# Chapter 1

# Introduction

An Electromyogram or Electromyographic (EMG) signal is the recording of the electrical activity produced by muscles. It measures the electric currents generated in muscles during their contraction. The EMG signal provides insight into the neural activation and dynamics of the muscles, and is therefore important for clinical investigations that attempt to diagnose neuromuscular deficiencies such as those caused by stroke and Parkinson's disease [50]. Other potential important applications exist in areas such as aging, exercise physiology, space medicine, and ergonomics, where it is of interest to understand whether the control of muscles is altered as a consequence of aging, exercise, exposure to microgravity, fatigue, and excessive and prolonged force production [38]. The EMG signal can also be used to identify hand gestures and finds its application in controlling prosthetic devices for rehabilitation or wearable devices for gaming [57]. In particular, the work in this thesis is motivated by rehabilitation for patients with spinal cord injury. The EMG signal is very important for researchers and practitioners to monitor and evaluate the effect of the rehabilitation training and the condition of muscles [24], as the EMG signal provides information that helps infer the neural activity in the spinal cord. Before the work in this thesis, EMG analysis in [24] requires significant amount of manual labeling of interesting signal features. The motivation of this thesis is to fully automate the EMG analysis tasks and yield accurate, consistent results.

The processing of the EMG signal is a hard problem in general. Similar to action potentials and electrocardiogram (ECG), the useful information (the actual muscle response other than noise) in an EMG recording is classified as a transient signal, as it is contained within a short time ($\sim 10-50ms$). Every EMG recording contains multiple muscle responses. The difficulty in processing EMG arises from the fact that the transient muscle response is a transient signal with unknown arrival time, unknown duration, and unknown shape. In addition, the EMG signal recorded from patients with spinal cord injury during rehabilitation is very different from the EMG signal of normal healthy people undergoing the same motions. For example, some of the muscle responses are very weak and thus hard to detect. Because of this, traditional EMG processing tools and methods are either not applicable or insufficient. In this thesis, two EMG processing tasks were addressed and effective methods were derived. The solutions proposed in this thesis are not only helpful to the rehabilitation after spinal cord injury, but applicable to other general processing tasks on transient signals (biological signals in particular).

The first task is to detect useful information (the transient muscle response) from EMG recording. The detection of transient signals with unknown arrival time, unknown duration, and unknown shape is still an active research problem. There is not a universal optimal detector. Usually, different detection schemes are employed to tackle transient signals from different applications. In this thesis, a peak-based transient detection algorithm was proposed. The methodology consists of a combination of several techniques stemming from multi-resolution wavelet decomposition, statistics, and detection theory. The detection method is totally unsupervised, and it's especially good at detecting trasient signals with low signal-to-noise ratio (SNR). Simulation results show that this method works much better than existing methods found in the literature. In addition, the simulation shows that this method is applicable to general transient signals.

The other task addressed in this thesis is to automatically segment and classify the EMG signal into different areas of interest. A theoretical framework is proposed to segment the EMG signal based on the detected peaks. The scale information of the detected peaks is used to derive a measure for its effective support. Several different techniques have been adapted together to solve the clustering problem. An initial hierarchical clustering is first performed to obtain most of the monosynaptic Motor Evoked Potentials (MEPs). Principal component analysis (PCA) is used to reduce the number of features and effect of the noise. The reduced feature set is then fed to a Gaussian mixture model (GMM) to further divide the MEPs into different groups of similar shapes. The method of breaking down a segment of multiple consecutive MEPs into individual MEPs is also derived.

## 1.1 Spinal Cord Injury

The work of this dissertation is motivated by the field of *spinal cord injury* (SCI). The goal of the spinal cord injury research is to help patients with severe spinal cord injury to recover their ability to stand, walk, and have voluntary movements [24]. According to the report published by the National Spinal Cord Injury Statistical Center in 2012, the number of people in the United States who have SCI has been estimated to be approximately 270,000, with approximately 12,000 new cases each year. In particularly, about 30% of the cases are *complete* SCI, which means all the sensation and voluntary control of some parts of the body is lost [58]. Figure 1.1-1 shows that SCI is mainly caused by motor vehicle crashes, followed by falls and violence. Because of this, SCI primarily affects young

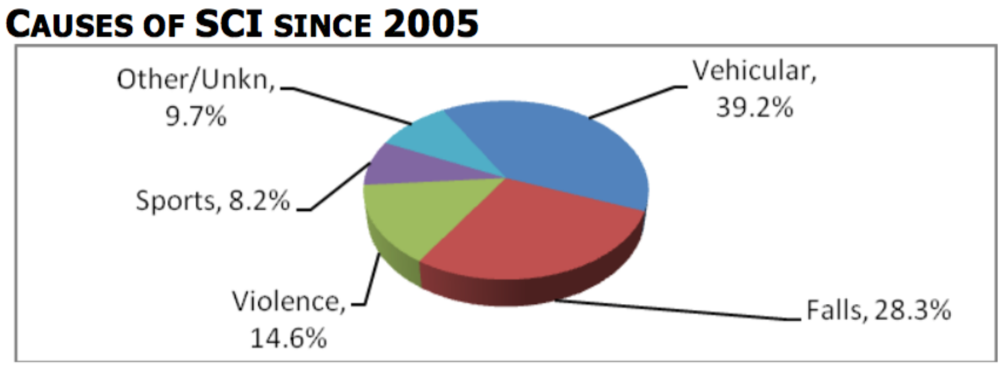



Figure 1.1-1: Causes of SCI since 2005 [58]

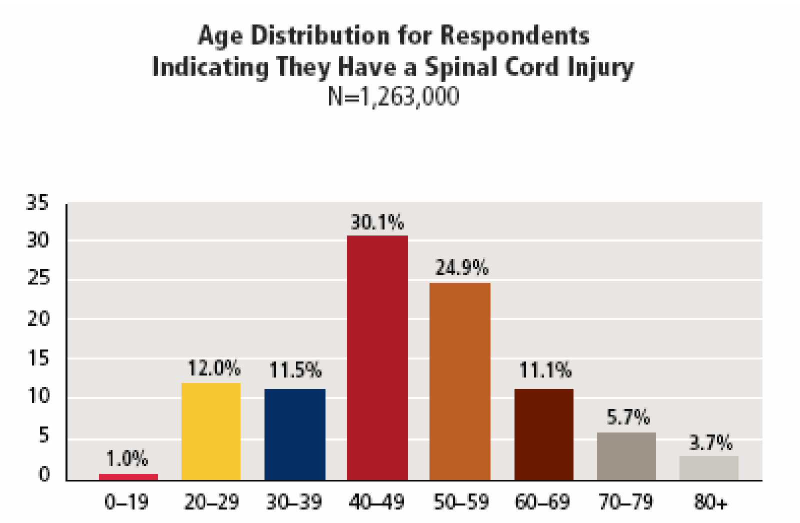
adults. According to the data from the Christopher & Dana Reeve Foundation, the average age of those who reported being paralyzed due to a spinal cord injury was 48, and the average length of time since the spinal cord injury occurred was 14 years, as shown in Figure 1.1-2 [9]. This indicates that those sustaining SCI tend to be young individuals suffering traumatic injury at the prime of their personal lives and economic earning potentials.

Right now, SCI cannot be cured, and rehabilitation can be very difficult. Besides the pain and extreme high cost of lifetime medical care, people with severe SCI cannot currently stand or walk, which can impact their social and work life. The work described in this thesis supports a collaborative effort between Caltech, University of Louisville, and UCLA to provide new therapies for SCI [24].

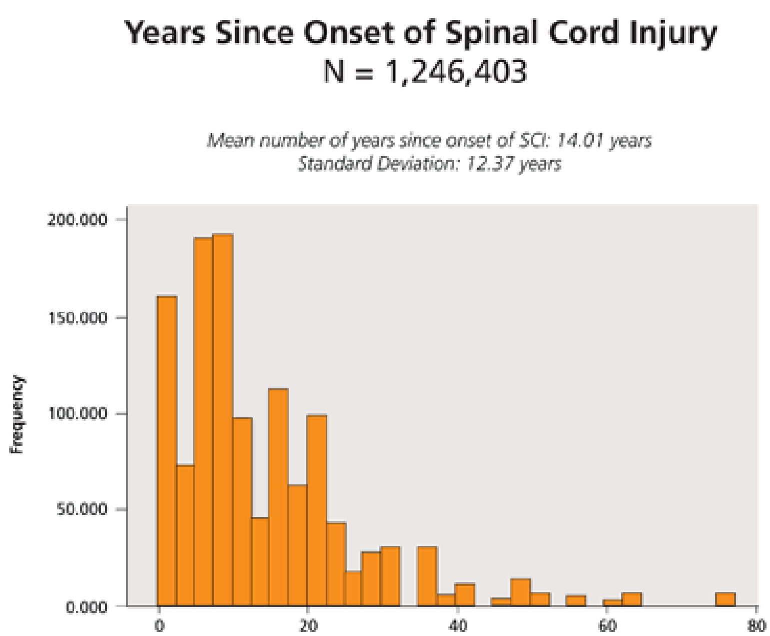
## 1.2 Recovery of Spinal Cord Injury

The spinal cord is the main pathway for information connecting the brain and the rest of the body. The human spinal cord is divided into 31 different segments: different segments connect to different parts of the body, as shown in Figure 1.2-1. When the spinal cord is injured, this pathway is blocked, and patients can lose sensation and voluntary control of some parts of the body depending on which level of the spinal cord is injured. In addition to the role of transmission of neural signals, the spinal cord also contains neural circuits that can independently control numerous reflexes and the central pattern generator. In the case of SCI, the spinal circuitry below the lesion remains intact, but loses supraspinal modulation from the brain.

Although there are no fully restorative treatments for SCI, various rehabilitative, cellular, and molecular therapies have been tested in animal models. In the cellular therapies, cells are transplanted, for example, to replace dead cells. In the molecular therapies, different pharmacological agents are used to restore biochemical imbalance after SCI. Rehabilitative training, such as loco-




(a) Age Distribution of SCI Patients [9]




(b) Years since onset of SCI [9]

Figure 1.1-2: Some statistics of SCI patients: age distribution as in (a); years since onset as in (b)

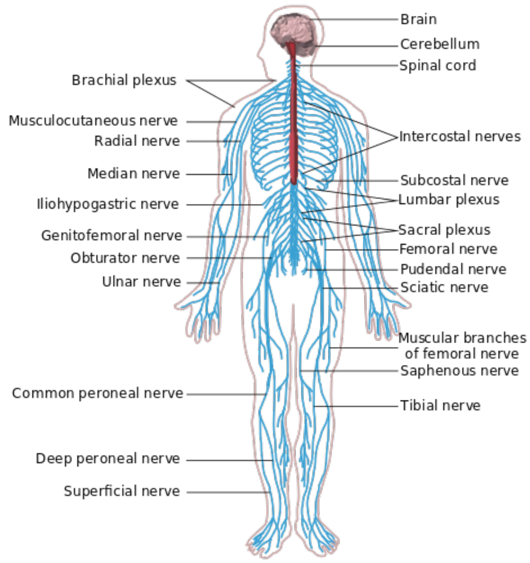



Figure 1.2-1: Human Spinal Cord [63]

motor training, improves the locomotor function. For a complete review, please see [59] and the references within. Our strategy for recovering from spinal cord injury is derived mainly from two important facts about the spinal circuitry.

First, basic posture and locomotion control depends significantly on the spinal circuitry, not the brain. In fact, one can stand without much conscious effort. The spinal circuitry needs to achieve some crucial level of excitability in order to work properly. For healthy humans, the brain sends neural signals (train of action potentials) down to the spinal cord to raise the excitation level of the spinal circuitry. For people with complete SCI, this modulation is lost, and hence the spinal circuitry cannot function properly to generate locomotion [12]. Our strategy is to use electrical stimulation to reactivate previously silent spinal circuitry.

The second fact is the plasticity of the spinal circuitry. The spinal circuitry can be trained to adapt to the loss of modulation from the brain after SCI. Therefore, task-specific training is performed to let the spinal circuitry adapt to the new modulation from the electrical stimulation [24].

Epidural Electrostimulation (EES) involves electrically stimulating the spinal cord via an electrode or multi-electrode array placed in the epidural space of the vertebral canal. To date, human studies of EES in [24] have been based on an epidural spinal cord stimulation unit from Medtronic (RestoreADVANCED, Medtronic, Minneapolis, MN, USA), a FDA-approved commercial product for back pain management. The electrode array was implanted over spinal cord segments L1–S1, the lumbosacral enlargement, which is responsible for lower-body movement.

The electrode array dimension is about 49mm by 10mm, and it contains 16 electrodes. Different

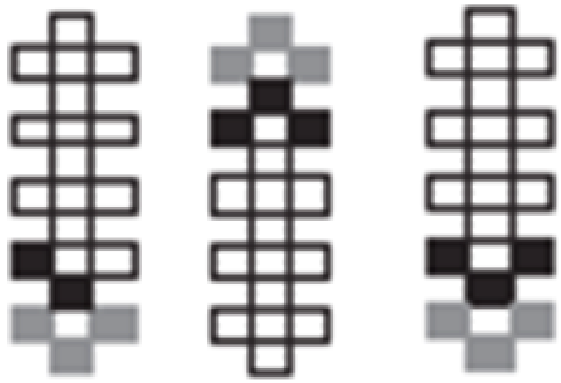


Figure 1.2-2: Examples of electrode configurations: anodes are black, and cathodes are gray, while the rest are floating electrodes (from Harkema [24]).

patterns of the stimulating electric field can be chosen by assigning anodes and cathodes to different electrodes (See Figure 1.2-2). The stimulation signal is a pulse train whose parameters, such as frequency, pulse width, and amplitude, can be adjusted.

EES must be coupled with physical training for maximum effect. The training currently employed includes standing, stepping, and voluntary movement control while lying supine. Various sensors are used during the training to analyze the performance of a subject under EES, and guide future training. In particular, surface electrodes are placed on various major muscle groups of the lower body, and the electric signal from the muscle, formally known as the Electromyogram or Electromyographic (EMG) signal, is recorded during the training. The EMG signal is one of the most important clinical data obtained in this project and many other neuromuscular research efforts. My research is to develop new tools for analyzing and processing the EMG signal.

## 1.3 Electromyographic (EMG) Signal Processing: Objective Statement

The Electromyographic (EMG) Signal is the electric potential generated by the muscle cells. In general, the EMG signal is very useful in many areas, such as detection of the medical abnormalities, and analysis of the biomechanics of movement. For spinal cord research, the EMG signal is one of the most important quantities to evaluate motor performance and training. Ideally, one would like to know how neurons in the spinal cord react to the epidural electrical stimulation. In practice, it's almost impossible to directly measure the neural activities in the spinal cord. On the other hand, when motor neurons in the spinal cord are activated, they send action potentials down to the specific muscles and activate the muscle fibers. EMG is a measurement of the electric activity happening around the muscle fibers. Therefore, the EMG signal can be used to infer the activity of the motor neurons in the spinal cord. In Section 2.1.1, a detailed review of the physiology of the EMG signal is presented. In summary, if one thinks of the spinal cord as a black box, by checking its output (the EMG signal), one can often infer the neural activity inside the spinal cord.

A major effort in the EMG processing with the application to the diagnosis of neuromuscular disorder is to decompose the EMG signal into its continuant Motor Unit Action Potential Trains (MUAPTs) [53, 38]. A fully description of MUAPTs is given in Section 2.1.1. Basically, EMG decomposition involves mainly two steps: Firstly, detect/segment the raw EMG signal to obtain individual Motor Unit Action Potentials (MUAPs), and then group MUAPs from the same motor unit to obtain MUAPTs. Another research topic on EMG processing is to detect the onset of muscle activity [50, 57]. In this problem, waveforms of individual MUAPs are no long desired; instead, the EMG signal is viewed on a higher level, and active periods of the EMG signal are determined.

In the following, a brief description of the two specific EMG processing tasks to be addressed in this thesis is given first. For detailed descriptions, please refer to Chapter 3 and Chapter 4, respectively. After that, a brief discussion on the challenges and motivations is presented. For full description on the challenges, please read Section 2.2. For a complete literature review, please read the sections within Chapter 3 and Chapter 4.

### 1.3.1 Peak-based Detection

There are many different kinds of processing we can do on the EMG signal. One very important aspect of the EMG signal is to measure the level of activation. More specifically, we'd like to detect the *Motor Evoked Potential* (MEP). MEPs are the action potentials along the muscle fibers generated by active motor neurons as a result of EES. MEPs are transient signals, and are corrupted by noise in the recording. Our goal is to detect all the peaks of the MEPs in the EMG signal (See Figure 1.3-1). The MEP peaks can potentially tell us what motor neurons are active, when they are active, and how strong the activation is. Throughout the thesis, EMG signal detection refer to the detection of the MEP peaks.

### 1.3.2 Segmentation and Classification

Another Important task is to segment and classify the EMG signal into various intervals of interest. Segmentation refers to the task of dividing the given EMG signal into various sections, and classification is to label the segmented sections according to their nature. In the example of the EMG signal obtained from EES, segmentation is supposed to find the sections of data that only contain

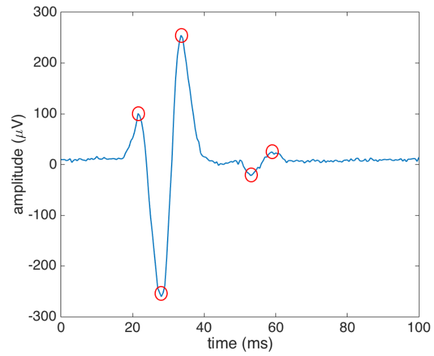



Figure 1.3-1: An example showing the detection of the peaks of MEPs in the EMG signal: red circles show the peaks of MEPs. (The example EMG signal is from the muscle of left medial gastrocnemius while the patient is lying in supine position under EES.)

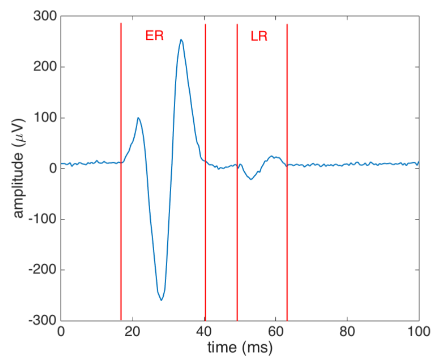



Figure 1.3-2: An example showing the segmentation of the EMG signal: ER (short for Early Response) is the label for the early MEP; LR (short for Late Response) is the label for the late MEP. (The example EMG signal is from the muscle of left medial gastrocnemius while the patient is lying in supine position under EES.)

MEP waveforms. In addition, the MEPs can be further classified into early MEPs and late MEPs (See Figure 1.3-2). An early MEP is the direct response after an electrical stimulus, and so it's normally strong. A late MEP is the indirect response after an electrical stimulus. It comes after the early MEP and is normally weak. Both early MEPs and late MEPs are important, and they correspond to different underlying biophysics. Therefore, it's crucial to identify and differentiate them. With this information, one can better infer the neural activities in the spinal cord. Please refer to Section 2.1.2 for a full description of early MEPs and late MEPs.

Before my work, EMG processing was entirely carried out manually by many technicians at University of Louisville. It's laborious to manually label all the features (including peaks, early and late MEPs), since the rehabilitation training generates enormous amount of data from every experiment. Typically, a training session lasts a couple of hours. There are about 20 EMG channels, each of which measures a specific muscle group in the lower body. The sampling rate is 2000Hz. As a result, the EMG data file from one experiment easily exceeds 100MB. In addition, manual inspection and processing of the EMG signal are very subjective and the quality of the results depends on the experience and the knowledge of the individuals. Therefore, an automatic tool to process the EMG signal will save lots of time for the practitioners, and give reliable, accurate results.

Another motivation behind our work is to facilitate the machine learning algorithm developed to optimize the stimulation parameters [11]. The EMG signal is a very important input signal to the algorithm. The machine learning algorithm is supposed to automatically adjust the stimulation parameters from all observations. Thus, an unsupervised, accurate EMG processing algorithm is needed to enable on-line learning and adjustment of the array stimuli.

There are mainly two challenges behind the EMG processing work in this thesis.

First of all, the EMG signal is in nature hard to process automatically. The muscle response (the signal carrying useful information other than noise) is a transient signal with unknown arrival time, unknown duration, and unknown shape. The characteristics of the muscle responses vary from time to time, from subject to subject. As a result, most of the EMG processing in clinical research is still carried out with lots of human supervision, and the methods employed for detection and segmentation are still rudimentary [34, 37, 15, 8, 7, 27]. See Section 3.1 for a complete review of the literature on EMG detection.

Second of all, the EMG processing tasks and the characteristics of the EMG signal in the research on spinal cord injury (in particular, for rehabilitation with epidural electro-stimulation as in [24]) are quite different from those in typical research on neuromuscular disorder. As a result, no existing tools can be used to solve the problems encountered in this project, and new tools and methodologies need to be invented.

## 1.4 Thesis Outline and Contributions

The primary contribution of this thesis is the development of a set of automatic, unsupervised tools for the analysis of Electromyogram (the EMG signal) for the purpose of studying electrical stimulation based rehabilitation on patients with spinal cord injuries. A wavelet-based, double-threshold algorithm was developed for the detection of the transient peaks in the EMG signal (Chapter 3). Based on the transient peak detection result, the EMG signal is further segmented and classified into various groups of monosynaptic MEPs and polysynaptic MEPs using techniques stemming from Principal Component Analysis (PCA), hierarchical clustering, and Gaussian mixture model (Chapter 4). A software with graphic user interface has been implemented in Matlab. The software implements the proposed peak detection algorithm, and enables the physiologists to visualize the detection results and modify them if necessary.

Chapter 2 gives an introduction to the necessary background in order to understand the thesis. It first talks about the physiology of the generic EMG signal and the EMG signal from patients with spinal cord injuries undergoing rehabilitation. With that, the characteristics of the EMG signal and the challenge of processing it are explained. After that, a basic introduction to the classical detection theory and the wavelet transform is given, because these two are the main theories behind the transient peak detection algorithm proposed in Chapter 3. In particular, two classical detectors, the matched filter and the energy detector, are introduced because they normally serve as the upper bound and the lower bound of any given detector. Frequency properties of the wavelet transform are studied here in order to help understand the derivation of the choices of scales later in Chapter 3.

Chapter 3 extends existing theories in the transient detection field. The application of the wavelet transform in the detection of transient signals has been studied extensively and employed successfully. However, most of the theories assumes certain knowledge about the shapes of the transient signals, which makes it hard to be generalized to the transient signals with arbitrary shapes. The proposed detection scheme focuses on the more fundamental feature of most of the transient signals (in particular the EMG signal): peaks. The continuous wavelet transform with Mexican Hat wavelet is employed. This thesis theoretically derived a framework for selecting a set of scales based on the frequency domain information. Ridges are identified in the time-scale space to combine the wavelet coefficients from different scales. By imposing two thresholds, one on the wavelet coefficient and one on the ridge length, the proposed detection scheme can achieve both high recall and high precision. A systematic approach for selecting optimal parameters via simulation is proposed and demonstrated. Comparing with other state-of-the-art detection methods, the proposed method in this thesis yields better detection performance, especially in the low Signal-to-Noise-Ratio (SNR) environment.

In Chapter 4, a method for automatically segmenting and clustering the EMG signal is derived. A theoretical framework is proposed to segment the EMG signal based on the detected peaks. The scale information of the detected peaks is used to derive a measure for its effective support. Several different techniques have been adapted together to solve the clustering problem. An initial hierarchical clustering is first performed to obtain most of the monosynaptic Motor Evoked Potentials (MEPs). Principal component analysis (PCA) is used to reduce the number of features and the effect of the noise. The reduced feature set is then fed to a Gaussian mixture model (GMM) to further divide the MEPs into different groups of similar shapes. The method of breaking down a segment of multiple consecutive MEPs into individual MEPs is derived.

Chapter 5 reviews the work and the contributions of this thesis and talks about some future directions and applications.