1.2.1. Invasiveness

BCIs use brain activities with processing that can provide a control signal (motor decision …etc.) that reflects the intention of the user. For the invasiveness, we can distinguish in generally two types in brain signal recording (invasive and non-invasive) depending on the level of measure on the brain and specifically to the scalp. The activity of the brain generates magnetic and electrical activities in the environment of the brain depending on the area of the used technic of measurement. Sensors can identify different types of variation in magnetic and electric activity over time and space. Depending on the needs of the used application (time and space precision), we can record the brain signal in different areas on the environment of the head with surgery for the invasive and without surgery for the non-invasive recording.

Most of non-invasive brain signal recording measure electrical current flow over the head using sensors placed on the skin, this measure deliver the EEG (Electroencephalography) which refers to the sum of currents come from neurons near the surface of the cortex. Currents in this case received many translations, on its form and energy, across the head layers from neurons to sensors. Non-invasive signal recording refers in general to the EEG which is clinically very well established and used in clinical research settings. EEG recording can be classified as a very good signal depending on the application because of its inexpensive, lightweight, and usable equipment. As the EEG has advantages it present also disadvantages like the spatial resolution and pollution with other electrical signals as EOG (Electrooculography) and EMG (Electromyography) of eye movement and blink. Other non-invasive technics of brain signal recording based on the measure of magnetic field related to the brain activity deliver the MEG (MagnetoEncephaloGraphy), other measures the changes in the BOLD (Blood Oxygenation Level-Dependent) related to the cortical activation deliver the fMRI (Functional Magnetic Resonance Imaging) and NIRS (Near Infrared Spectroscopy). These technics beyond EEG are not useful for most of BCIs based systems because of their impractical, dimension, and price of equipment [20, 21].

Invasive brain signal recording measures the brain neural activity implanting electrodes under the scalp. These methods require surgery to implant the sensors necessary to the record. In the surgery, the user in this case receives an opening of its skull through a craniotomy (surgical procedure on the skull) and a cut of membranes covers the brain. Depending on the needs, sometimes electrodes are placed on the surface of the cortex to record ECoG (electroencephalography) with no penetration of the brain so with no damage of any neuron. When electrodes penetrate the brain tissue it can measure in deeper than the surface of the cortex to get the LFP (Local Field Potential), MUA (Multi-Unit Activity), and SUA (Single-Unit Activity). Invasive brain signal recording delivers the best signal quality in spatial resolution and higher frequency range.

1.2.2. Synchronicity

The synchronicity of a based BCI system can be defined in the generation and realization time of a desired action or command in the cortex. In synchronous BCIs based systems responses are evaluated at specific times or easily when the user reacts to a stimulus (visual and/or sound) generated by a BCI during many precised periods of time or in fixed time windows. We can also see it when the brain responds to any stimulus, the potential linked to an event is evoked, and in this case, the system analyzes the neural response corresponding to this stimulus. These types of systems are very limited in learning from the side of users [22]. As an example of synchronous interfaces, let’s talk about the most famous example of "speller matrix", in which the user must focus his intention on a particular character and count the number of intensifications corresponding to the character.

Counting stimuli at specific moments results in the appearance of event-related potential (ERP) type P300. This interface is synchronous because the ERP signal is generated by a response of the brain to an external visual stimulus.

In synchronous BCIs based systems, users act with the system in voluntary moments without any reaction to any external stimulus. Synchronous BCIs require important training and a self-voluntary control of brain activity [23, 24]. Most of other BCIs based systems are asynchronous because users are free at the time of their voluntary decision except the need of a priori training making them able to activate in any manner their neurons responsible for the realization of the desired actions.

1.2.3. Dependence

In any BCI based system, the brain generates messages and/or neural commands and sends them to the external world without passing through the normal pathways from the cortex to muscles (in down BCIs) passing by the cerebellum and spinal cord using peripheral neurons. In a communication based on ECoG brain signal, the messages of commands are encoded in ECoG brain signal recorded instantaneously doing the desired actions in where the BCI system offers to users alternative methods to act on the external world. BCIs fall into dependent and independent.

Dependent BCI does not use the normal brain output pathway to carry out the neural command, but the activity in these pathways is necessary to generate the brain activity that caries it, or more simply, dependent BCI uses secondary brain signals generated by the brain to transmit a neural command. An example of that is explained in [25, 26], where the user is asked to watch a matrix of letters flashing one by one, and the user selects a specific letter by fixing his gaze directly on it so that the visual evoked potential (VEP) recorded from the scalp on the visual cortex when this letter flashes is much larger than the VEPs produced when other letters flash. In this case, it should be noted that the measured brain signal is generated by the brain (EEG) but its origin depends on the gaze direction and especially on nerves responsible for the activation of extra-ocular muscles. [24]. Brain muscles can also be a source of the recorded EEG.

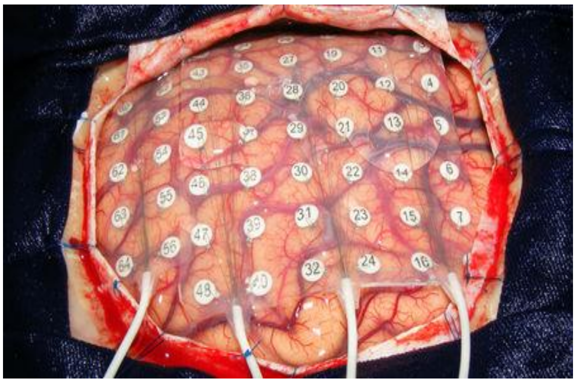
Independent BCIs in contrast, the output is generated by the brain depends solely on the individual's attention and not on nerves and muscles responsible of the gaze direction. The message is not carried by the nerves and/or peripheral muscles and, moreover, the activity in these pathways is not necessary to generate the brain activity (eg EEG) which conveys the message. A similar example of that is seen in the dependent BCI but asks in this case the user to count the number of flashes of a letter on a matrix. The P300 evoked at the individual depends on its intention [27, 28]. Independent BCIs are widely used and more suitable for people with motor disorders [24].

1.3. Architectures

BCIs use brain signals and translate them into desired acts of the patient in a way of assisting, augmenting, and repairing human cognitive or sensory-motor cortex to ensure the communication and make the bridge of information flow possible with the outside world. The design and realization of those ways necessities can be a challenge in technical engineering started from the signal recording to the information extraction and exploitation either the adaptation of the user with the BMI. Developing practical applications of BCIs technology requires being in face of higher technological challenges depending on the application and means. Down brain computer interface consists of five principal components including (1) brain signal recording for reading the inter-brain activities, (2) preprocessing for signal enhancement and noise reduction, (3) spikes sorting and representation to separate and represent all implicated information in a simple and smoothed data easy to be used by a dynamic decoder, (4) decoding information comprising the feature selection with the mapping between the neural signal with the desired actions and finally, (5) the application or the practicing on an artificial robot arm or screen … etc.

1.3.1. Signal recording

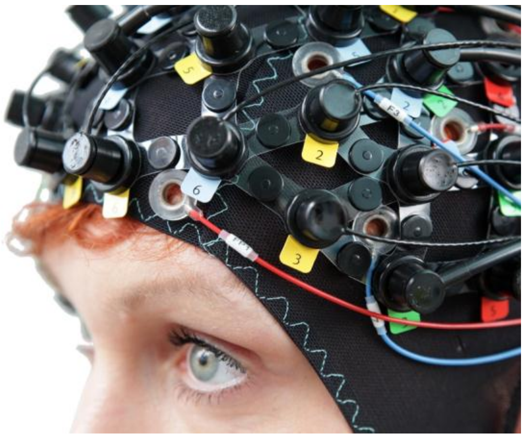
BCIs use brain signals to be translated generally into motor commands or desired actions passing by the five stages of BCI started from reading brain behavior. Measuring brain activities is one of the main components of all BCIs based systems. Brain signal recording is the bridge to the brain-behavior, generates oscillations that refer to currents flow between brain neurons. The first brain signal was recorded by the Austrian Dr. Hans Berger in 1929 and published on a chart paper [9]. Next research shows the possibility of measuring brain signals from different levels on the brain where it is possible to record the (1) EEG (ElectroEncephaloGraphy) from the surface of the head skin, (2) ECoG (ElectroCorticoGraphy) in which sensors are implanted directly on the surface of the cortex under the scalp, this technic is invasive where surgical intervention is necessary to make the direct contact of implants with the surface of the cortex possible, (3) other invasive brain signal recordings are possible deeper the brain providing more spatial and temporal precision on the signal processing, these signals are known LFP (Local Field Potential), MUA (Multi-Unit Activity) and SUA (Single Unit Activity), and other techniques based on the measure the fluxes of electric charge exchange fields between neurons, theses technics are non-invasive we cite the fMRI (functional Magnetic Resonance Imaging) and NIRS (Near InfraRed Spectroscopy).



(a) ECoG

(b) EEG

(c) fMRI



(d) NIRS

Figure 1.2: Brain signals recording

1.3.2. Preprocessing

The aims of the preprocessing are to make next steps easier and possible to be realized using operations on the signal without losing information embodied in the brain signal improving the obtained signal quality, we talk here on SNR (Signal to Noise Ratio). A small or bad SNR means that the information is buried in the rest of the signal which makes relevant information hard and difficult to be detected.

Once the brain signal is recorded, it should be amplified and digitized then subjected to one or more of a variety of features extraction such as temporal and spatial filtering to make the next steps of the BCI based system possible. Filtering the brain signal offers a cleaning by removing muscular and ocular artifacts present in the signal. This cleaning makes the SNR better where the information will be easier for the decoding process. Many technics for the enhancement of the brain signals were applied we cite:

- **Temporal filtering:** these kinds of filters aim to attenuate or remove non-interesting frequencies bands, or more easily temporal filters are designed to enhance the useful frequencies containing the desired information and attenuate other frequencies modulating non-desired information.
- **Spatial filtering:** where in first applied to remove the random fluctuation from the intensity on an image. In Brain Computer Interfaces based systems, spatial filters are necessary for the enhancement of the brain signals where we need to focus the sours of the decoder on such regions of the recording without others, spatial filter makes it possible by attenuating signal comes from the non-interesting regions of the brain.

1.3.3. Spikes sorting and representation

Or single-neuron brain signal separation and representation. It is a classification of all recorded stereotyped electrical pulses named also action potentials or spikes following their shapes. The recorded brain signal is measured using microelectrodes implanted on the cortex (or on the surface of the head) picking up signals of many neighbor neurons of the local region. The spikes sorting, starting from the detection and isolation of spikes and finished by the classification process, offers singles neurons spikes activity which will be used as an input of the decoder after a conventional representation. Spikes' sorting simulates the spatial filters applied to processes of a summation of many sources and separates them into singles sources [29].

Spikes' sorting process delivers trains (series) of spikes that can be represented by ones, separated by long silences of zeroes where at least there will be a respect of relaxing times. This representation can't be used directly as an input of the decoder process and a continuous representation is necessary by using a convolution of these series with sliding windows.

Other researches pass directly to the decoding stages without spikes sorting and representation but they use a set of single thresholds across all channels and fitting the resultant events with a spline tuning function, a control signal was extracted from this population using a Bayesian particle-filter extraction algorithm [30] or obtaining spike train collected at recording electrodes from thresholding the band-passed voltage signal [31].

1.3.4. Decoding of information

When brain signals are recorded and separated into spikes trains, a special representation is necessary to the mapping related the brain event to the desired actions or any intent of the user. Decoding brain signals is a challenge because of the non-linearity of the relationship between the desired actions (and/or intention) with the extracted signal from the most complex part of the body. The mapping, in general, is realized after the preprocessing, features extraction, and detection and classification, and sometimes data reduction. This stage is called also the mapping or the translation algorithms that convert the brain signal or the firing rate's amplitude that translates the density of spikes into dependent variables of any kind of desired action or intention of the user.

The decoding process contains in general two parts, feature selection and data mapping related the command to brain intentions hidden in the brain signals. An algorithm of translation can be judged by its success which is determined by the correctness of its features selection, and by an effective translation of the mapping output in device control. A good feature selection makes the translation algorithm perfect and easier for training; feature selection can be based on standard guidelines mentioning the location or the region from where the brain signal should be recorded and the spatial and temporal frequencies separation. These strategies are enhanced by operator inspection of initial topographical and spectral data from each user [32, 33, 34]. Mapping algorithms might use linear methods for the decoding problems using classical statistical analyses like in [35] or nonlinear methods as in [2, 6] where authors used the time-delay neural network to decode the rat's hand trajectory.

1.3.5. Practicing

Results of four past stages of BCIs based system are the same in most of BCIs and finally, the application can be different depending on the user needs, we can find the practicing on a computer when users try to command a cursor on a screen for the pursuit and the selection of icons, letters or any other target on the screen [28, 36, 37]. Other applications use the brain signal in order to repair functional motor disabilities for handicaps and the output of the last stages of BCIs based system will be used as a command of an artificial arm or other systems as in our work this signal open water to the rat. We can find many other applications and we summarize them in reading the brain intentions.

Many protocols are applied for guiding the BCIs processes where they should take into account the time of work of these systems and moments of starting and stopping the work and the kind of continuous or discontinuous communication.

1.4. Applications

In this part is presented details on the applications of the brain computer interfaces with their types (but we concentrated mostly on the down BMIs) with used maps of the signal through. We first talk about the realization stared from the information processing to the brain introducing or reading, depending on the way of the used BCI. As known on the up BCI is a way for reading information from the external world and try to introduce it in a manner or another to known signal by the brain which is called the coding of information. The most important questions that can be asked in this case are: (i) how can the brain accept and (ii) be able to be adapted to the new input and (iii) in where should we introduce it? The answer of these three questions is situated in the brain's abilities to manage its inputs and outputs. Medical and biological studies improve that the brain is separated into regions and each one is responsible for the control of a member (Homunculus) until recent research where authors can detect precised points on the cortex in direct relation to the control of muscles.

The down BMIs based systems used the brain signal (state) to decode an internal decision of a motor command or simple act and try to restore the disorder in the action responsible members by using artificial arms in case of amputation and ensure the command transfer generally in the case of spinal cord (and/or neural fibers) lesion or injury. The down BMIs can be summarized in three stages stared from the recording of the brain signals, the processing level in where our work is focused on a brain signal of a rat and know the time of the desired action on the lever to bring water. The final step is the realization of the action using artificial elements to deliver the need of the brain like bringing water to the rat in our work.

BCIs can't be used only in medical applications to restore motor disordered actions like an artificial arm control and a wheelchair movement, cochlear for hearing repair, and bionic retina for vision. We can find BCIs in many beyond medical applications which can be described as a challenge of the BCIs applications. In these cases, BCIs can improve the quality of life allowing the control of computers and electronic devices prepared specially for the application. We can also talk about social networks and spelling devices [28] and controlling computer games [38], navigating a computer mouse on a screen [39]. All these applications need special signal processing steps for coding and/or decoding the information to and/or from a brain signal respectively.

In [40] are summarized the medical and non-medical applications of BCIs based systems.

1.4.1. Medical Applications

Brain computer interfaces have contributed to a wide beach of medical applications depending on the need of the patients. The healthcare field has various applications that may take advantages of brain signals in all linked stages including prevention, detection, diagnosis, rehabilitation, and restoration.

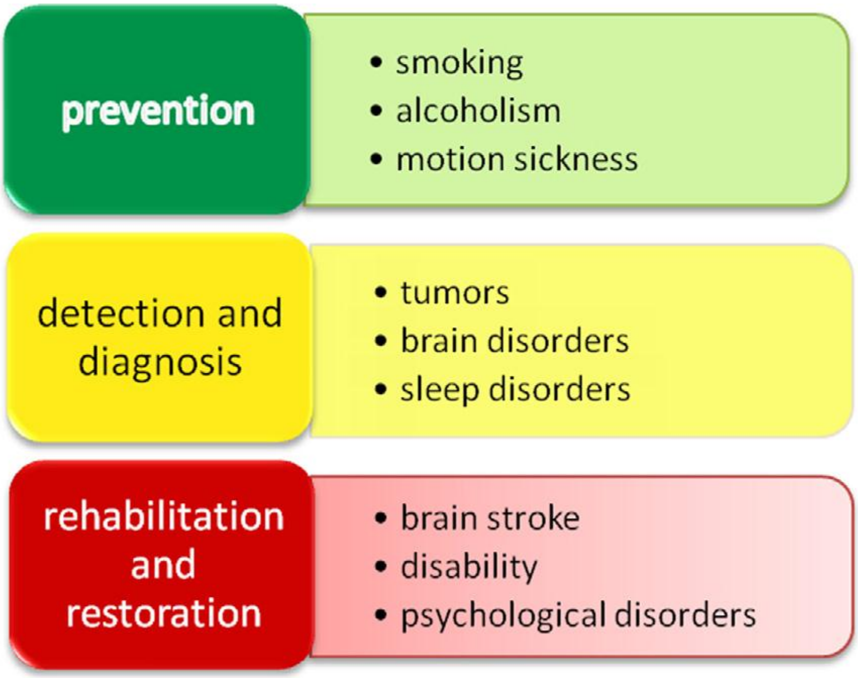


Figure 1.3: Medical applications of BCIs [40].

1.4.1.1. Prevention

Various systems for determining the level of consciousness as well as their studies related to the brain have been developed. The attention influences of smoking and alcohol on brain waves have been illuminated in [41, 42]. The degree of importance of any medical prevention studies lies in the possibility of losing function and decreasing the level of alertness resulting from smoking and/or alcohol consumption, while the authors of [43] studied parts of the brains most susceptible to alcoholism. Road accidents are considered to be the leading cause of death or serious injury, as noted in [44]. The analysis of the causes for further prevention is becoming a concern for research in various fields. Thus, the level of concentration of people with motion sickness, and especially drivers, was studied. Motion sickness, which happens as a result of the sending of contradictory sensory information comes from the body, inside ear, and eyes to the brain, usually occurs on moving transport media.

1.4.1.2. Detection and diagnosis

The BCI systems mental state monitoring function has also aided in the prediction and detection of health problems such as abnormal brain structure (such as brain tumor), seizure disorder (such as epilepsy), a sleep disorder (such as narcolepsy), and swelling of the brain (such as encephalitis). Tumors generated from uncontrolled self-division of cells could be diagnosed using EEG as a secondary alternative for MRI and CT-SCAN.

Detection systems of brain tumors based on EEG have been the main subject in [45, 46]. Authors in [47] have been interested in the identification of breast cancer using the EEG. Sharanreddy and Kulkarni, in [48], proposed a system that recognizes EEG abnormalities associated with brain tumors and epileptic seizures. The early detection of epileptic seizures is one of the most common neurological disorders, and the monitoring of its effects is presented [49].

1.4.1.3. Rehabilitation and restoration

Hand trajectory decoding from brain signals is the aim of this work and it can be classified in the physical and motor rehabilitation and restoration applied on the rat for bringing water which is similar and maybe more difficult with the human user. The purpose of the BCIs used for motor rehabilitation is designed for (i) people suffering from motor disorders where we know that more than 230 million people in the world are suffering from handicaps and the number keeps growing rapidly. Motor neural prosthetics are designed to restore sensory-motor functions lost due to an injury or a disease by activating paralyzed muscles or move artificial arms using brain signals to know the desired action. (ii) On the other side, rehabilitation by introducing an external signal reconstructed from information (sound and images) from the external world of the body and translated into a brain signal. It can be seen easily in the cochlear and the bionic retina.

1.4.2. Non-Medical Applications

BCI was firstly desired for medical applications, but actually, it covers a wide beach of applications that can be useful for those who don't seek medical uses. In [50, 40] are described several high-level non-medical applications of BMI with examples. The rate of merits can be calculated using the five following criteria: (1) the quality and quantity of relevant research, (2) the ability of technology contribution to solve societal challenges, (3) the number of potential applications and users (economic), (4) price sensitivity and (5) the time to market. Non-medical applications can be summarized in:

a- *Neuro-ergonomics and smart environment:* Smart homes, workplaces or transport, and all smart environments could also harness brain-computer interfaces to provide more security, luxury, and physiological control to human everyday life using internet of things [51]. in [52], authors proposed a self-adjusting control system of the intelligent living environment based on the Brain interface. It is a cognitive controller system. It observes the user's mentality and adapts the nearby components accordingly. It has prolonged its functionality with the participation of the Universal Plug and Play home network. On the other side, they take into account the nearby environmental influence in the enhancement of domestic applications based on BCI through the consciousness of the context. In [53], authors developed an application that modifies automatically, depending on the current context, the available options accessible by the user. Brain signals also help improve working conditions by assessing the cognitive state of an operator [54].

b- *Neuro-marketing and advertisement:* The field of marketing has also been of interest for BCI research. Research in [55] has explained the benefits of using EEG assessment for television commercials related to both commercial and political fields. BCI-based valuation measures the attention generated accompanying observational activity [56]. On the other hand, the researchers of [57] considered the impact of another cognitive function in the field of neuro-marketing. They were interested in estimating the memorization of television commercials, thus providing another technique of advertising assessment.

c- *Educational and self-regulation:* brain signals can improve the internal state of the user and can be an initial factor to measure the degree of clearness of a study. In [58] authors present the use of BCIs based systems in an educational context where they used the EEG in identifying the workload of a learner which is a crucial factor of successful learning. BCIs based systems record and process brain activity allowing the individual whose activity is recorded to monitor this activity at the same time. Assistance units are included in BCI applications for severely paralyzed patients to help them in control or communication with external devices, as well as brain biofeedback to self-regulation of brain activities for treatment of epilepsy, attention deficit hyperactivity disorder (ADHD), anxiety, and other psychiatric conditions to improve cognitive performance in healthy individuals.

d- *Gaming and entertainment:* BCIs was first designed to applications devoted to disabled peoples. Development and progress of BCIs research make healthy users look for other demands as devices with a limit of physical movement which require a very long time of training, for that we found both of gaming and entertainment opened the market of BCIs based systems and prototypes of BCIs appear in gaming and entertainment that adapt and offer users the ability to control games using brain signals. In [59] authors talked about control of fly of helicopters to any point in 2 and 3D virtual world.

e- *Security and authentication:* Authentication is a process used to identify users, security systems include knowledge based on biometric authentication using (i) morphological identifiers like the fingerprints, hand and finger shape, vein pattern, the eye (iris and retina), and face’s shape, or (ii) biological analyses of DNA, blood, saliva, and urine. These identifiers sometimes can’t be sufficient for security authentication and research derived to explore other factors based on the information hidden in brain waves [60].

1.5. BMI Realization

The BMIs based system design and realization take into account the direction of the information from where recorded and in where will be exploited. As our work is focused on the down brain computer interface in which we hope to make a controlled device (arm) by brain signals.

1.5.1. Up BMI realization

The up BMI is generally summarized in the cochlear or bionic retina. It can be realized by reading external signals which can be a sound or images and translated into brain signals. Complications are situated in the signal introduction into the brain where users are not adapted to the new input in their brain and need rehabilitation for which takes a real long time. The map of an up BMI knows three stages started from reading the external information (sound, video, …etc) and translate the recorded signal into brain signals (spikes) in a form of action potentials with a good priori knowledge of the ear (for cochlear) and eye (for bionic retina) functions or models simulated their map. The final step is the generated signal introduction into the brain by surgical operation.

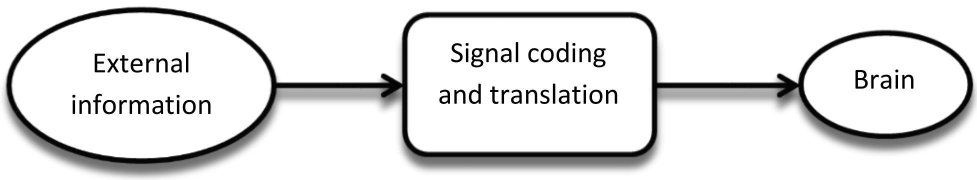


Figure 1.4: Up BCI map

1.5.2. Down BMI realization

In the down BMI (our topic), the aim is to read a brain signal from the temporal cortex (responsible of arms movement region on the brain) and translate it into desired position of the arm terminal. Decoding and realizing desired action requires several intensive tasks stared from the signal recording to the action accomplishment. Recording brain signals can be realized in different ways on the human (and/or rat) head depending on the invasiveness. One used electrode can record the neural activity of such number of neighbor neurons and the recorded signal is given by the sum of the flowed currents, a separation of this channel into several channels (one channel for each neighbor neuron) called spikes sorting is used to simplify the complexity of the brain signal. After spikes sorting, the obtained signal is composed of trains of long silences separated spikes. Information is embodied in the distance of time between spikes so a special representation is required to simplify the work of the decoding algorithm, this representation is called firing rate and can be realized by convolving the channel with a sliding window. The next step is the decoding problem named also the mapping between the brain activities with the desired action in where many highly nonlinear technics are applied. The final step is the realization in real-time of the desired action by a device with a feedback for correction.

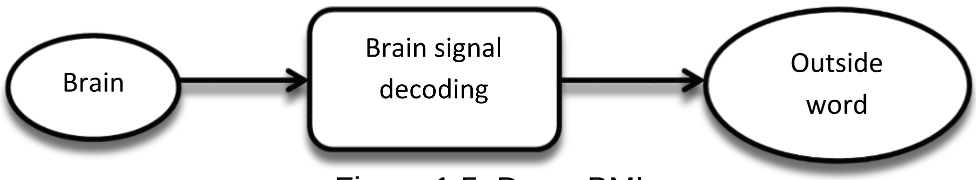


Figure 1.5: Down BMI map

1.6. Conclusion

In this chapter, we presented the most important descriptions and applications of brain computer interface based systems started from definitions of BCIs and ideas and means of realizations of any communication pathway between the brain and external world from residing brain-behavior to signal processing and decoding and to the final step of application on any device (arm, screen, …), next we presented the medical and non-medical applications. We finished with the realization of BCIs based systems.

2. BMI SIGNALS

2.1. Introduction

When it comes to the brain signals in this chapter, it is better to understand the brain anatomy with the neural communication to know better why and how we can record the brain signals; next we'll present the origins and relationship between the brain signals and the used technique for recording depending on the invasiveness. Development of technology and biology specialty anatomy from medicine offers the possibility of recording single-neuron activity which makes the recorded signal more precise in temporal and spatial.

2.2. Central Nervous System

The alternate contractions of muscles allow the movement of the ensemble of the body parts and also the body in the space, where yet studies in neurophysiology have shown that the execution of a movement is the result of muscular activity of all the body segments. These acts are organized in general into two kinds of movement: (i) reflex which is an action without the decision of the brain, and (ii) voluntary movement which requires the activation of the motor cerebral areas of the cerebral cortex by different mechanisms of initiation and programming. The action potential released by the motor area will spread by different paths to the muscles throughout this journey from the cerebral cortex to muscles; different structures will control the motor nervous impulses of the voluntary movements. In our work, we are interested in a voluntary movement where the brain is the first responsible of actions by activating such a group of motor neurons in a special behavior to realize a precise action. So we have to know more about the central nervous system and its components next.

The central nervous system (CNS) is made up of the brain and spinal cord; it is vital to human (and/or animal) life and functioning. The CNS maintains the fastest way for the body communication ensured by the transfer of electric fields (action potentials) from the brain to the other parts of the body for command and vice-versa for a sensation. Compared to the endocrine way, ensured by the transfer of information and commands by chemical fields (hormones), it is an instantaneous way for direct commands. At the level of the CNS are processed all sensations and thought that the human experiences. In BCIs based systems, it is important to know better CNS anatomy with a special view and analysis of the brain function and architecture. Next, we will present the brain and its architecture, function, and neurons which is its principal component:

2.2.1. Brain (Cerebral Cortex)

The brain (cerebral cortex) is the most important part of the human (also animals) body; it represents the main part of the nervous system which allows the control and movement of parts, feeling, thinking …, by processing, integrating, and coordinating the information received from organs and making decisions as instructions to the rest of the body. The brain is the most complex organ of the body, it is made up of several billion neural cells to transmit and receive messages from different parts of the body, and these cells are neurons. The brain is protected within the skull bones of the head [61].

The brain is made up of three principal specialized parts that working together:

- *Cerebrum:* is the largest part of the brain. The Cerebrum can be divided into two hemispheres; left and right; joined by a bundle of fibers that transmits messages between sides called the corpus callosum. The right hemisphere controls the left or the opposite side of the body and the same for the left hemisphere with its opposite side of the body, if any stroke occurs on the left hemisphere, it may make a weak or paralyze on right arm or leg. The entire outermost layer of the brain is made up of a thin layer of nerve tissue called the cerebral cortex. Each Hemisphere can be divided into four distinct fissures called lobes, we cite the frontal lobe responsible for voluntary motor coordination and thinking, the parietal lobe which assimilates information coming from the senses relating to touch and memory, the occipital lobe is located on the back of the head, it represents the visual center, that is to say, it makes it possible to recognize the orientations and contours of images and finally the temporal lobe which plays an important role in several cognitive processes. It contains the centers of hearing, taste, and memory [62].

- *Cerebellum:* is located under the cerebrum in the base at the back of the brain, it plays an important role in motor control, coordination and balance. It may also be involved in some cognitive functions such as attention, language, and emotions like the regulation of pleasure and fear responses [63].

- *Brainstem (or brain stem):* is the posterior part of the brain, it acts as a relay center connecting the other two parts (cerebrum and cerebellum) with the spinal cord. It is also the responsible for automatic functions in the body such as breathing, heart rate, regulation of the body temperature, sleeping and waking up cycles, etc.

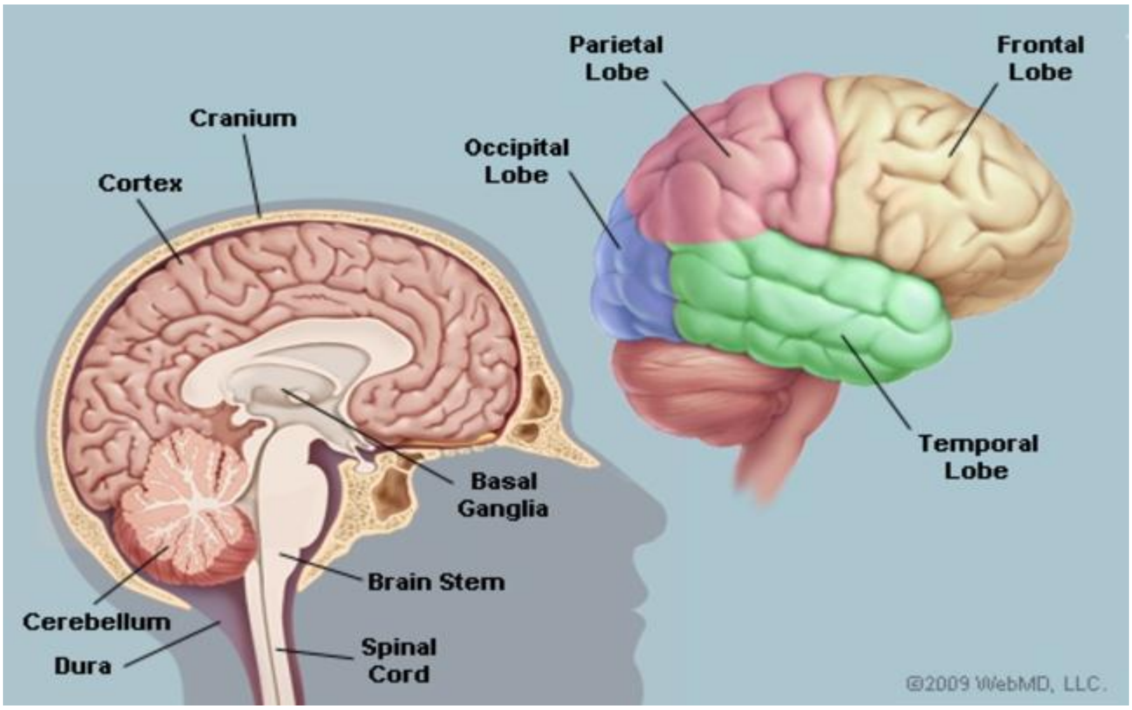


Figure 2.1: Brain architecture

2.2.2. Neurons

Human brain is made up of several billions of nervous cells called neurons. Neurons (or nerves) are the basic working biological unit (cells) of any nervous system; these cells are electrically excitable to provide communication with other cells through synapses by transmitting action potentials. Neurons can be classified into three types depending on their functions. (i) Sensory neurons responding to any stimuli affect sensory organs like sound on the ear, touch, and heat on skin or light on eyes. Once the stimulus is realized, these cells send the signal to the brain to decode the quality of the outcome stimuli. (ii) Motor neurons transmit neural commands from the brain; directly in the level of the head or through the spinal cord on other parts of the body, and expose it to the muscles to realize the desired action by contractions. (iii) Interneurons are neurons responsible to ensure the communications between two regions and generally between sensory and motor regions.

Typically, the neuron is composed of the soma (or the body of the neuron), dendrites, and axons. The soma is the control center of the neuron, which makes the sum of the information which reaches it. It then processes the information and sends the result back in the form of electrical signals (action potential), from the cell body to the input of other neurons through its axon. The axon is the responsible for connections between each neuron to others. The functions of dendrites are to receive signals from other neurons and especially from synapses of other neurons, to process these signals, and to transfer the information to the soma of the neuron.

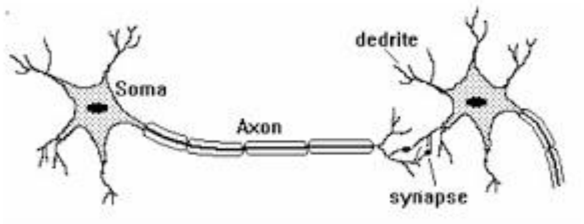


Figure 2.2: Principal neuron structure with its main components.

2.2.3. Brain and Neuron communication

Neurons work together or at least by dependent groups of neurons to each other for processing until accomplishing any task of the brain on the body. Neurons realize links between them and form chains of neurons constituting fantastic networks by receiving and transmitting signals of electrochemical nature through dendrites and axons. Dendrites receive the chemical fields that represent the electrical signal which will be transmitted by axons. Each neuron can receive information from a number of neurons and transmit it to a number of neurons that go to 250 000 connections. At the level of synapses, the information can be transmitted from a neuron to another. The membranes of neurons "in communication" are close to each other but separated by a synaptic space. Since electrical signals cannot cross this space, communication between two neurons takes place through chemicals called neurotransmitters. These substances differ depending on the synapses. They are produced by the neuron which sends the message and recognized by the neuron which receives it. In the upper level of the brain, communication can be realized between regions or lobes through interneurons.

2.2.4. Movement physiology

The origin of voluntary movement for any animal (or man) involves several nervous and motor units. In general, for a contraction or extension of an arm; whether hand or foot (muscle); the decision is made at the level of the brain and exactly from the thalamus to the temporal brain, at this level the information is very complex and difficult to acquire or process. This information will be transferred to the cortex (surface of the brain) and each part of the body is controlled by a region of the cerebral cortex (for a human there is the Homunculus to describe the regions responsible for parts of the body). Through pyramidal neurons, the information can be transmitted to the muscle passed by the cerebellum and the spinal cord to be realized [64].

.

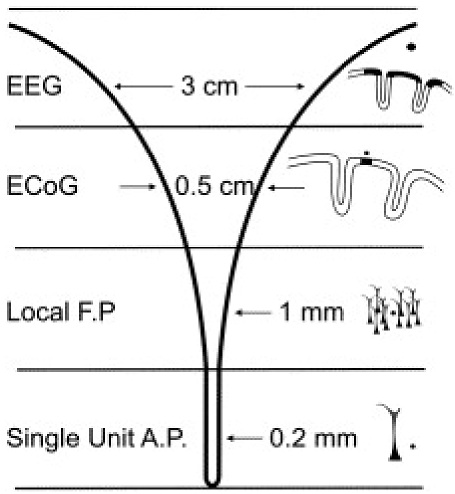


Figure 2.3: Levels of electrical brain signal recording with spatial precisions.

2.3. Brain Signal Recording

Brain computer interfaces use brain signals recorded in different manners including electric and magnetic fields; functional near-infrared imaging (fNIR); positron emission tomography (PET); and functional magnetic resonance imaging (fMRI) [20]. Brain signal recording is the bridge to brain-behavior which provides accessibility to the neural activity recording. In reality, most of the used methods for recording are not practical for clinical use because of their intricate technical demands, prohibitive expense, and limit of real-time capabilities. Actually, the most used is the electric field recording (sometimes fNIR) [65].

Electric field recording measures the flow of current charges or densities in/on the membrane of synapses, neurons, and axons of the central nervous system when receiving and transmitting information with neighbors. When a neuron receives an electric charge; on dendrites; it integrates it on the synaptic inputs and transmits it throw its axon to the soma in a form of an action potential. Synaptic and action potentials make changes in the flow and density of ions across and the two sides of the membrane. The produced electric field, in this case, can be measured capturing a voltage fluctuations as EEG (ElectroEncephaloGraphy) on the scalp, ECoG (ElectroCorticoGraphy) on the cortical surface, or in an intra-cortical neuron recording inside the brain which can be separated into three levels LFPs (local field potentials), SUA (Single-Unit

Activity) and MUA (Multi-Unit Activity). In figure 2.3 is illustrated the electric field measurements on the brain levels.

The biggest problem with the invasive techniques is to obtaining a stable long-term recording after the implementation of microelectrodes; this has to be adapted with the immune response of the brain tissues. Where, after the insertion of the probes, the body forms protective scar tissues that limit the recording quality. Over time, the signal-to-noise ratio of the measured signal becomes too low in the extraction of reliable information. Figure 2.4 shows the response to inserted microprobes over 12 weeks. Many research groups are focused on overcoming this problem including investigations of different materials and shapes of electrodes and improving the surgical procedures of implantation [66, 67].

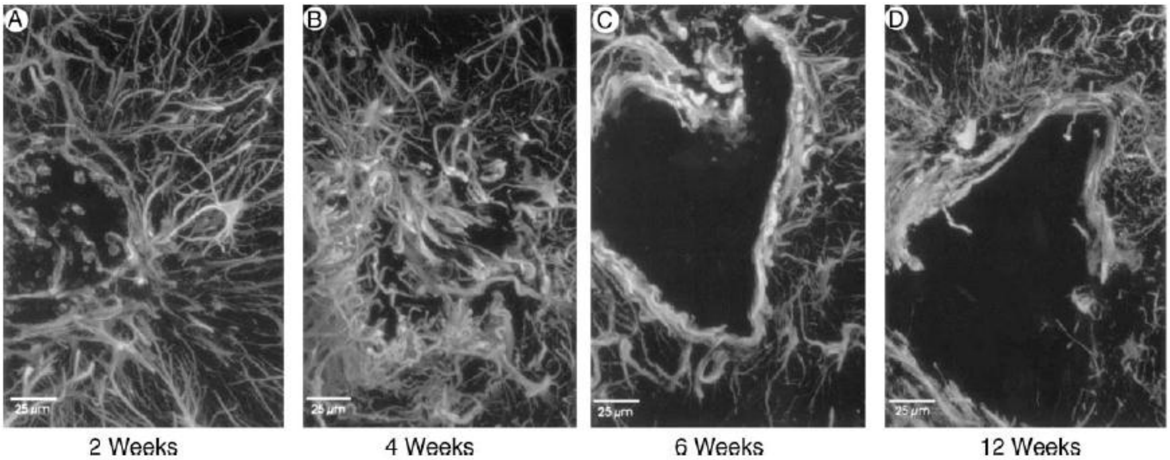


Figure 2.4: Tissue response over time (12 weeks) [67].

In this part, one presents and discusses the most existing signals in BCI in its electrical form and magnetic form.

2.3.1. EEG (or ElectroEncephaloGraphy)

In this case, users measure the electrical brain activity along the scalp generated by the voltage fluctuations associated with the neural transmission activity in the brain caused by the flow of electric currents during the synaptic excitation of dendrites. Sometimes EEG looks the best signal if we don't talk about its disadvantages where the suitable words for this 'why should we measure the velocity of each molecule in a room to determine the temperature of the room?'

EEG is recorded with electrodes placed on the scalp without surgery for that this method was used in clinical and research for decades. This technic of measure has unique usability advantages over other types of brain signal recording that recommend it for commercial use. It is the easiest to use, expensive and portable. EEG recording provides high temporal resolution. However, its signal-to-noise ratio and spatial resolution represent a limitation compared to other methods.

2.3.2. ECoG (or ElectroCorticoGraphy)

The ECoG is an electrical technique for measuring brain activity, it signifies the measure of the brain signal as an activity from the cerebral surface using direct micro-electrodes placed on the surface of the cortex, electrodes are implanted inside the skull but outside the gray matter. ECoG is classified with invasive (sometimes classified in partially invasive) technics of brain signal recording which requires a craniotomy for the implementation of the electrodes. This recording offers an advanced temporal and spatial resolution with high amplitude and low vulnerability to artifacts such as blinks and eye movement. The first invasive experiment of brain signal recording was performed in 1992 using ECoG for a BCI framework mainly with animals. Early studies on animals assessed for long-term stability of the signals from regions where the ECoG is acquired and electrodes provide a stable signal over several months. ECoG can be classified as one of the best signals that can be used in BCIs based systems, it provides high-level stability of the signal for long-term over several months with good spatial and temporal stability if we don’t talk about the invasiveness. We actually find the use of ECoG in many research in decoding brain intentions and motor decisions as in controlling of cursor or hand trajectory decoding.

2.3.3. Intracortical Neuron Recording

In this case, we talk about a group of three similar neuroimaging techniques based on the measure of electrical activity inside the brain and especially in the level of the grey matter. These techniques are invasive that require the implementation of an electrode (or array of electrodes) inside the grey matter of the brain to measure the spikes and local potential from neurons. We can describe the SUA (Single-Unit Activity) which can be obtained by applying high-pass filtering of frequency over to 300Hz on the signal of a single neuron, the MUA

(Multi-Unit Activity) which is obtained using the same technique of SUA but applied on a group of neurons and the LFP (Local Field Potential) is extracted with a frequency less than 300Hz (low-pass filtering) in the vicinity of the electrode tip [68].

Intracortical neuron recording provides the highest spatial and temporal resolution compared to the other existing techniques of brain signal recording. Like all techniques, the intracortical neuron recording has its disadvantages situated in the invasiveness and the reaction of the neighboring tissues of the brain to the microelectrodes.

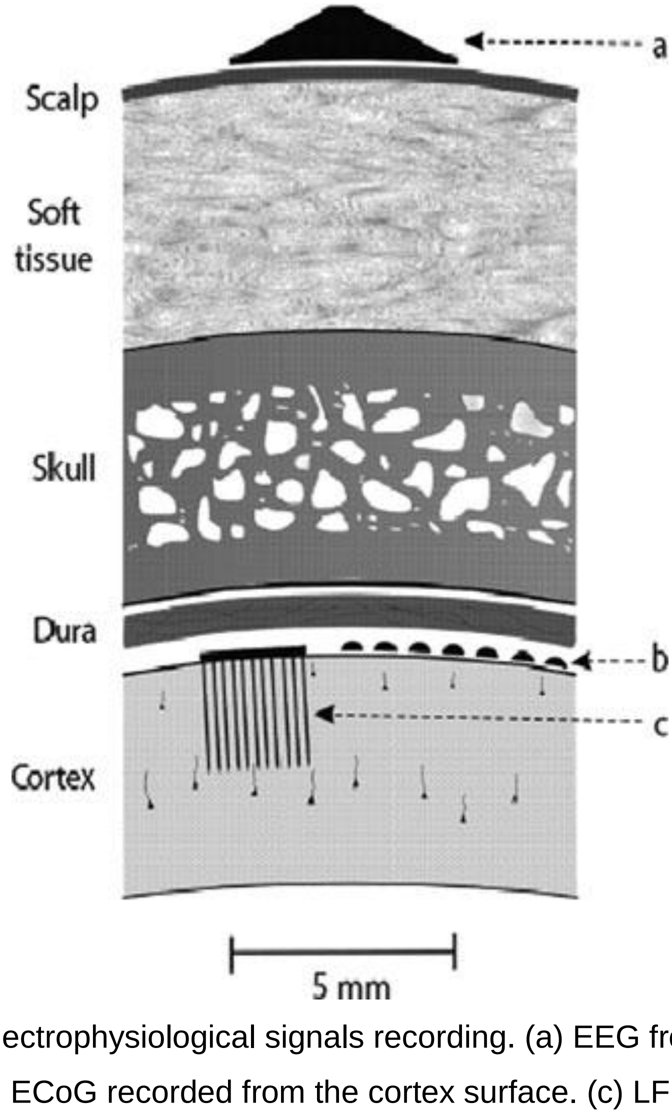


Figure 2.5: Electrophysiological signals recording. (a) EEG from the scalp. (b) ECoG recorded from the cortex surface. (c) LFPs

2.3.4. MEG (or MagnetoEncephaloGraphy)

This signal comes from the brain's magnetic activity by means of magnetic inductions and it is recorded non-invasively by the imaging technique. The MEG signal is similar to the EEG signal in their neurophysiological processes of production. Nevertheless, while MEG is extremely more sensitive to the primary current sources, the EEG is still sensitive to those of the secondary currents [10]. Sensors used to record the MEG should be implanted outside of the head providing the measure of magnetic fields produced by the intracellular currents flowing inside dendrites [47]. Superconducting quantum interference devices are used to detect magnetic fields. These devices are extremely sensitive to magnetic disturbances produced by neuronal activity [49]. The MEG signal is difficult because of many factors depending on the sensitivity of sensors to the magnetic field come from external sources which require their implementation inside a magnetically shielded room and make the recording process so big compared to the size of the head, and can't be used in outside a laboratory. Another problem of the use of MEG in BMIs based systems situated in the requirement of superconductivity of sensors which should be cooled to nearly -273 degrees Celsius. Also, the technology of MEG is bulky and expensive to be used for longtime and researches are deviated to other technics. With these difficulties of using the MEG signal in BCIs, it is still one of the best signals because it offers a higher spatiotemporal resolution compared to EEG, and makes it able to detect and localize active regions on the brain. The decoding problem using MEG becomes easier with reduced training time [50].

2.3.5. fMRI (Functional Magnetic Resonance Imaging)

The fMRI is a non-invasive technic classified with neuroimaging methods that are associated with the changes in the local cerebral blood volume and flow, oxygenation levels during neural excitation by the mean of the electromagnetic field. Typically, BCI systems with fMRI as brain signal measure the Blood Oxygen Level Dependent (BOLD) during neural activation but with no direct relationship but a correspondence between both does exist. As all recording technics, fMRI has its advantages situated in the high spatial resolution and its ability to localization of active regions inside the brain [69]. However, the use of fMRI presents disadvantages as the low temporal resolution of about 1 to 2 seconds, and in addition, the hemodynamic response introduces a physiological delay of about 3 to 6 seconds [70]. Actually, non-clinical fMRI applications are not expected because fMRI requires overly bulky and expensive hardware.

2.3.6. NIRS (Near InfraRed Spectroscopy)

NIRS is a non-invasive brain data recording method based on optical spectroscopy that uses infrared light to illustrate fluctuations acquired metabolism in the brain during neural activity. The main idea of NIRS is to measure the attenuations of light intensities with alterations in oxyhemoglobin and deoxyhemoglobin concentrations using infrared light that penetrates the skull to a depth of 1 to 3 cm.

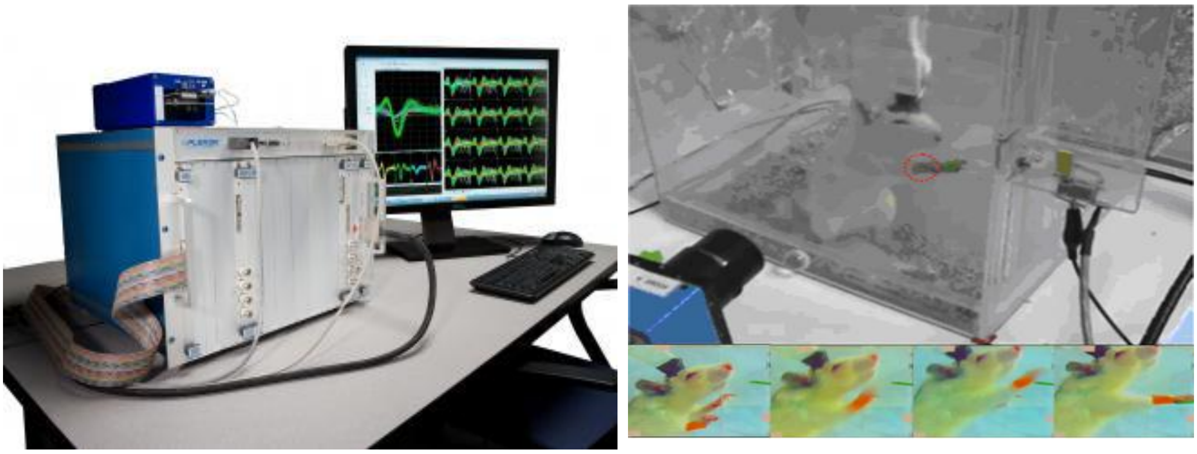
Due to the low penetration of light into the brain, this neuroimaging optical technique is limited to the superficial cortical layer. Similar to fMRI, one of the major limitations of NIRS is the nature of the hemodynamic response, as vascular changes occur, a number of seconds after its associated neuronal activity. Another disadvantage of the use of NIRS in BCIs systems is the spatial resolution which is quite low in an order of 1cm.

2.4. Experiment and data recording

The study approved by the Institutional Animal Care and Use Committee at the National Chiao Tung University was conducted according to the standards established in the Guide for the Care and Use of Laboratory Animals (GCULA). Four male Wistar rats weighing $250 - 300\ g$ (BioLASCO Taiwan Corp., Ltd.) were individually housed on a 12h light/dark cycle, with access to food and water *ad libitum.*

Dataset was recorded from the motor cortex of awaked animal performing a simple reward task. In this task, male rats (BioLACO Taiwan Corp., Ltd.) were trained to press a lever to initiate a trial in return for a water reward. The animals were water restricted 8-hours/day during training and recording sessions but the food was always provided to the animal *ad lib* every day.

Data recording: The animals were anesthetized with *pentobarbital* (50 mg/kg i.p.) and placed on a standard stereotaxic apparatus (Model 9000, David Kopf, USA). The dura was retracted carefully before the electrode array was implanted. The pairs of 8 microwire electrode arrays (no.15140/13848, 50m in diameter; California Fine Wire Co., USA) are implanted into the layer V of the primary motor cortex (M1). The area related to forelimb movement is located anterior 2-4 mm and lateral 2-4 mm to Bregma. After implantation, the exposed brain should be sealed with dental acrylic and a recovery time of a week is needed. During the recording sessions, the animal was free to move within the behavior task box (30 cm_30 cm_ 60 cm). During the recording sessions, the animal was free to move within the behavior task box (30 cm_30 cm_ 60 cm), where rats only pressed the lever via the right forelimb for receiving 1-ml water reward as shown in Figure. 2.6 (a). A Multi-Channel Acquisition Processor (MAP, Plexon Inc., USA) was used to record neural signals. The recorded neural signals were transmitted from the head-stage to an amplifier, through a band-pass filter (spike preamp filter: 450-5 kHz; gain: 15,000- 20,000), and sampled at 40 kHz per channel as depicted in Figure 2.6 (b). Simultaneously, the animal's behavior was recorded by the video tracking system (CinePlex, Plex Inc., USA) and examined to ensure that it was consistent for all trials included in a given analysis [71].



(a) (b)

Figure 2.6: Protocol of ECoG collection and recording by invasive microelectrodes implanted on the primary motor cortex from a free rat, (a) MAP, Plexon for rat's brain signal recording and (b) Free moving rat when recording ECoG.

The neural activity during a daily research session was commonly recorded in a series of short epochs lasting 2-6 minutes during which the subjects performed one behavioral task, interleaved by resting periods when no activity was recorded.

2.5. Conclusion

The brain computer interface measures the brain activity generated by neural communication. We presented the origins of brain signals for down BMI based on two main techniques depending on the invasiveness. Sensors can detect different changes in electrical and magnetic activity in real-time over different areas of the brain. Invasive brain signal recording is based on measuring local activity with a high temporal and spatial resolution for that the ECoG is the best used for BCIs based systems. The used database is the ECoG recorded from the rat’s primary motor cortex of the rat pushing a lever to bring water.

3. SPIKES SORTING

3.1. Introduction

As we have seen before in the BCIs systems, spikes detection and sorting appear as the first step before the decoding process. Spikes detection and sorting is the main way to single-neuron activity monitoring and processing before all studies to understand the brain mechanism. It contains three principal levels where the signal in first should be prepared in the preprocessing stage of the BCI architecture. In this chapter is presented the detection of spikes done to locate their times of the action potential from neighbor neurons and spikes here should be isolated. The feature extraction is done by using PCA (Principal Component Analyses). We proposed the clustering by using the self-training SVM (Support Vector Machine) [1] applied on other data. In the SVM, data came from 30' multiunit recording from a human epileptic patient. This data was collected at the lab of Itzhak Fried at UCLA, using a Neuralynx System (Tucson Arizona). The spikes have been detected and also clustered using the Waveclus toolbox by the University of Leicester and this clustering has been accepted as our ground truth [72]. In the last part of this chapter in presented another online technic based on template matching. We talk here on the brain signal of the rat for the online spikes detection and sorting started by an online detection using an energy operator and a variable cluster based on the template matching for spikes identification and classification. All data were recorded and band-pass of (400 to 3 kHz) filtered.

3.2. Spikes and Firing

Firing action potential signal is the way of communication in the brain neurons and from the brain with other parts of the body, it is assisted by the transfer of brief electrical charges from transmitter neurons to receivers, after the processing receivers become a transmitter to others through their axons. The brief voltage (electrical charge) of the firing action potential takes the form of quick impulses called spikes. Electro physiologists have already been recording the brain signal by the use of microelectrodes implanted on the head to record the EEG or on the surface of the brain or cortex to record the ECoG and either deeper the brain to get LFPs signals. Microelectrodes record the spikes of many neighbor neurons or a group of neurons in different places thus many forms of spikes of the local region of neurons can be recorded simultaneously superposed with the addition of current flow.

3.3. Spikes detection and sorting

The spikes detection and sorting (mostly called spikes sorting) regroup three principal stages functioning in the order where the output of the first passes to the third through the second. The aim of the spike sorting stage in most down BCIs based systems is to make the decoding step easier by following brain signals individually for every single neuron. Another advantage of spikes sorting is the data reduction by the substitution of irrelevant and redundant channels from databases.

- ***a- Spikes detection:*** The detection process consists of localizing the spike’s maximal energy to make it possible to separate spikes from background noises.
- ***b- Feature extraction:*** extraction of features from spike is the realization of a vector containing characterizations of the spike shape as the spike heights and widths or from the Principal Component Analyses (PCA) in two or three dimensions (first and second components). In this case, each spike is represented and processed using its corresponding vector of features.
- ***c- Clustering:*** this stage is the main of the spikes sorting where the output of this stage is a group of channels containing the spikes of each detected neuron.

3.4. Spikes Detection

Spikes sorting can’t be realized without detection and isolation. The detection step involves in general two main operations of emphasis and comparison to a threshold. The quality of spike detection and sorting depends on how the brain signal is emphasized and how the threshold is determined. The spikes detection called also extraction and localization of spikes for sorting which requires the separation of shapes from neural activity in a noisy signal. The detection function is a nonlinear time-varying filter and reports the evidence of spikes position. By applying an energy operator on the brain signal, values over the chosen threshold give the spikes locations.

The amplitude of spikes shapes is sufficiently larger if we compare it to noises, thus the application of an appropriate energy operator makes it easy to locate the spike's position only by the use of a fitting threshold. Spikes detection and isolation regroup three main steps. Emphasis spikes by the use of the NEO (Nonlinear Energy Operator) or TEO (Teager Energy operator) to emphasize the difference between spikes and noise [73]. The TEO was first applied in acoustic for AM and FM signal processing. Actually, TEO takes a strange place in biomedical engineering such as in brain signal processing for spikes detection. The formulation of TEO is driven by a second-order differential equation [74]. The total energy of oscillation can be obtained by equation 3.1:

$$E = \frac{1}{2}kx^2 + \frac{1}{2}mx^2 \tag{3.1}$$

Where m is the mass of the oscillation body and k is the spring constant. The Nonlinear Energy operator can be defined as:

$$\psi[x(t)] = [x'(t)]^2 - x(t)x''(t) \tag{3.2}$$

Where $x''(t)$ and $x'(t)$ are the second and first-order derivatives of $x(t)$. The discret-time NEO can be given by:

$$\psi[x(n)] = x^2(n) - x(n-1)x(n+1) \tag{3.3}$$

The NEO provides a direct instantaneous dependence on the power of the signal (ECoG). It delivers a positive constant reducing the squared product of amplitude and frequency; it is large only when the signal is high in power (squared samples $x^2(n)$ and high in frequency ($x(n-1)x(n+1)$ is small). Spikes are located at high frequencies and increased energy. Other technics of spikes emphasize based on DWT (Discrete Wavelets Transform) like [75, 76]. The use of TEO in spikes detection is better than other technics because of its advantage of implementation in the discrete and analog domain. In [77], authors found that the

TEO provides more accurate spikes detection than other detectors exploring prediction error and in addition low computational requirements.

The threshold used after the NEO application on the ECoG can be realized and makes the spikes locations clear and by a simple algorithm. The threshold can be set to a scaled variety of the mean of the TEO and determined using the formula:

$$Thre = C\frac{1}{N}\sum_{n=1}^{N}\psi[x(n)] \tag{3.4}$$

Where N determines the number of samples of the signal and C is a constant. The detection of spikes should be completely autonomous, so the threshold has to be automatically calculated. The value of the threshold can be described by a multiple of the standard deviation of the noise to minimize the probability when noises exceed the threshold. Calculating the standard deviation of the whole signal (counting spikes) is one method of estimating the standard deviation of noises, assuming that the spikes are sparse. In [78], authors defined the threshold depending on standard deviation given by:

$$Thre = 4\sigma_n; \qquad \sigma_n = median\left\{\frac{|\psi[x(n)\ ]|}{0.6745}\right\} \tag{3.5}$$

Where σ_n represents the standard deviation. The threshold is given by a multiple of the standard deviation of the signal.

In figure 3.1, is shown the original ECoG recorded from the brain's rat. The signal was filtered by a band-pass filter of 300 to 2kHz. The ECoG is emphasized by the NEO and compared to the threshold. It can be seen the accuracy of the detection step by taking the standard deviation of the signal (with spikes) drive to an important threshold when we have a high number of spikes in a short time registration and with large spikes amplitude and duration, but the case of using the median to calculate the threshold, the interference of spikes decrease. It can be seen on the time segment registration of ECoG of 0.2sec with 8 different spikes shape, we get 8 crossings with the threshold. The result obtained by the crossing of the NEO with the threshold is described by a square window weighed with the duration of the crossing with the threshold; the best positions of the spikes can be chosen in the middle of windows or corresponded to the highest values of the NEO in each window.

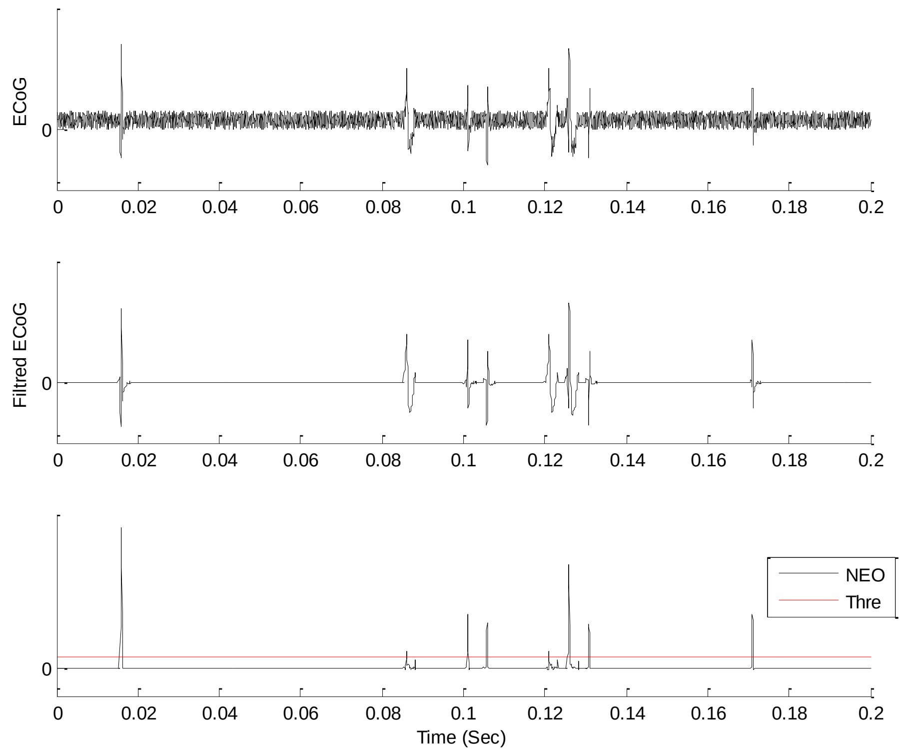


Figure 3.1: ECoG filtering with application of NEO and threshold.

Once spikes locations are detected, the shape of each wave of spikes can be isolated and extracted from each row of the signal. When the NEO crosses the threshold, a location window is captured and means the existence of a spike. The pretended spike can be isolated taking into account the maximal spikes duration of less than 3ms. Practically, we take 4ms as the maximal duration of spikes to be able to cover the form of the action potential of all spikes until the stability of the ECoG.

In figure3.2, is illustrated the spikes isolation on one channel of recorded and filtered ECoG. The isolation depends only on the detected positions taking 1.5ms after the location and 2.5ms after the pretended position. Spikes in this case are aligned to the points of crossing threshold and exactly in maximal values of the NEO in each window. Spikes in this case are misaligned such as maximum and minimum which makes the clustering easier. In the same point of the temporal misalignment, increase the spread of point in feature space. The alignment is realized following the maximum amplitude of action potentials voltage. Usually, this process is started by an up-sampling (not done in our work) to get the continuity dynamic events, the second step consists of the alignment following some events in time (maximum amplitude in this work).

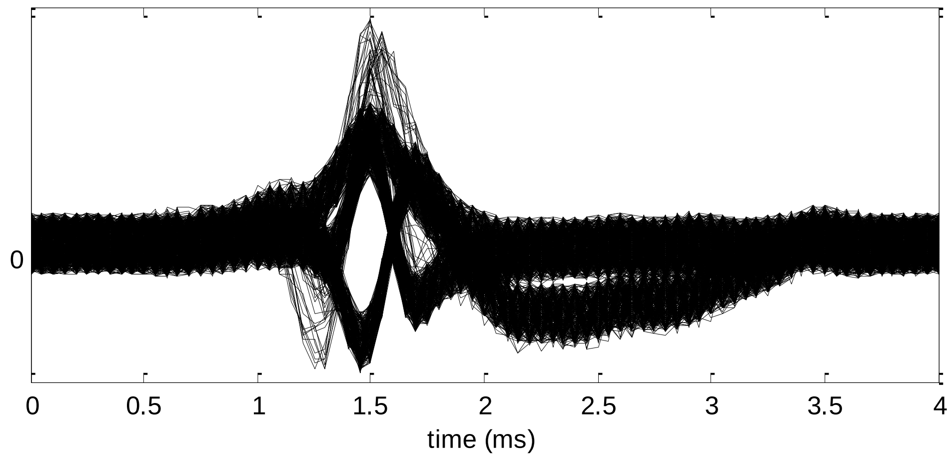


Figure 3.2: Spikes isolation

In biomedical engineering, the maximum slope in the action potential has a biological signification. This method was proposed in [79]. In other works like [80], authors proposed the use of the maximum values of the NEO used in spikes detection as a time reference of the alignment. Figure 3.3 (B) shows the alignment based on the maximum value of the maximum slope in the action potential. In the detection using NEO, automatically, spikes are aligned to the maximum value of the NEO as shown in figure 3.3 (A).

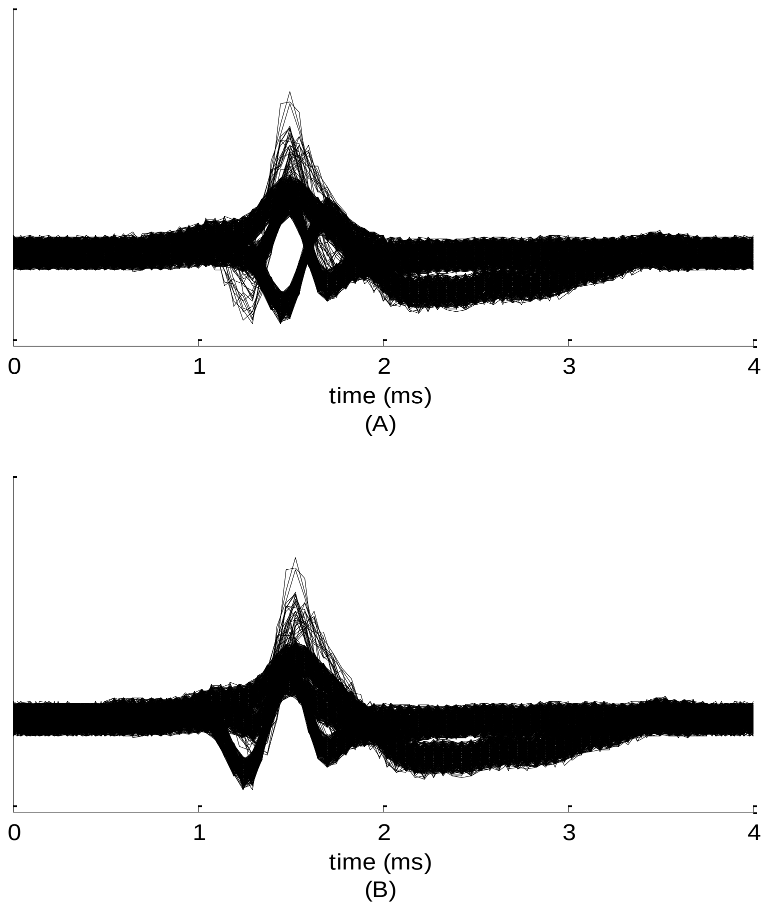


Figure 3.3: Spikes alignment depending on time events. (A) Alignment following the detection time on NEO. (B) Alignment following the maximum amplitude of the action potential.

3.5. Spikes Sorting using SVM

Spike sorting is the task of assigning each waveform of detected spikes to its origin neuron from where it is generated. This task is typically realized by the clustering process. In this approach, each spike shape is characterized by a vector in the feature space. By the supposition of waveform consistency during the extracellular recording process and well-chosen features, it is possible to assign each waveform to its neuron.

In recent years several classifiers have been used in spikes sorting. Among them, SVM (Support Vector Machine) has been modified several times to be employed in this context because of its generalization power [81, 82]. In addition, Ding et al. in [82] showed the ability of a multi-class clustering done by SVM to overcome superposed spikes problem. In their proposed method, they ignored the spikes that were not classified in their one-against-all approach.

A standard SVM originally is designed for two-class classification problems. In a spike sorting study, we usually counter with a mixture of neurons in each channel. One-against-all or one-against-one could be employed to expand a standard two-class SVM to a multi-class case. In [82], it is shown that the one-against-all technique has superiority over the others. In this part, we modified the self-training semi-supervised algorithm proposed by [83] and applied it to spikes sorting tasks using the one-against-all technique.

The theory of SVM was introduced and widely distributed by Vapnik in 1995 [84]. The SVM permits the calculation the separation in the two-class case from a labeled training set.

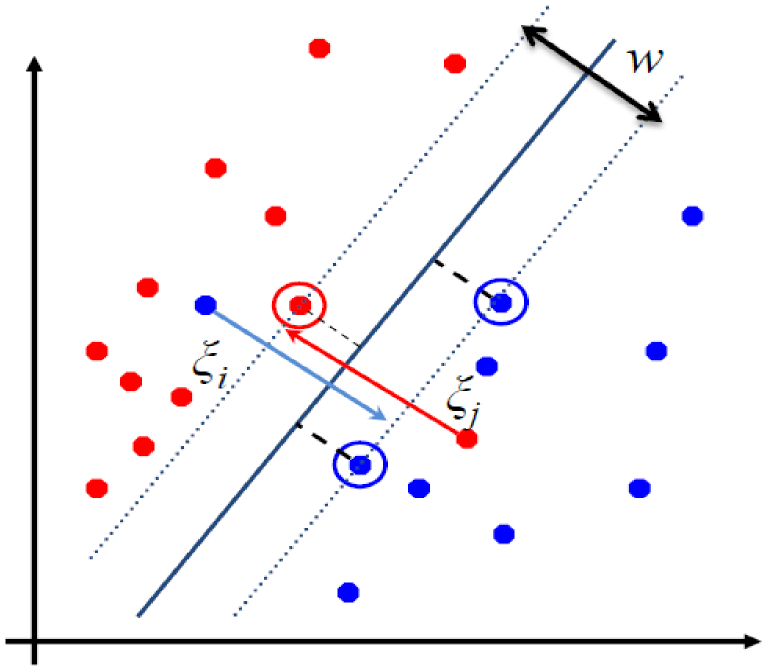


Figure 3.4: Two classes SVM separation with misclassified individuals.

3.5.1. Self-training Semi-supervised Algorithm

Let $X = x_1, x_2, \dots, x_l$ a distributed feature vector of learning population into two classes where $x_i \in \mathbb{R}^l$ is the vector of i^{th} individual descriptor. Let too $Y =$

$y_1, y_2, \dots, y_l$ be the set of associated labels describe the two classes. The Objective function of two-class standard C-SVM is defined as follow:

$$f(w,\xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\xi_i \tag{3.6}$$

Where ξ and C are respectively the relaxation variable parameter for misclassified individuals and the penalty importance coefficient (see Figure 3.4). We aim to find the normal vector of the separating hyper-plane minimizing the objective function (3.6) subject to:

$$\begin{gathered} y_i(w^T x_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0, \qquad i = 1,2,\dots,m \end{gathered} \tag{3.7}$$

In a standard C-SVM, C is a constant value during the minimization process, later we will discuss how we will choose this constant value in each iteration. A self-training semi-supervised algorithm uses this C-SVM classifier. Suppose that F_I is a small training set that contains N_I sample $\{x_i,\ i = 1, 2, \dots, N_I\}$ which has been labeled by an operator $\left[y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(N_I)\right]$, and a test set F_T containing $N_T$samples with $\left\{x_{N_I+i}, i = 1,2,\dots,N_T\right\}$ unknown labels. The steps of the self-training semi-supervised C-SVM algorithm are as follows:

- In the first step, we train the standard SVM with the labeled data and use the designed classifier to label the test set F_T. The parameters of the SVM classifier are denoted $W^{(0)} \in \mathbb{R}^l, \epsilon^{(0)} \in \mathbb{R}^{N_T}$and $b^{(0)} \in \mathbb{R}$, the predicted labels are $\left[y^{(0)}(N_I+1), y^{(0)}(N_I+2), \dots, y^{(0)}(N_I+N_T)\right]$. The upper subscript denotes the iteration number.
- The second step is a loop. In each iteration k, we define a new training set $F_N = F_I + F_T$, with the labels predicted earlier, $\left[y^{(k-1)}(N_I+1), y^{(k-1)}(N_I+2), \dots, y^{(k-1)}(N_I+N_T)\right]$. Then by this augmented training set, we train the SVM and perform a new classification on F_T. We do this step for all C values in a predefined subset and find the optimal C as it is described later on model selection. For this optimum C value, the parameters of the SVM are denoted as $W^{(k)} \in \mathbb{R}^l, \epsilon^{(0)} \in \mathbb{R}^{N_I+N_T}$ and $b^{(k)} \in \mathbb{R}$. We next apply this classifier to the data and update the labels of the test set $\left[y^{(k)}(N_I+\right.$

$1), y^{(k)}(N_I + 2), \dots, y^{(k)}(N_I + N_T)\big]$. Train set labels will remain the same in each iteration. Finally in this step, we calculate the objective function by obtained parameters in each iteration :

$$F\left(w^{(k)}, \xi^{(k)}\right) = \frac{1}{2}\left\|w^{(k)}\right\|^2 + C \sum_{i=0}^{N_I+N_T} \xi_i^{(k)} \tag{3.8}$$

- In the third step, we check the termination criteria by calculating the absolute value of the two consecutive iterations and the algorithm will be stopped if this criteria is less than a predefined positive value:

$$\left|f\left(w^{(k)}, \xi^{(k)}\right) - f\left(w^{(k-1)}, \xi^{(k-1)}\right)\right| < \delta_0 \tag{3.9}$$

Where δ is the least change we expect from our objective function. This also means that the algorithm reaches its local minimum.

The convergence of the original algorithms has been theoretically proved in [83]. This convergence can be observed in our results. Another modification of this algorithm has been proposed by [85] which enters the scattering matrix into the SVM's objective function and claimed better results.

3.5.2. Model Selection

In the standard algorithm of SVM described previously, we should set the penalty parameter C of the SVM through model selection. In [83], it was suggested to use a one-dimensional Fisher ration and it was not mentioned how we can employ it in a high-dimensional space. Here we used a robust Fisher ratio which can be used in our high-dimensional feature space. It is also common to use the cross-validation on the training data set and search in a specific range for C to find the best possible C which gives us the maximum accuracy in our training set. Here because our training set is so small we could not rely on these results. In [85] cross-validation has been used in self-training semi-supervised algorithms however it was prohibited in [83]. The Fisher discriminant ratio is calculated to search a finite set of C values in a predefined subset $C_1, C_2, \dots, C_i$:

$$FDR(k, C_i) = \frac{u^T(\mu_1 - \mu_2)(\mu_1 - \mu_2)^T u}{u^T(\sum_1 + \sum_2)u} \tag{3.10}$$

Where μ and $\sum$ denote the mean and covariance of each class respectively, and u is a discriminant vector. The lower subscript denotes the class label.

$$u = (\textstyle\sum_1 + \sum_2)^{-1}(\mu_1 - \mu_2) \tag{3.11}$$

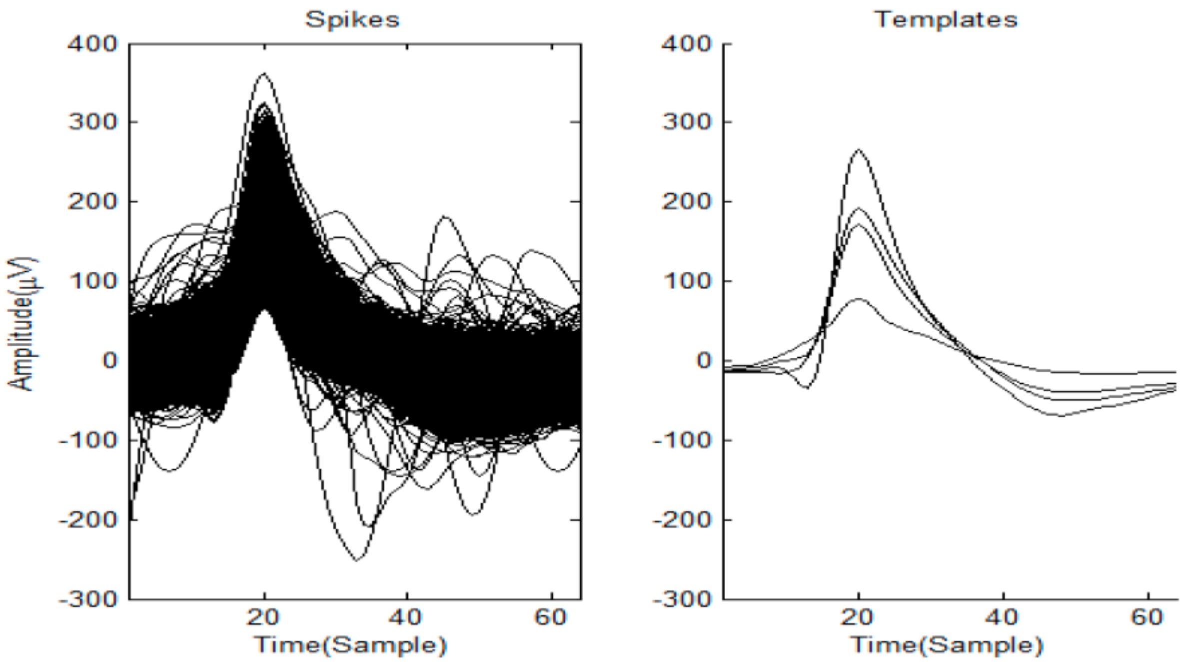
After computing Fisher ratio for all C_l we define C as:

$$C_{opt}^{(k)} = \max FDR(K, C_i) \tag{3.12}$$

Where $C_{opt}^{(k)}$ is the penalty parameter applied to the augmented training set in the k^{th} iteration in the second step of the algorithm. This model selection method is based on the fact that the original Fisher ratio represents the reparability of the cluster, i.e., the Fisher ration increase implies high reparability. Therefore in each iteration, we choose a C that gives us higher discrimination.

3.5.3. SVM in Spikes Sorting

Data in this part came from a 30' multiunit recording from a human epileptic patient. This data was collected at the lab of Itzhak Fried at UCLA, using a Neuralynx System (Tucson Arizona). The spikes have been detected and also clustered using the Waveclus toolbox by the University of Leicester and this clustering has been accepted as our ground truth [72]. The cluster membership number for spikes also contains some zero values that were considered spikes and were not assigned to any cluster by the Waveclus contributor. An overall view of the detected spikes is represented by 64 sample points and their corresponding templates are plotted in figure 3.5. These templates are calculated by averaging the spikes in the same class. The template with the lowest peak corresponds to the non-clustered spikes.

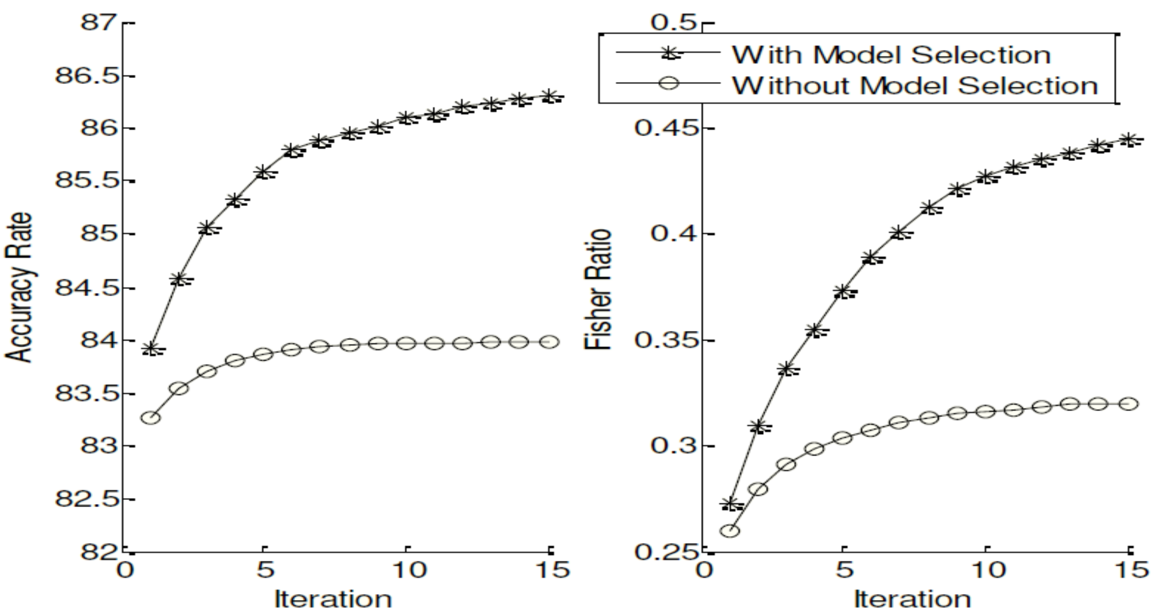


(a) (b)

Figure 3.5: (a) Overall view of the spikes and (b) the corresponding templates.

Since we had a multiclass problem in spike sorting and a two-class classifier, on the other hand, we had to choose between one-against-one, one-against-all, and Direct Acyclic Graph SVM. In [82] Ding et al. claimed one-against-all showed a better performance in neural activity gathered by micro-electrode arrays. Accordingly, we employed this approach by combining three binary classifiers. We could have considered the unlabeled data as a separate class but we ignored it in our results. To do the clustering we need to extract features from our spike waveforms and project each spike on the feature space. Here we used Principal Component Analyses (PCA) and project the spike signal to an orthogonal space as the first component has the biggest projection of the signal. The succeeding components also have the most value under the constraint that they are orthogonal to the preceding components. The original spikes were sampled in 64 points, so we selected the first 25 components of PCA and used them as our features. These 25 components contain almost 90 percent of the spike's energy. To apply the algorithm to the dataset here we chose 20 labeled samples as our training set. For the test set, we used 980 spikes to a role in the iterative scheme in the second step. Choosing all the 10000 remained spikes in our 30' dataset although could improve the accuracy but it causes a severe computational burden.

In figure 3.6(a) is illustrated the accuracy rate of the algorithm in each iteration. The result is the mean value for 100 runs of the algorithm with a random permutation of training and test sets. It can be seen from the curves that the algorithm converges in 10 iterations. In the same figure 3.6(b), is demonstrated the monotonic increase of the Fisher discriminant ratio. The accuracy rate which we obtained here is relative to the Waveclus classification. Therefore if we assume the classification done by Waveclus is correct, which is not, we showed that we reached a better classification in each iteration. For a better comparison of the method, we should use an intracellular recording along with the extracellular recording for all the neurons located nearby the recording area. In that case, we have the actual ground truth, and as it is shown here the algorithm increases the accuracy and converges to the actual ground truth by iterating.



(a) (b)

Figure 3.6: (a) Accuracy comparison and (b) discriminant measurement of the simulated spike data.

Extracellular recording of neural activity usually suffers from noise existence. Here to model this noise in the recording, we assumed it doesn't harm the spike detection process. So we added a different level of noise to our already detected spikes to have different SNR. Figure 3.7 shows that in our particular spikes with a peak value of around 300 µV both SVMs work properly in their ideal situation till the SNR is less than 20. For noise power bigger than this value the relative accuracy of the algorithm drops till the random cluster numbers. This figure shows that the SVM with the model selection outperforms the other method without model selection.

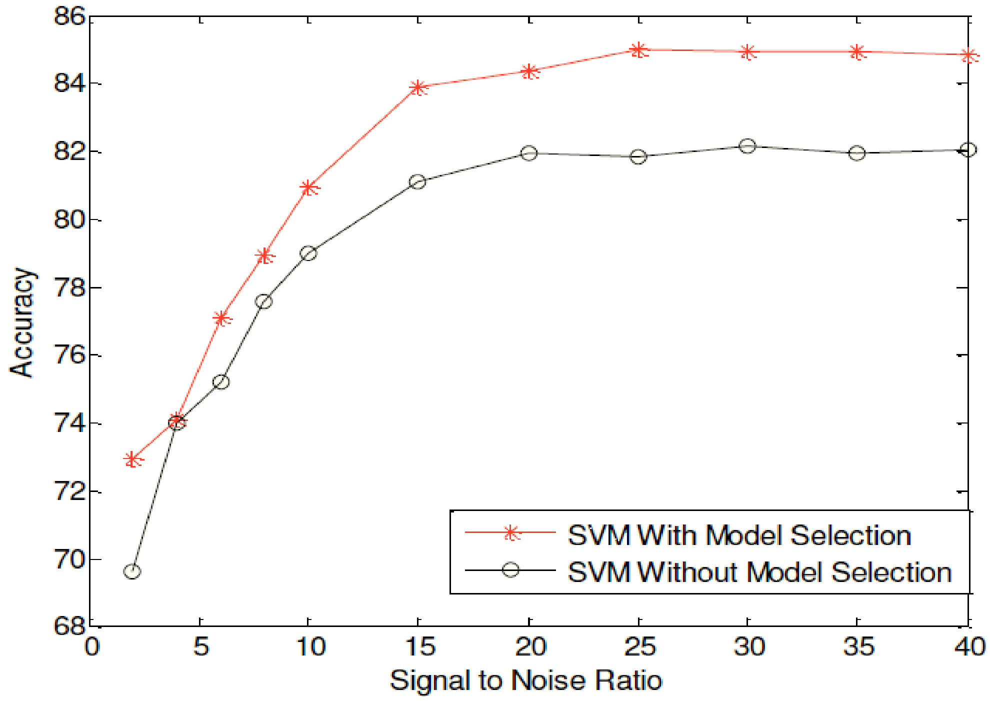


Figure 3.7: Accuracy measurement of the proposed method and the method without the model selection in different SNRs.

3.6. Online Spikes Sorting

The online spike sorting (detection and isolation) is based on the use of a process composed of two mains steps, (i) detector of spike position in an online operation providing stems defining all spikes positions in time and (ii) an online classifier based on the use of template matching. Figure 3.8 shows the methodology of this technic.

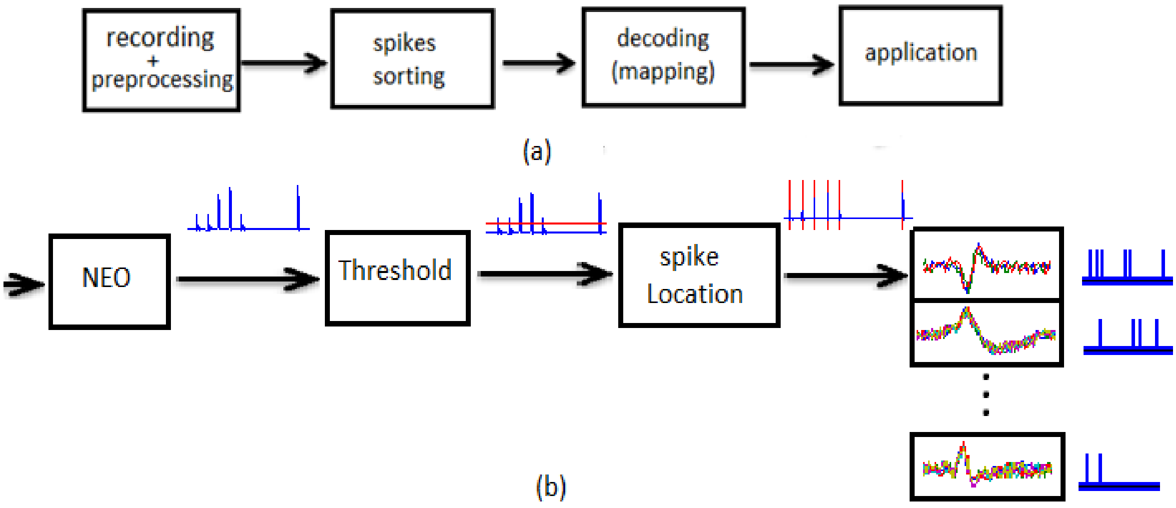


Figure 3.8: (a) Architecture of the used brain computer interface with (b) the online spikes detection and sorting.

The algorithm of the online spikes detection and sorting is composed of four principal stages where the first three stages consist of the localization and isolation of spikes and the last stage is the clustering with templates matching. The detection is firstly started by the application of the NEO given by equation 3.3 which uses the last three samples of the filtered brain signal and carries out the NEO with a late of one sample (one period of sampling of the brain signal). The NEO allows and makes the spikes positions so clearly to be localized with precision. Once the NEO is applied, easily by using a simple threshold given by equation 3.5, results obtained by the use of the threshold are squared signals with durations of crossings of the NEO signals with the threshold. The center of the detection signal is considered to be the exact spike position. In figure 3.9, is presented the technic of online detection using NEO and threshold crossing for spikes positions. The dynamic of some spikes are so slow in time and their action potential is figured in a low-frequency spectrum, the corresponding NEO vanishes with weekly form. A good threshold makes the detection possible as seen in the fourth spike in the same figure. With a good detection and localization of spikes, the next process of clustering has to wait for 2ms (the duration of the second part of the spike form) to be able to compare the spike to the calculated template.

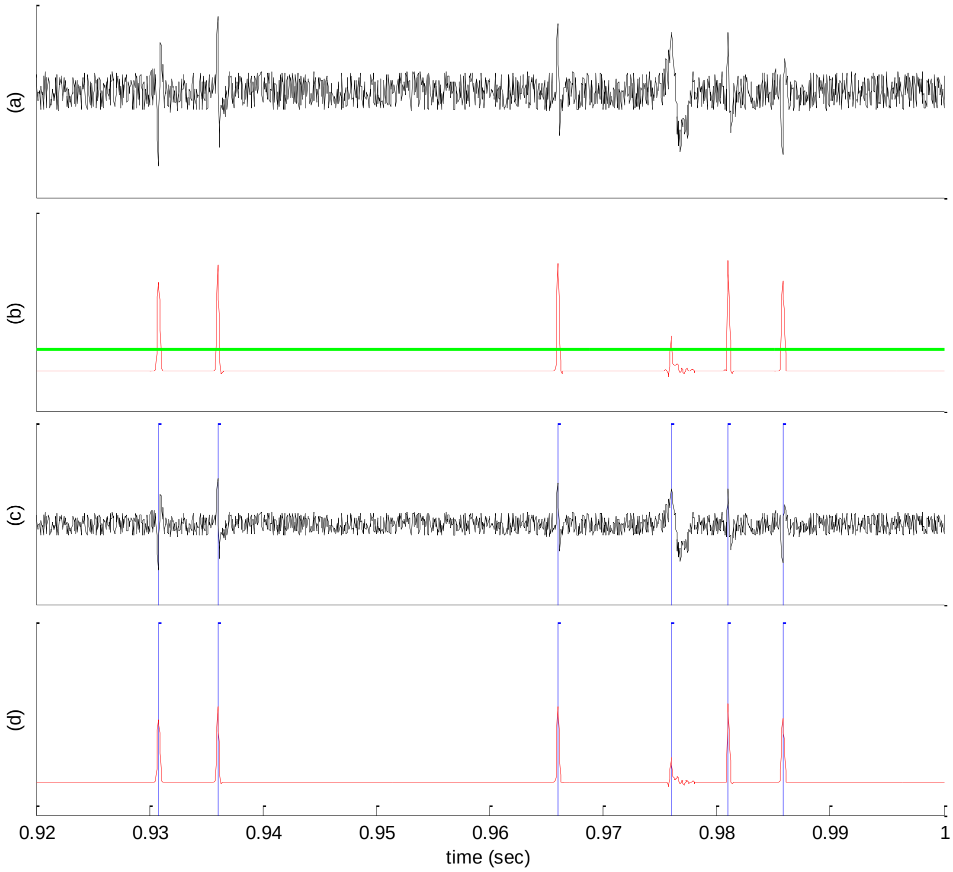


Figure 3.9: Method of spike detection, (a) original ECoG, (b) NEO stage output and (c and d) the results of comparison of the NEO with the threshold and spikes localization.

Template matching will be presented in next (chapter 4). After spikes detection and isolation, simultaneously recorded spiking activities are represented as a sequence of zeros “0” and ones “1”. The existence of spikes is designed by one “1”. The principle of clustering a shape of a spike is situated in the measure of the distance between older corresponding forms with the new detected form. If the new spike exists in form and is related to the template, a new spike on the corresponding neuron (channel) is generated and the template should be adjusted to get an update of the form because of the dynamic of the neural systems and the mobility of the implanted electrodes on the cortex. If the detected spike doesn’t correspond to any existing templates, this means a new form is detected so a new channel is appeared and should be added to the group of channels.

After longtime of the process, some channels will be empty and no spikes are detected for a time, this means that the neuron is offline to the recording process and its channel should be eliminated.

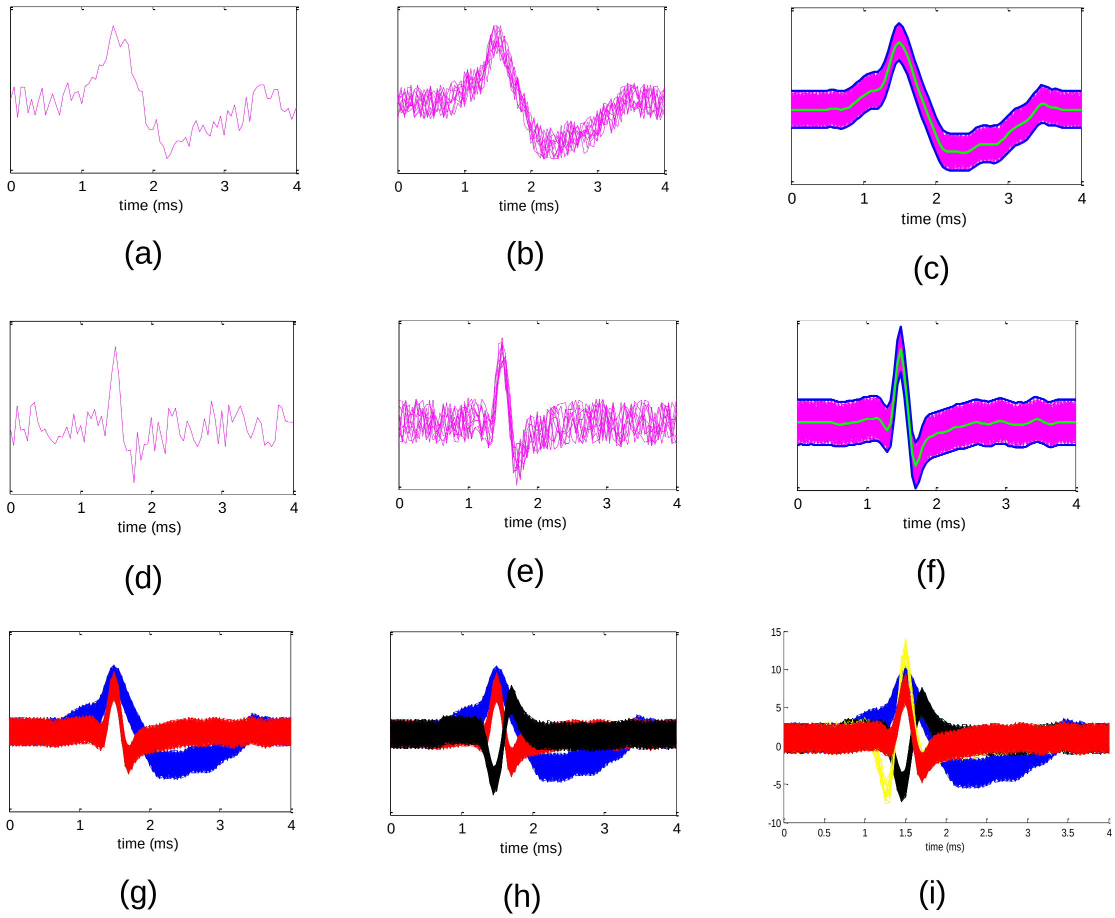


Figure 3.10: Spike cluster based on template matching. (a, d): the first detected spike's shape. (b, e) next first ten detected spikes. (c, f): templates of two forms of spikes with the maximal and minimal limits. In the middle is given the template. (g, h, i): spikes sorting of two, three and five spikes forms (templates).

In figure 3.10(a, b, c, d, e, f) is presented two different templates with the first and the first ten detected spikes of the same form (the same source or neuron) and the used templates at the end of the clustering algorithm. The template is calculated by using the mean of the weighed template with the new detected form using a trust coefficient in the template depends on the deviation, we can use the mean of the last detected forms. The last parts of figure 3.10(g, h,

i) present the cluster results applied on a channel of ECoG contains more than 900 spikes. Results of classification present five templates with a different number of spikes in each one.

3.7. Brain Signal Representation

The spike sorting gives many (48 in rat's data) binary time series channels of neural firing rate represented by '1' separated by a long silence represented by '0'. The real information corresponding to the rat's hand position (lever) is involved in the time between spikes. A special representation taking into account times between spikes is important to translate the binary signal to a continuous signal where the decoder will be easier working with a continuous signal unless the binary one. The continuous representation allows the synchronization of the time samples in hand trajectory with the neural signal. The firing rate of spikes is one of the most used technics in the brain signal representation after spike sorting which calculates the number of spikes in fixed time windows.

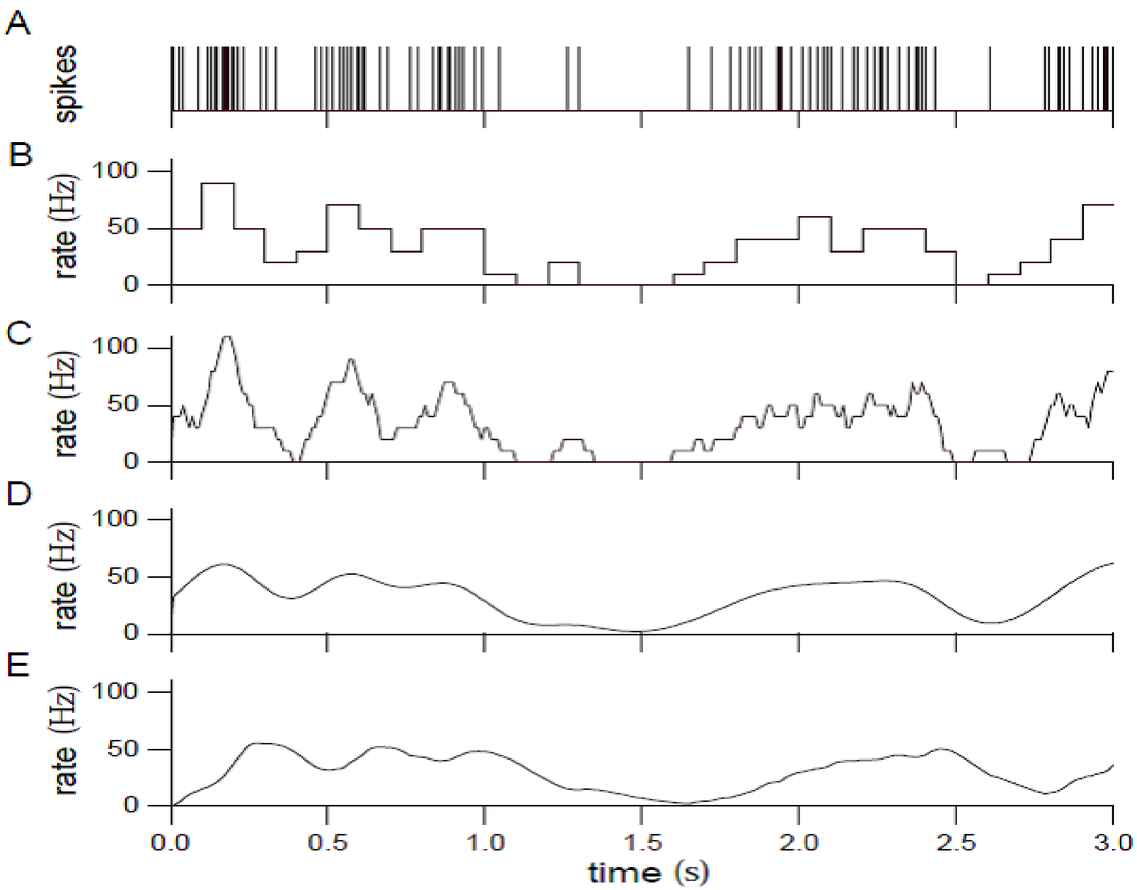


Figure 3.11: Continuous spikes representation

In figure 3.11 is shown different techniques and comparison of continuous spikes train representations. In the part of figure 3.11.A is given three seconds of the response of a neuron in the inferior temporal cortex recorded while a monkey watched a video. Neurons in the region of the cortex from where this recording was made are selective for complex visual images including faces. A simple way of extracting an estimate of the firing rate from a spike train like this is to divide time into discrete bins of duration t, count the number of spikes within each bin, and divide by t. Figure 3.11.B shows the approximate firing rate computed using this procedure with a bin size of 100ms. Note that, with this procedure, the quantity being computed is really the spike-count firing rate over the duration of the bin, and that the firing rate $r(t)$ within a given bin is approximated by this spike-count rate. The binning and counting procedure illustrated in figure 3.11.B generates an estimate of the firing rate that is a piecewise constant function of time, resembling a histogram. Because spike counts can only take integer values, the rates computed by this method will always be integer multiples of $1/\delta t$, and thus they take discrete values. Decreasing the value of t increases temporal resolution by providing an estimate of the firing rate at more finely spaced intervals of time, but at the expense of decreasing the resolution for distinguishing different rates. One way to avoid quantized firing rates is to vary the bin size so that a fixed number of spikes appear in each bin. The firing rate is then approximated as that fixed number of spikes divided by the variable bin width. In figure 3.11.C is shown an approximate of the firing rate determined by sliding a rectangular window function along the spike train with $\delta t = 100ms$ and in figure 3.11.D and E is the same as in C but with two different Gaussian functions. In figure 3.12 is shown the spiking rate representation after the application of a moving Gaussian window of 100ms time width.

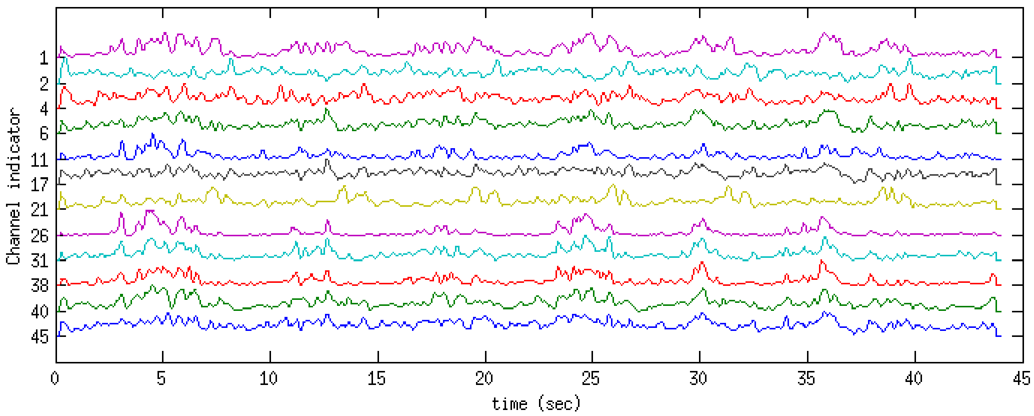


Figure 3.12: Spiking rate of some neurons

3.8. Conclusion

In this chapter, we presented the essentials of spikes detection and classification based on the SVM and with an online technic based on template matching. The use of the NEO was very important to emphasize the spikes and separate between spikes and noise. Once spikes are located in time, another technic of isolation to prepare the process of clustering is applied.

Templates cluster shows its capability to detect and classify each spike in its corresponding channel. The other advantages of template cluster situated in the ability to add new channels (neurons) when new forms are detected and to substitute disappeared shapes after a defined time. A special representation involves the desired information of the hand trajectory (position).

4. FEATURE SELECTION

4.1. Introduction

The number of inputs of such decoder scales the number of its parameters that creates problems with model generalization. Moreover, the size of training data required for a good approximation increases with the number of parameters. To outcome the problem of model order, many technics of data selection and reduction should be applied for an optimal decoder. In this chapter, we concentrated on the data organization and feature selection to prepare the most important stage in the BCI process which consists of the data selection and time lag identification using template matching.

4.2. Data Reduction

Linear and nonlinear modeling of time series data is a practical technique for making a direct mapping that allows the tracking of the hand motions (coordinates: position, velocity, and acceleration) from recorded ECoG. The recorded brain signal has a complicated data representation because of the used matrix of wires which capture the brain signal from not only neurons in direct contact with electrodes but also flows of currents from neighbor neurons. However, the hand position (*ie:* the output of the estimation model) has not only a direct relationship to its own nearest corresponding previous value of the recorded neural signal, but it has also a relationship to previous neural activity with different importance in time, thus some of them are redundant or even irrelevant and are not important for the decoding process, so it is better to delete them from databases to reduce parameters of the used estimation's system and simplify the estimation proposed models. Results with the use of all variables input with redundant and irrelevant will be poorer than if only with required input variables. Variable selection is used to reduce the number of inputs and parameters of the estimation model, so reduce the complexity of the prediction model.

Because the rat lives in an ever-changing environment, an intelligent system must encode patterns over time, recognizing and generating temporal patterns. A question perhaps as old as modeling is “which variables are important?” Because the need to select a model applies to more than just variable selection in regression models, there is a rich variety of answers. In the field of brain-machines, the selection of relevant features is considered absolutely necessary for the ECoG dataset, since correlated neurons are not known in detail. The time delay between the hand position coordinates (output the vector Y_{t+k} to predict) and the intra-cortical activity (the input vector X_t) is a key point in this work, but cannot be determined exactly and considered as a shifting value in time.

Assume that the training data is given by a set of observations $\{\mathbf{x}(t), \mathbf{y}(t)\}_{t=1}^{N}$ with $\mathbf{x}(t) = \left[x_1(t), \ldots, x_q(t)\right]^T \in \mathbb{R}^q$ and $\mathbf{y} \in \mathbb{R}^2$. At time instant t, $\mathbf{u}(t) = \left[x_1(t), \ldots, x_1(t-(n-1)\Delta), \ldots, x_q(t), \ldots, x_q(t-(n-1)\Delta)\right]$ is defined as the input set with all variables and all possible time-lags. The goal of the variable and delay selection procedure is to select a subset $S \in \mathcal{X}$ of the most significant variables for the prediction set. Reducing dimension may also improve the accuracy of an inferred predictive model. Moreover, reducing the number of features may give insights into the working of the system itself. Inference and/or prediction may be computationally costly in high dimensions.

A question can be asked here: How Time is embodied in temporal patterns? Reducing dimension may improve the accuracy of an inferred predictive model because the inference or the prediction may be computationally costly in high dimensions. Moreover, reducing the number of features may give insights into the working of the system itself. Temporal processing of neural data is a challenge because the information is embedded in (i) the *temporal order* which refers to the ordering among the components of a sequence and by (ii) the *time duration*, *i.e.* the time that separates two timestamps [86]. Generally, the selection of the best variables and time lag selection for the prediction set are done by manual selection using experts. Methods for input variable selection can be classified into two classes: *filter* methods and *wrapper* methods (see Guyon and Elisseeff [87] for a review). Filter methods use statistical measures to classify the variables, according to their influence and relevance on the target variable. On the other hand, wrapping methods use the learning model as the basis for selection. Notice that the latter may often achieve more accurate prediction results because variables selection will take into account the approximation model. The correlation method, partial correlation method, Mallows coefficient, and PLS-based method are often used for time-lag selection. Shakil *et al.* propose a linear time-lag model optimized by a genetic algorithm for delay selection [88]. Performing the selection of the time-lags of each variable before variable selection improves the final prediction performance. For better knowledge about the variables x that affect the target variable y, the best choice is by means of filter methods because they use only the information content of data rather than methods based on the correlation coefficient. But for systems involving nonlinear interactions among variables, linear filter methods fail. The linear correlation value which can measure nonlinear interaction among candidates of input variables and the target output is a better choice [89]. But the main difficulty sets in the estimation of the mutual information in multidimensional space which is computationally expensive in addition [90].

Building a predictor from a selection of the most relevant variables is usually suboptimal; conversely, a subset of useful variables may exclude many redundant and relevant variables. This contrasts with the problem of ranking all potentially relevant variables.

4.3. Feature Selection

In current brain signal uses; the desired recorded signal is confounded by other sources that come from electrical charges of neighbor neurons. To obtain clean features, many techniques can be applied to reduce the number of channels used as inputs of the mapping and by eliminating irrelevant and redundant channels. In the input vector, we have to eliminate the redundant elements which are correlated with other inputs and irrelevant which have no hidden information of the hand position. We cite the reduction techniques based on data reduction using correlation functions and coefficients as a first application. Another way depending on the hand trajectory form, we can call to the template matching on all channels when the rat action on the lever comes, and by analyzing the signal forms, we will be able to choose which channel can be a feature of the decoder.

4.3.1. Selection method based on correlation

The correlation is a mathematical-statistical operation that closely resembles convolution means that the two signal sequences are involved in correlation. Our objective in computing the correlation between the two signals is to measure the degree to which the two signals are similar and thus to extract some information that depends to a large extent on the application (similarity and time-shifting). The correlation is often used in radar, sonar, digital communication, geology, and other areas in science engineering. The correlation function in our work is used to minimize the number of inputs of the decoding process (NNT, KF) for tracking the hand motion. To be specific, let supposing the two signals sequences $u(n)$ and $y(n)$ that we wish to compare using the correlation between them. In our work $u(n)$ is the input of the decoding process represented by the recorded neural activity (composed by a such number of channels equal to the number of neurons) and $y(u)$ represents the hand trajectory or motion (can be represented by a vector of one or more component like the position, velocity, and acceleration). In our work, the output is the vertical rat's hand position. Let S be the equivalent hidden function of the used decoder (depending on the application of TDNN and KF) or the transformation relays the output $y(n)$ to the input $u(n)$, the output can be represented by a delayed and shifted version of the decoder (example of a TDNN).

$$y(n) = S\left(\sum_{i}\sum_{j} a_{ij}u_i(n-j)\right) \tag{4.1}$$

Where i represents the number of channels in ECoG (neuron indicator or position) and j is the discrete time-shifting or position in time. a_{ij} are coefficients of validity (selection) of the i^{th} neuron at j^{th} time-shifting. These coefficients have a direct linear relation to the correlation function of the output y to the input u_i by using different shifting of the input. The problem with this technique is situated in the great number of calculus where it is necessary to calculate the correlation coefficient for each of these cases as a combination between the number of channels and the maximal time-shifting. To remedy to this problem, let separate the two dimensions and use, for the first time, correlation coefficients to channel selection for the goal to reduce the number of inputs by reducing the number of ECoG lines taking large windows to take into account the actual moment and previous moment. Second, we use the correlation function to detect the time-shifting between the action realization at muscles (hand position) and its decision at the level of the cortex.

4.3.2. The coefficient of correlation (Bravais-Pearson)

The correlation coefficient can be used to determine the degree of dependence of such variable to another where this can be applied in our case for the selection of which channels of neural activity can be used as input to the machine learning based on a neural network to make a good prediction for tracking the vertical hand trajectory. The correlation coefficient of two measured vectors of data $Y = [y(1),\ y(2),\ ,y(n)]$ and $X_i = [u_i(1), u_i(2),\ , u_i(n)]$ is calculated using the equation 4.2:

$$R_k = \frac{\sigma_{xy}}{\sigma_x \, . \, \sigma_y} \tag{4.2}$$

Where σ_{xy} represents the covariance of the two vectors X and Y. σ_x, σ_y are the variances of the two variables. The equation can be rewritten by:

$$R_k = \frac{\sum_{i=1}^{n} (X_k(i) - \overline{X}_k).(Y(i) - \overline{Y})}{\sqrt{\left(\sum_{i=1}^{n} (X_k(i) - \overline{X}_k)^2\right)\left(\sum_{i=1}^{n} (Y(i) - \overline{Y})^2\right)}} \tag{4.3}$$

R_k takes values in $[-1, 1]$. To interpret the dependence of two variables, we have to compare the correlation coefficient with the position to the center of the interval and its ends. If the correlation coefficient R_k of neural activity with the vertical hand position is nearest to zero, that means the two signals are not correlated or there is poor information in the neural activity which can represent the vertical hand motion. Else, if the correlation coefficient is nearest to the higher extremity (+1), that means there is the neural activity that has a direct linear relationship between the two signals and a reverted relationship if it goes nearest to (-1). With the use of this technique, we can select variables from the nearest to +1 or -1.

4.3.3. The correlation function

The correlation function is different than the correlation coefficients because it provides as result the correlation and allows also the known of dependencies between vector variables with different shifting. The correlation function at a given moment is calculated by the sum of the product of one signal with the other shifted. The correlation function of two vectors is given by the equation 4.4:

$$R_{xy}(k) = \sum_{i=1}^{n} x(i)y(k-i) \tag{4.4}$$

The correlation is represented by the distribution function that represents the degrees of similarity of the two signals with shifting time. The amplitude of the correlation function can be used as weight and the important distribution can all be used as input.

4.3.4. Interpretation with rat's data

The correlation coefficient is used to determine which from the neural activity channel is rich in information and can be selected as input of the mapping used to decode the vertical hand trajectory motion from the brain signal. The distribution function of the coefficients is given in figure 4.1.

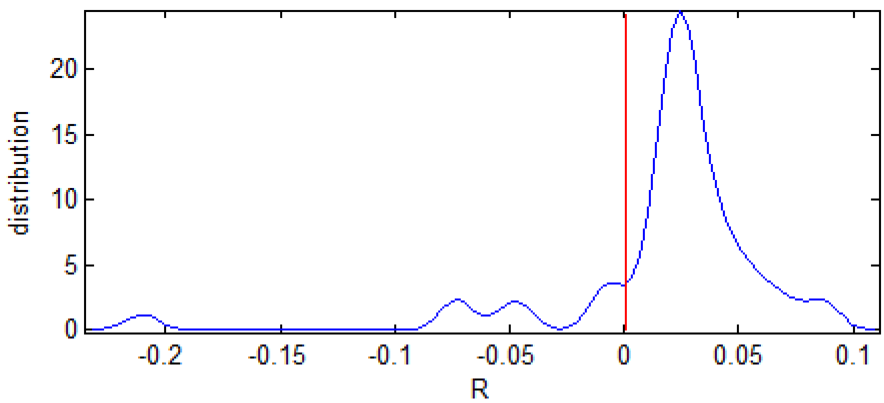


Figure 4.1: Distribution of the correlation coefficient of the neural activity and the hand position

The information of the hand position has a poor linear relationship to the neural activity where the maximal absolute values go to 0.2103 and most of the information is hidden in one channel. We compare the most representable channel of the used information channel to other channels by using local surface or the ratio of each one to the total given information. We can also use other channels but it is clear that the other channels are really very poor looking to the correlation coefficient.

We analyze now with the degree of correlation coefficient and we can select which channel can be used as input of the decoder for making the hand position. The correlation function is now calculated to detect the used shifting samples from the actual state to some past samples where we can prune some samples.

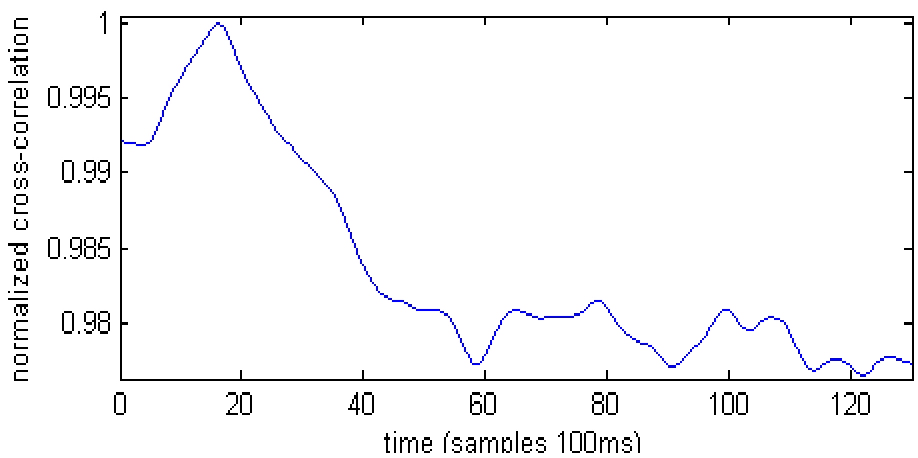


Figure 4.2: Correlation functions for time lag interpretation.

In movement physiology, we know that the information began from the brain (neural activity), so we don’t use the negative values of the correlation function because of the causality of the process. It is clear from the graph of the correlation function that the actual sample has less information than backward with 5 to 27 samples. And also to be precise, we use more than one sample from the selected neurons to get more information on the hand position. Another way is used to define the time lag based on template matching can be used also.

4.4. Template Matching

Template matching was generally used in image processing for certain features (single-pixel, lines, edge, or a complete object) detection, localization, and isolation from original images. In literature, we can find it under other names like matching vision, pattern recognition. The idea of template matching is to have a template to find in time series (either image). The existence of a template can be illustrated by measuring (i) the Euclidean distance between the template and the signal parts [91] or (ii) the convolution or cross-correlation template matching [92, 93].

The Euclidean distance between a template ξ_i and a signal (data) $X(t)$ at time t is given by:

$$D_i^{Eucl}(t) = \|X(t) - \xi_i\|^2 \tag{4.5}$$

The same distance using cross-correlation is given by:

$$D_i^{cs-corr}(t) = X^T(t)\xi_i \tag{4.6}$$

The process of template matching works easily like filtering but seeking other goals. It is realized only by the two-dimensional correlation between the templates (window) with the original images. Results will be an image with high values in where the templates match with parts of the image (strong correlation) and elsewhere values are low. The idea of template matching is to detect events buried in a signal and locate closely resemble the template events. In brain signal processing (in our work), we use the template matching for (i) spikes classification by using time dynamic templates of detected spikes (chapter 3: online spikes detection and sorting) and (ii) detection of time lags between the brain signal and gesture realization.

In figure 4.3 is shown a superposition of an instantaneous time event related to the hand position with the spiking rate recorded from the primary motor cortex of the rat. The dynamic of the two variables shows the resemblance between them and makes it possible to realize a detection of the time arriving of the events on the lever before its realization from the brain signal.

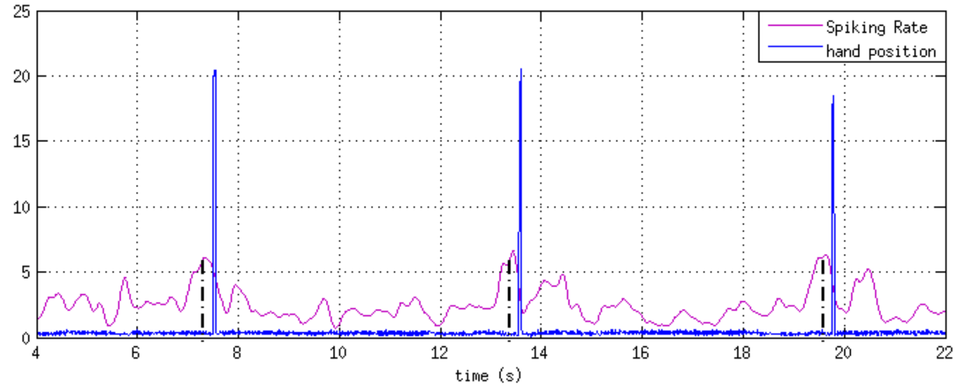


Figure 4.3: Coincidence events (dotted lines) are detected based on superposition of spiking rate (*magenta*) with time events on the lever (hand trajectory in *blue*).

The decoding process should essentially regroup its main stages to get the best hand trajectory decoding using any kind of decoder. The time lag selection is very important for the delayed process of decoding because of the late between the time decision of the action on the lever and the time of event realization.

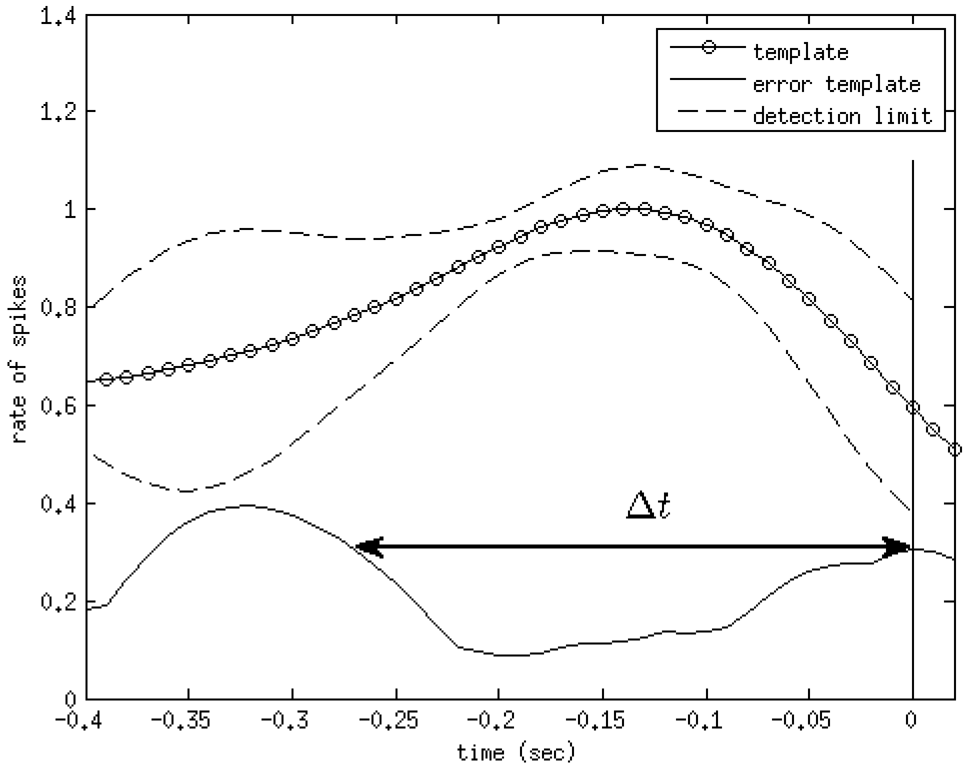


Figure 4.4: Templates for time lag definition.

Figure 4.4 shows the template of one channel (21) of brain signal and the time lag definition from the dynamic of the template.

A classification can be used based on the measure of performance of these brain signal detection techniques by computing the sensitivity and the specificity of the channels for a range of descriptors. Table 4.1 presents the ten "best" neural channels for which six descriptors have been computed: the template variance and mean resp. μ_i and σ_i, the variation coefficient $cv_i = \frac{\sigma_i}{\mu_i}$, the early firing time before the behavior event is detected Δt_i, the maximum rate of missing events ME_i and the maximal amplitude of the template MA_i. The ideal channel is a channel with a detection time Δt_i as large as possible, a small variability cv_i and a low missing detection rate ME_i.

i^{th} ch.	$\boldsymbol{\sigma_i}$	$\boldsymbol{\mu_i}$	$\boldsymbol{cv_i}$	$\boldsymbol{\Delta t_i}$	$\boldsymbol{ME_i}$	$\boldsymbol{MA_i}$
21	**0.2090**	0.8530	**0.2450**	**0.500**	**0.1930**	6.6144
25	0.2575	0.8792	0.2928	0.360	0.4044	7.8553
43	0.2395	0.8108	0.2954	0.280	0.3116	4.1574
2	0.2753	0.9088	0.3029	0.468	0.3911	5.8860
42	0.2849	0.8688	0.3279	0.250	0.4095	**7.9383**
29	0.2800	**0.7923**	0.3534	0.365	0.4985	4.1551
1	0.3524	0.9324	0.3779	0.299	0.5307	5.8253
22	0.3202	0.8095	0.3955	0.430	0.4298	6.4084
14	0.3391	0.7997	0.4241	0.350	0.4638	2.6330
5	2.0931	0.4095	3.3851	–	5.6899	0.2919
13	2.3192	0.4407	1.6367	–	5.5854	0.2598

Table 4.1: Channel ranking obtained from a range of descriptors computed on individual channel. The two last channels are randomly chosen from the non-selected group of channels: there is no time detection because of the non-regularity of templates.

This electrode selection strategy, originally suggested in [89], was based on our motivating rationale that channels eliminated before the main processing would further simplify the decoding of motor cortical recordings. Our experiment retains the following 4 channels at maximum.

4.5. Conclusion

Features selection by reducing data and determining the time lag between the decision time of events and its realization time offers a simplification of the decoder by reducing the number of its parameters. Correlation coefficients and correlation functions were used respectively to reduce the number of redundant channels by minimizing channels involving the same hidden information and eliminate irrelevant channels with no relation to the hand position.

5. DECODING HAND TRAJECTORY

5.1. Introduction

The mapping algorithm is a series of computations that decode the BMI selected inputs derived by any technique of feature selection into a device neural command. The feature selection (or data reduction) becomes more important when the input vector contains many irrelevant inputs. We are interested in the correlation method and template matching for the input features. Used mapping is based firstly on a linear and non-linear modeling using linear Regressor and the Kalman filter with many procedures of stat and measure variables, in the second time, the Neural Network (NN) and Time Delay Neural Network (TDNN) will be introduced, another architecture in melange between the Kalman Filter and Neural Network is introduced called Neural Network aided Kalman Filter. The application of the spikes sorting and detection gives 48 channels of neural activity represented by firing represented by ’1’ resembles to spikes separated by long silences represented by ’0’. A moving Gaussian window of 100ms time width is used to compute the spikes rate and also to get a simultaneous brain signal with the hand trajectory because of the used camera frequency to follow the lever. The decoding process uses the firing rate extracted from the spike sorting stage. It can be implemented in different ways depending on the type of intent labels. In the case of discrete labels, due to the restriction of the number of intents or tasks, the decoding process becomes a pattern recognition problem, and often a classifier is used to determine the intent. For instance, the determination of an image among three to a monkey is a discrete decoding problem [94]. Kim et al. [95] employed the support vector machine (SVM) and multi-layer perceptron (MLP) to extract the direction of hand movement. According to their results, SVM outperforms MLP when the spike detection or sorting step is imprecise or imperfect. Continuous labels are usually associated with motion data recorded from the motor cortex such as determination of the position, velocity or acceleration of an organ of the body [96, 11]; they can be categorized into stationary and non-stationary. For stationary labels, sudden and drastic changes do not exist. Most of the decoding problems such as the arm trajectory control from a prosthetic system are in this category. In the non-stationary case, the intent or task exposes sudden presents in time we typically would like to detect with the shape of these changes. These changes prevent the algorithms that use past and actual observations to estimate accurately position and shape. Hence, the methods which are just dependent on current points of observations, such as linear methods, cannot estimate perfectly; because the changes in neural activities are not as strong as the target task. In [97], a non-stationary problem was solved to find vigorous changes in the movement of the rat's hand.

5.2. Linear Regressor

Linear Regressor is the way for making a direct relationship between two dependent variables defined as the hand position and the neural activity. We talk about a linear regression between the two dependent variables. Let U and Y be respectively the selected neural activities and the hand position. The output to be estimated is the hand position Y and the input is the neural activities U. The linear regression of the input and output can be written as follow (equation 5.1):

$$y_n = \sum_{i=1}^{l} h_i u_{n-i} + w_n \tag{5.1}$$

Where l denotes the memory of the Regressor and w_n is a white Gaussian noise. The mapping in this case can be augmented to equation 5.2:

$$\begin{pmatrix} y_k \\ y_{k-1} \\ \vdots \\ y_{l+1} \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & \cdots & h_l \end{pmatrix} \times \begin{pmatrix} u_{k-1} & u_{k-2} & \cdots & u_l \\ u_{k-2} & u_{k-3} & & u_{l-1} \\ \vdots & \vdots & \ddots & \vdots \\ u_{k-l} & u_{k-l+1} & \cdots & u_1 \end{pmatrix} + \begin{pmatrix} w_k \\ w_{k-1} \\ \vdots \\ w_{l+1} \end{pmatrix} \tag{5.2}$$

We can rewrite the equation 5.3:

$$Y = HU + W \tag{5.3}$$

For the hand trajectory decoding from brain signal where it is represented by Y, the input U represents the neural activity selected by using the correlation coefficient (channel 21) and the time lag is defined by using the correlation.

The pseudo-inverse can be applied for parameters learning of the Regressor solving equation 5.4:

$$min\, J_H = \sum \|HU - Y\|^2 \tag{5.4}$$

The solution of minimization of this equation is given by equation 5.5:

$$H = YU^T(UU^T)^{-1} \tag{5.5}$$

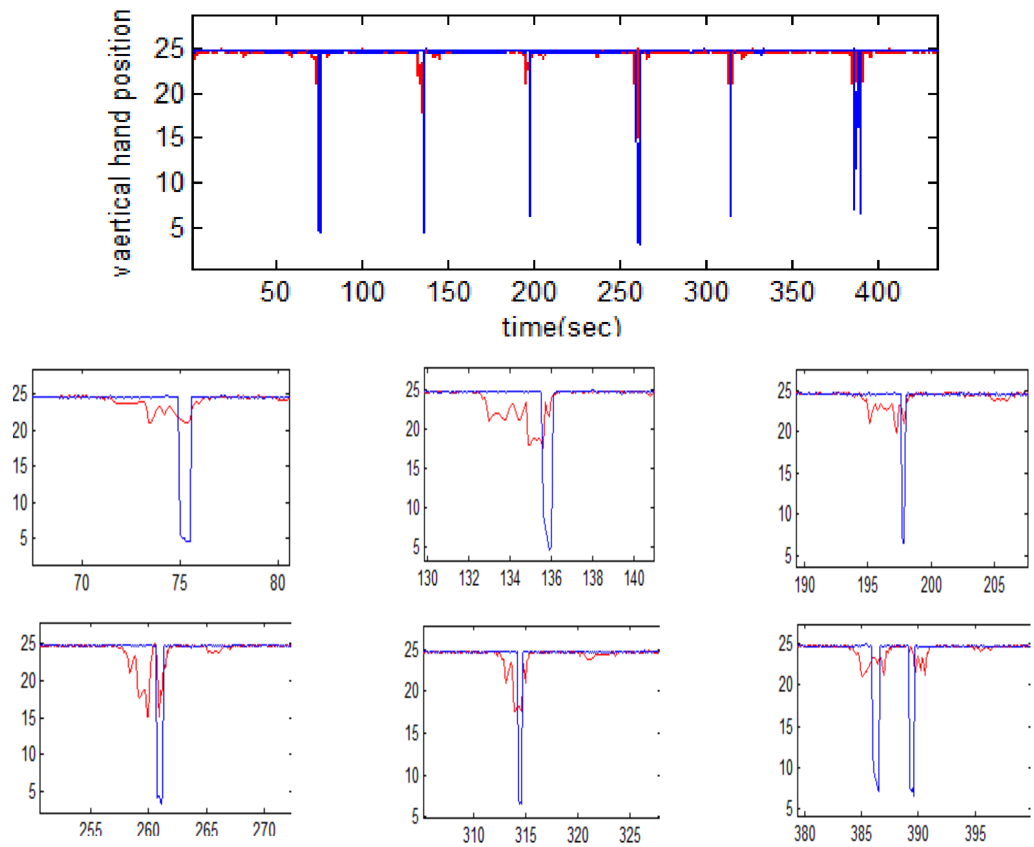


Figure 5.1: Decoding hand trajectory from primary motor cortex using a linear Regressor.

The half data of simultaneous brain signal (Channel 21) and vertical hand trajectory were used for the training and the other half was used for testing. Results on the hand trajectory decoding from primary motor cortex using linear Regressor were really poorer with no resemblance to the real trajectory. These results show the disability of pursuit of the hand position because the pseudo-inverse takes a smooth on all training and takes into account all data for defining the parameters.

Another technic based on the conjugate gradient descent (CGD) was used for the training of the linear Regressor; results, in this case, are shown in Figure 5.1. It can be seen that the Regressor is trying to realize the tracking at the true time but with a poor pursuit. The training using CGD was better for the linear regressor but the nonlinearity of the mapping limits its ability for the good making of the hand trajectory.

5.3. Kalman Filter

The Kalman filter is one of the most used observers in signal and parameter estimation. Its origin is extracted from probability and linear systems by Rudolf Email Kalman in the late 60’s; its first use was in spatial research (APOLO 2). The Kalman filter is defined as a recursive mathematical framework that operates on noisy data to the estimation of state variables; it was used in the Brain Machine Interface in many works [3, 98]. because it is used on a non-precise model on the brain functions for coding and decoding neural commands, also it is used because of its algorithm of decoding based on simple equations, another point of view on dawn BMI hidden in the time-shifting between the decision on the arm action and its realization which can be a benefit in the time calculus. The Kalman filter uses as input the output of the process to estimate (reconstruct) the state. So we can consider the hand coordinates as state’s variables and neural activities as measured variables. The Kalman filter operates two principal steps:

Prediction: The prediction step means the use of stochastic equations of the state and measurement where we apply a linear combination of the new state with its previous value and another linear combination calculated state value with the measurement, in this step, the calculus of the error co-variance is necessary.

Update: In the update step, one calculates the Gain and fixes the error on the predicted state and the error covariance.

5.3.1. Linear Modeling

Let X and Y be respectively the state and observation variables where $X \to \mathbb{R}^n$ and $Y \to \mathbb{R}^p$. The linear state space non-commanded model relays the state and the output, in this case, can be represented at an operating space by its linear discrete equations given by the equation 5.6:

$$\begin{gathered} x_{k+1} = A_k x_k + w_k \\ y_k = C_k x_k + v_k \end{gathered} \tag{5.6}$$

Where $x_k \in \mathbb{R}^n$is the state vector at k^{th}sample time $y_k \in \mathbb{R}^p$is the observation at K^{th} sample time. Matrices $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{p\times n}$ are respectively the state evolution (transitional) matrix and measure matrix. We use $S(A,C)$ to denote the system given by equation. Vectors $w_k$and v_k are respectively the state and the measurement noises. We consider these noises as independent Gaussian and:

$$\begin{gathered} w_k \sim N(0, Q\delta_k) \\ v_k \sim N(0, R\delta_k) \\ E[v_k w_k^T] = 0 \end{gathered} \tag{5.7}$$

Matrices A and C depend on time, so the system given by equation 5.6 is linear time-varying, therefore, we need to have A and C for each sample time. For the first time; to simplify the study; we consider that the system is linear time-invariant thus:

$$\begin{gathered} A_k \simeq A \\ C_k \simeq C \end{gathered} \tag{5.8}$$

If we take l sample time registration of variables, we can construct two big matrices of data which are $X_{1:l}$and $Y_{1:l}$ respectively the matrix of hand coordinates with $n \times l$ and matrix of neural activity with $m \times l$. $n = 6$ in the case where the state vector contains the $2d$ hand position, velocity and acceleration. For the second matrix, $m = number\ of\ neurons$. Because noises are zero means, so A and C are obtained applying Least Square estimation and the solution is calculated by minimizing the error estimation functions of the state and measure:

$$\min J_A = \sum_l \|x(k+1) - Ax(k)\|^2$$
$$\min J_C = \sum_l \|y(k) - Cx(k)\|^2 \tag{5.9}$$

The solution of J_A is given by:

$$A = X_{2:l} X_{1:l-1}^T (X_{1:l-1} X_{1:l-1}^T)^{-1} \tag{5.10}$$

Now we can calculate the covariance of the state noise Q which is given by:

$$Q = \frac{(X_{2:l} A X_{1:l-1})(X_{2:l} A X_{1:l-1})^T}{(l-1)} \tag{5.11}$$

The same thing is applied for the calculus of C and R:

$$C = YX^T(XX^T)^{-1}$$
$$R = \frac{(YCX)(YCX)^T}{l} \tag{5.12}$$

The big matrix Y contains a small quantity of information depending on the hand movement but in the calculus of the measurement matrix, we oblige it to take into account the non-interesting part of data. For the reverse calculus in the Kalman filter, we will face the ill-conditioned (ill proposed) problem because there are many solutions given the same result, and also we are faced with the non-linearity of the process. To reduce the number of used neural activities for the modeling we need a technique based on the projection of the matrices depending on their singular values. We use now Y_r instead of Y with r neural activity instead of all recorded neural activities. The new matrix Y_r will be calculated in such a space of distribution of singular values of its original matrix Y.

5.3.2. Kalman Filter Formulation

The Kalman filter was designed to estimate the state of a linear process given by equation using as input the measurements and the command (there is no command in our proposed model) of the process, we derive the estimated state of the process in the output.

$$\hat{x}_{k+1} = A_f \hat{x}_k + G_k y_k \tag{5.13}$$

And this will become:

$$\hat{x}_{k+1} = A_f \hat{x}_k + G_k C_k x_k + G_k v_k \tag{5.14}$$

Where $\hat{x}$ indicates the estimation of the state of the filter, A_f and G represent respectively the transitional and the gain matrices of the filter. The error time evolution is given by:

$$\begin{aligned} \hat{e}_{k+1} &= x_{k+1} - \hat{x}_{k+1} \\ &= A_k x_k + w_k - \left(A_f \hat{x}_k + G_k C_k x_k + G_k v_k\right) \\ &= (A_k - G_k C_k) x_k - A_f \hat{x}_k + w_k - G_k v_k \\ &= \left(A_k - G_k C_k - A_f\right) x_k - A_f \hat{e} + w_k - G_k v_k \end{aligned} \tag{5.15}$$

The Kalman filter algorithm is designed to make the error of estimation go to zero, so by introducing the mean function on the last equation we obtain:

$$A_k - G_k C_k - A_f = 0 \Leftrightarrow A_f = A_k - G_k C_k \tag{5.16}$$

Let P be the matrix of the covariance of the estimation error so:

$$P_k = E[e_k e_{k-\tau}^T] = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] \tag{5.17}$$

And

$$P_{k+1} = E[e_{k+1}^T e_{k+1}] = (I - G_k C_k) P_k (I - G_k C_k)^T + G_k V_k G_k^T \tag{5.18}$$

The Kalman filter minimizes the covariance of the estimation error. The solution that minimizes this equation is given by:

$$G_k = P_k C^T (C_k P_k C_k^T + V_k)^{-1} \tag{5.19}$$

So the evolution of the covariance estimation error is given by:

$$P_{k+1} = A_k P_k A_k^T + Q_k \tag{5.20}$$

Kalman filter algorithm: finally, the Kalman filter algorithm for the time-invariant system can be formulated by rewriting past equations, started by variables initialization in the first iterations of the algorithm than a combination between the two other steps of the prediction and update:

Algorithm:

Initialization:
1:$k=0$
2: A, C
3: $\hat{x}_k$, P_k
Repeat:
Prediction
1: $P_{k+1/k}=A_kP_{k/k}A_k^T+Q_k$
2: $\hat{x}_{k+1/k}=A_k\hat{x}_{k/k}$
Update
1: $G_{k+1}=P_{k+\frac{1}{k}}C_k^T\left(C_kP_{k+1/K}C_k^T+V_k\right)^{-1}$
2: $\hat{x}_{k+1/k+1}=\hat{x}_{k+1/k}+G_{k+1}\left(y_k-C_k\hat{x}_{k+1/k}\right)$
3: $P_{k+1/k+1}=(I-G_{k+1}C_k)P_{k+1/K}$

The extended Kalman filter has the same algorithm as the Kalman filter but it is applied for nonlinear systems. By a linearization, the obtained model of equation 5.6 becomes an augmented to a linear model with variable parameters. These equations are obtained by using the Jacobean applied on the nonlinear process. The new augmented model of the linearized system is given by equation 5.21:

$$\begin{cases}\begin{bmatrix}x_{k+1}\\ \theta_{k+1}\end{bmatrix}=\begin{bmatrix}A_k & 0\\ 0 & I\end{bmatrix}\begin{bmatrix}x_k\\ \theta_k\end{bmatrix}+W_k\\ y_k=\begin{bmatrix}c_k & 0\end{bmatrix}\begin{bmatrix}x_k\\ \theta_k\end{bmatrix}+V_k\end{cases} \tag{5.21}$$

Where θ is the vector of stat and measurement parameters.

5.3.3. Decoding the hand trajectory using the Kalman filter

The Kalman filter with its versions can be also applied in the decoding of the hand trajectory from the primary motor cortex. In [99], the authors proposed the use of the Kalman filter to decode the hand motion (position, velocity and acceleration) on the 2D plan. The proposed model to be estimated contains as state the hand motions and the measurement is the brain signal. The Kalman filter uses as inputs the measurements (brain signal) and estimates the state of the model (hand motion). The decoding of the hand trajectory in the plan using the extended Kalman filter on the model augmented to the stats and measurements parameters of the model from the primary motor cortex of the rat is realized in our work. The state variables are the hand position (x,y). The horizontal hand position of the rat was stable at the same place denoted by y and variations were only on the vertical x because the rat pushes (only up and down) a lever for bringing water.

Decoding results are shown in Figure 5.2. We presented only the vertical hand position with the best-selected channel (21). The Kalman filter follows much better the hand trajectory compared to the linear regressor. Regarding the events, the Kalman filter is classified as a good detector of pushing events on the lever.

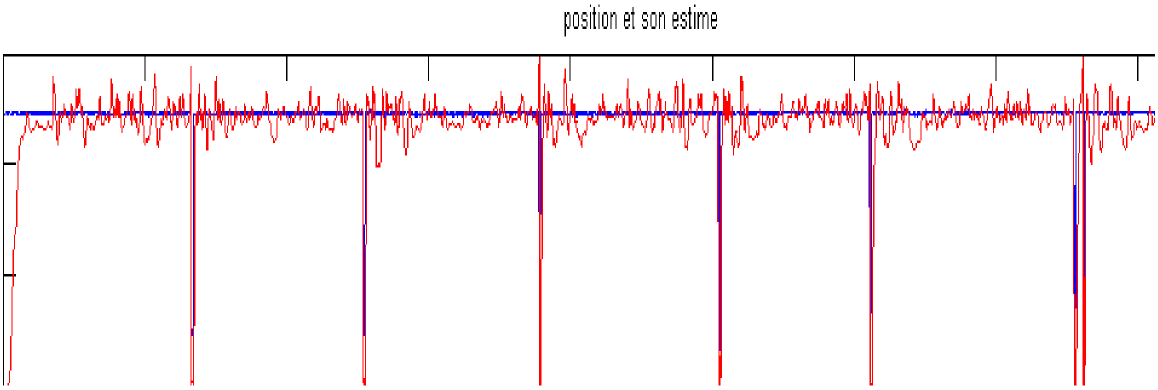


Figure 5.2: Vertical hand position estimation using the extended Kalman Filter.

Parameters given by the Kalman filter are also given in Figure 5.3. As we see, there is a big variation in the state's and measurement's parameters of the model related to the vertical hand position. The Kalman filter decodes better the events on the lever with low precision on the hand trajectory tracking.

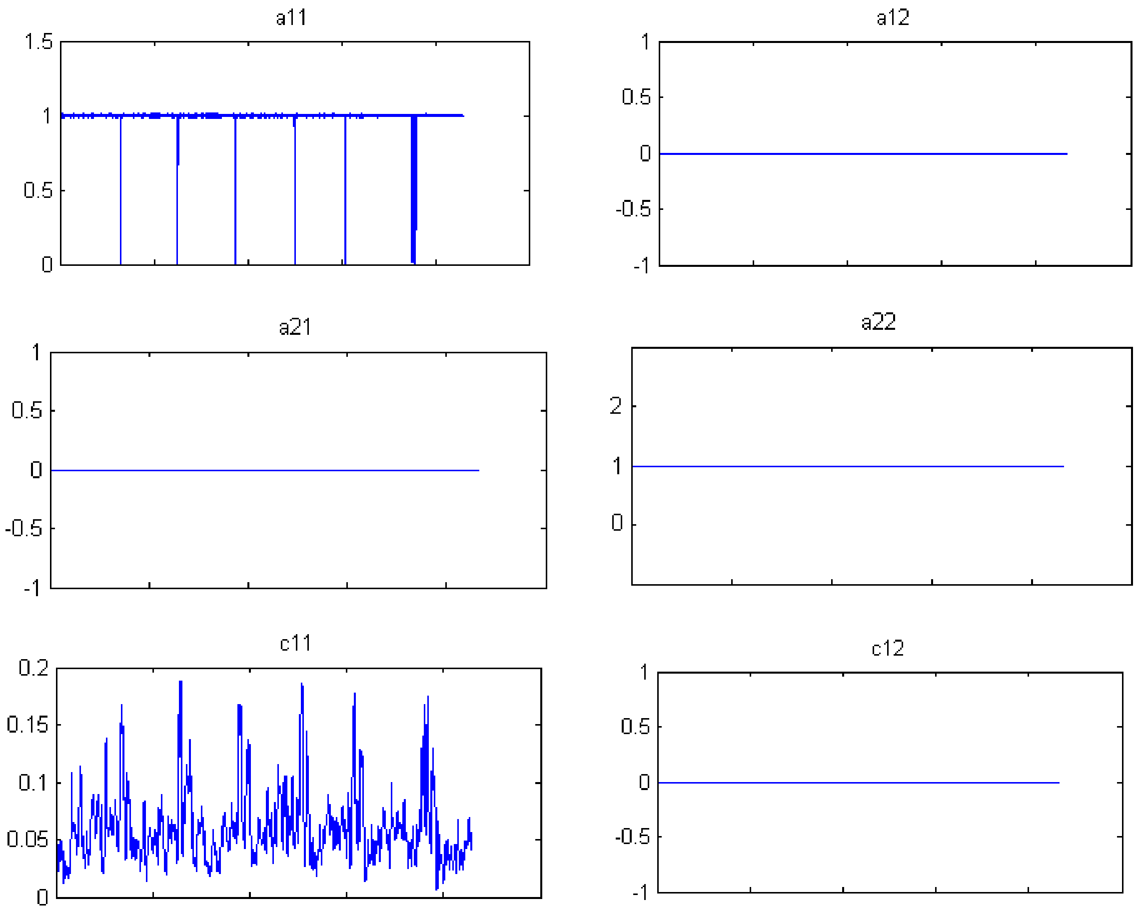


Figure 5.3: Stat's and measurement's parameters estimation with EKF.

5.4. NN and TDNN

Neural Networks (NN) can be considered as black boxes able to simulate systems behaviors with high performance. NNs are extracted from the simulation of real brain neurons. The inspiration for neural networks came from the examination of central nervous systems. In an artificial neural network, simple artificial nodes, called ”neurons”, ”neurodes”, ”processing elements” or ”units”, are connected together to form a network that mimics a biological neural network. Artificial neural network (ANN) can be represented as a black box composed of several neurons interconnected in different cases depending on the ANN architecture. The black box has a great capacity in predictive modeling and mapping of the output in the function of input of processes. The result of estimation becomes better after a good train of the neural network. Many researches were devoted to the application of ANN in process modeling, variable and/or parameter estimation, state prediction, control, and diagnosis of process.

These researches aim to solve various problems proposing many architectures of ANN depending on the difficulties of application in the complexity, linearity or none of the process, optimality, memory, data reduction. We cite the most used ANN architectures as the Perceptron [100], ADALINE, i.e. adaptive linear element [101], Hopfield network [102], Kohonen network [103], and Boltzmann machine [104], multi-layer feed-forward neural networks learned by back-propagation algorithm [105]. The description of these methods can be found in various books such as Freeman and Skapura [107], Gallant [107], Smith [108], Ripley [109], Bishop [110], etc. The choice of the type of network depends on the nature of the problem to be solved. At present, two popular ANNs are (i) multilayer feed-forward neural networks trained by the backpropagation algorithm, i.e. backpropagation network (BPN), and (ii) Kohonen self-organizing mapping, i.e. Kohonen network (SOM). The BPN is most often used, but other networks have also gained popularity. To know better neural networks, it is important to know the process of the calculus of one neuron. The architecture of a neuron with n input data is composed of two phases of calculus known to a neuron system:

- **Summation:** the corresponding equation in this level to a weighted and biased inputs:

$$s_k = \sum_{i=1}^{n} w_{k,i} u_{k,i} + b_k \qquad 5.22$$

- **Activation function:** each neuron is characterized by an activation function applied on the calculated sum. The activation function provides the final state of the neuron's output at each iteration.

$$y_k = f_k(s_k) \qquad (5.23)$$

A neural network is a group of neurons interconnected in different manners where the information propagated throwing the used architecture of the network from the input to the output. Any neural network is characterized by its number of hidden layers known a number of neurons composed each layer with the corresponding activation functions. Artificial neural network or the black box as defined before is composed of several neurons interconnected in different cases depending on the used architecture. The black box, from its architecture and good training, inspires its great capacity in predictive modeling and mapping processes and improves it in modeling, variable and/or parameter estimation, state prediction, control, and diagnosis of process. Many architectures were proposed to solve various problems depending on their needs and difficulties of application in the complexity, linearity or none of the process, optimality, memory, data reduction.

5.4.1. Decoding hand Trajectory with TDNN

Multilayer perceptron (MLP) offers a popular approach to temporal pattern learning. TDNN is MLP architecture of an artificial neural network perceptron with a feed-forward. Another version of the TDNN with feed-forward resembles to the NARX applied for input-output modeling of nonlinear dynamical systems.

5.4.1.1. The NARX models

Despite that linear filters such as wiener filters [86] deliver a simple analytic solution with low complexity training and online feasibility, the reconstructed trajectories may be suboptimal since the output is limited to mappings in the input space and mapping of control features to motor behavior may include nonlinear translations. Therefore, herein, we study modeling hand movements using artificial neural networks. The NARX (Nonlinear AutoRegressive with eXogenous inputs) model is an essential type of nonlinear discrete-time system for time-series forecasting [111, 112]. The recurrent NARX can be represented:

$$y(t) = f(u(t-\Delta), \ldots, u(t-n\Delta); y(t-\Delta), y(t-2\Delta), \ldots) \tag{5.24}$$

Where the regressor is represented by $u(t) = \big(u(t-\Delta), \ldots u(t-n\Delta)\big)^T \in \mathbb{R}^{q\times n}$and $y(t)$ represents the output of the system at time t ; $\Delta, 2\Delta, \ldots, n\Delta$ are the time lags of the system variables and $f(.)$ an unknown nonlinear function. The general prediction (1) computes the next value of time series $y(t)$ (output) from the past observation u and the past output y. Note that u is multidimensional and sized by the number q of neural channels. Regarding this way, the prediction becomes a function estimation problem, and so the aim of the technique is focused on the approximation of the function f, like a multi-layer perceptron (MLP). The obtained system is a compact embedded memory model called a NARX network as sketched in Figure 5.4 a first tapped delay line of inputs with delayed connections from the output to the input.

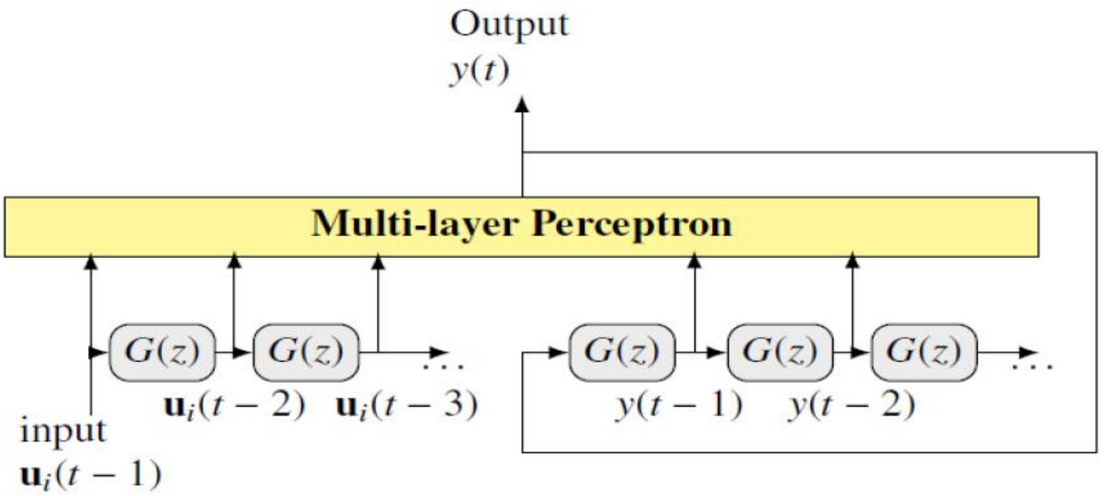


Figure 5.4: Architecture of the NARX network: Only the i^{th} regressor $u_i(t) = \left(u_i(t-1), \ldots, u_i(t-n_i)\right)^T$ is shown (among the q possible neural channels). The inputs are fully connected to the hidden layer of the MLP. Note that the transfer functions $G(z) = z^{-1}$ reduces the structure to a tap-delay-line

A Time Delay Neural Networks (TDNN) is a NARX with zero output-memory order [113] given by the form:

$$y(t) = f\left(u(t-\Delta), \ldots, u(t-n\Delta)\right) \tag{5.25}$$

This formulation is simplified but with a significant restriction of the NARX from its representational abilities like a dynamic network. Simulated results given by Lee *et al*. [114] show that NARX networks are often much better than conventional recurrent neural networks at discovering long-time dependences. An explanation of why output delays can help long-term dependences can be found by considering how gradients are calculated using gradient-descent learning algorithms: for gradient based training algorithms, the information about the gradient contribution m steps in the past vanishes for large m, i.e. mathematically $\lim_{m\to\infty} \frac{\partial y(t)}{\partial y(t-m)} = 0 \; \forall t$ where y is the state variable (and output neuron) and t the time index.

5.4.1.2. Learning algorithms

The number of inputs scales the number of parameters of the model and thus creates problems with model generalization. Moreover, the size of the training data required for good approximation increases with the number of parameters. To overcome the problem of model order, a dynamic back-propagation algorithm is required to compute the gradients, which defer the memory structure to the hidden, recurrent layer instead of time-embedding but is more computationally intensive than static back-propagation and takes more time. In addition, training is more likely to be trapped in local minima [111]. A simple back-propagation algorithm is inappropriate since the parameter convergence needs long iterations and a learning rate to adjust. Hence we propose to use the conjugate gradient algorithm (CGA) to accelerate the weight update. The weight vector of a feed-forward neural network is a point in the real Euclidean space $\mathbb{R}^N$ defined by

$$w = \left(\dots, w_{i,j}^{(l)}, w_{i+1,j}^{(l)}, \dots, w_{n_l,j}^{(l)}, \dots\right) \tag{5.26}$$

where $w_{i,j}^{(l)}$ is the weight from unit i in the layer l to unit j is layer $l+1$, n_l is the number of units in layer l. We assume a global error function $E(w)$ depending of the weight and biases attached to the network and chosen here as the standard least square function.

$$E(w) = \sum_{t=1}^{T} \big(y(t) - f(u(t-1), \dots, u(t-n); y(t-1), y(t-2), \dots; w)\big)^2 \tag{5.27}$$

$E(w)$ is evaluated with one pass forward and the gradient $E'(w)$ with one pass backward:

$$E' = \left(\dots, \sum_{t=1}^{T} \frac{\partial E(t)}{\partial w_{i,j}^{(l)}}, \sum_{t=1}^{T} \frac{\partial E(t)}{\partial w_{i+1,j}^{(l)}}, \dots, \sum_{t=1}^{T} \frac{\partial E(t)}{\partial w_{n_l,j}^{(l)}}, \dots\right)^T \tag{5.28}$$

Back propagation update rule is given by

$$w^{(k+1)} = w^{(k)} - \alpha_k E'\left(w^{(k)}\right) \tag{5.29}$$

At a time instance k, where $-E'\left(w^{(k)}\right)$ is the gradient vector and α_k is the learning rate that should be chosen carefully to make the algorithm is based on the linear approximation $E(w+h) \approx E(h) + E'(w)^T h$, the algorithm showed poor convergence. Hence, a Conjugate Gradient Algorithm (**CGA**) is generally a better choice and results in a faster convergence to reach the minimum of a $E(w)$. The optimization strategy consists to choose a search direction p_k and then deciding how far to go in the specified direction, i.e. to determine the step size α_k. **CGA** chooses the search direction and the step size by using a second-order approximation.

$$E(w+h) \approx E(w) + E'(w)^T h + \frac{1}{2} h^T E''(w) h \tag{5.30}$$

where h is the line search direction [115]. From the current search direction, the next search direction is determined so that it is conjugate to previous search directions. The critical points of $E(w)$ are the solutions of the linear system (8).

$$E_q'(h) = E''(w)^T h + E'(w)^T = 0 \tag{5.31}$$

if $E_q'\ (h)$ denote the quadratic approximation to E in the neighborhood of a point w so that $E(w+h) = E_q\ (h)$. Let the vector basis $p_1, \dots, p_N$ be a conjugate system of $\mathbb{R}^N$, the step from a starting point h_1 to a critical point h^*can be expressed as a linear combination of $p_1, \dots\ , p_N$

$$h^* - h_1 = \sum_{i=1}^{N} \alpha_i p_i \, \alpha_i \in \ \mathbb{R} \tag{5.32}$$

Multiplying (5.32) with $p_j^T E''\ (w)$ and substituting $E'\ (w)$ for $-E''(w)^T h^*$ from (5.31) gives:

$$p_j^T(-E'(w) - E''(w)h_1) = \alpha_j p_j^T E''(w) p_j \tag{5.33}$$

$$\Rightarrow \alpha_j = \ \frac{p_j^T(-E'(w) - E''(w)h_1)}{p_j^T E''(w) p_j} = \frac{-p_j^T E'(h_1)}{p_j^T E''(w) p_j} \tag{5.34}$$

The intermediate points $h_{k+1} = h_k + \alpha_k p_k$ given the iterative determination of $h^*$are in fact minima for $E(h)$ restricted to the plane P_k described by $h = h_1 + \alpha_1 p_1 + \cdots + \alpha_k p_k$. The conjugate weight vector $p_1, \dots\ , p_N$ are determined recursively:

$$p_{k+1} = r_{k+1} + \beta_k p_k \tag{5.35}$$

Where

$$r_{k+1} = -E'(h_{k+1}),\ \ \beta_k = \frac{|r_{k+1}|^2 - r_{k+1}r_k}{p_k^T r_k} \tag{5.36}$$

and $h_{k+1} = k_k + \alpha_k p_k$, $\alpha_k = \mu_k/\delta_k$, $\mu_k = p_k^T E_q'\ (h_k)$ and $\delta_k = p_k^T E''(w)p_k$. For each iteration the above described algorithm is applied to the quadratic approximation E_q of the global error function. In most cases, preconditioning is necessary to ensure fast convergence of the conjugate gradient method that takes the following form in algorithm 1.

Algorithm1: Scaled CGA
Require: choose initial weight vector w^1 (can be 0);
Ensure:$r^{(1)} = p_1 - E'(w^{(1)})$;
1: $k \leftarrow 1$
2: **repeat**
3: Calculate second-order information:
$s_k = E''(w_k)p_k$
$\delta_k = p_k^T s_k$
4: Calculate step size:
$\mu_k = p_k^T r_k$
$\alpha_k = \frac{\mu_k}{\delta_k}$
5: Update weight vector:
$w_{k+1} = w_k + \alpha_k p_k$
$r_{k+1} = -E'(w_{k+1})$
6: **If** k mod $N = 0$ then restart algorithm: $p_{k+1} = r_{k+1}$
Else create new conjugate direction:
$\beta_k = \frac{|r_{k+1}|^2 - r_{k+1}r_k}{\mu_k}$
$p_{k+1} = r_{k+1} + \beta_k p_k$
7: until the steepest descent direction $r_k \approx 0$; then set $k = k + 1$ and go to 2
Else terminate.
8: **return** w_{k+1}

The conjugate gradient algorithm works very well on well-conditioned matrices; however, in reality, most matrices are ill-conditioned, reducing the efficiency of the algorithm. Môller showed that algorithm 1 might restart when $r_{k+1} \approx r_k$ or $p_k \approx r_{k+1}$. Hence he proposed in [116] different versions of the conjugate gradient methods in regulating indefiniteness of $E''(w_k)$ with Lagrange multiplier λ_k. This is done by setting

$$s_k = \frac{E'(w_k + \sigma_k p_k) - E'(w_k)}{\sigma_k} \tag{5.37}$$

and for each iteration adjusting λ_k looking at the sign of δ_k. This results in the following algorithm 2

Algorithm2: Preconditioned CGA
Require: choose initial weight vector w^1 (can be 0) and scalars
$\sigma = 0, \lambda > 0$and $\overline{\lambda} = 0$;
Ensure: $\boldsymbol{r^{(1)} = p_1 - E'(w^{(1)})}$**;**
1: $k \leftarrow 1$ and $success = true$
2: **if**$success = true$ **then**
3: **repeat**
4: Calculate second-order information: $\sigma_k = \frac{\sigma}{|p_k|} s_k = \frac{E'(w_k+\sigma_k p_k)-E'(w_k)}{\sigma_k} \delta_k = p_k^T s_k$
5: Scales$_k$: $s_k = s_k + (\lambda - \bar{\lambda}_k) p_k \delta_k = \delta_k + (\lambda - \bar{\lambda}_k)|p_k|^2$
6: **if** $\delta_k \leq 0$**then**
make the Hessian matrix positive definite:
$s_k = s_k + \left(\lambda_k - 2\frac{\delta_k}{|p|^2}\right) p_k \bar{\lambda}_k = 2\left(\lambda_k - 2\frac{\delta_k}{|p|^2}\right) \delta_k = -\delta_k + \lambda_k |p|^2 \lambda_k = \bar{\lambda}_k$
end if
7: Calculate step size: $\mu_k = p_k^T r_k \alpha_k = \frac{\mu_k}{\delta_k}$
8: Calculate the comparison parameter: $\Delta_k = \frac{2\delta_k(E(w_k)-E'(w_k+\alpha_k p_k))}{\mu_k^2}$
9: **if**$\Delta_k \geq 0$**then**
a successful reduction in error can be made: $w_{k+1} = w_k + \alpha_k p_k r_{k+1} = -E'(w_{k+1})$
10: **if**k mod $N = 0$ **then** restart algorithm: $p_{k+1} = r_{k+1}$
11: **else** create new conjugate direction: $\beta_k = \frac{|r_{k+1}|^2 - r_{k+1} r_k}{\mu_k} p_{k+1} = r_{k+1} + \beta_k p_k$
end if
12: **if** $\Delta_k \geq 0{,}75$ **then**
reduce the scale parameter: $\lambda_k = \lambda_k/2$
13: **else**
a reduction in error is not possible : $\overline{\lambda_k} = \lambda_k$
$success = false$
end if
14: **if** $\Delta_k \leq 0{,}25$**then** increase the scale parameter: $\lambda_k = 4\lambda_k$
end if
end if
15: **until** the steepest descent direction $r_k \approx 0$; then set $k = k + 1$ and go to 2 **else** terminate.
8: **return** w_{k+1}
end if

5.4.1.3. Decoding results

We calculated and used correlation coefficients to choose channels that will be used as features of the brain machine interface decoder based on the TDNN and also to remove the redundant and non-correlated input variables with the output. Best cross-correlation values give the delay of the explanatory variables that will be used as network inputs. We used 10 delayed cortical signals from the 3 most substantial neurons (among 48 neurons having been recorded):

a. **Experiment assessment with one neuron (channel of ECoG):** First, we start our simulation with a single input as in a similar work presented in [2], where the best channel is selected with the correlation method (channel 21) which was the unique input to the TDNN but with a trial-error time regression. In this time, the best selected channel is given using the variation coefficient cv and the time shifting is defined by the template matching. Table 5.1 gives the time training and its log, and the estimation error with the corresponding neurons in the hidden layer. A single input channel with different neurons of the hidden layer, respectively (2; 4; 7; 9; 12; 15; 20; 25), we measure the training time and the estimation error.

Neurons	log(time)	time(sec)	error(mm)
2	4.8114	122.8988	8793
4	5.0279	152.6142	2779.2
7	5.4143	224.6053	28.4411
9	5.6386	281.0812	0.7123
12	5.9613	388.1077	0.7656
15	6.2793	533.4191	0.2185
20	6.7274	834.9931	0.4679
25	7.0821	1190.5	0.6063

Table 5.1: Results simulation with one neuron

b. **Experiment with multiple neurons:** For multiple neurons (channels), we use the first two and three selected neurons with respect to their order.

b.2. Two inputs (Two channels of ECoG): For the double neurons, we started our simulation using the Time Delay Neural Network with (2; 4; 7; 9; 12; 15; 20; 25; 30) neurons. The best selected neurons are the two channels 21 and 25 keeping their order. In Table 5.2 is given the training time of the TDNN

Neurons	log(time)	time($10^3 s$)	error(mm)
2	4.8519	0.1280	0.0697
4	4.9234	0.1375	0.3333
7	6.3729	0.5858	0.0979
9	6.6495	0.7724	0.2668
12	7.0341	1.1347	0.2020
15	7.6010	2.0002	0.2133
20	8.1951	3.6232	0.2535
25	8.6910	5.9490	0.2927
30	9.0956	8.9161	0.0989

Table 5.2: Results simulation with two inputs

b.1. Three inputs (Three channels of ECoG): For this case, we have three selected inputs which are the past two selected for the case of double input (21 and 25) with the channel 43 in the third rank. The TDNN was trained on half data and tested on the rest; in Table 5.3 are given results.

Neurons	log(time)	time($10^3 s$)	error(mm)
2	4.8519	0.0129	0.2713
4	5.4893	0.0242	0.0847
7	6.8829	0.0975	0.4108
9	7.4274	0.1681	0.0814
12	7.9518	0.2841	0.4487
15	8.4373	0.4616	0.1535
20	9.1152	0.9092	0.3883
25	9.6679	1.5803	0.1362
30	10.1438	2.5432	0.1375

Table 5.3: Results simulation with three inputs

c. Discussion: The training was terminated when the cross validation error continuously increased for more than 10 steps. The network specifications are listed in Table 5.4. While the choice for the right architecture is mainly intuitive and implies arbitrary decisions, an attempt to apply ANN directly fails due to the dimensionality of the inputs. Therefore the dimension of the inputs has been reduced drastically by feature selection.

Architecture	#parameters	CGA		Pre-conditioned CGA	
		x	y	x	Y
27-2-2	62	3.691	4.723	2.700	3.720
27-3-2	92	3.394	**0.426**	**1.406**	3.435
27-4-2	122	**2.732**	2.822	2.762	**2.799**
27-5-2	152	5.066	5.148	3.089	5.132
27-6-2	182	5.105	6.162	8.117	6.133
27-7-2	212	6.531	9.570	11.541	9.550

Table 5.4: channel ranking obtained from a range of descriptors

The change in the number of units in the TDNN affects the decoding performance as shown in Table 5.4, the mean AIC across the 50 data sets is used to quantify the performance of each model for the two decoding algorithms. For clarity here, we have 7 peaks which represent 10% of data. In Figure 5.5, by the red line is described the model output. It can be seen that the hand vertical coordinate on a plan is tracked and the trajectory is well reconstructed along the experiences. When movements are sufficiently complex (presence of peaks), the algorithm shows the ability to catch well these peaks as shown by a zoom onto the fit in Figure 5.5.

The **Akaike information criterion** (**AIC**) is an estimator of prediction error and thereby the relative quality of statistical models for a given set of data. Given a collection of models for the data, AIC estimates the quality of each model, relative to each of the other models. Thus, AIC provides a means for model selection. The Akaike information criterion is used for the measurement of performances. As a mean for model selection, the AIC is used. The AIC is calculated $AIC = 2p - 2\ln(L)$, where p represents the number of model's parameters, and L denotes the maximized value of the likelihood model's function. AIC deals with the compromise between the quality of the model's fit and its complexity. The best model is characterized with the lowest AIC value.

When using in decoding a reasonable number of units (more than three), the preconditioned CGA shows its superiority to the classic CGA. In Figure 5.5 is illustrated the result obtained from the MLP 27-5-2, where the trial-error technique is used for the MLP training. The performance of the MLP 27-5-2 was much worse than that of the CGA. The performance differences between the CGA and preconditioned CGA were larger than 5 units.

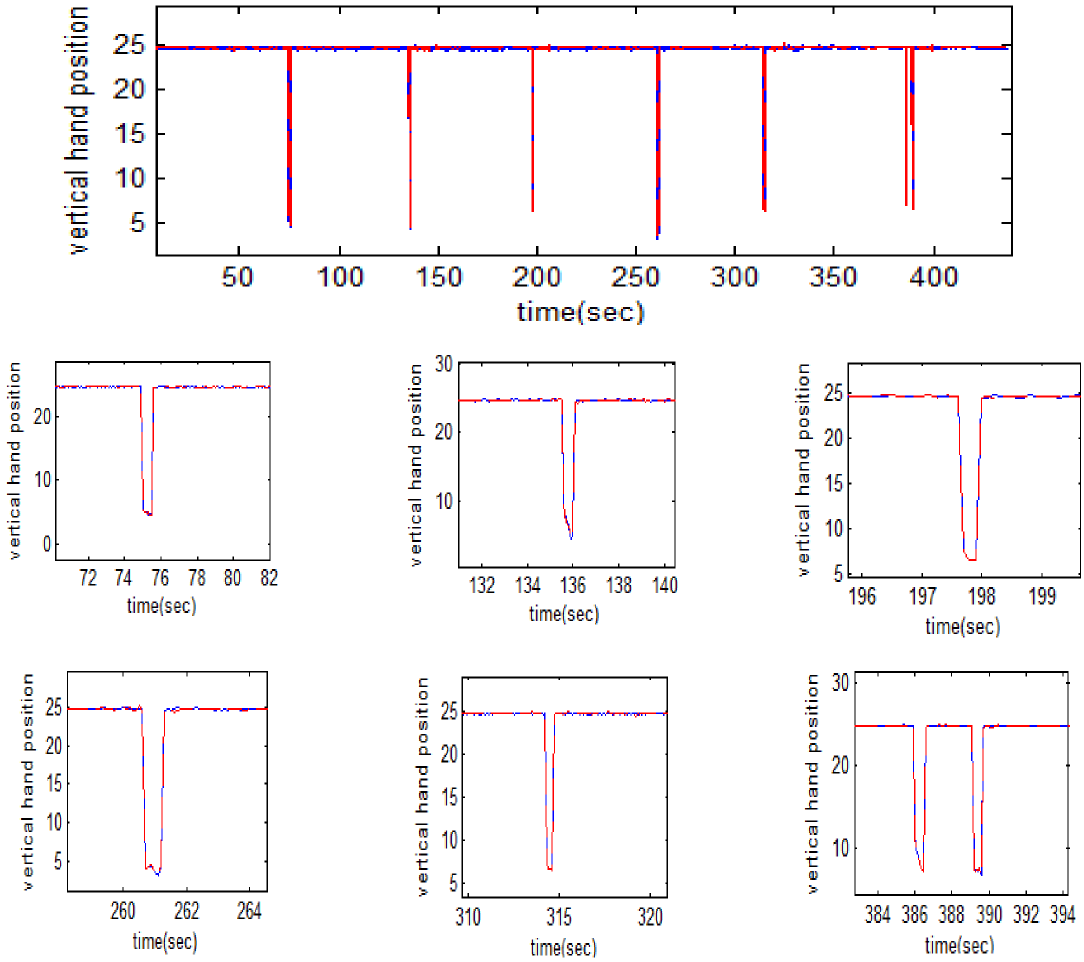


Figure 5.5: Decoding hand trajectory using TDNN

5.4.2. TDNN Aided Kalman Filter in Hand Trajectory Decoding

Time delay neural network can be used for aiding the Kalman filter by identifying the adjustable state's and measurement's coefficients where the Kalman filter is used only to decode the vertical hand trajectory of the rat using brain signals. The Kalman filter here should not be extended to the parameters because the TDNN estimates them and the Kalman filter will be used as a simple observer applied to estimate the state of the process. Many procedures were used combining ANN and KF. The true definition of this technique is generally “ANN aided Kalman filter” where some authors use the ANN to calibrate the sensor measurement in first, and take them into the Kalman filter to improve its performance [117, 118]. Others employ the ANN to approximate the uncertainty of the model [119].

In our work, the TDNN estimates the state's and measurement's variables at the same time where the Kalman filter uses them in hand trajectory decoding. The same algorithm presented in the Kalman filter is used but applied with variable states and measurement matrices provided by the TDNN.

Results of hand trajectory are exhibited in figure 5.6. The TDNN aided the Kalman filter shows its ability for the best hand trajectory decoding from primary motor cortex ECoG.

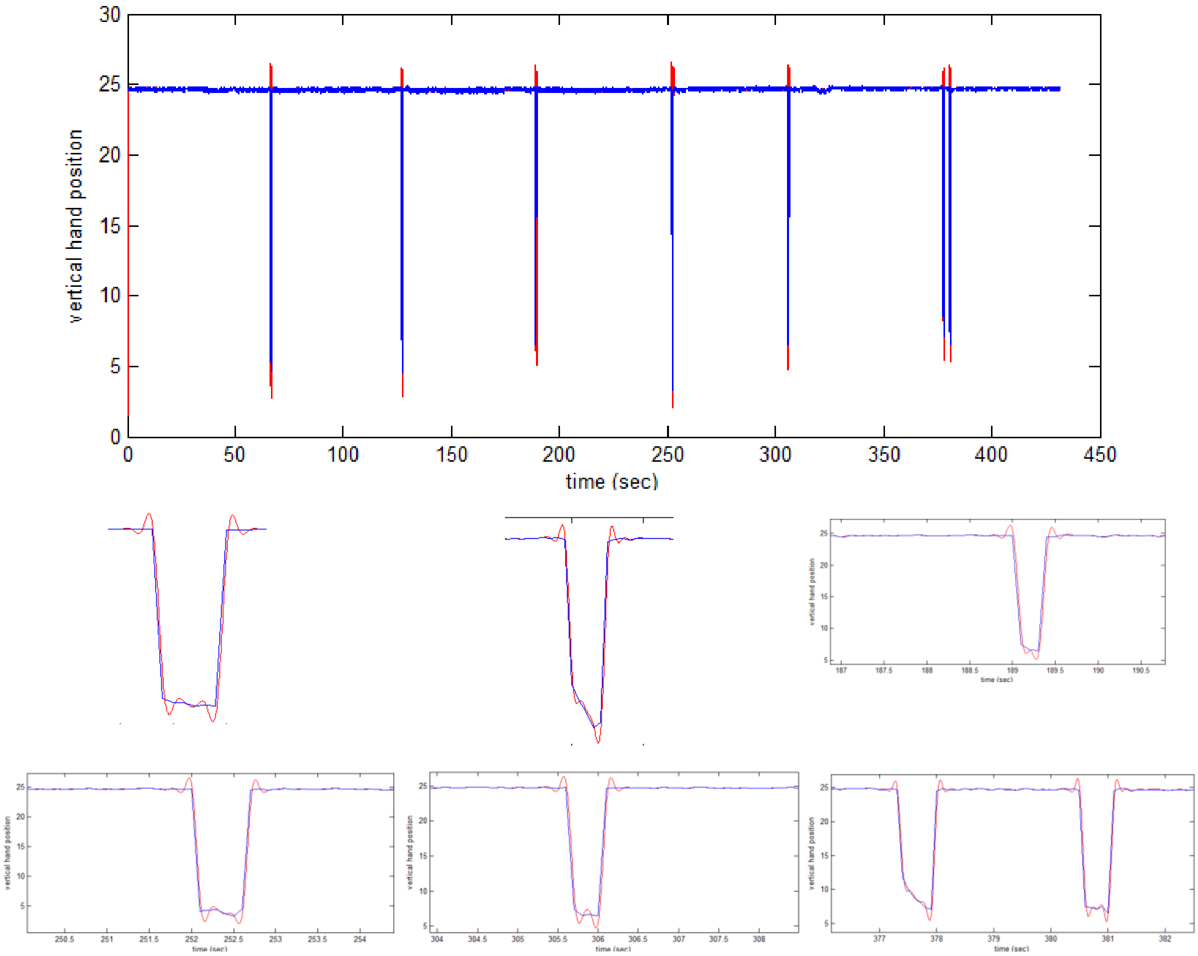


Figure 5.6: Results of decoding hand trajectory using TDNN aided the KF.

5.5. Conclusion

Linear decoding of the hand trajectory of the rat from its primary motor cortex ECoG by using the linear regressor and linear Kalman filter was poorer and the tracking was weak. This improves the nonlinearity of the mapping related to the hand trajectory to the brain signals. Nonlinear decoding introducing the time delay neural network alone and aiding the Kalman filter shows its ability to detect events and reconstruct a good hand trajectory.

The TDNN architecture was proposed for fast hand trajectory decoding from the primary motor cortex ECoG. Since the application of TDNN in BMIs yields too many free parameters, the approach performs faster with the conjugate gradient algorithm. However, a TDNN with 20 hidden neurons and 10 time delays of 46 channels of neuronal signal produces more than 1,160 free parameters and we only have 10,000 samples for training and cross-validation. Another issue is the possibility to use a Kalman filter for detecting pulses and grouping them, hence simplifying the activation measurement and this appears in the TDNN aiding the Kalman filter.

CONCLUSION

During this thesis work, we were interested in an invasive BCI-based system applied for decoding fast hand movement trajectory using ElectroCorticoGraphy (ECoG) signal recorded from the primary motor cortex. We started this work under the joint supervision of Mr. A. Guessoum, professor at the University of Blida and Mr. V. Vigneron, professor at the University d'Evry Val d'Essone, France. This work is complementary to the work of M. Vigneron in collaboration with Hsin Chen from the University of Tsing Hua Taiwan and biomedical engineering of the Iran University. It does a part of the project PRONEUR. One of our major contributions is the realization of the decoding process from brain signal before the application on real systems which does complement the last works realized before in the same topic and the same application on rat's hand.

Brain computer interface is the way for making a direct communication pathway between the brain and the external world (devices) of the body in order to assisting, augmenting, and repairing human cognitive or sensory motor functions. Brain computer interface improves its ability to decode brain signals following four main stages started from the brain signal recording to the application through the two middle parts of spikes sorting and task decoding. Recording brain signals can be invasively my measuring electrical fields by a direct implementation of sensors on the brain with surgery where sensors can be on the surface of the cortex to record the ECoG or deeper for single neuron recording. The other technique of brain signal is the non-invasive brain signal recording which measures electrical fields for EEG or magnetic fields for NIRS, MEG, and fMRI. The best signal that can be used as a signal taking into account the health of the user, economy, and portability is the ECoG and EEG. The EEG is more secured for health but with less spatio-temporal precision. The used signal in this work is the ECoG recorded from the primary motor cortex of a free moving rat in a box with the instantaneous hand actions on a lever caption with a camera.

The recorded ECoG is captured using microelectrodes implanted on the surface of the rat's cortex. Each electrode measures the electrical activity of such a group of neighbor neurons with different forms of their action potentials. The action potentials of the same neuron have a similar shape if we don't talk about noises and the stability of electrodes over recording time. Spikes-sorting allows the separation and classification of these action potentials to several groups of corresponding shapes. Each group represents a neuron. The spikes sorting is realized using two technics based on (i) self-training semi-supervised support vector machine which makes the classifier takes a small time for training, this method was used in an offline training and classification. The second technique was the (ii) online spikes detection and sorting with no features extraction where this method used the template of each detected channel as a reference for extraction and classification. This technique was also able to detect new channels that appear recently and eliminate absent neurons.

Spikes sorting applied on all recorded brain signals from the primary motor cortex of the rat provides 48 neural channels represented by sequences of ones "1" separated by long suites of zeros "0". Each "1" in a channel represent the spike arrival from the corresponding neuron or it means that the neuron fire at that moment. Information in spikes trains is hidden in the distance between successions on ones and especially in the temporal distance. A technique making these distances more clear is to translate them into the amplitude of the signal by using a correlation with a moving window of 100ms width.

The number of inputs of such decoder scales the number of its parameters that makes the processing more difficult with a need of big training data for a good approximation of all parameters. To outcome the problem of the model order, we proposed two main technics for data reduction. We used the correlation coefficients and the correlation functions respectively to reduce the number of redundant channels by minimizing channels involving the same hidden information and eliminate the irrelevant channels with no relation to the hand position.

Template matching was used to determine the maximal time-lag for the recurrence of the decoder and the classification of the best channel can be used as an input of the decoder. The used classification is based on the measure of performance of these brain signal detection techniques by computing the sensitivity and the specificity of the channels for a range of descriptors.

Decoding algorithms are series of computations that decode the rat intentions (actions on the lever to bring water) from the best selected inputs derived by the feature selection with the order of priority of each channel. Decoding the hand trajectory from the primary motor cortex ECoG was realized using linear and nonlinear techniques. The linear regressor was used to decode the hand trajectory from the best ECoG. The nonlinearity of the mapping limits the ability of the linear regressor where results were so poorer with no resemblance to the real hand trajectory. The extended Kalman filter was also proposed to decode the hand trajectory from the brain signals where the used model assumes the hand coordinates as state variables and the brain signal as the output. The model is extended to the states and measures parameters. The Kalman filter uses the output (brain signal) to estimate the states of the model (hand motions). Decoding the hand trajectory using the Kalman filter was better than the linear regressor but it was still weak. If we seek to detect the actions on the lever, the Kalman filter answers the need with bad tracking of the real hand position.

Neural networks with all their versions were all-time considered as black boxes able to model systems with high nonlinearity. The time delay neural network was proposed to decode the hand trajectory with fast movements from the ECoG recorded from the primary motor cortex of the rat. Results obtained using the TDNN were better than all other proposed techniques and with one channel we were able to follow the hand position in real time. The problem with the use of the TDNN is situated in the number of parameters. However, a TDNN with 20 hidden neurons and 10-time delays of 46 channels of neuronal signal produces more than 1,160 free parameters and we only have 10,000 samples for training and cross-validation. Another issue is the possibility to use a Kalman filter for detecting pulses and grouping them, hence simplifying the activation measurement and this appears in the TDNN aiding the Kalman filter.

Future work:

The work presented in this thesis can be continued where several paths can be envisaged as prospects for continuation. In first, the choice of the used signals and as we work on signals come from animals with more or less intelligence and consciousness than those of human, the practice on human beings and more efficient with more intelligence to human response moments with the machine.

On the other hand, in the part of the separation of spikes, the use of a fixed threshold allows false alarms of the existence of spikes whereas it is necessary to think carefully about the application of an adaptive threshold for the localization of spikes so as not to get soaked in false alarms for the presence of spikes or a loss of spikes with a very low NEO.

Deep learning can be applied for the separation of spikes and decoding because its networks have the ability to apply in large systems and embedded systems for several successive tasks including feature extraction, decoding, and decision making.

EFERENCES

1. **A. Ghanbari, M. B. Shamsollahi, V. Vigneron, A. Kifouche. "Neural spike sorting with a self-training semi-supervised support vector machine".** ***35th Annual International Conference of Engineering in Medicine and Biology Society (EMBC 2013)*****, Jul 2013, Osaka, Japan. pp.6007—6010.**
2. **A. Kifouche, V. Vigneron, M. B. Shamsollahi, A. Guessoum, "Decoding Hand Trajectory from Primary Motor Cortex ECoG Using Time Delay Neural Network. In Mladenov V., Jayne C., Iliadis L. (eds) Engineering Applications of Neural Networks. EANN 2014. Communications in Computer and Information Science, vol 459. Springer, Cham.**
3. W. Wu, M. Black, Y. Gao, E. Bienenstock, M. Serruya, and J. Donoghue. Inferring hand motion from multi-cell recordings in motor cortex using a kalman filter. In *Workshop on Motor Control in Humans and Robots: On the Interplay of Real Brains and Artificial Devices*, pages 1–8, 2002.
4. G. J. Gage, K. J. Otto, K. Ludwig, and D. R. Kipke. Co-adaptive kalman filtering in a naïve rat cortical control task. In *IEEE Conference on Engineering in Medicine and Biology Conference*, volume 6, pages 4367–70, 2004.
5. E. Brockwell, L. Rojas, and R. E. Kass. Recursive Bayesian decoding of motor cortical signals by particle filtering. *Journal of neurophysiology*, 91(4):1899–907, Apr. 2004.
6. **A. Kifouche, A. Guessoum. "Tracking times in temporal patterns embodied in intra-cortical data for controlling neural prosthesis an animal simulation study."** ***International Journal of Electrical and Computer Engineering*** **vol. 10, No. 5, October 2020, pp. 4721~4737.**
7. Brown, E., Frank, L., Tang, D., Quirk, M., and Wilson, M., "A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells," The Journal of Neuroscience, vol. 18, no. 18, pp. 7411-7425, 1998.
8. Tudor M, Tudor L, Tudor KI. Hans Berger (1873-1941)--povijest elektroencefalografije [Hans Berger (1873-1941)--the history of electroencephalography]. Acta Med Croatica. 2005;59(4):307-13. Croatian. PMID: 16334737.

9. Berger, H. Über das Elektrenkephalogramm des Menschen. Archiv f. Psychiatrie 87, 527–570 (1929).
10. Takahashi, S., Anzai, Y., and Sakurai, Y., “Automatic sorting for multi-neuronal activity recorded with tetrodes in the presence of overlapping spikes,” Journal of Neurophysiology, vol. 89, no. 4, pp. 2245-2258, 2002.
11. Wessberg, J., et al., “Realtime prediction of hand trajectory by ensembles of cortical neurons in primates,” Nature, vol. 408, no. 6810, pp. 361-365, Nov 2000.
12. Dickey, A., Suminski, A., Amit, Y., and Hatsopoulos, N., “Single-unit stability using chronically implanted multielectrode arrays,” Journal of neurophysiology, vol. 102, no. 2, pp. 1331-1339, 2009.
13. Vigneron, V., and Chen, H., “Sparse data analysis strategy for neural spike classification,” Computational Intelligence and Neuroscience, 2014.
14. M.T. Kani, A.H. Jafari, A. Khoshnevisan, H. Arabalibeyk, and M. J. Abolhasani. Neuronal spike sorting based on radial basis function neural networks. TUMJ, 68(11):638–643, 2011
15. K.H. Kim and S.J. Kim. Neural spike sorting under nearly 0-db signal-to-noise ratio using nonlinear energy operator and artificial neural-network classifier. IEEE Transactions on Biomedical Engineering, 47:1406–1411, 2000.
16. C. Yang, Y. Yuan, and J. Si. High performance spike detection and sorting using neural waveform phase information and SOM clustering. In International Joint Conference on Neural Networks, pages 1–7, 2010.
17. A. Madanymamlouk, H. Sharp, K. Menne, U. Hofmann, and T. Martinetz. Unsupervised spike sorting with ica and its evaluation using GENESIS simulations. Neurocomputing, 65-66:275–282, 2005
18. C. Pouzat, M. Delescluse, P. Viot, and J. Diebolt. Improved spike-sorting by modeling firing statistics and burst-dependent spike amplitude attenuation: A markov chain monte carlo approach. J Neurophysiol, 91,:2910–2928, 2004.
19. Kostov, A., Andrews, B., Popovic, D., Stein, R., and Armstrong, W., “Machine learning in control of functional electrical stimulation (fes) for locomotion,” IEEE Trans. Biomed. Eng., vol. 42, pp. 541-551, 1995.
20. J.R. Wolpaw, G.E. Loeb, B.Z. Allison et al., BCI Meeting 2005 – workshop on signals and recording methods, IEEE Trans Neural Syst Rehabil Eng: A Pub IEEE Eng Med Biol Soc. 14, Jun., 138–141, (2006).
21. G. Dornhege, J.D.R. Millan, T. Hinterberger, D.J. McFarland, K. Müller, and T.J. Sejnowski, *Toward Brain-Computer Interfacing*, The MIT Press, Cambridge, MA, (2007).

22. Blankertz B, Dornhege G, Krauledat M, Müller K-R, KunzmannK-r V, Losch F, and Curio G. ”The berlin brain computer interface : Eeg-based communication without subject training.”, 2006.
23. Borisoff JF, Mason S G, and Birch G E. ”Classification of imaginary movements in ECoG with a hybrid approach based on multi-dimensional Hilbert-SVM solution.”. IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society, 14(2): 160–4, June 2006.
24. Wolpaw. ”Brain computer interfaces for communication and control”. Clinical Neurophysiology, 113(6) :767–791, June 2002.
25. Sutter EE. The brain response interface: communication through visually induced electrical brain responses. J Microcomput Appl 1992;15:31–45.
26. Sutter EE, Tran D. The field topography of ERG components in man. Vision Res 1992;32:433–446.
27. Hoffmann U, Mesin V-J, Ebrahimi T, and Diserens K. ”An efficient P300-based brain-computer interface for disabled subjects.”. Journal of neuroscience methods, 167(1) :115–25, January 2008.
28. Farwell L and Donchin E. ”Talking off the top of your head : toward a mental prosthesis utilizing event-related brain potentials.”. Electroencephalography and Clinical Neurophysiology, 70(6) :510–523, December 1988.
29. Sarah Paige Gibson, Neural Spike Sorting in Hardware: From Theory to Practice, PHD thesis. University of California, Los Angeles 2012.
30. Fraser, George & Chase, Steven & Whitford, Andrew & Schwartz, Andrew. (2009). Control of a brain-computer interface without spike sorting. Journal of neural engineering. 6. 055004. 10.1088/1741-2560/6/5/055004.
31. Ventura, Valérie. (2008). Spike Train Decoding Without Spike Sorting. Neural computation. 20. 923-63. 10.1162/neco.2008.02-07-478.
32. Rockstroh B., Elbert T., Lutzenberger W., Birbaumer N. “Operant Control of Slow Brain Potentials: A Tool in the Investigation of the Potential’s Meaning and Its Relation to Attentional Dysfunction”, (1984).
33. Elbert T., Rockstroh B., Lutzenberger W., Birbaumer N. (eds) Self-Regulation of the Brain and Behavior. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-69379-3_16.
34. McFarland, D J., Lefkowicz, T., Wolpaw, J R. “Design and operation of an EEG-based brain-computer interface (BCI) with digital signal processing technology,” Behavioral Research Methods Instruments and Computers, vol. 29, pp. 337-345, 1997a.

35. Jain AK, Duin RPW, and Mao J, 2000. Statistical pattern recognition: a review. IEEE Trans Patt Anal Mach Intell 22:4–37.
36. Wolpaw JR, McFarland DJ, Neat GW, Forneris CA. An EEG-based brain computer interface for cursor control. Electroenceph clin Neurophysiol 1991; 78:252-259.
37. Perelmouter J, Kotchouby B, Kübler A, Taub E, Birbaumer N. Language support program for thought-translation devices. Automedica 1999;18:67–84.
38. Finke, A, Lenhardt, A, & Ritter, H. The MindGame: A P300-based brain-computer interface game," Neural Networks, November (2009). , 22(9), 1329-1333.
39. Citi, L, Poli, R, Cinel, C, & Sepulveda, F. P300-Based BCI Mouse With Genetically-Optimized Analogue Control," IEEE Trans. Neural Syst. Rehabil. Eng., February (2008). , 16(1), 51-61.
40. Sarah N. Abdulkader, Ayman Atia, Mostafa-Sami M. Mostafa, Brain computer interfacing: Applications and challenges, Egyptian Informatics Journal, Volume 16, Issue 2, 2015, Pages 213-230,
41. Hanafiah ZM, Taib MN, Hamid N. Eeg pattern of smokers for theta, alpha and beta band frequencies. In: Research and Development (SCOReD), 2010 IEEE Student Conference on. IEEE; 2010. p. 320–23.
42. Malar E. A novel approach for the detection of drunken driving using the power spectral density analysis of EEG. Int J Comput Appl 2011; 21.
43. Shooshtari MA, Setarehdan SK. Selection of optimal eeg channels for classification of signals correlated with alcohol abusers. In: Signal Processing (ICSP), 2010 IEEE 10th International Conference on. IEEE; 2010. p. 1–4.
44. Fan X, Bi L, Wang Z. Detecting emergency situations by monitoring drivers' states from eeg. In: Complex Medical Engineering (CME), 2012 ICME International Conference on. IEEE; 2012. p. 245–48.
45. Selvam VS, Shenbagadevi S. Brain tumor detection using scalp eeg with modified wavelet-ica and multi layer feed forward neural network. In: Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE. IEEE; 2011. p. 6104–09.
46. Sharanreddy M, Kulkarni P. Detection of primary brain tumor present in eeg signal using wavelet transform and neural network. Int J Biol Med Res 2013;4(1).
47. Poulos M, Felekis T, Evangelou A. Is it possible to extract a fingerprint for early breast cancer via eeg analysis? Med Hypotheses 2012;78(6):711–6.
48. Sharanreddy M, Kulkarni P. Automated eeg signal analysis for identification of epilepsy seizures and brain tumour. J Med Eng Technol 2013;37(8):511–9.

49. Liang S-F, Shaw F-Z, Young C-P, Chang D-W, Liao Y-C. A closed-loop brain computer interface for real-time seizure detection and control. In: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE; 2010. p. 4950–53.
50. Jan B.F. Van Erp, Fabien Lotte, Michael Tangermann. Brain-Computer Interfaces: Beyond Medical Applications. Computer -IEEE Computer Society-, IEEE, 2012, 45 (4), pp.26-34.
51. Domingo MC. An overview of the internet of things for people with disabilities. J Netw Comput Appl 2012;35(2):584–96.
52. Lin C-T, Lin B-S, Lin F-C, Chang C-J. Brain computer interface-based smart living environmental auto-adjustment control system in UpnP home networking. IEEE Syst J 2012
53. Navarro AA, Ceccaroni L, Velickovski F, Torrellas S, Miralles F, Allison BZ, R.Scherer R, Faller J. Context-awareness as an enhancement of brain-computer interfaces. Ambient assisted living. Springer; 2011
54. Venthur B, Blankertz B, Gugler MF, Curio G. Novel applications of BCI technology: psychophysiological optimization of working conditions in industry. In: Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on. IEEE; 2010. p. 417–421.
55. Vecchiato G, Astolfi L, De Vico Fallani F, Salinari S, Cincotti F, Aloise F, Mattia D, Marciani MG, Bianchi L, Soranzo R. The study of brain activity during the observation of commercial advertising by using high resolution EEG techniques. In: Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE. 2009. p. 57–60.
56. Yoshioka M, Inoue T, Ozawa J. Brain signal pattern of engrossed subjects using near infrared spectroscopy (nirs) and its application to tv commercial evaluation. In: Neural Networks (IJCNN), The 2012 International Joint Conference on. IEEE; 2012. p. 1–6.
57. Vecchiato G, Babiloni F, Astolfi L, Toppi J, Cherubino P, Dai J, Kong W, Wei D. Enhance of theta EEG spectral activity related to the memorization of commercial advertisings in chinese and italian subjects. In: Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on, vol. 3. IEEE; 2011. p. 1491–94.
58. Spüler, M., Krumpe, T., Walter, C., Scharinger, C., Rosenstiel, W., and Gerjets, P. (2017). “Brain-computer interfaces for educational applications,” in Informational Environments (Cham: Springer), 177–201. doi: 10.1007/978-3-319-64274-1_8.

59. Royer AS, Doud AJ, Rose ML, He B. Eeg control of a virtual helicopter in 3-dimensional space using intelligent control strategies. Neural Syst Rehabilitation Eng, IEEE Trans 2010;18(6):581–9.

60. J. Chuang, H. Nguyen, C. Wang, and B. Johnson, “I think, therefore i am: Usability and security of authentication using brainwaves,” in Financial Cryptography and Data Security, 2013, pp. 1–16.

61. Frigeri, T.; Paglioli, E.; De Oliveira, E.; Rhoton Jr, A. L. (2015). "Microsurgical anatomy of the central lobe". Journal of Neurosurgery. 122 (3): 483–98. doi:10.3171/2014.11.JNS14315.

62. D. TRAD, “Interface Cerveau-Machine, Application au handicap moteur”, Thèse Doctorat, October 7, 2012, Université de Versailles Saint-Quentin en Yveline.

63. Wolf U, Rapoport MJ, Schweizer TA (2009). "Evaluating the affective component of the cerebellar cognitive affective syndrome". Journal of Neuropsychiatry and Clinical Neurosciences. 21 (3): 245–53.

64. Freund HJ. Mechanisms of voluntary movements. Parkinsonism Relat Disord. 2002 Oct;9(1):55-9. doi: 10.1016/s1353-8020(02)00044-5. PMID: 12217622.

65. G. Bauernfeind, R. Leeb, S.C. Wriessnegger, and G. Pfurtscheller, Development, set-up and first results for a one-channel near-infrared spectroscopy system, Biomedizinische Technik, 53, 36–43, (2008).

66. Vadim S Polikov, Patrick A Tresco, and William M Reichert, “Response of brain tissue to chronically implanted neural electrodes.” Journal of neuroscience methods, vol. 148, no. 1 pp. 1–18, Oct 2005.

67. J N Turner, W Shain, D H Szarowski, M Andersen, S Martins, M Isaacson, and H Craighead, “Cerebral astrocyte response to micro-machined silicone implants.” Experimental neurology, vol. 156, no. 1 pp. 33–49, Mar 1999.

68. Waldert, S.; Pistohl, T.; Braun, C.; Ball, T.; Aertsen, A.; Mehring, C. A review on directional information in neural signals for brain-machine interfaces. *J. Phisiol. Paris* 2009, *103*, 244–254.

69. deCharms, R.C.; Christoff, K.; Glover, G.H.; Pauly, J.M.; Whitfield, S.; Gabrieli, J.D.E. Learned regulation of spatially localized brain activation using real-time fMRI. *Neuroimage* 2004, *21*, 436–443.

70. Weiskopf, N.; Mathiak, K.; Bock, S.W.; Scharnowski, F.; Veit, R.; Grodd, W.; Goebel, R.; Birbaumer, N. Principles of a brain-computer interface (BCI) based on real-time functional magnetic resonance imaging (fMRI). *IEEE Trans. Biomed. Eng.* 2004, *51*, 966–970.

71. Van Staveren, G., Buitenweg, J., Heida, T., Ruitten, W.: Wave shape classification of spontaneaous neural activity in cortical cultures on micro-electrode arrays. In: Proceedings 24th Annual Conference of the EMBS/BMES society, vol. 3, pp. 2010–2011. TX, USA (2002).
72. R. Q. Quiroga, Z. Nadasdy, and Y. Ben-Shaul, “Unsupervised spike detection and sorting with wavelets and super-paramagnetic clustering”, Neural Computation, Vol. 16, no. 8, pp. 1661-87, Aug. 2004.
73. J. F. Kaiser, “On a simple algorithm to calculate the `energy' of a signal”, in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP '90), vol. 1, Albuquerque, NM, 1990, pp. 381-384.
74. R. S. Holambe and M. S. Deshpande, “Nonlinear measurement and modeling using Teager energy operator,” in Advances in Non-Linear Modeling for Speech Processing. SpringerBriefs in Electrical and Computer Engineering, pp. 45–59, Springer Science & Business Media, Berlin, Germany, 2012.
75. E. Hulata, R. Segev, Y. Shapira, M. Benveniste, and E. Ben-Jacob, “Detection and Sorting of Neural Spikes Using Wavelet Packets”, Phys. Rev. Lett., vol. 85, no. 21, pp. 4637-4640, 2000.
76. R. J. Brychta, S. Tuntrakool, M. Appalsamy, N. R. Keller, D. Robertson, R. G. Shiavi, and A. Diedrich, \Wavelet Methods for Spike Detection in Mouse Renal Sympathetic Nerve Activity," IEEE Trans. Biomed. Eng., vol. 54, no. 1, pp. 82-93, 2007.
77. Mukhopadhyay S, Ray G (1998) A new interpretation of nonlinear energy operator and its efficacyin spike detection. IEEE Trans Biomed Eng 45(2):180–187.
78. Donoho, D., Johnstone, I. M. Ideal spatial adaptation by wavelet shrinkage. Biometrika (1994), 81, 425–455.
79. R. Chandra and L. M. Optican, “Detection, classification, and superposition resolution of action potentials in multiunit single-channel recordings by an online real-time neural network,” IEEE Trans. Biomed. Eng., vol. 44, no. 5, pp. 403-412, May 1997.
80. J. H. Choi, H. K. Jung, and T. Kim, “A New Action Potential Detector Using the MTEO and Its Effects on Spike Sorting Systems at Low Signal-to-Noise Ratios,” IEEE Trans. Biomed. Eng., vol. 53, no. 4, pp. 738-746, 2006.
81. R. Jacob Vogelstein, K. Murari, P. Thakur, C. Diehl, S. Chakrabartty and G. Cauwenberghs, “Spike sorting with support vector machines”, Annual International Conference of the IEEE Engineering in Medicine and Biology Society., vol. 1, pp. 546-9, Jan. 2004.

82. W. Ding and J. Yuan, “Spike sorting based on multiclass support vector machine with superposition resolution”, Medical & biological engineering & computing, vol. 46, no. 2, pp. 139-45, Feb. 2008.

83. Y. Li, C. Guan, H. Li, and Z. Chin. A self-training semi-supervised svm algorithm and its application in an eeg based brain computer interface speller system. Pattern Recognition Letters, 29(9):1285–1294, 2008.

84. Corinna Cortes and Vladimir Vapnik. 1995. Support-Vector Networks. Mach. Learn. 20, 3 (Sept. 1995), 273–297. DOI: https://doi.org/10.1023/A:1022627411411.

85. Y. Jin, C. Huang and L. Zhao, “A Semi-Supervised Learning Algorithm Based on Modified Self-Training SVM”, Journal of Computers, vol. 6, no. 7, pp. 1438-1443, Jul. 2011.

86. Chen, Z., Haykin, S., Eggermont, J., Becker, S.: Correlative Learning: A Basis for Brain and Adaptive Systems. John Wiley & Sons (2008)

87. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. Journal of Machine Learning Research pp. 1157–1182 (2003).

88. Shakil, M., Elshafei, M., Habib, M., Maleki, F.: Soft sensor for nox and o2 using dynamic neural networks. Computers & Electrical Engineering 35(4), 578–586 (2009)

89. Sameni, R., Vrins, F., Parmentier, F., Herail, F., Vigneron, V., Verleysen, M., Jutten, C., Shamsollahi, M.: Electrode selection for noninvasive fetal electrocardiogram extraction using mutual information criteria. In: MaxEnt2006 proceedings - 26th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, vol. 872, pp. 97–104 (2006).

90. Vrins, F., Lee, J., Verleysen, M., Vigneron, V., Jutten, C.: Improving independent component analysis performances by variable selection. In: NNSP’2003 proceedings - Neural Networks for Signal Processing, pp. 359–368. Toulouse (France) (2003).

91. Sato T, Suzuki T, Mabuchi K. Fast automatic template matching for spike sorting based on Davies-Bouldin validation indices. *Annu Int Conf IEEE Eng Med Biol Soc. 2007; 2007():3200-3.*

92. Friedman DH. Detection of Singals by template matching. Baltimore: Johns Hopkins Press; 1968.

93. Kim S, McNames J. Automatic spike detection based on adaptive template matching for extracellular neural recordings. *J Neurosci Methods. 2007 Sep 30; 165(2):165-74.*

94. Y. Salimpour and Y. Soltanian-Zadeh. Particle filtering of point processes observation with application on the modeling of visual cortex neural spiking activity. In *4th International IEEE/EMBS Conference on Neural Engineering*, pages 718–721, Apr. 2009.

95. K. H. Kim and S. S. J. Kim. Advantage of support vector machine for neural spike train decoding under spike sorting errors. In *Proceedings of the 27th Engineering in Medicine and Biology Conference*, volume 5, pages 5280–5283, 2005.

96. U. Eden, W. Truccolo, M. Fellows, J. Donoghue, and E. Brown. Reconstruction of hand movement trajectories from a dynamic ensemble of spiking motor cortical neurons. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 6, pages 4017–20, 2004.

97. X. Zheng, Q. Zhang, S. Zhang, Y. Yu, W. Chen, and Y. Zhao. Real-time decoding algorithm in brain machine interface. In *IEEE/ICME International Conference on Complex Medical Engineering*, pages 79–84, Gold Coast, Australia, July 2010.

98. W.Wu, A. Shaikhouni, J. P. Donoghue, andM. J. Black. Closed-loop neural control of cursor motion using a kalman filter. In *26th Annual International Conference of the IEEE Engineering inMedicine and Biology Society*, volume 6, pages 4126–9, 2004.

99. Wu W, Kulkarni JE, Hatsopoulos NG, Paninski L. Neural decoding of hand motion using a linear state-space model with hidden states. IEEE Trans Neural Syst Rehabil Eng. 2009

100. F. Rosenblatt. The perceptron, a probabilistic model for information storage and organization in the brain. *Psychol. Rev*, 65:386–408, 1958.

101. B.Widrow and M.E. Hoff. Adaptive switching circuits. In *IRE WESCON Convention Record*, volume 4, pages 96–104, August 1960.

102. J.J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. In *Proc. Natl. Acad. Sci*, volume 79, pages 2554–2558, 1982.

103. T. Kohonen. Self-organized formation of topologically correct feature maps. *Biol. Cybern.*, 43:59–69, 1982.

104. D.H. Ackley, G.E. Hinton, and T.J. Sejnowski. A learning algorithm for boltzmann machines. *Cogn. Sci.*, 9:147–169, 1985.

105. D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Learning representations by backpropagating errors. *Nature*, 323:533–536, 1986.

106. A. Freeman and D. M. Skapura. *Neural Networks - Algorithms, Applications, and Programming Techniques*. Addison–Wesley, New York, USA, 1991.

107. S.I. Gallant. *Neural network learning and expert systems*. The MIT Press, Massachusetts, USA, 1993.

108. M. Smith. *Neural networks for statistical modelling*. an Nostrand Reinhold, NY, USA, 1994.
109. B.D. Ripley. Neural networks and related methods for classification. *J. R. Stat. Soc.*, B56(53):409–456, 1994.
110. M.C. Bishop. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford, UK, 1995.
111. Hagan, M., De Jesus, O., Schultz, R.: Training recurrent networks for filtering and control. In: L. Medsker, E. L.C. Jain (eds.) chap. 12 in Recurrent Neural Networks: Design and Applications, pp. 311–340. CRC Press (1999).
112. Narendra, K., Mukhopadhyay, S.: Adaptive control using neural networks and approximate models. *IEEE* Transactions on Neural Networks 8, 475–485 (1997).
113. Lang, K., Waibel, A., Hinton, G.: A time delay neural network architecture for isolated word recognition. Neural Networks 3(1), 23–44 (1990).
114. Tsung-Nan Lin, C. L. Giles, B. G. Horne and Sun-Yuan Kung, "A delay damage model selection algorithm for NARX neural networks," in *IEEE Transactions on Signal Processing*, vol. 45, no. 11, pp. 2719-2730, Nov. 1997.
115. Hestenes, M., Stiefel, E.: Methods of conjugate gradients for solving linear systems. Journal of Research of the National Bureau of Standard 49(6) (1952).
116. Møller, M.: Original contribution: A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. 6(4), 525–533 (1993).
117. Kramer, K. A., S. C. Stubberud and J. A. Geremia (2007). Sensor calibration using the neural extended kalman filter in a control loop. In: *Proc. of IEEE Intl. Conf. on Computational Intelligence for Measurement Systems and Applications 2007*. Ostuni, Italy. pp. 19–24.
118. Zou, A., Z. G. Hou, L. J. Zhang and M. Tan (2005). A neural network-based camera calibration method for mobile robot localization problems. In: *Proc. of Second Intl. Symp. on Neural Networks*. Chongqing, China. pp. 277–284.
119. Choi, M., R. Sakthivel and W. K. Chung (2007). Neural network-aided extended kalman filter for slam problem. In: *Proc. of IEEE Intl. Conf. on Robotics and Automation 2007*. Roma, Italy. pp. 1686–1690.