UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH

Programa de Doctorat en Enginyeria Biomèdica
Departament d'Enginyeria de Sistemes,
Automàtica i Informàtica Industrial
Centre de Recerca en Enginyeria Biomèdica

Tesis Doctoral por compendio de publicaciones

Muscular pattern based on multichanel surface EMG during voluntary contractions of the upper-limb

Mislav Jordanić

September 2017

Directores:
Miguel Angel Mañanas Villanueva
Mónica Rojas-Martínez

Abstract

Extraction of neuromuscular information is an important and extensively researched issue in biomedical engineering. Information on muscle control can be used in numerous human-machine interfaces and control applications, including rehabilitation engineering, e.g., prosthetics, exoskeletons and rehabilitation robots.

Neuromuscular information can be extracted at the brain level, peripheral nerves, or muscles. Among these options, muscle interface is the only viable way of information extraction in everyday life. Although brain and nerve recordings are promising, they usually require invasive measurement and achieve relatively low extraction speed which prevents real time control. Even though in electromyographic (EMG) recordings information is not obtained directly from neural cells, it contains similar information as nerve recording. Information contained in action potential of the innervated muscle fibers (MUAP) is equivalent to the information contained in the action potential of corresponding motor neurons. Moreover, muscles contain multiple motor units that activate simultaneously so their electrical activity sums on the surface of the skin, resulting in a relatively high amplitude compared to the other bioelectrical signals. Therefore, due to the richness of neural information, noninvasiveness and high signal-to-noise ratio, the surface EMG is extensively used for man–machine interfacing, especially in commercial/clinical upper-limb prosthetic control.

Motivation and merit of this thesis lies in the fact that information associated with muscular pattern during exercises can be very useful in different applications such as monitoring patients' control strategies during recovery, personalizing rehabilitation processes to increase their effectiveness or to provide information to be used for control of external devices (EMG based control of prosthesis or exoskeletons).

Within this doctorate a pattern recognition approach was used to assess neuromuscular information and to identify subjects' intended motion based on multichannel surface electromyographic recordings. Research was focused on control strategies of upper-limb, both in normal subjects and in patients with impaired mobility caused by incomplete spinal cord injury. Methods which are proposed can be used for the design and monitoring of rehabilitation therapies intended for patients with neuromuscular impairment, as well for the control of external devices like rehabilitation robots, exoskeletons, prostheses and even virtual games. However, that is in the domain of future applications and is not the scope of the thesis.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Performing a movement is a complicated process that involves many physiological entities working in high coherence. It involves bones, tendons, nerves, and many other working in perfect harmony. Even the simplest movements are rarely performed using just one muscle. Everything we do involves high muscular coordination and constant and precise regulation. While standing, muscles of legs and trunk are constantly simultaneously co-contracting, maintaining balance. Muscle is a body tissue capable of transforming chemical energy to force. There are several muscle types: smooth, building internal organs, cardiac, building the heart, and skeletal. Only skeletal muscles can be controlled voluntarily and are used in locomotion. They are usually connected to bones with tendons (collagen fibers), as shown in figure 1.1.

The neurons controlling the movement are organized in hierarchical fashion (Widmaier et al., 2014). In the highest level of hierarchy, the movement is conceived. Here the complex plan of intention is made. Very little is known about the exact location of neurons responsible for this task. Higher centers then transmit this command to the middle level structures, where the task is elaborated. Simultaneously, this middle level neurons receive the information from the receptors in muscles, skin, tendons, and join, but also from the visual system. Planning of the movement that is about to be performed is performed with respect to the space this movement will occupy, and detailed control signals for each muscle involved in the movement are generated. Centers involved in this tasks are located in cerebral cortex, cerebellum, subcortical nuclei, and brainstem. The information is then transmitted to the lowest level of the motor hierarchy:

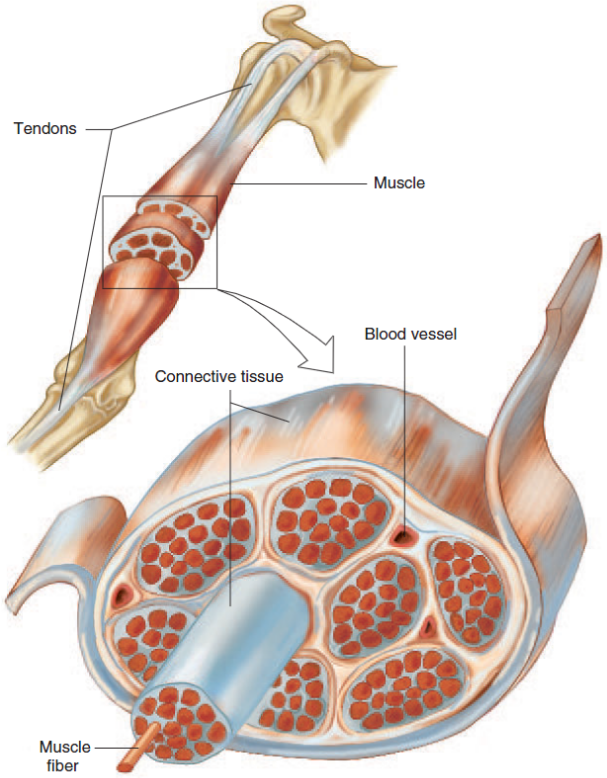


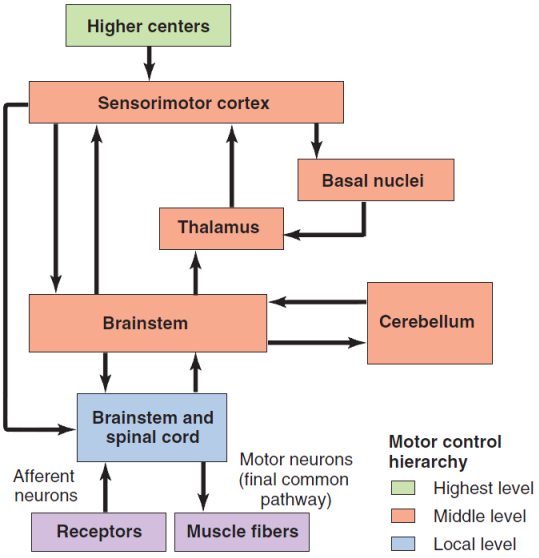
Figure 1.1: Organization of skeletal muscle with attachment to the bone. Retrieved from (Widmaier et al., 2014)

spinal cord and brainstem. Here information is transmitted over motor neurons to the muscles. The selection of motor neurons involved in the task and timing is performed at this level. Organization and locations of the neural system for motor control can be seen in figure 1.2.

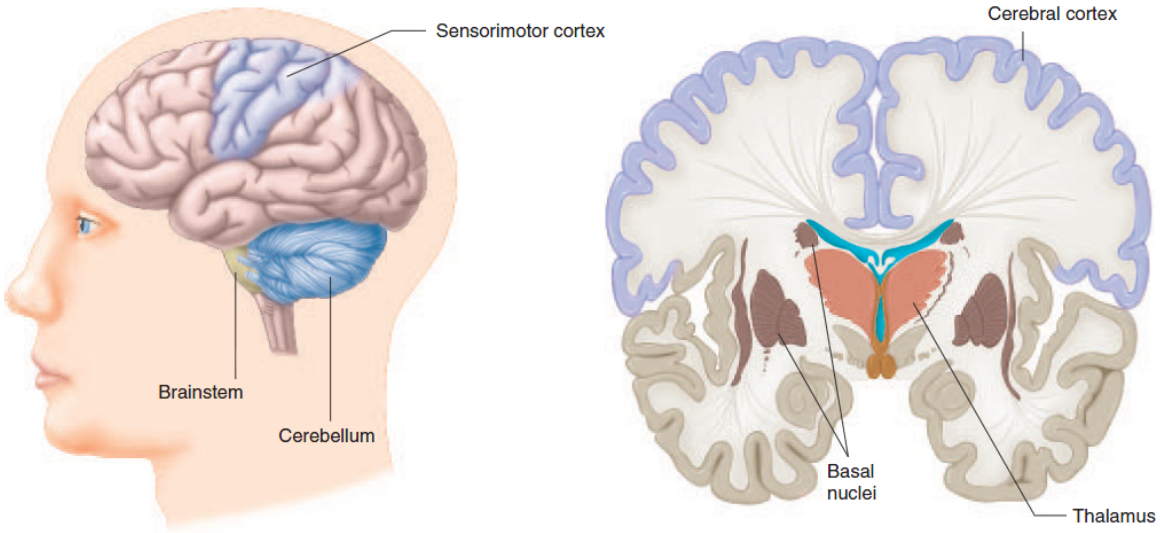
1.1 Muscle physiology

Elementary building block of a muscle is muscle cell, or muscle fiber - *myocyte*. They are ensheated by *endomysium*, a connective tissue that contains nerves and capillaries. Myocytes are organized in bundles of 10 to 100 fibers, which are called *fascicles*, and they are surrounded by sheath of connective tissue, *perymisium*. Group of fascicles is finally grouped together and enveloped by *epimysium*, forming a muscle.

Sarcolema is the cell membrane of myocyte, consisting of a lipid bilayer that contains intracelular liquid, *myoplasma*. In the myoplasma, thin and thick filaments are serially connected, forming



(a)



(b)

Figure 1.2: Figure describes **a)** hierarchical organization of neural system for motor control and **b)** side view and cross section of the brain showing motor control centers. Retrieved from (Widmaier et al., 2014)

sarcomeres, which are longitudinally connected in *myofibrils* that extend through entire length of the myocyte. During shortening of muscle fibers, thin and thick filaments of sarcomeres are pulled together by cross-bridges between them. Total shortening of myofibril is summation of shortenings of sarcomeres of which it is composed.

During the stable state when there are no stimuli, i.e., in the resting state, the interior of the myocyte is at higher electrical potential that the exterior. This difference in potential is usually

around 80 mV and it is caused by the higher concentration of positive ions, namely Na+, outside of the sarcolema (Nazmi et al., 2016).

Motor neurons transfer nerve impulses that control the muscle from spinal cord to neuromuscular junction. At the nerve endings, action potentials induce the opening of calcium channels, which enables calcium from extracellular fluid to enter axon terminals and trigger the release of the neurotransmitter *acetylcholine*. Acetylcholine is released to the narrow space between the axon and sarcolema of the myocyte, and causes sodium channels in sarcolema to open and allow the flow of Na+ and K+ ions in both directions. Na+ ions now flow into the myoplasma by diffusion due to higher concentration of Na+ ions outside of the membrane, but because of similar gradient, concentrations of the K+ ions don't change a lot. This process causes depolarization of sarcolema during which the outside potential of the muscle cell is at lower voltage than inside potential by around 30 mV. Depolarization is immediately followed by repolarization, a process during which the electrochemical balance and the resting potential of the cell are restored. It is achieved by flushing the Na+ ions outside of the sarcolema by the *ion pump*. The process can be seen in figures 1.3 and 1.4.

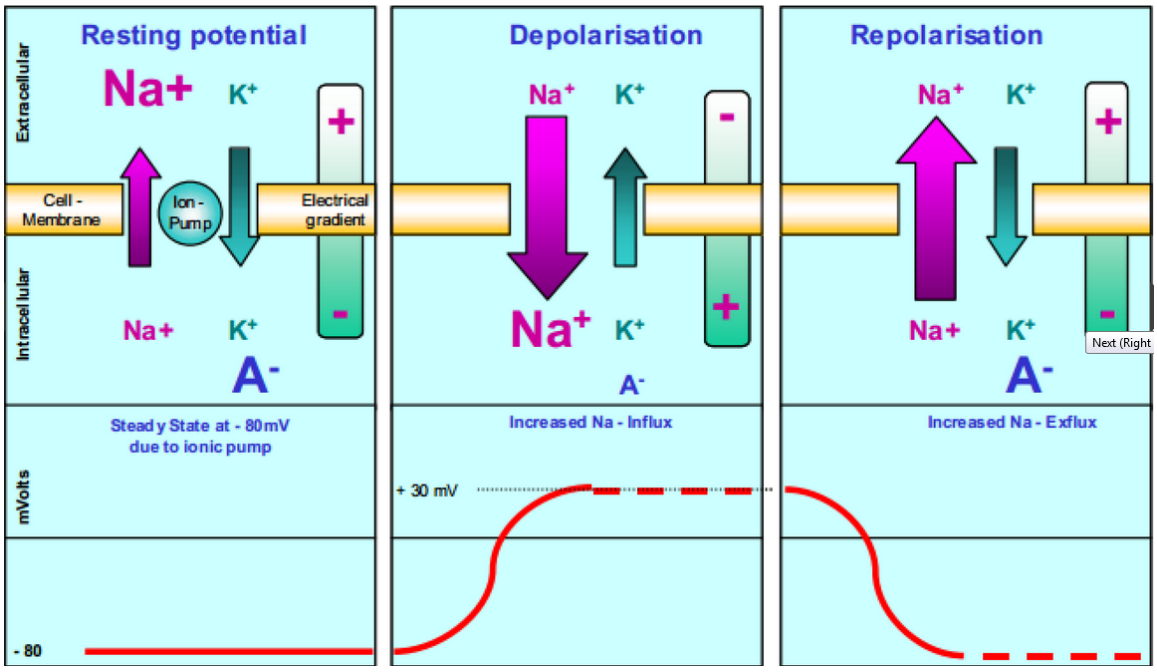


Figure 1.3: Ilustration of depolarization/repolarization of the muscle fiber. Adopted from (Nazmi et al., 2016).

If the amount of acetylcoline is sufficient for the excitation, depolarization/repolarization wave,

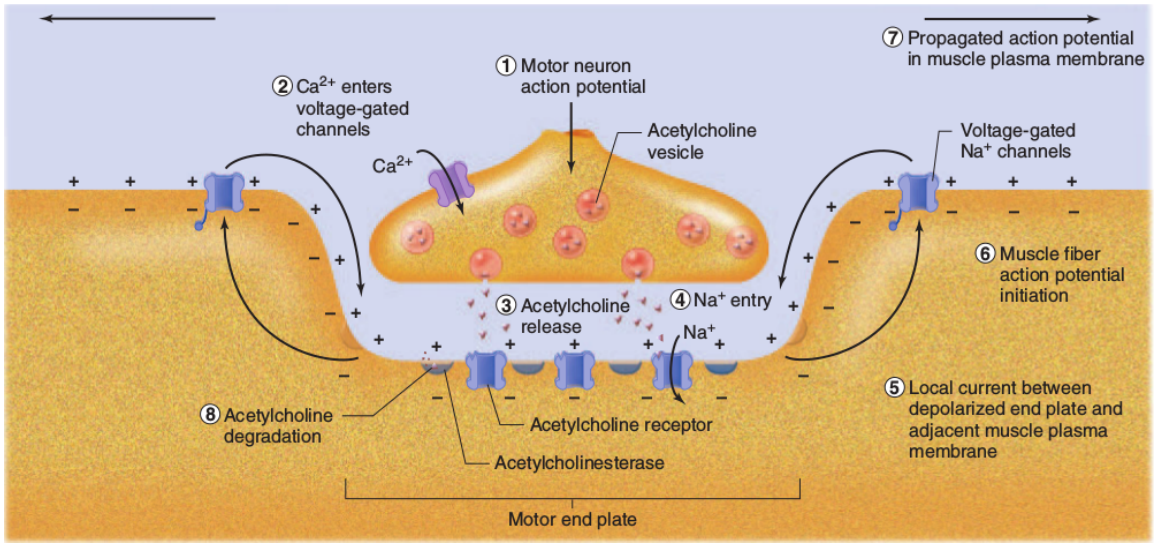


Figure 1.4: Illustration of generation of action potential. Retrieved from (Widmaier et al., 2014).

that is, action potential, propagates longitudinally from the neuromuscular junction towards the ends of the muscle fiber causing contraction (Henneberg, 1999). Speed of action potential propagation is called *conduction velocity* and typically ranges around 4 m/s.

Detailed analysis of muscle physiology can be found elsewhere (Squire, 1986; Widmaier et al., 2014).

1.2 Muscle activation

Each motor neuron at the neuromuscular junction innervates several muscle fibers, forming the smallest functional unit called *motor unit*. It was firstly defined by Liddell and Sherrington in 1925 (Liddell and Sherrington, 1925; Sherrington, 1925) and is composed of motor neuron with axon and dendrites, and muscle fibers that axon innervates (Duchateau and Enoka, 2011). Since motor neuron with a single action potential usually evokes action potentials simultaneously in all belonging muscle fibers, by observing action potentials of the muscle fibers, information on activity of motor neurons in spinal cord or brain stem can be inferred (Merletti and Farina, 2016).

Pool of motor neurons that innervates entire muscle generally ranges from ten to thousand, depending on the muscle (Merletti and Farina, 2016).

By the characteristics of muscle fiber, there are three main types of muscle fibers:

Fast twitch, fatigable fibers (FF, or type IIb): They have high levels of ATP (source of energy) for anaerobic energy supply, and are dominantly present in pale muscles. They are glycolytic and work well in ischemic or low oxygen conditions. Regarding contraction properties, they are characterized by fast twitch, large forces and high nerve conduction velocity, but they get fatigued faster than the other muscle fiber types.

Fast twitch, fatigue-resistant (FR, or type IIa): They are oxidative glycolytic fibers, characterized by fast twitch and are resistant to fatigue. They have intermediate conduction velocity.

Slow twitch, very resistant to fatigue (S, or type I): They are oxidative fibers and do not work well in low oxygen conditions. They generate small forces, have slow twitch and resistant to fatigue. lower nerve conduction velocity. This fiber type is very resilient to fatigue because of high oxidative metabolism and energy efficiency. They are present in high percentage in red muscles, such as soleus.

Muscle fibers innervated by the same motor neuron have similar histochemical and contractile characteristics.

Force that muscle fibers generate depends on firing frequency of the action potentials (rate coding) innervating the neuromuscular junction, and the recruitment strategy by which the motor units are activated, i.e., the number of activated motor units. Firing frequency and the recruitment strategy depend on the speed and force of contraction. Muscle units with low threshold are activated firstly, resulting in low force and high endurance, i.e., resistance to fatigue. If greater force is required, muscle units with higher threshold that are prone to fatigue are activated (Freund et al., 1975; Merletti and Parker, 2004). This was firstly proposed by Henneman et al. in 1965 (Henneman et al., 1965), who state that order of recruitment of motor neurons is based on size principle, that is, neurons with smaller axons are recruited at lower effort levels and with increase in force, larger motoneurons are recruited. Therefore, S type muscle units, which have the smallest motoneurons are recruited first, followed by FR type units, and finally FF units. The recruitment strategy and resistance to fatigue can be seen in

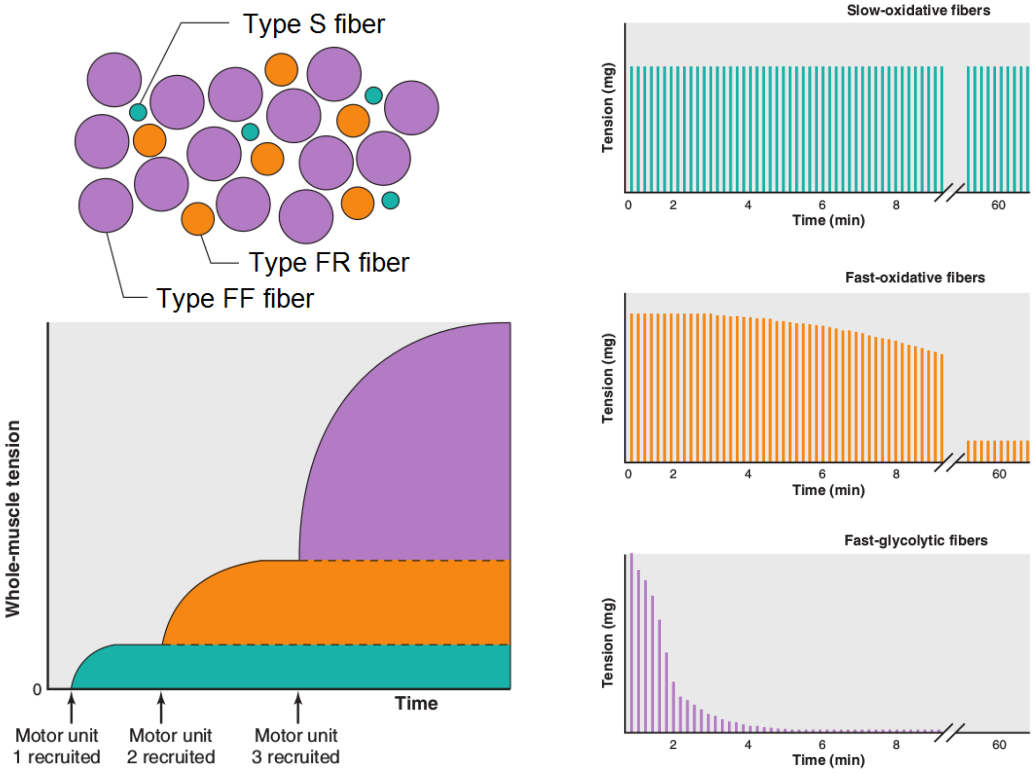


Figure 1.5: Figure describes characteristics of different types of muscle fibers. In **a)** is a diagram of different muscle fibers in muscle cross section (top), and muscle tension produced by recruitment of different types of muscle fiber (bottom), whereas in **b)** is the illustration of the time interval during which specific muscle fibers can remain tension. It can be noted that type S fibers are activated first, generate low force level, and are resistant to fatigue. On the other hand, type FF fibers are activated last, generate high forces, and develop fatigue fastest. Retrieved from (Widmaier et al., 2014)

figure 1.5.

1.3 Muscle fatigue

Muscle during contraction develops muscle fatigue. It is characterized by decrease of conduction velocity **Wdimaier**. Although muscle fatigue begins to develop at the beginning of contractions, it is a challenge to grade it during contraction. What

1.4 Surface electromyography

Muscle unit action potentnial (MUAP) is the combination of action potentials generated by the single motor unit. Myoelectric signal is a superposition of electrical activity (propagating action potentials) produced by the muscle fibers while contracting.

EMG signals could be recorded either non-invasively (surface EMG, sEMG) or invasively with needle and wire electrodes (intramuscular EMG, iEMG) (Marateb et al., 1999). Although iEMG signal is usually with higher quality (in terms of signal-to-noise ratio), it was shown that both approaches provide a similar identification rate of upper-arm motor task (Hargrove et al., 2007).

Surface electromyographic signal (sEMG) is the sum of the electrical activity of the muscle fibers recorded on the surface of the skin. Since muscle fibers are activated by the impulse train of the innervating motor neurons, i.e. neural drive to the muscle, sEMG is the convolution of motor neuron spike trains by the motor unit action potential recorded on the electrodes (Farina et al., 2010, 2014a):

$$sEMG(t) = \sum_{i=1}^{M} \sum_{j=-\infty}^{+\infty} MUAP_i(t)\ \delta(t - t_{i,j}) \tag{1.1}$$

, where M is the number of active motor units, $MUAP_i(t)$ is the action potential waveform of the i^{th} motor unit recorded by the electrodes, and $t_{i,j}$ is the time of the discharge of the i^{th} motor neuron. This model assumes there is no interference and that neuromuscular junction never fails, which is not the case. In the equation, $MUAP_i(t)$ is related to the electrophysiological state of the muscle fiber membranes and conduction properties of the tissue through which the potential propagates, whereas neural information is contained in motor neuron spike trains $\delta(t-t_{i,j})$ (Farina et al., 2014b). It is important to notice that following this model, sEMG reflects all information that is present in motor neuron. Therefore, it is more appropriate to extract motor control information carried by motor neurons using sEMG, than directly by invasive measurement of electrical potential of the motor neuron. The advantage of the sEMG is that multiple fibers are activated simultanously, generating bioelectrical signal with relatively high SNR, which can be measured on the surface of the skin. In this context, sEMG can be considered as the amplified neural signal, whereas muscle can be considered as biological amplifier of nerve activity (Farina et al., 2014a). Origin of sEMG signal can be seen in figure 1.6.

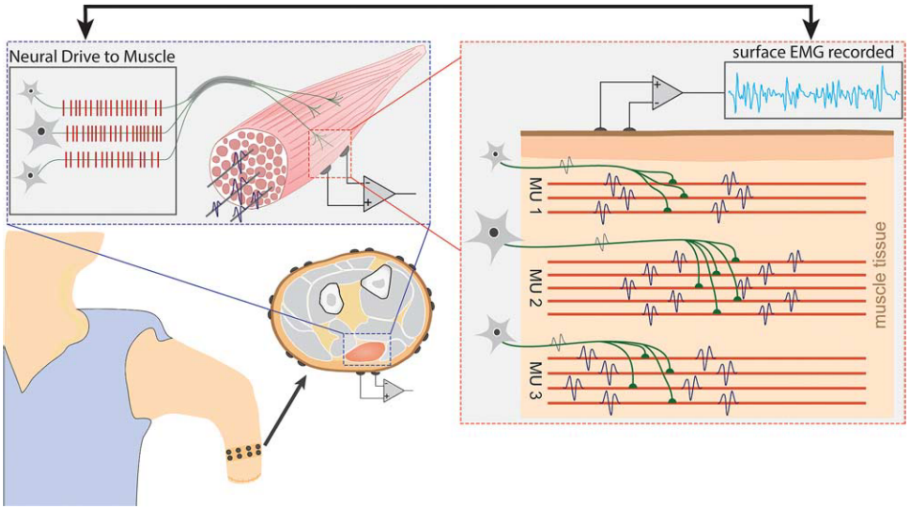


Figure 1.6: Origin of sEMG signal. SEMG signal is a sum of each motor unit action potential recorded on the electrodes convoluted by belonging motor neuron spike train. Retrieved from (Farina et al., 2014a).

Depending on number of electrodes used for the recording. the following classification exists: monopolar, bipolar, linear electrode array, and high-density EMG (see Figure 1.7)

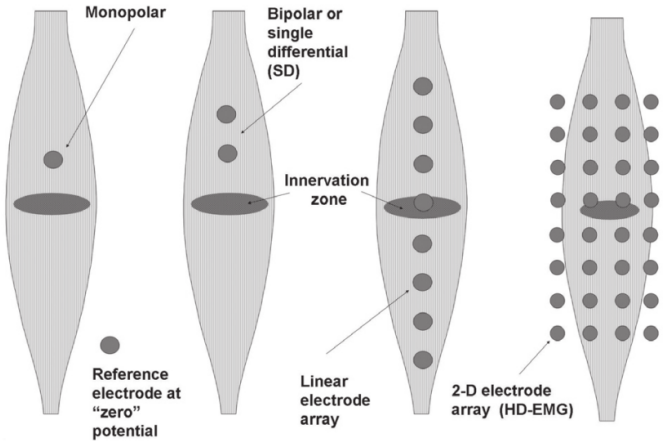


Figure 1.7: Four types of recording surface EMG signal: monopolar, bipolar, linear electrode array, HD-EMG. Figure was modified from (Merletti et al., 2010)

Technological advancement of EMG acquisition systems enables use of high-density electromyography (HD-EMG) (Zwarts et al., 2004). Using an array of closely spaced electrodes organized in a quadrature grid, a wide muscle area is recorded. This technology is the only one that allows insights into spatial distribution of motor units in a muscle. By observing the amplitude or intensity of signals recorded in different channels, it is possible to analyze how different muscle regions activate depending on joint position (Vieira et al., 2010), contraction level (Holtermann et al., 2005), and duration of movement and fatigue (Tucker et al., 2009; Staudenmann et al.,

2014).

In addition, activation of individual motor units, i.e. individual motor neuron spike train, can be extracted from the HD-EMG recordings using Blind Source Separation methods (Holobar and Zazula, 2007; Holobar et al., 2010), which can be a valuable information in force estimation because motor unit recruitment and firing frequency depend primarily on force level (Merletti and Parker, 2004). Several authors have used this approach instead of the traditional one based on intramuscular (invasive) EMG. One of the obvious advantages of this method is that is safe and not painful, although it has not been implemented in clinical practice yet. Using this technique, authors in (Holobar et al., 2010) were able to extract 6 to 7 motor units starting from contractions at 5% MVC and up to 20% MVC with associated discharge rates between 10 pps and 12 pps. However, one of the current limitations is that the intensity of isometric contraction must remain constant during the measurement.

Similar algorithms can also be used to separate EMG activity of adjacent muscles (Farina et al., 2004; Holobar and Farina, 2014). This method can be a powerful tool in task identification (Naik et al., 2007), because it could minimize crosstalk effect from nearby muscles. Consequently, extracted features would characterize only the target muscles. However, HD-EMG can be corrupted by low quality channels, which are a common issue in measurements due to well-known artifacts, such as: electrode displacement, bad electrical contact between skin and the electrode, movement of cables, electromagnetic interference, etc. (Clancy et al., 2002). Affected channels differentiate themselves in amplitude and spectral content, which makes them outliers that compromise classification. To cope with this problem, authors in (Rojas-Martínez et al., 2012) developed an expert system for detection, removal and interpolation of HD-EMG channels corrupted by artifacts.

1.5 Task identification

The central nervous system (CNS) is responsible for processing information received from all parts of the body. The two main organs of the CNS are the brain and the spinal cord and are entirely composed of two kinds of specialized cells: neurons and glia. The brain is the most complex part of the human body and exerts a centralized control over the other organs.

Neurons, the basic working units of the brain, are designed to transmit information within the brain to other nerve cells and to communicate with muscles and gland cells. The complex architecture of the brain is built on the extensive number of interconnected neurons sharing information through specialized connections called synapses. This connection allows neurons to communicate through an electrical or chemical signals, producing ionic currents that generate electric and magnetic fields.

The CNS is organized in multiple levels, from simple connections between cells to coordinated cell populations, building a complex architecture of interconnected brain regions. The neural processes at this last level are produced by the dynamic coordination of smaller elements. In the cerebral cortex, all this brain activity is summed and its electric and magnetic fields can be measured on the scalp surface.

In most of the commercial prosthesis (Parker and Scott, 1986), sEMG of two muscles is recorded. In this simple scheme a single Degree-of-Freedom (DOF) can be controlled: the EMG amplitude of one muscle controls the output of one direction, whereas the EMG amplitude of the other muscle controls the other direction. If prosthesis needs to operate in multiple DOFs, a subject needs to switch between currently active DOF either by co-contraction or by pressing a switch button. In any case, the method is not intuitive nor efficient for the user (Farina et al., 2014a).

Pattern recognition is an alternative to conventional control algorithms. The prerequisite of using pattern recognition for task identification is the presence of a pattern that can be extracted from the EMG signal. Major advancement over conventional conventional switching myocontrol is the possibility of.

Pattern recognition approach doesn't support proportional and simultaneous control for multiple motor tasks. Therefore, tasks need to be performed sequentially. This type of control prevents the user from achieving a fluid movement, but also demands planning of movement execution. Although pattern recognition improves the possibility, it has serious limitations.

This implies sequential control which prevents the subject from doing fluent, e.g. Davidge et al. designed a system where movements that combine DoFs are labeled as unique classes in LDA problem, whereas Young et al. (Young et al., 2013) propose system of parallel LDA classifiers that use conditional probabilities to separate between combination of tasks.

Proportional review (Fougner et al., 2012) Force can be estimated based on the EMG : (Staudenmann et al., 2010)

In pattern recognition, there is still a large gap between industry and practice (Jiang et al., 2012).

On the other hand, one of the disadvantages of pattern recognition is the fact that in spite of the high accuracy, an error could lead to the completely unwanted task. Also, although identification rate is usually very high during the stationary task, errors often occur during transition between tasks. This problems can be partially prevented by employing the e.g. majority voting principle (Englehart and Hudgins, 2003) (300ms, LDA), or decision-based velocity ramp that attenuates the velocity of a movement after the change of a task (Simon et al., 2011).

Challenges in pattern recognition are electrode shift (Hargrove et al., 2008; Young et al., 2011), change in arm posture (Fougner et al., 2011), slow time dependent changes (Farina et al., 2014a) such as fatigue (Tkach et al., 2010), and change in electrode-skin impedance (**?**).

Future works: dynamic system, hybrid system

1.5.1 Pattern recognition

Given the one to one relationship between the neural commands and the activation of motor units in the muscles, surface electromyography (sEMG) has been used for more than a half of century as a noninvasive and natural way of extracting motor control information for identification of motion intention. Such information is used in numerous applications in rehabilitation engineering, e.g., prosthetics (Li et al., 2010; Young et al., 2013; Stango et al., 2015), exoskeletons (Vaca Benitez et al., 2013) and rehabilitation robots (Dipietro et al., 2005; Marchal-Crespo and Reinkensmeyer, 2009). Ideally, an identification system should fulfill the following criteria (Farina et al., 2014a):

- Intuitive control: simultaneous and proportional
- Insensitive to changes in electrode - skin impedance,

- Adaptive to changes during the use, i.e. fatigue, electrode-skin impedance change due to sweating and drying of conductive gel
- Insensitive to precise position of electrodes
- Fast and easy training procedure (ideally none)
- Real time identification, i.e. time delay less than 300 ms (Oskoei and Hu, 2007)
- Low computation complexity which enables implementation in battery-powered device

Pattern recognition – based control strategy enables proportional usage of multiple DoFs without switching between states, which makes it more intuitive. According to Oskoei et al. (Oskoei and Hu, 2007), this strategy includes four main modules:

Data segmentation: Comprises various techniques and methods that are used to handle data before feature extraction

Feature extraction: This module computes and presents preselected features for a classifier. Features, instead of raw signals, are fed into a classifier to improve classification efficiency. Selection or extraction features is one of the most critical stages in myoelectric control design.

Classification: A classification module recognizes signal patterns, and classifies them into predefined categories. Due to the complexity of biological signals, and the influence of physiological and physical conditions, the classifier should be adequately robust.

Controller: Generates output commands based on signal patterns and control schemes. Post-processing methods, such as majority voting, which are often applied after classification to eliminate destructive jumps and make a smooth output, are included in this module too.

The main drawback of this method is that only one movement can be activated at the time. Any task that requires more than one DoF must be performed sequentially. However, several authors recently proposed solutions which enable simultaneous control (Young et al., 2013; Kamavuako et al., 2013; Baker et al., 2010). A variety of classifiers (e.g. hidden Markov model, support vector machine, artificial neural network, fuzzy logic and linear discriminant analysis) (Oskoei

and Hu, 2007) has been used in myocontrol research. Nevertheless, multiple authors agree that the identification does not significantly depend on the classifier type (Hargrove et al., 2007; Zhang and Zhou, 2012; Hakonen et al., 2015). Therefore, simple and easy to train classifiers like linear discriminant analysis (LDA) are preferred (Li et al., 2010; Englehart et al., 1999; Tkach et al., 2010; Li et al., 2014; Hakonen et al., 2015). On the other hand, finding an appropriate set of features is challenging (Englehart et al., 1999; Tkach et al., 2010; Liu and Zhou, 2013). In literature, a lot of feature types were considered:

Time domain features: mean absolute value (Hudgins et al., 1993), integrated EMG (Park and Lee, 1998), variance (Park and Lee, 1998; Zardoshti-Kermani et al., 1995), root mean square (Farrell and Weir, 2008), waveform length (Hudgins et al., 1993), zero crossing (Hudgins et al., 1993), log detector (Tkach et al., 2010), Wilson amplitude (Zardoshti-Kermani et al., 1995), slope sign change (Hudgins et al., 1993), autoregressive coefficients (Hargrove et al., 2007), Cepstral coefficients (Park and Lee, 1998), mean absolute value slope (Phinyomark et al., 2012a), histogram of EMG (Phinyomark et al., 2012a; Zardoshti-Kermani et al., 1995)

Frequency domain features: mean frequency (Phinyomark et al., 2012b), median frequency (Phinyomark et al., 2012b), modified mean frequency (Phinyomark et al., 2009)

Time-frequency domain features: short time Fourier transform (Englehart et al., 2003, 2001), continuous wavelet transform (Englehart et al., 2003, 2001), discrete wavelet transform (Englehart et al., 2003), stationary wavelet transform (Englehart et al., 2003), wavelet packet transform (Englehart et al., 2003, 2001; Chu et al., 2006)

Spatial domain features: Experimental periodogram (Stango et al., 2015), center of gravity (Rojas-Martínez et al., 2012, 2013)

Time domain features are commonly used (Hakonen et al., 2015) because they achieve high identification accuracy and are computationally efficient.

However, Zwartz et al. (Zwarts and Stegeman, 2003) pointed out that single channel EMG disregards important spatial aspects of MUAP propagation, which are essential for the force-generating capacity of the muscle, and, if not well addressed, can lead to incorrect conclusions.

Moreover, since muscles do not activate homogeneously, single bipolar channel EMG has some serious drawbacks, which can be overcome by using 2D electrode arrays: high density EMG (HD-EMG).

In HD-EMG measurements, multiple EMG channels are recorded using an array of closely spaced electrodes placed over the wide area of the muscle. This type of recording is more reliable because it can record activations in different parts of the muscle and increase redundancy. Commonly, authors in literature report identification based on HD-EMG and time domain features or autoregressive features calculated for each channel (Hakonen et al., 2015). Zhang et al. (Zhang and Zhou, 2012), for example, used combination of these features with dimensionality reduction to identify 20 wrist and hand movements, and Muceli and Farina (Muceli and Farina, 2012) performed estimation of hand kinematics during more than 20 movements using EMG envelopes as features with reduced dimensionality of channels.

But HD-EMG recordings also allow calculation of two-dimensional activation maps where intensity of each pixel represents the intensity of a corresponding EMG channel (see figure 1.8). Consequently, information on spatial distribution of EMG intensity over the muscle is provided. Recent studies show that changes in spatial activation pattern are related to duration of movement and fatigue (Tucker et al., 2009; Staudenmann et al., 2014), position of joint (Vieira et al., 2010) and the level of contraction (Holtermann et al., 2005). Since spatial distribution contains a lot of information on the muscle, it is acknowledged as a valuable feature in identification of motion intention (Stango et al., 2015; Hakonen et al., 2015; Rojas-Martínez et al., 2013). For example, Stango et al. (Stango et al., 2015) used spatial characteristics of HD-EMG recording of the forearm muscles to identify 8 hand and wrist tasks (4 degrees of freedom). They fed support vector machine classifier with a statistical measure of spatial correlation, i.e. variogram and achieved high identification results (95% accuracy). Furthermore, they proved that proposed spatial features are robust to electrode shift.

Most of pattern recognition identification methods are subject-specific. They usually achieve very high identification results, but require time consuming training procedure for every patient individually. This could be avoided by building a single identifier for a group of patients, i.e. group-specific identifier. However, inter-subject variability is a big concern in design of a

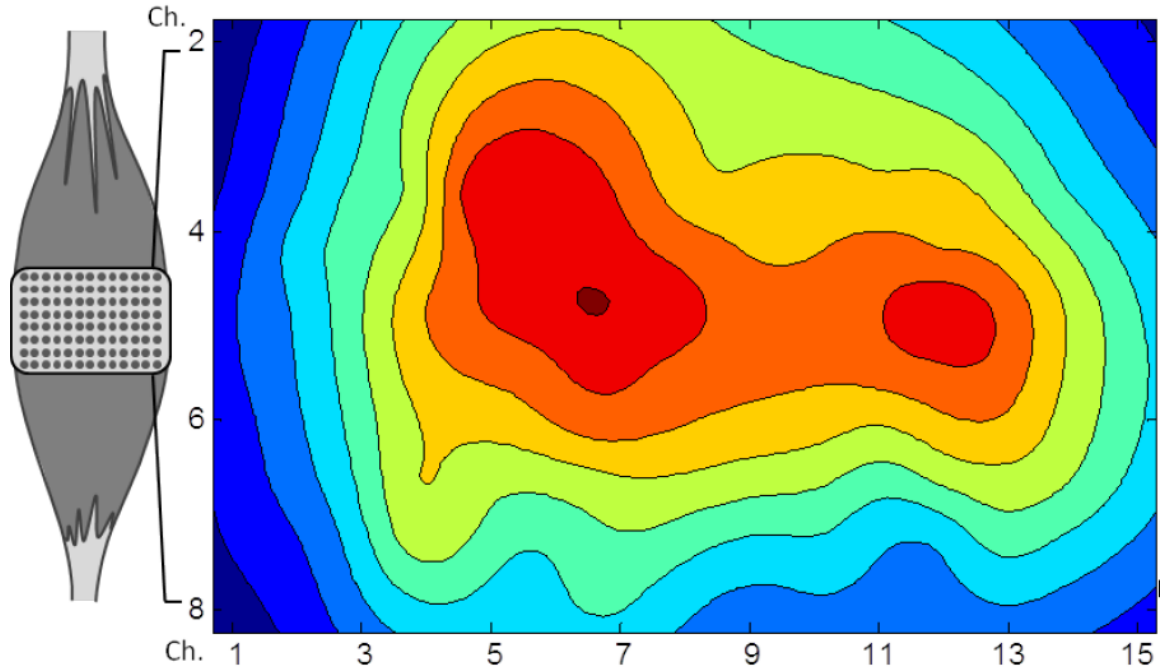


Figure 1.8: The figure represents the HD-EMG activation map recorded on the biceps brachii muscle during flexion. Distinct activation of the two heads can be noticed in the map. Modified from Monica

group-specific pattern recognition-based identifier. Individuals differ from each other in a lot of physiological parameters, e.g., conductivity of subcuntaneous tissue, and limb dimension. Nevertheless, by comparing HD-EMG activation maps between normal subjects it has been shown that inter-subject activation patterns exists for different tasks and levels of contraction (Rojas-Martínez et al., 2012).

In (Rojas-Martínez et al., 2013) authors demonstrate that by using intensity and spatial features extracted from activation maps it is possible to construct an inter-subject identification method based on LDA classifier not only for different tasks, but also for different effort levels. Authors reported that in healthy subjects identification performance improves by adding spatial features in the identification, which proves that spatial distribution is less sensitive to inter-subject variability. They achieved sensitivity higher than 75% for identification of four upper-limb tasks at three different effort levels and more than 90% sensitivity when identifying only four tasks and no effort level. Also, they report higher classification results when using classification in two steps (in first step task is classified, and in the second step level of effort), rather than a single step classification.

1.5.2 Application to patients with neuromuscular impairment

According to World Health Organization, each year there are 500 000 spinal cord injuries [56] and 15 million strokes (of which 5 million result with death and 5 million with permanent disability) [57] every year. Furthermore, number of people who are older than 60 years will increase to 22% of the world population by 2050 and will count 2 billion people [58]. Unfortunately, in affected patients motor control can be impaired as a result of damaged nerves and they often suffer from uncoordinated movements, lack of force, and spasticity. During recovery process, rehabilitation robots that stimulate neuroplasticity are commonly used (Vaca Benitez et al., 2013; Dipietro et al., 2005; Marchal-Crespo and Reinkensmeyer, 2009).

Patients can still have uncoordinated movements, and lack of force, or, in more difficult cases, they can weakly activate their muscles, but cannot perform the movement. If their motion intention could be extracted in real time, it would allow them to control assistive devices and maximize the benefits of robotic-aided therapies where it has been proved that the active participation improves the medical condition of the patient (Hogan et al., 2006).

It is already shown that intensity-related and task-specific activation patterns exist in patients with neurological disorders and that motion intention can be extracted from EMG. In other words, movement that patient is trying to perform can be predicted using the recorded myoelectric activity. Liu and Zhou (Liu and Zhou, 2013) were able to successfully perform identification of tasks using time domain and autoregressive model features in patients with incomplete spinal cord injury, whereas Zhang and Zhou (Zhang and Zhou, 2012) identified tasks in patients with stroke using a similar feature set.

Physical injury to the brain, spinal cord, or nerves, is usually the cause of neurological disorders. Stroke is a serious life-threatening condition that occurs when the blood supply to the brain is interrupted, resulting in severe disability among survivors. Brain damage due to stroke can affect important areas that control everything we do, including how we move different parts of our body.

In disabling neurological disorders, treatment and rehabilitation should start as soon as possible after the diagnosis. This situation is critical in stroke patients where rehabilitation usually

begins two days after the stroke has occurred. Early intervention can improve bodily functions and even achieve remarkable recoveries, and should be continued as necessary after release from the hospital to become as independent as possible. Rehabilitation can help regain control of weak limbs, or learn new ways of using them again, that is, to relearn skills that are lost when part of the brain is damaged, mainly associated with motor capabilities.

Human-machine interfaces cannot only translate brain signals to control targets, but also can combine with a muscle-based output. This hybrid approach is composed of multimodal data from the brain (EEG) and from the muscles (EMG), especially important for patients with residual muscular activity. Additionally, corticomuscular coherence assessed through both EEG and EMG signals is a promising measure to evaluate the motor recovery of stroke patients. This Project is on monitoring the patient's progress during the rehabilitation program, and biomarkers composed of EMG and EEG information will be very interesting. For example, Transcranial Magnetic Stimulation (TMS) has been used to probe corticospinal physiology and to map the primary motor cortex (M1) representations of upper limb muscles following stroke.

its application to pattern recognition to provide a control signal to interfaces like prostheses or rehabilitation robots, particularly for stroke or other neuromuscular disorders

Common manifestations of upper extremity motor impairment include muscle weakness or contracture, changes in muscle tone, joint laxity, and impaired motor control. These impairments induce disabilities in common activities such as reaching, picking up objects, and holding onto objects.

The aim of this research line is to analyze biological signals of patients undergoing stroke rehabilitation and extract measures that reflect the current degree of recovery and neuromuscular ability, and to define and evaluate expert-based quantitative indices for predicting the final clinical outcome of the standard 6-month rehabilitation. This latter aim would be very useful for the physicians leading the therapy and would save a lot of time and resources. For example, if a patient is not able to achieve true recovery of the affected motor function, he could start with an alternative treatment immediately and learn a compensatory motor task, improving its quality of life faster. According to rehabilitation professionals, these particular cases are very difficult to identify early on, and it takes months before realizing that a different approach is needed.

On the other hand, if the measure indicates a good recovery potential, it would be possible to tailor a patient-oriented rehabilitation program that would maximize its effect.

Neural data can be inferred from HD-EMG signal using HD-EMG decomposition to spike train of individual motor units (Holobar et al., 2014, Negro et al., 2016). This technique enables measuring of valuable information about muscle unit recruitment: muscle fiber conduction velocity, location of the innervation zones, estimation of muscle fatigue, and estimation of number, type and the spatial distribution of muscle fibers (Marateb et al., 2016). Other potential measures that can be found in the literature are the clustering index (Zhang et al., 2017), measures based on textural and spatial analysis of HD-EMG activation maps (Rasool et al., 2017), the coherence between muscle units (Dai et al., 2017), or measures based on muscle synergies (Li et al., 2016). The advantages of the HD-EMG lie in the large amount of recorded information, which enables minimizing the effect of electrodes shift and allows choosing an appropriate subset of channels for further analysis. Features will be designed by optimizing its robustness to electrode shift and minimizing the number of electrodes, which is an important issue in HD-EMG analysis (Pan et al., 2015), but also in EEG analysis and BCI systems (Alotaiby et al., 2015, Tam et al., 2011).

Stroke patients often do not have the ability to achieve a specific task, even though they maintain correct neuromuscular activation, due to spasticity or insufficient contraction (i.e. insufficient force) (Liu et al., 2016). However, it can be possible to observe both, the neural and muscular response (potentials) to motor intention in absence of joint movement. Lack of movement can be a misleading factor in clinical assessment and the design of the rehabilitation program. Accurate task identification would provide clinicians with the patient's real capabilities and potential.

Pattern recognition is a classification technique and the principle by which it is performed is learned independently from the data, i.e., training set. There are two main types of pattern recognition: supervised and unsupervised. Supervised pattern recognition implies that the classes of the training set are known and are used to obtain the model. New inputs are identified as one of the predetermined classes. On the other hand, unsupervised pattern recognition is used when no labels are available and samples are assigned to unknown classes. This technique is more appropriate for the clustering problem because the classes are determined automatically by the system, whereas supervised approach is more appropriate for the classification because

classes are defined by the system designer, and, therefore, it is usually used in task identification based on EMG.

In statistical pattern recognition, each sample is composed of m measures that form the pattern, i.e., features $(x_0, x_1, \ldots, x_{m-1})$. The objective of the algorithm is to obtain a decision rule, i.e., the decision boundary which separates well samples of different classes. There are many *state-of-the-art* classifiers that use various principles to construct these boundaries. However, many researchers agree that the fidelity of the classification in EMG applications depends mostly on selection of features. In other words, with appropriate selection of features, all classifiers will give similar classification result. A short introduction is provided on the two methods used in the thesis: Linear Discriminant Analysis and Support Vector Machine.

Linear Discriminant Analysis

All models are wrong; some models are useful.

George E. P. Box

Linear Discriminant Analysis is a computationally simple and efficient classifier with linear decision boundary and it is based on the Bayesian equation. In a classical problem with n samples in training set, which consist of m features, the dataset of available samples is a matrix of dimension $[n \times m]$, whereas the label that describes the belonging of each sample to one of the classes is y, where $y \in (0, 1, 2, ..., K-1)$.

According to Bayesian equation, the probability that a sample $\mathbf{x}_0$ belongs to a class k is equivalent to the:

$$P(y = k \mid \mathbf{x} = \mathbf{x}_0) = \frac{P(\mathbf{x} = \mathbf{x}_0 \mid y = k)\ P(y = k)}{P(\mathbf{x} = \mathbf{x}_0)} \tag{1.2}$$

, where k represents the class. Term $P(\mathbf{x} = \mathbf{x}_0 \mid y = k)$ is called the *class-conditional* probability and describes the probability that the sample with exact features $\mathbf{x}_0$ is encountered within the group of samples belonging to the class k. Term $P(y = k)$ is called the *a priori* probability and describes the probability that the sample belonging to the class k is found within the group of all samples, regardless of the features. Finally, the term $P(\mathbf{x} = \mathbf{x}_0)$ is called the *marginal* probability and describes the probability of finding the sample with exact set of features in the

dataset, regardless of the class. Marginal probability can be written as a sum of class-conditional probabilities multiplied by the a priori probabilities for each class:

$$\begin{aligned} P(\mathbf{x} = \mathbf{x}_0) = P(\mathbf{x} = \mathbf{x}_0 \mid y = 1)\ P(y = 1) + \\ P(\mathbf{x} = \mathbf{x}_0 \mid y = 2)\ P(y = 2) + \cdots + \\ P(\mathbf{x} = \mathbf{x}_0 \mid y = K)\ P(y = K) \end{aligned} \tag{1.3}$$

Following the Bayesian theory, the hypothesis, i.e., the predicted class of a sample $\mathbf{x}_0$ is chosen as the class which has the highest probability $P(y = k \mid \mathbf{x} = \mathbf{x}_0)$:

$$h(\mathbf{x}_0) = \underset{k}{\operatorname{argmax}}\, P(y = k \mid \mathbf{x} = \mathbf{x}_0) \tag{1.4}$$

Statistically speaking, this is the best possible classifier. The problem arises in the implementation. The exact probability density functions are unknown and have to be estimated from the available data, which is the source of error. Estimated version of the stated probabilities will be marked with a different symbols to stress out the fact they are just an estimates:

$$p_k(\mathbf{x}) \coloneqq P(y = k \mid \mathbf{x}) \tag{1.5}$$

$$g_k(\mathbf{x}) \coloneqq P(\mathbf{x} \mid y = k) \tag{1.6}$$

$$\pi_k \coloneqq P(y = k) \tag{1.7}$$

Linear Discriminant Analysis estimates marginal probability term (π_k) as a ratio of number of samples belonging to class k and the total number of samples, whereas the class-conditional probability term in the Bayesian equation is estimated as a multivariate Gaussian function:

$$g_k(\mathbf{x}) = \frac{1}{(2\pi)^{m/2}\,|\Sigma_k|^{1/2}}\ e^{-1/2(\mathbf{x}-\mu_k)^T \Sigma_k^{-1} (\mathbf{x}-\mu_k)} \tag{1.8}$$

, where m is the dimensionality of the feature space, i.e., number of features representing each

sample. Function g_k is estimated class-conditional probability of class k, and μ_k and Σ_k are the mean and co-variance matrix for class k, respectively, and they are estimated from the available data as:

$$\mu_k = \frac{1}{n_k} \sum_i \mathbf{x}_i \Big|_{\forall \mathbf{x} \in k} \tag{1.9}$$

$$\Sigma_k = \frac{1}{n_k - K} \sum_i \Big(\mathbf{x}_i - \mu_k\Big)\Big(\mathbf{x}_i - \mu_k\Big)^T \Big|_{\forall \mathbf{x} \in k} \tag{1.10}$$

, where n_k represents the number of samples belonging to a class k. To simplify the model, LDA assumes that the co-variance matrices Σ_k are the same for all classes:

$$\Sigma_0 = \Sigma_1 = \cdots = \Sigma_{K-1} = \Sigma \tag{1.11}$$

and they are usually calculated using the weighted average:

$$\Sigma = \frac{\sum_{k=1}^{K} n_k \Sigma_k}{\sum_{k=1}^{K} n_k} \tag{1.12}$$

The consequence of this assumption is the linearity of the decision boundary. Without this assumption the same calculus would lead to quadratic discriminant analysis, which has non-linear boundary.

In a two class example ($y \in \{0, 1\}$), all samples on the decision boundary will have the same probability of belonging to class 0 or 1:

$$D.B. = \Big\{ \mathbf{x} \;\Big|\; P\big(y = 0 \,\big|\, \mathbf{x} = \mathbf{x}_0\big) = P\big(y = 1 \,\big|\, \mathbf{x} = \mathbf{x}_0\big) \Big\} \tag{1.13}$$

Following this idea, the decision boundary can be estimated by solving the equation:

$$\frac{g_0(\mathbf{x})\, \pi_0}{\cancel{\sum_{k=1}^{K} g_k\, \pi_k}} = \frac{g_1(\mathbf{x})\, \pi_1}{\cancel{\sum_{k=1}^{K} g_k\, \pi_k}} \tag{1.14}$$

$$\frac{1}{(2\pi)^{m/2}\, |\Sigma_0|^{1/2}}\; e^{-1/2(\mathbf{x}-\mu_0)^T \Sigma_0^{-1} (\mathbf{x}-\mu_0)}\; \pi_0 = \frac{1}{(2\pi)^{m/2}\, |\Sigma_1|^{1/2}}\; e^{-1/2(\mathbf{x}-\mu_1)^T \Sigma_1^{-1} (\mathbf{x}-\mu_1)}\; \pi_1 \tag{1.15}$$

If making the assumption on the equal co-variance matrices for both classes ($\Sigma_0 = \Sigma_1 = \Sigma$), and taking the logarithm, the equation takes the form:

$$-\frac{1}{2}\Big(\mathbf{x}-\mu_0\Big)^T\Sigma^{-1}\Big(\mathbf{x}-\mu_0\Big)+\log\Big(\pi_0\Big) = -\frac{1}{2}\Big(\mathbf{x}-\mu_1\Big)^T\Sigma^{-1}\Big(\mathbf{x}-\mu_1\Big)+\log\Big(\pi_1\Big) \tag{1.16}$$

, which can be written in the form of the linear function $x^T\beta+\alpha=0$ as:

$$\mathbf{x}^T\Big(\Sigma^{-1}\mu_0-\Sigma^{-1}\mu_1\Big)+\frac{1}{2}\Big(\mu_1^T\Sigma^{-1}\mu_1-\mu_0^T\Sigma^{-1}\mu_0\Big)+\log\Big(\frac{\pi_0}{\pi_1}\Big)=0 \tag{1.17}$$

This equation represents the decision boundary between two classes, i.e., all samples lying on this line will have equal probability of belonging to class 0 and class 1. It should be noted that the slope of the line depends only on the class means and co-variance matrix, whereas a priori probabilities (which are the result of number of samples belonging to class 0 or 1) have effect only on the y-intercept term, i.e., the offset of the function. This is an interesting point that demands caution. If groups are unbalanced, that is, number of samples of one group is higher than in the other group, y-intercept of the decision boundary will be affected and the classifier will be biased by this disproportion. If groups are unbalanced because of the incomplete or missing data, whereas in reality they are balanced, this can have a negative effect.

When considering multiclass classification problem, probability of a sample belonging to each class is firstly estimated by the equation:

$$p_k=-\frac{1}{2}\log|\Sigma|-\frac{1}{2}\Big(\mathbf{x}-\mu_k\Big)^T\Sigma^{-1}\Big(\mathbf{x}-\mu_k\Big)+\log\Big(\pi_k\Big) \tag{1.18}$$

and then the class is estimated as the one with the highest probability as:

$$h(\mathbf{x})=\underset{k}{\operatorname{argmax}}\, p_k(\mathbf{x}) \tag{1.19}$$

Support Vector Machine

Try to solve the problem directly and never solve a more general problem as an intermediate step.

Vladimir Vapnik

Support vector machine is nowadays known as a very powerful classifier with a lot of different applications. The big advantage over LDA is the fact that it is a *non-parametric* classifier. The model is not obtained using assumptions of the form of the class density function and estimation of it's parameters, which is inevitably erroneous. Instead, SVM forms the decision boundary using the samples (not their density estimates) by maximizing the distance between samples and the boundary. This was the idea Vladimir Vapnik, the inventor of this method stood for. It is better to try to solve the problem directly and simply, without many intermediate steps that can be complicated and inaccurate.

In pattern recognition, the decision rule (h) is usually obtained by multiplying the sample ($\mathbf{x}$) by predefined weights (Θ):

$$\Theta^T \mathbf{x} + \Theta_0 \tag{1.20}$$

, where Θ_0 is a constant. If samples $\mathbf{x}_0$ and $\mathbf{x}_1$ lay on the decision boundary, following statements are true:

$$\Theta^T \mathbf{x}_0 + \Theta_0 = \Theta^T \mathbf{x}_1 + \Theta_0 \tag{1.21}$$

$$\Theta^T (\mathbf{x}_0 - \mathbf{x}_1) = 0 \tag{1.22}$$

This result implies that Θ is perpendicular to the boundary:

$$\Theta \perp (\mathbf{x}_0 - \mathbf{x}_1) \tag{1.23}$$

The goal of the SVM is to find the decision boundary between two classes so that the distance between the samples and the decision boundary, i.e., the margin is maximized. The distance (d) from a sample to the decision boundary can be defined as the distance between the sample $\mathbf{x}$ and any point lying on the boundary, $\mathbf{x}_0$, projected onto the vector Θ.

$$d = \frac{\Theta^T \big(\mathbf{x} - \mathbf{x}_0\big)}{\big|\Theta\big|} \tag{1.24}$$

Term $|\Theta|$ is introduced to normalize the vector Θ. Without the normalization the distance would depend on the norm of Θ.

Since $\mathbf{x}_0$ is on the decision boundary, the expression $\Theta^T\mathbf{x}_0 + \Theta_0 = 0$ is valid, and, therefore, the expression for the distance can be written as:

$$d = \frac{\Theta^T\mathbf{x} + \Theta_0}{|\Theta|} \tag{1.25}$$

Margin (M) can be defined as the distance from the boundary to the closest sample:

$$M = \min_i d_i \tag{1.26}$$

Depending on which side of the boundary the sample is located, the distance can be positive or negative. In order to keep it strictly positive, term y is introduced, where $y \in \{-1, 1\}$:

$$M = \min_i \left\{ y_i d_i \right\} \tag{1.27}$$

$$M = \min_i \left\{ \frac{y_i \left(\Theta^T\mathbf{x}_i + \Theta_0\right)}{|\Theta|} \right\} \tag{1.28}$$

The objective is to maximize the margin M. Since Θ can be rescaled, a certain Θ exists so that $y_i \left(\Theta^T\mathbf{x}_i + \Theta_0\right) = 1$, which implies

$$\exists\Theta,\ y_i \left(\Theta^T\mathbf{x}_i + \Theta_0\right) = 1 \quad \Rightarrow \quad M = \min_i \left\{ \frac{1}{|\Theta|} \right\} \tag{1.29}$$

Therefore, to maximize the margin, a separating hyperplane should be found such that a norm of vector orthogonal to the hyperplane (Θ) is minimal.

For every point not on the boundary the following term is valid:

$$y_i \left(\Theta^T\mathbf{x}_i + \Theta_0\right) > 0 \tag{1.30}$$

Value C can be selected such that:

$$y_i \left(\Theta^T \mathbf{x}_i + \Theta_0\right) > C \tag{1.31}$$

$$y_i \left(\frac{\Theta^T \mathbf{x}_i}{C} + \frac{\Theta_0}{C}\right) > 1 \tag{1.32}$$

Since Θ and Θ_0 can be rescaled, it can be written:

$$\Theta := \frac{\Theta}{C}, \quad \Theta_0 := \frac{\Theta_0}{C} \tag{1.33}$$

, and, therefore:

$$y_i \left(\Theta^T \mathbf{x}_i + \Theta_0\right) > 1 \tag{1.34}$$

Finally the optimization problem states:

$$\min \frac{1}{2}|\Theta|^2, \quad s.t. \;\; y_i \left(\Theta^T \mathbf{x}_i + \Theta_0\right) > 1. \tag{1.35}$$

L_2 norm is preferred because it has continuous derivative, whereas constant $1/2$ is introduced for the mathematical convenience. The optimization is solved using Lagrangian method as:

$$L(\Theta, \Theta_0, \alpha_i) = \frac{1}{2}|\Theta|^2 - \sum_{i=1}^{n} \alpha_i \left[y_i \left(\Theta^T \mathbf{x}_i + \Theta_0\right) - 1\right] \tag{1.36}$$

$$\frac{\partial L}{\partial \Theta} = \Theta - \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i = 0 \quad \Rightarrow \quad \Theta = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i \tag{1.37}$$

$$\frac{\partial L}{\partial \Theta_0} = \sum_{i=1}^{n} \alpha_i y_i = 0 \tag{1.38}$$

By rewriting the problem in 1.36 in terms of dual variable α, the following expression can be obtained:

$$L(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_j \sum_i \alpha_j \alpha_i y_j y_i \mathbf{x}_i^T \mathbf{x}_j \tag{1.39}$$

Since this function depends only on dual variable α, the solution can be obtained by maximization:

$$\max L(\alpha) \quad s.t. \quad \begin{cases} \alpha_i \geq 0 \\ \sum_i \alpha_i y_i = 0 \end{cases} \tag{1.40}$$

In this optimization problem, the objective has the form of quadratic function, whereas constraints are linear. This problem is typically solved using quadratic programming. Since it is a convex problem, the solution will always be global maximum. Once the dual variable α is found, the primal variable Θ can be calculated using the equation 1.37.

In the optimization, Karush-Kuhn-Tucker conditions need to be satisfied (Boyd and Vandenberghe, 2004). One of this condition is *complementary slackness*, stating that in the optimal point (the solution of the problem), the product of dual variable and the constraint must be zero:

$$\alpha_i \left[y_i \left(\Theta^T \mathbf{x}_i + \Theta_0 \right) - 1 \right] = 0 \tag{1.41}$$

This condition explains well the principle of SVM. Since the dual variable must be greater or equal to zero ($\alpha \geq 0$), there are two possibilities:

1. If α is greater than zero, $\left[y_i \left(\Theta^T \mathbf{x}_i + \Theta_0 \right) - 1 \right]$ must equal one:

$$\alpha_i > 0 \quad \Rightarrow \quad y_i \left(\Theta^T \mathbf{x}_i + \Theta_0 \right) = 1 \tag{1.42}$$

2. If $\left[y_i \left(\Theta^T \mathbf{x}_i + \Theta_0 \right) - 1 \right]$ is greater than zero, α must be zero:

$$y_i \left(\Theta^T \mathbf{x}_i + \Theta_0 \right) > 1 \quad \Rightarrow \quad \alpha = 0 \tag{1.43}$$

Since for all samples lying on the margin, the statement

$$y_i \left(\Theta^T \mathbf{x}_i + \Theta_0 \right) = 1 \tag{1.44}$$

holds, α will be greater than zero only for the samples lying on the decision hyperplane, whereas for the samples further away from the hyperplane, α will be zero. Given the fact that Θ

depends on linear combination of samples weighted by α (eq. 1.37), only the samples lying on the boundary will have effect in the calculation of Θ (where $\alpha > 0$), and they are called *support vectors*. The inconveniency of this approach is the fact that the data need to be linearly separable, i.e., there should not be any data on the other side of the margin, which is rarely the case in practice. For this reason it is called the *hard margin SVM*. Margin has distance one from the boundary and all points have to be distanced more or equal (constraint in eq. 1.34). To relax this constrain, variable β_i is introduced for every sample $\mathbf{x}_i$, such that $\beta_i \geq 0$:s

$$y_i \left(\Theta^T \mathbf{x}_i + \Theta_0\right) \geq 1 - \beta_i \tag{1.45}$$

For points lying on the other side of the margin, β will be positive, whereas for the points on the margin or on the correct side of it, it will be zero. This is the ground assumption for *soft margin SVM*. The new optimization problem states:

$$\max \frac{1}{2}|\Theta|^2 + \gamma \sum_{i=1}^{n} \beta_i \quad s.t. \quad \begin{cases} y_i \left(\Theta^T \mathbf{x}_i + \Theta_0\right) \geq 1 - \beta_i \\ \beta_i \geq 0 \end{cases} \tag{1.46}$$

The term $\gamma \sum_{i=1}^{n} \beta_i$ is introduced to minimize this effect, whereas γ is the constant of penalization. The procedure of solving the problem is the same as in hard margin SVM, using the Lagrangian method:

$$L(\Theta, \Theta_0, \beta_i, \alpha_i, \lambda_i) = \frac{1}{2}|\Theta|^2 + \gamma \sum_{i=1}^{n} \beta_i - \sum_{i=1}^{n} \alpha_i \left[y_i \left(\Theta^T \mathbf{x}_i + \Theta_0\right) - 1 + \gamma\beta_i\right] - \sum_{i=1}^{n} \lambda_i \beta_i \tag{1.47}$$

$$\frac{\partial L}{\partial \Theta} = \Theta - \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i = 0 \quad \Rightarrow \quad \Theta = \sum_{i=1}^{n} \alpha_i y_i \mathbf{x}_i \tag{1.48}$$

$$\frac{\partial L}{\partial \Theta_0} = \sum_{i=1}^{n} \alpha_i y_i = 0 \tag{1.49}$$

$$\frac{\partial L}{\partial \beta_i} = \gamma - \alpha_i - \lambda_i = 0 \tag{1.50}$$

By rewriting the optimization problem in terms of dual variable α, the same term can be obtained

as in eq. 1.39:

$$L(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_j \sum_i \alpha_j \alpha_i y_j y_i \mathbf{x}_i^T \mathbf{x}_j \tag{1.51}$$

, and the new optimization problem states:

$$\max L(\alpha) \quad s.t. \quad \begin{cases} \alpha_i \geq 0 \\ \lambda_i \geq 0 \end{cases} \tag{1.52}$$

However, since objective function $L(\alpha)$ does not depend on dual variable λ_i, the substituation can be made following the expression in eq. 1.50 and the new optimization problem states:

$$\max L(\alpha) \quad s.t. \ 0 \leq \alpha_i \leq \gamma. \tag{1.53}$$

This is the only diference between hard margin SVM and soft margin SVM.

It is important to state that the optimization problem does not depend on $\mathbf{x}$, but on $\mathbf{x}^T\mathbf{x}$. This allows the use of *kernel trick* and implicitly enables nonlinear transform of the feature space at little additional cost. Usually, non-linear decision boundary can be achieved by nonlinear transform of features:

$$\mathbf{x} \rightarrow \Phi(\mathbf{x}) \tag{1.54}$$

However, this operation is computationally expensive. The solution can be achieved using kernel functions. Kernel is a function $K(x, y)$ for which:

$$K(\mathbf{x}, \mathbf{y}) = \Phi(\mathbf{x})^T \Phi(\mathbf{y}) \tag{1.55}$$

Since in the equation 1.38 $\mathbf{x}$ does not appear by itself, but in a form of dot product $\mathbf{x}^T\mathbf{x}$, non-linear transform can be used in a form of kernel trick:

$$L(\alpha) = \sum_i \alpha_i - \frac{1}{2} \sum_j \sum_i \alpha_j \alpha_i y_j y_i K(\mathbf{x}_i, \mathbf{x}_j) \tag{1.56}$$

Most often used kernel is a radial basis kernel ($K_{RBF}(\mathbf{x}_i, \mathbf{x}_j)$):

$$K_{RBF}(\mathbf{x}_i, \mathbf{x}_j) = e^{\frac{-\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}} \tag{1.57}$$

Although SVM is conceptually designed as a two-class classifier, techniques for multiclass classification also exist, e.g. *one-versus-one* or *one-versus-all*.

Chapter 2

Problem statement

2.1 Introduction

Voluntary movements are achieved by the contraction of skeletal muscles controlled by the Central and Peripheral Nervous system. The contraction is initiated by the release of a neurotransmitter that promotes a reaction in the walls of the muscular fiber, producing a biopotential known as Motor Unit Action Potential (MUAP) that travels from the neuromuscular junction to the tendons. The surface electromyographic signal records the continuous activation of such potentials over the surface of the skin and constitutes a valuable tool for the diagnosis, monitoring and clinical research of muscular disorders. Moreover, the use of electrode arrays facilitate the investigation of the peripheral properties of the active Motor Units such as: conduction velocity and fatigue (Soares et al., 2015); anatomical characteristics in terms of location of the innervation zones (Beck et al., 2012), the spatial composition of the muscle, that is, muscle compartmentalization (Vieira et al., 2010); and change in spatial distribution of MUAPs with exercise and pain (Madeleine et al., 2006). This last property of the muscles has proven to be very useful to infer motion intention not only regarding the direction of the movement but also its power (Rojas-Martínez et al., 2013).

2.2 Task identification

Something on task identification Problems

2.3 Objectives

Main objective

This doctoral thesis addresses the problem of extraction of information from muscular patterns obtained from multichannel surface electromyography and associated with different movement directions. The aim of the thesis is to analyze the muscular pattern of upper-limb muscles during isomeric contractions and its relationship to neuromuscular disorders, particularly to incomplete spinal cord injury. This information can be useful for the identification of motion intention, i.e. identification of intended motor task and force based on EMG.

Specific objectives

To achieve the main objective, this thesis strives for the following specific objectives:

I To develop a pattern recognition-based procedure for identification of task and force of isometric contractions (i.e. extract motor intention). For this assignment, different features and methods for their selection as well as classification techniques will be evaluated and compared in order to choose the best-suited solution. Special attention will be paid to features related to spatial distribution of myoelectric intensity recorded over the surface of the muscle. This is a new and unexplored approach that has proven to have high potential in identification.

II To test stability and robustness of extracted features regarding physiological and non-physiological changes which are consequences of long-term contractions (i.e. myoelectric fatigue and gel drying).

III To publish the obtained results and conclusions in high-impact journals, as well as in international and national conferences.

Methods were applied to control subjects as well as to patients with incomplete spinal cord injury with reduced mobility of the upper-limb.

2.4 Thesis framework

This thesis and the published articles that provide its content as a compendium were developed in the *Department of Automatic Control (ESAII)* of the *Universitat Politècnica de Catalunya (UPC)* under the framework of the brain research line of the *BIOsignal Analysis for Rehabilitation and Therapy Research Group (BIOART)*, which belongs to the *Biomedical Signals and Systems* division of the *Biomedical Engineering Research Centre (CREB)* of UPC that belongs to the Biomedical Research Networking Center in Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN). The research was done with the collaboration of the Institut Guttman in Badalona (Spain) and the Laboratory of Engineering of Neuromuscular System and Motor Rehabilitation at the Politecnico di Torino.

Furthermore, this work has been supported by multiple funding projects:

1. Ayudas para la contratación de personal investigador novel (FI-DGR 2014). *Agencia de Gestión de Ayudas Universitarias y de Investigación (AGAUR) - Generalitat de Catalunya.*

2. Sistemas multicanal de análisis y sensorización para rehabilitación y monitorización clínica. (DPI2011-22680) *Ministerio de Economia, Industria y Competitividad (MINECO)*

3. Design of methods for assessing processes of neurological and neuromuscular decline associated with aging. (DPI201459049R) *Ministerio de Economia, Industria y Competitividad (MINECO)*

Chapter 3

Conclusion

3.1 Summary

Magnetoencephalography is a noninvasive technique that provides excellent temporal resolution and a whole-head coverage that allows the spatial mapping of cerebral sources. These characteristics make MEG an appropriate technique to localize the epileptogenic zone (EZ) in the preoperative evaluation of pharmacoresistant epilepsy. There is great interest in using interictal biomarkers to determine the EZ because these can occur more frequently than seizures (**?**) and reduce the discomfort of the patient during the recordings.

Presurgical evaluation with MEG can guide the placement of invasive electrodes, the current gold standard in the clinical practice, and even supply sufficient information for a surgical intervention without invasive recordings, reducing invasiveness, discomfort, and cost of the presurgical epilepsy diagnosis (**?**). However, MEG signals have low signal-to-noise ratio compared with intracranial EEG and can sometimes be contaminated by noise that mask or distort the brain activity. This may prevent the detection and the localization with interictal epileptiform discharges (IEDs) and high frequency oscillations (HFOs), two important biomarkers used in the preoperative evaluation of epilepsy.

To achieve the main goal of this thesis, namely the development and validation of methods for noninvasive localization of interictal biomarkers with MEG, its signal-to-noise ratio must

improve. The reduction of two kinds of interference was aimed: metallic artifacts that affect exclusively MEG recordings and mask the activity of IEDs; and the high-frequency noise, produced mainly by muscular interferences that mask HFO activity. Considering the large number of MEG channels and the long time of the recordings, reducing noise and marking events manually is a time-consuming task. The algorithms presented in this thesis provide automatic solutions aimed at the reduction of interferences and, the detection of HFOs.

In chapter **??**, an automatic BSS-based algorithm to reduce metallic interference was developed and validated using simulated and real signals (**?**). Three methods were evaluated: AMUSE, a second-order BSS technique; and INFOMAX and FastICA, based on high-order statistics. To objectively evaluate the effectiveness of BSS and the subsequent interference reduction, simulated signals consisting in real artifact-free data mixed with real metallic artifacts were generated. Subsequently, an automatic detection of the artifactual components was proposed, exploiting the known characteristics of metallic-related interferences. Results indicated that AMUSE performed better when recovering brain activity and allowed an effective removal of artifactual components.

In chapter **??** the influence of metallic artifact filtering using the previously developed algorithm was evaluated in the source localization of IEDs in patients with refractory focal epilepsy (**?**). In this study, a comparison between the resulting positions of equivalent current dipoles (ECDs) produced by IEDs was performed: without removing metallic interference, rejecting only channels with large metallic artifacts and after BSS-based reduction. The results showed that a significant reduction on dispersion was achieved using the BSS-based reduction procedure, yielding feasible locations of ECDs in contrast to the other two approaches.

In chapter **??** an algorithm for the automatic detection of epileptic ripples in MEG using beamformer-based virtual sensors was developed (**?**). The automatic detection of ripples was performed using a two-stage approach. In the first step, beamforming was applied to the whole head to determine a region of interest where more high frequency activity was taking place. In the second step, the automatic detection of ripples was performed using the time-frequency characteristics of these oscillations. The performance of the algorithm was evaluated using intracranial EEG recordings as gold standard. Furthermore, the number of events that were

detected inside the region of interest were significantly higher than the number of events found outside.

The main conclusions of the different studies and its relationship with the objective of the thesis are summarized in the following section.

3.2 Main conclusions

Nine subjects with iSCI performed four isometric forearm tasks (flexion, extension, supination, and pronation) at three levels of effort (10% MVC, 30% MVC, and 50% MVC). High density EMG was measured on five muscles of forearm and upper arm in monopolar configuration. Intensity maps were calculated for each muscle and three different feature sets were extracted: the average intensity of an HD-EMG map (I), the intensity and center of gravity of an HD-EMG maps (I + CG), and the intensity of a single differential channel (Diff) (gold standard). Using the extracted feature sets and LDA-based classification, both task and effort level were identified, and the influence of fatigue and other time-dependent changes (e.g. drying of conductive gel) on identification was evaluated. Since the goal of this study was to analyze different feature sets rather than classification methods, LDA was utilized given that this method is the most commonly used, and is generally recommended for myoelectric interfaces (Hakonen et al., 2015). Although it assumes normal distribution of patterns in each class, it has proven to have good performance even when the normality assumption does not hold (Grouven et al., 1996).

When identification using the different features was tested on signals recorded in short time intervals, the combination of I + CG outperformed the other feature sets. The results show that a muscular co-activation pattern exists not only for the task intention ($Acc = 98.7\%$; $S = 96.8\%$; $P = 97.0\%$; $SP = 99.2\%$), but also for the force intention ($Acc = 98.8\%$; $S = 92.5\%$; $P = 93.2\%$; $SP = 99.4\%$).

Although the identification based on the features Diff has slightly better performance in average than the identification based on the features I, a repeated measures ANOVA showed that there is no significant difference in their distributions. Moreover, a small displacement in the position of bipolar electrodes can have a great effect on signal intensity, as well as on spectral content.

Consequently, if using Diff as features in classification, a small displacement can have a high influence on the identification performance. This effect does not exist in feature I, making it more robust to small changes in the position of the electrodes. On the other hand, the identification based on the combination of intensity and spatial features significantly outperforms both of them. This result was obtained both for identification of tasks and identification of tasks and effort levels. Furthermore, it has been shown that the classifier based on I + CG discriminates between types of tasks at low levels of effort (10% MVC) significantly better than the classifiers based on the other feature sets (Figure **??**).

The impedance between electrodes and skin changes during time on account of several causes, e.g., drying of conductive gel and sweating. Consequently, the identification performance deteriorates as the time between the training of the classifier and the identification increases. When the identification is performed long after the training of classifier, the results show that the identification based on I + CG performs just slightly better than the identification based on I features, while the identification based on Diff features is much worse ($S_{I+CG} = 94\%$, $P_{I+CG} = 95\%$; $S_I = 93\%$, $P_I = 94\%$; $S_{Diff} = 83\%$, $P_{Diff} = 83\%$). Although it may seem that, in average, spatial features do not improve the classification with respect to using only the intensity of an HD-EMG map, it is important to outline that these results were obtained on contractions of high levels of effort (50% MVC), where performances were similar even when contractions were recorded at the same time (see **??**).

Muscle fatigue also affects the recorded EMG signal both in the time and spectral domains and therefore the identification performance deteriorates with fatigue. The results of this work show that the classifier based on intensity and spatial features is less sensitive to fatigue than classifiers based on the other feature sets. The proposed classifier shows a very good performance in task identification even at the final stage of fatigue ($Acc = 91.3\%$, $S = 84.3\%$, $P = 87.0\%$, $SP = 93.5\%$).

The proposed method could significantly improve the human-machine interface technology and can be used in numerous applications: computer games, exoskeletons, automatic wheelchairs, rehabilitation robots, prostheses, etc. As suggested by Müller-Putz et al. (Muller-Putz et al., 2015), non-invasive hybrid brain-computer interfaces (BCI) can be designed as EEG-based BCI

supplemented with other biological and mechanical signals. For example, they reported significantly higher identification results for motion intention when using a hybrid BCI system composed of EEG and EMG sensory systems than when using only one of them. EMG usually has higher SNR ratio than EEG and it is widely used in the identification of the motion intention, however, it is prone to malfunction due to fatigue. When fatigue occurs, the supplemented EEG input keeps the identification stable, and increases the robustness of the system. Thus, advances in obtaining methods more robust to fatigue or time effect are very interesting.

Some patients with neuromuscular impairment can weakly activate their muscles, but insufficiently to generate a movement. In these patients, as well as in patients that can generate only weak movements, HD-EMG maps can be generated and used in identification of motion intention, as demonstrated in this study. This approach could supplement the existing BCI or inertial sensors based prostheses and result in a device with a better performance. For example, Rohm et al. (Rohm et al., 2013) performed a very interesting study with a single SCI patient. Their neuroprosthesis consisted of a functional electrical stimulation of the forearm and upper arm muscles, and a semiactive elbow orthosis. Using BCI and a shoulder joystick, the patient was able to perform complex hand and elbow tasks from everyday life (e.g. eating an ice cream cone). The reported performance of that study was 70%, which was remarkable considering the fact that the patient did not have any control over involved muscles. However, performance of similar patients could be increased using hybrid BCI if myoelectric activation exists.

Furthermore, compared to inertial signals, which are also used as input to control devices, EMG has a major advantage because myoelectric activation precedes the actual movement, which can save valuable response time.

However, it should be noted that although this study represents an improvement in the identification of motion intention, additional experiments should be considered in the future. Firstly, HD-EMG recordings were carried out during controlled isometric submaximal contractions, i.e. patient's arm was fixed and supported by a mechanical brace. Since the methodology was capable to successfully and automatically differentiate between none, very low, low and medium effort levels, we might hypothesized that the method can be useful in prediction without the support of the brace. However, more experiments without the brace and the analysis of the

recorded HD-EMG signals would be necessary to confirm and quantify this hypothesis.

In this study, the spatial distribution of EMG intensity was evaluated for identification of tasks and different levels of effort in patients with iSCI. Results show that the spatial activation of motor units is dependent on the type of exercise and contraction intensity, and that related features can improve identification performance.

Although results show that spatial features also enhance the robustness of the identification to time effect and fatigue, additional experiments need to be performed to test robustness to temporal dependent changes more thoroughly and to determine when the classifier fails by further tests done on fatigue.

The center of gravity was used as a figure of merit to describe the spatial distribution. Although it shows a significant improvement in classification, by definition it is insensitive to fine changes in the distribution of muscle units. Therefore, in future works, more appropriate measures of spatial distribution should be analyzed in order to better describe the spatial distribution of muscle intensity. Also, additional features as those related to the frequency content could be considered to improve even more the classification performance.

In order to demonstrate the existence of distinguishable group-specific patterns in HD-EMG, the identification of different tasks was performed. Within-group identification of motion intention at different effort levels was tested on nine patients with iSCI performing four upper limb tasks (flexion/extension of the elbow and supination/pronation of the forearm) at three different effort levels (10%, 30%, and 50% MVC).

Although a single type of a classifier would be sufficient to demonstrate the existence of different patterns, for an additional verification two types of classifiers were evaluated in the identification of motion intention: LDA and SVM. The former is a classical, simple, and computationally efficient classification method, whereas the latter is a more powerful classifier that can employ a nonlinear transform of features to improve their separability among classes. In this paper, a SVM with radial kernel was considered. Although the SVM is superior in classification performance, the LDA is commonly used in myocontrol applications because of its simplicity and performance in real-time. However, with the increasing computational power of new computer generations, SVM could become more common in these applications.

The identification of tasks was tested using two feature sets: 1) the average intensities of HD-EMG activation maps (I) of five muscles and 2) the combination of average intensities and centers of gravity (I+CG) of the activation maps of five muscles. On the other hand, a conjoint identification of tasks and effort levels was designed as two-step classifier, following the procedure described by Rojas et al. (Rojas-Martínez et al., 2013) and tested on a healthy population. The first step comprised the identification of tasks using a combination of intensity and spatial features of all five muscles, whereas in the second step the levels of effort were identified separately for each task. The effort levels were identified using a combination of the intensity and spatial features of agonist-antagonist muscle pairs involved in the task (Rojas-Martínez et al., 2013).

HD-EMG activation maps were calculated for all exercises and compared among patients.

Rojas-Martínez et al. (**?**) calculated the relative standard deviation between maps within a group of healthy subjects (17.4% in average), reporting an increase in standard deviation between maps with increasing effort levels (12.1%, 16.6%, and 23.6% for 10%, 30%, and 50% MVC, respectively). As expected, the dispersion between maps of iSCI patients was considerably higher (56% in average), but the variability was similar in the case of patients with iSCI (Table 1). However, when maps were compared among patients with the same level of injury, the standard deviation between maps was greatly reduced (19% in average). Moreover, the variability was higher for muscles of the upper-arm (biceps and triceps) than for forearm muscles. This reduction could be either due to a distinct activation, specific to the level of injury, or because during the rehabilitation process patients developed similar activation patterns. This is an important finding that has to be taken into account when training a classifier for a group of patients. Muscle activation patterns in patients differed from those of healthy subjects in (Rojas-Martínez et al., 2012): the Biceps Brachii was more active during supination than during flexion; the Pronator Teres was more active during supination and especially during flexion than during pronation. This could be because both muscles are particularly affected by the iSCI at the level of C4 (Young).

Furthermore, the results using the LDA showed much better identifications within the group of patients with a C4 level of injury than within the group of all patients. These findings could be

related to a higher homogeneity among patients with the same level of injury. The combination of intensity and center of gravity performed better than only intensity features. These results showed that similar patterns exist in spite of the diverse nature of their injuries. This correlation exists not only in the average intensity of the HD-EMG activation maps, but also in the spatial distribution of EMG intensity, which justifies the choice of these intensity and spatial features for automatic identification.

Finally, a considerable improvement was observed when using the SVM instead of the LDA, reaching the following results: 1) excellent automatic task identification even in the group of all patients ($Acc = 99.0\%$, $S = 97.9\%$, $P = 98.0\%$, and $SP = 99.3\%$), 2) a good combined classification of four tasks and three effort levels also in the group of all patients ($Acc = 97.5\%$, $S = 85.2\%$, $P = 85.3\%$, and $SP = 98.7\%$) which is even better in 3) conjoint identification of four tasks and low or moderate effort levels ($Acc = 98.0\%$, $S = 92.0\%$, $P = 92.4\%$, and $SP = 98.9\%$). In spite of the previous reports suggesting the greater importance of selection of the features than the selection of the classifier, our results have shown that both have considerable impact on the identification.

Several array subsets corresponding to 3×3 square grids of channels (IED = 10 mm) located at different positions were also used to evaluate the possibility of task identification using a much smaller number of electrodes. In this case, the results were considerably worse, especially when using the LDA classifier. Due to the small region covered by electrodes in each muscle, the spatial information could not be extracted and it was not possible to increase the performance as in the case of using all the electrodes.

Although this study presents an important improvement in the identification of motion intention, it is important to mention that the recordings were carried out during highly controlled isometric contractions. Therefore, even though the findings are promising, they are only a step towards final real-time applications involving free movements and multiple DoFs.

The results show that the use of a SVM-based classifier is indeed a promising approach in myocontrol-oriented pattern recognition applications. Moreover, even though a different activation pattern can be expected in subjects with neurological impairment, as in the present case, such pattern can still be associated with task and level-dependent changes in the spatial distri-

bution of the intensity, as has been previously observed in non-injured subjects (Rojas-Martínez et al., 2012).

Group-specific identification of motion intention in impaired patients has a potential to improve the translation of pattern recognition techniques to clinical practice. Unfortunately, group-specific design is a difficult topic because it assumes strong task-related and level of effort-related co-activation patterns among patients, but given the diverse nature of injuries and the high inter-patient variability, co-activation patterns are weak.

This study shows that muscular co-activation patterns in intensity and spatial distribution indeed exist. Furthermore, it shows that stronger co-activation patterns can be found between patients of the same level of injury. Whether because of the rehabilitation process or the level of injury, muscle control strategies are similar for the group of patients with an injury at C4, which makes them a more homogenous population and enables the control of universal assistive devices with higher reliability. In summary, in spite of the difficulty to identify both task and effort level in patients with iSCI, very promising results were found to provide a useful estimation of motion intention.

This study showed that the combination of intensity and spatial information is useful for the extraction of neuromuscular information. The spatial information was calculated from the RMS activation maps using the mean shift algorithm. Results were evaluated using the 70% repeated holdout method and stratified sampling as to have sufficient number of samples of each class in the sets. To prevent the type III statistical error (Mosteller, 1948; Mohebian et al., 2017), a repeated hold-out was used. Sensitivity and precision, as appropriate unbiased measures in analyzing imbalanced multi-class problems (Jordanić et al., 2016; Rojas-Martínez et al., 2013), were used to quantify the identification.

IMS features achieved very good results compared to other feature sets during task identification when the task was performed at very low effort level. Moreover, the Friedman test showed no significant differences in task identification using IMS when tasks were performed at 10% MVC, 30% MVC, or 50% MVC. This can be a very important quality in everyday applications where subject could not need to contract muscles at moderate effort level to complete the task. It can be a step toward more natural control where even slight contractions can be successfully

identified. In fact, only activations with low level of intensity are sometimes possible in patients with neuromuscular impairments.

A high identification rate is not the only factor important in the extraction of neural information from sEMG. The system should also be robust to slow time-dependent changes such as fatigue and electrode-skin contact impedance (Farina et al., 2014a). Therefore, the robustness of the proposed features was tested with respect to time and fatigue. When evaluating the time effect, no significant differences in performance were found between IMS, ICG, and I feature sets and IMS significantly outperformed TD and Diff features. However, time effect was evaluated only when test set was composed of contractions recorded at 50% MVC and, as shown in Figure **??**, all features perform similarly for the identification of that effort level. This phenomenon was already remarked and described in (Jordanic et al., 2016) where authors noted that adding spatial features to intensity features significantly improved the identification of tasks recorded at low effort levels, whereas improvement is not significant at moderate effort levels. On the other hand, the proposed features are particularly robust in task identification during fatiguing exercises and show significantly higher identification rate when compared to other features. Further improvements in reliability of the identification during the long-term contractions and fatiguing contractions can be achieved by using adaptive identification models that are being constantly updated during the usage (e.g., (Vidovic et al., 2016; Hahne et al., 2015; Sensinger et al., 2009)).

In the current work, features were extracted from the RMS activation maps of the HD-EMG. Although these features proved to be very effective, by describing the EMG signal with its RMS value, i.e., the estimator of variance, the information is partially lost. Since the gradient of the probability density function of raw EMG is a useful feature in task identification, statistical measures (e.g., modes) of the raw HD-EMG, i.e., joint distribution of instantaneous EMG amplitude over the electrode array, could provide valuable information. Moreover, in literature, features were often calculated for each channel separately and selected using the simple sequential method prior to classification (Hargrove et al., 2009; Li et al., 2017). On the other hand, Geng et al. recently proposed a more advanced channel selection method based on common spatial patterns (Geng et al., 2014). Modes of the HD-EMG density function could be correlated with the channels with discriminative information and could be a useful tool in channel

selection.

Finally, the mean shift algorithm can be used for clustering and, since it was shown that the algorithm is most effective in low-dimensional data, image segmentation is one of its most successful applications (Comaniciu and Meer, 2002). A mode of the density estimate, or in this case, a channel selected by the mean shift algorithm, can be considered as a cluster representative (Hennig et al., 2015), related to the possible image segments, where spatial (pixel locations) and range features (the intensity of the grayscale value) are considered. The advantage of the mean shift is that it can be used for clustering non-convex shapes, albeit, it could segment complex non-convex regions in the activation maps. Since segmentation of the muscle activation map can improve the neuromuscular activity estimation (Vieira et al., 2010), this could be a reason why mean shift features improved the performance of the movement detection system compared with previously published attributes. In addition, the algorithm only requires setting one parameter, bandwidth (h) and, unlike in the similar methods, it is not necessary to define the number of expected clusters. This is a big advantage because it does not require a priori knowledge on the number of clusters.

As a limitation of the study, it should be noted that the proposed features were tested only in highly controlled conditions of isometric contractions. The experiments during non-isometric contractions should be performed in order to validate the quality of the features in dynamic and more natural movements. Also, the experiment included only four tasks related to the elbow joint. Further analysis should include higher number of more complex tasks related to hand and shoulder. Moreover, all results were obtained during offline analysis. To evaluate practical aspects of the features, the experiment should be repeated using online identification and considering multiple transitions between tasks.

In conclusion, a new set of features for the identification of isometric motor tasks of upper limb was proposed. It was based on the combination of intensity and the spatial distribution of intensity of HD-EMG. These new features were evaluated using the LDA classifier and the results showed they improve the identification of tasks. Moreover, robustness of the features was tested under the influence of slow time-dependent changes of the EMG. They proved to be particularly useful for task identification when muscles were fatigued. The proposed methods

could be used for the design and monitoring of rehabilitation therapies intended for patients with neuromuscular impairment, as well as for the control of external devices like exoskeletons, and prostheses.

3.3 Main contributions

The original contributions provided by the compendium of publications of this thesis are:

- The definition of a novel algorithm for task and force identification, and its validation in terms of robustness during slow time dependent changes, such as fatigue and drying of conductive gel, on patients with incomplete spinal cord injury. Specific co-activation pattern exists both in intensity and spatial distribution for each patient.
- The co-activation pattern in intensity and its spatial distribution of HD-EMG was identified for the group of patients with spinal cord injury. Furthermore, greater similarity was found within the group of patients with similar level of injury. This result implies the possibility of building assistive/rehabilitation device for the group of patients with significantly lower training time.
- Novel spatial feature derived from the HD-EMG was used for identification of task and effort level. This feature is based on the probability density function of the HD-EMG activation map.

3.4 Future Work

The work developed in this thesis open new possibilities in the brain research line of the *BIOsignal Analysis for Rehabilitation and Therapy Research Group (BIOART)* to which the candidate belongs. Some of the most interesting further possibilities are the following:

-
-

-
-
-
-
-

3.5 Publications derived from the thesis

3.5.1 Journal papers

- Jordanić, M., Rojas-Martínez, M., Mañanas, M.A., Alonso, J.F., Marateb, H.R. A Novel Spatial Feature for the Identification of Motor Tasks Using High-Density Electromyography. *Sensors*, 17(7): 1597, 2017, JCR 2.077, Q1 in Instruments and instrumentation (10/58)

- Rojas-Martínez, M., Alonso, J.F., Jordanić, M., Romero, S., Mañanas, M.A. Identificación de tareas isométricas y dinámicas del miembro superior basada en EMG de alta densidad. *Revista Iberoamericana de Automática e Informática Industrial*, Accepted for publication 2017, JCR 0.390, Q4 in Automation and Control Systems (57/60)

- Jordanić, M., Rojas-Martínez, M., Mañanas, M.A., Alonso, J.F. Prediction of isometric motor tasks and effort levels based on high-density EMG in patients with incomplete spinal cord injury. *Journal of Neural Engineering*, 13(4): 46002, 2016, JCR 3.465, Q1 in Biomedical Engineering (13/77)

- Jordanić, M., Rojas-Martínez, M., Mañanas, M.A., Alonso, J.F. Spatial distribution of HD-EMG improves identification of task and force in patients with incomplete spinal cord injury. *Journal of NeuroEngineering and Rehabilitation*, 13(1): 41, 2016, JCR 3.222, Q1 in Rehabilitation (3/65)

Bibliography

Baker, J. J., Scheme, E., Englehart, K., Hutchinson, D. T., and Greger, B. Continuous detection and decoding of dexterous finger flexions with implantable myoelectric sensors. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(4):424–432, 2010.

Beck, T. W., DeFreitas, J. M., and Stock, M. S. Accuracy of three different techniques for automatically estimating innervation zone location. *Computer Methods and Programs in Biomedicine*, 105(1):13–21, 2012.

Boyd, S. and Vandenberghe, L. *Convex optimization*. Cambridge University Press, 2004.

Chu, J. U., Moon, I., and Mun, M. S. A Real-Time EMG Pattern Recognition System Based on Linear-Nonlinear Feature Projection for a. *IEEE Transactions on Biomedical Engineering*, 53(11):2232–2239, 2006.

Clancy, E., Morin, E. L., and Merletti, R. Sampling, noise-reduction and amplitude estimation issues in surface electromyography. *Journal of Electromyography and Kinesiology*, 12(1):1–16, 2002.

Comaniciu, D. and Meer, P. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002.

Dipietro, L., Ferraro, M., Palazzolo, J. J., Krebs, H. I., Volpe, B. T., and Hogan, N. Customized interactive robotic treatment for stroke: EMG-triggered therapy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 13(3):325–334, 2005.

Duchateau, J. and Enoka, R. M. Human motor unit recordings: Origins and insight into the integrated motor system. *Brain Research*, 1409:42–61, 2011.

Englehart, K. and Hudgins, B. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 50(7):848–854, 2003.

Englehart, K., Hudgins, B., Parker, P. a., and Stevenson, M. Classification of the myoelectric signal using time-frequency based representations. *Medical Engineering and Physics*, 21(6-7): 431–438, 1999.

Englehart, K., Hudgins, B., and Parker, P. A. A wavelet-based continuous classification scheme for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 48(3): 302–11, 2001.

Englehart, K., Hudgins, B., and Chan, A. D. C. Continuous multifunction myoelectric control using pattern recognition. *Technology and Disability*, 15(2):95–103, 2003.

Farina, D., Févotte, C., Doncarli, C., and Merletti, R. Blind separation of linear instantaneous mixtures of nonstationary surface myoelectric signals. *IEEE Transactions on Biomedical Engineering*, 51(9):1555–67, 2004.

Farina, D., Holobar, A., Merletti, R., and Enoka, R. M. Decoding the neural drive to muscles from the surface electromyogram. *Clinical Neurophysiology*, 121(10):1616–1623, 2010.

Farina, D., Jiang, N., Rehbaum, H., Holobar, A., Graimann, B., Dietl, H., and Aszmann, O. C. The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4):797–809, 2014a.

Farina, D., Merletti, R., and Enoka, R. M. The extraction of neural strategies from the surface EMG: an update. *Journal of applied physiology*, 117(11):1215–30, 2014b.

Farrell, T. R. and Weir, R. F. F. A comparison of the effects of electrode implantation and targeting on pattern classification accuracy for prosthesis control. *IEEE Transactions on Biomedical Engineering*, 55(9):2198–211, 2008.

Fougner, A., Scheme, E., Chan, A. D. C., Englehart, K., and Stavdahl, Ø. Resolving the Limb Position Effect in Myoelectric Pattern Recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 19(6):644–651, 2011.

Fougner, A., Stavdahl, O., Kyberd, P. J., Losier, Y. G., and Parker, P. a. Control of upper limb prostheses: Terminology and proportional myoelectric control - A review. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(5):663–677, 2012.

Freund, H. J., Büdingen, H. J., and Dietz, V. Activity of Single Motor Units from Human Forearm Muscles during Voluntary Isometric Contractions. *Journal of Neurophysiology*, 38 (4):933–46, jul 1975.

Geng, Y., Zhang, X., Zhang, Y.-T., and Li, G. A novel channel selection method for multiple motion classification using high-density electromyography. *Biomedical Engineering Online*, 13:102, 2014.

Grouven, U., Bergel, F., and Schultz, A. Implementation of linear and quadratic discriminant analysis incorporating costs of misclassification. *Computer Methods and Programs in Biomedicine*, 49(1):55–60, jan 1996.

Hahne, J. M., Dähne, S., Hwang, H. J., Müller, K. R., and Parra, L. C. Concurrent Adaptation of Human and Machine Improves Simultaneous and Proportional Myoelectric Control. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(4):618–627, 2015.

Hakonen, M., Piitulainen, H., and Visala, A. Current state of digital signal processing in myoelectric interfaces and related applications. *Biomedical Signal Processing and Control*, 18: 334–359, 2015.

Hargrove, L., Englehart, K., and Hudgins, B. A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control. *Biomedical Signal Processing and Control*, 3(2):175–180, 2008.

Hargrove, L., Li, G., Englehart, K., and Hudgins, B. Principal Components Analysis Preprocessing for Improved Classification Accuracies in Pattern-Recognition-Based Myoelectric Control. *IEEE Transactions on Biomedical Engineering*, 56(5):1407–1414, 2009.

Hargrove, L. J., Englehart, K., and Hudgins, B. A comparison of surface and intramuscular myoelectric signal classification. *IEEE Transactions on Biomedical Engineering*, 54(5):847–853, 2007.

Henneberg, K. Principles of Electromyography. In Bronzino, J., editor, *The Biomedical Engineering Handbook*. CRC Press, Boca Raton, second edition, 1999.

Henneman, E., Somjen, G., and Carpenter, D. O. Functional Significance of Cell Size in Spinal Motoneurons. *Journal of Neurophysiology*, 28(3), 1965.

Hennig, C., Meila, M., Murtagh, F., and Rocci, R. *Handbook of Cluster Analysis*. CRC Press, 2015.

Hogan, N., Krebs, H. I., Rohrer, B., Palazzolo, J. J., Dipietro, L., Fasoli, S. E., Stein, J., Hughes, R., Frontera, W. R., Lynch, D., and Volpe, B. T. Motions or muscles? Some behavioral factors underlying robotic assistance of motor recovery. *Journal of Rehabilitation Research and Development*, 43(5):605–618, 2006.

Holobar, a. and Farina, D. Blind source identification from the multichannel surface electromyogram. *Physiological measurement*, 35(7):143–165, 2014.

Holobar, A. and Zazula, D. Multichannel Blind Source Separation Using Convolution Kernel Compensation. *IEEE Transactions on Signal Processing*, 55(9):4487–4496, sep 2007.

Holobar, A., Minetto, M. A., Botter, A., Negro, F., and Farina, D. Experimental Analysis of Accuracy in the Identification of Motor Unit Spike Trains From High-Density Surface EMG. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(3):221–229, 2010.

Holtermann, A., Roeleveld, K., and Karlsson, J. S. Inhomogeneities in muscle activation reveal motor unit recruitment. *Journal of Electromyography and Kinesiology*, 15(2):131–137, 2005.

Hudgins, B., Parker, P., and Scott, R. N. A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 40(1):82–94, 1993.

Jiang, N., Dosen, S., Muller, K.-R., and Farina, D. Myoelectric Control of Artificial Limbs—Is There a Need to Change Focus? [In the Spotlight]. *IEEE Signal Processing Magazine*, 29(5): 152–150, 2012.

Jordanic, M., Rojas-Martínez, M., Mañanas, M. A., and Alonso, J. F. Spatial distribution of HD-EMG improves identification of task and force in patients with incomplete spinal cord injury. *Journal of NeuroEngineering and Rehabilitation*, 13(1):41, 2016.

Jordanić, M., Rojas-Martínez, M., Mañanas, M. A., and Alonso, J. F. Prediction of isometric motor tasks and effort levels based on high-density EMG in patients with incomplete spinal cord injury. *Journal of Neural Engineering*, 13(4):046002, 2016.

Kamavuako, E. N., Rosenvang, J. C., Horup, R., Jensen, W., Farina, D., and Englehart, K. B. Surface versus untargeted intramuscular EMG based classification of simultaneous and dynamically changing movements. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(6):992–998, 2013.

Li, G., Schultz, A. E., and Kuiken, T. A. Quantifying pattern recognition- based myoelectric control of multifunctional transradial prostheses. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(2):185–192, 2010.

Li, X., Samuel, O. W., Zhang, X., Wang, H., Fang, P., and Li, G. A motion-classification strategy based on sEMG-EEG signal combination for upper-limb amputees. *Journal of NeuroEngineering and Rehabilitation*, 14(1):2, 2017.

Li, Y., Chen, X., Zhang, X., and Zhou, P. Several practical issues toward implementing myoelectric pattern recognition for stroke rehabilitation. *Medical Engineering and Physics*, 36(6): 754–760, 2014.

Liddell, E. G. T. and Sherrington, C. S. Recruitment and some other Features of Reflex Inhibition. *Proceedings of the Royal Society of London B: Biological Sciences*, 97(686):488–518, 1925.

Liu, J. and Zhou, P. A novel myoelectric pattern recognition strategy for hand function restoration after incomplete cervical spinal cord injury. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 21(1):96–103, 2013.

Madeleine, P., Leclerc, F., Arendt-Nielsen, L., Ravier, P., and Farina, D. Experimental muscle pain changes the spatial distribution of upper trapezius muscle activity during sustained contraction. *Clinical Neurophysiology*, 117(11):2436–45, 2006.

Marateb, H. R., McGill, K. C., and Webster, J. G. Electromyographic (Emg) Decomposition. In *Wiley Encyclopedia of Electrical and Electronics Engineering*. John Wiley & Sons, Inc, 1999.

Marchal-Crespo, L. and Reinkensmeyer, D. J. Review of control strategies for robotic movement training after neurologic injury. *Journal of Neuroengineering and Rehabilitation*, 6:20, 2009.

Merletti, R. and Farina, D. *Surface Electromyography: Physiology, Engineering, and Applications*. Wiley-IEEE Press, Hoboken, New Jersey (USA), 2016.

Merletti, R. and Parker, P. *Electromyography : physiology, engineering, and noninvasive applications*. Wiley-IEEE Press, 2004.

Merletti, R., Aventaggiato, M., Botter, A., Holobar, A., Marateb, H., and Vieira, T. M. M. Advances in surface EMG: recent progress in detection and processing techniques. *Critical Reviews in Biomedical Engineering*, 38(4):305–45, 2010.

Mohebian, M. R., Marateb, H. R., Mansourian, M., Mañanas, M. A., and Mokarian, F. A Hybrid Computer-aided-diagnosis System for Prediction of Breast Cancer Recurrence (HPBCR) Using Optimized Ensemble Learning. *Computational and Structural Biotechnology Journal*, 15:75–85, 2017.

Mosteller, F. A k-Sample Slippage Test for an Extreme Population on JSTOR. *The Annals of Mathematical Statistics*, 19(1):58–65, 1948.

Muceli, S. and Farina, D. Simultaneous and proportional estimation of hand kinematics from EMG during mirrored movements at multiple degrees-of-freedom. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(3):371–378, 2012.

Muller-Putz, G., Leeb, R., Tangermann, M., Hohne, J. H., Kubler, A. K., Cincotti, F., Mattia, D., Rupp, R., Muller, K. R., and Millan, J. D. R. Towards Noninvasive Hybrid Brain–Computer Interfaces: Framework, Practice, Clinical Application, and Beyond. *Proceedings of the IEEE*, 103(6):926 – 943, 2015.

Naik, G. R., Kumar, D. K., and Weghorn, H. Performance comparison of ICA algorithms for Isometric Hand gesture identification using Surface EMG. In *3rd International Conference on Intelligent Sensors, Sensor Networks and Information*, pages 613–618. IEEE, 2007.

Nazmi, N., Abdul Rahman, M., Yamamoto, S.-I., Ahmad, S., Zamzuri, H., and Mazlan, S. A Review of Classification Techniques of EMG Signals during Isotonic and Isometric Contractions. *Sensors*, 16(8):1304, 2016.

Oskoei, M. A. and Hu, H. Myoelectric control systems-A survey. *Biomedical Signal Processing and Control*, 2(4):275–294, 2007.

Park, S. H. and Lee, S. P. EMG pattern recognition based on artificial intelligence techniques. *IEEE Transactions on Rehabilitation Engineering*, 6(4):400–405, 1998.

Parker, P. A. and Scott, R. N. Myoelectric control of prostheses. *Critical reviews in biomedical engineering*, 13(4):283–310, 1986.

Phinyomark, A., Limsakul, C., and Phukpattaranont, P. A Novel Feature Extraction for Robust EMG Pattern Recognition. *Journal of Computing*, 1(1):71–80, 2009.

Phinyomark, A., Phukpattaranont, P., and Limsakul, C. Feature reduction and selection for EMG signal classification. *Expert Systems with Applications*, 39(8):7420–7431, 2012a.

Phinyomark, A., Thongpanja, S., Hu, H., Phukpattaranont, P., and Limsakul, C. *Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges*. InTech, 2012b.

Rohm, M., Schneiders, M., Müller, C., Kreilinger, A., Kaiser, V., Müller-Putz, G. R., and Rupp, R. Hybrid brain-computer interfaces and hybrid neuroprostheses for restoration of upper limb functions in individuals with high-level spinal cord injury. *Artificial Intelligence in Medicine*, 59(2):133–142, 2013.

Rojas-Martínez, M., Mañanas, M. a., and Alonso, J. F. High-density surface EMG maps from upper-arm and forearm muscles. *Journal of Neuroengineering and Rehabilitation*, 9:85, jan 2012.

Rojas-Martínez, M., Mañanas, M. a., Alonso, J. F., and Merletti, R. Identification of isometric contractions based on High Density EMG maps. *Journal of Electromyography and Kinesiology*, 23(1):33–42, 2013.

Sensinger, J., Lock, B., and Kuiken, T. Adaptive Pattern Recognition of Myoelectric Signals: Exploration of Conceptual Framework and Practical Algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 17(3):270–278, 2009.

Sherrington, C. S. Remarks on some Aspects of Reflex Inhibition. *Proceedings of the Royal Society of London B: Biological Sciences*, 97(686):519–545, 1925.

Simon, A. M., Hargrove, L. J., Lock, B. a., and Kuiken, T. a. A decision-based velocity ramp for minimizing the effect of misclassifications during real-time pattern recognition control. *IEEE Transactions on Biomedical Engineering*, 58(8):2360–2368, 2011.

Soares, F. A., Carvalho, J. L. A., Miosso, C. J., de Andrade, M. M., and da Rocha, A. F. Motor unit action potential conduction velocity estimated from surface electromyographic signals using image processing techniques. *BioMedical Engineering Online*, 14(1):84, 2015.

Squire, J. *Muscle: design, diversity, and disease.* Benjamin/Cummnigs, Menlo Park, CA, USA, 1986.

Stango, A., Negro, F., and Farina, D. Spatial Correlation of High Density EMG Signals Provides Features Robust to Electrode Number and Shift in Pattern Recognition for Myocontrol. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 23(2):189–198, 2015.

Staudenmann, D., Roeleveld, K., Stegeman, D. F., and van Dieen, J. H. Methodological aspects of SEMG recordings for force estimation - A tutorial and review. *Journal of Electromyography and Kinesiology*, 20(3):375–387, 2010.

Staudenmann, D., van Dieën, J. H., Stegeman, D. F., and Enoka, R. M. Increase in heterogeneity of biceps brachii activation during isometric submaximal fatiguing contractions: a multichannel surface EMG study. *Journal of neurophysiology*, 111(5):984–90, 2014.

Tkach, D., Huang, H., and Kuiken, T. a. Study of stability of time-domain features for electromyographic pattern recognition. *Journal of Neuroengineering and Rehabilitation*, 7:21, 2010.

Tucker, K., Falla, D., Graven-Nielsen, T., and Farina, D. Electromyographic mapping of the erector spinae muscle with varying load and during sustained contraction. *Journal of Electromyography and Kinesiology*, 19(3):373–9, 2009.

Vaca Benitez, L. M., Tabie, M., Will, N., Schmidt, S., Jordan, M., and Kirchner, E. A. Exoskeleton technology in rehabilitation: Towards an EMG-based orthosis system for upper limb neuromotor rehabilitation. *Journal of Robotics*, 2013:13, 2013.

Vidovic, M. M.-C., Hwang, H.-J., Amsuss, S., Hahne, J. M., Farina, D., and Muller, K.-R. Improving the Robustness of Myoelectric Pattern Recognition for Upper Limb Prostheses by Covariate Shift Adaptation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(9):961–970, 2016.

Vieira, T. M. M., Merletti, R., and Mesin, L. Automatic segmentation of surface EMG images: Improving the estimation of neuromuscular activity. *Journal of Biomechanics*, 43(11):2149–58, 2010.

Widmaier, E. P., Raff, H., and Strang, K. T. *Vander's Human Physiology: The mechanisms of Body Function.* McGraw-Hill Higher Education, 2014.

Young, A. J., Hargrove, L. J., and Kuiken, T. A. The Effects of Electrode Size and Orientation on the Sensitivity of Myoelectric Pattern Recognition Systems to Electrode Shift. *IEEE Transactions on Biomedical Engineering*, 58(9):2537–2544, 2011.

Young, A. J., Smith, L. H., Rouse, E. J., and Hargrove, L. J. Classification of simultaneous movements using surface EMG pattern recognition. *IEEE Transactions on Biomedical Engineering*, 60(5):1250–1258, 2013.

Young, W. Spinal Cord Injury Levels & Classification, W M Keck Center for Collaborative Neuroscience.

Zardoshti-Kermani, M., Wheeler, B., Badie, K., and Hashemi, R. EMG feature evaluation for movement control of upper extremity prostheses. *IEEE Transactions on Rehabilitation Engineering*, 3(4):324–333, 1995.

Zhang, X. and Zhou, P. High-Density Myoelectric Pattern Recognition Toward Improved Stroke Rehabilitation. *IEEE Transactions on Biomedical Engineering*, 59(6):1649–1657, 2012.

Zwarts, M. J. and Stegeman, D. F. Multichannel surface EMG: basic aspects and clinical utility. *Muscle & nerve*, 28(1):1–17, 2003.

Zwarts, M. J., Lapatki, B. G., Kleine, B. U., and Stegeman, D. F. Surface EMG: how far can you go? *Supplements to Clinical Neurophysiology*, 57:111–9, 2004.