

INTERNATIONAL BURCH UNIVERSITY
FACULTY OF ENGINEERING AND INFORMATION
TECHNOLOGIES
INFORMATION TECHNOLOGIES DEPARTMENT



EVALUATING EFFECTS OF DENOISING AND FEATURE
EXTRACTION METHODS ON CLASSIFICATION OF EMG
SIGNALS

PhD DISSERTATION

Ercan GOKGOZ

Supervisor

Prof. Dr. Abdulhamit SUBASI

SARAJEVO

June, 2014

**EVALUATING EFFECTS OF DENOISING AND FEATURE
EXTRACTION METHODS ON CLASSIFICATION OF EMG
SIGNALS**

Ercan GOKGOZ

Dissertation Submitted in Fulfillment of Requirement for the Degree of Doctor of
Philosophy in Information Technology

INTERNATIONAL BURCH UNIVERSITY

2014

APPROVAL PAGE

Student Name & Surname : Ercan GOKGOZ
Faculty : Faculty of Engineering and IT
Department : Information Technologies
Dissertation Title : Evaluating effects of denoising and feature extraction methods on classification of EMG signals
Date of Defense : 19.06.2014

I certify that this final work satisfies all the requirements as a PhD Dissertation for the degree of Doctor of Philosophy in Information Technology.

.....
Assist. Prof. Dr. Elmedin SELMANOVIC
Head of Department

I certify that this final work satisfies all the requirements as a PhD Dissertation for the degree of Doctor of Philosophy in Information Technology.

.....
Prof. Dr. Abdulhamit SUBASI
Supervisor

Examining Committee Members

Assoc.Prof. Dr. Kemal HAJDAREVIC	UNSA	
Prof. Dr. Abdulhamit SUBASI	IBU	
Prof. Dr. Meliha HANDZIC	IBU	
Assist. Prof. Dr. Elmedin SELMANOVIC	IBU	
Assist. Prof. Dr. Nejdett DOGRU	IBU	

EVALUATING EFFECTS OF DENOISING AND FEATURE EXTRACTION METHODS ON CLASSIFICATION OF EMG SIGNALS

ABSTRACT

Machine Learning techniques and algorithms are extensively using in biomedical signal processing field to produce diagnosis system for definition of neuromuscular diseases. Different approaches have been applied for quantitative analysis of biomedical signals. Electromyography (EMG) biomedical signal data is used in this study to define neuromuscular diseases which consist of electrical currents from muscles during its contraction. Motor Unit Action Potentials (MUAPs) in EMG signals provides an important source of information for the diagnosis of neuromuscular disorders.

This study shows how machine learning techniques are used for detection of neuromuscular diseases. EMG signals were recorded under usual and real conditions for quantitative analysis. The EMG signals were recorded from five places in the muscle at three levels of insertion. The dataset consist of control, a group of patients with Myopathy and Amyotrophic Lateral Sclerosis (ALS) data.

EMG signals contain noise while traveling on different tissues. Moreover, the EMG signal acquisition collects signals from motor units at a time which may be effected by different signals. Numerous feature extraction and denoising methods applied to remove noise from signal and to extract features from signal.

Feature extraction methods such as Autoregressive (AR) Burg, covariance (COV), multivariate covariance (MCOV), Eigen value, Multiple Signal Classification (MUSIC), discrete wavelet transform (DWT) and classifiers such as k-Nearest Neighbor (k-NN), Artificial neural network (ANN), Radial Basis Function (RBF), Support vector machine (SVM), Regression Trees (CART), C4.5 and Random Forest classifiers are utilized in this study. Multiscale Principal Component Analysis (MSPCA) denoising method applied to remove noise from EMG signal.

The effect of the MSPCA denoising and feature extraction methods is discussed on EMG signal classification by applying different classifiers. The comparisons between the developed classifiers were based on a number of scalar performances such as sensitivity, specificity, accuracy, F-measure and area under ROC curve (AUC). The results show that MSPCA de-noising has considerably increased the accuracy as compared to EMG data without MSPCA de-noising. Combination of DWT feature extraction and Random Forest achieved best performance for k-fold cross validation with 96.67% total classification accuracy. These results demonstrate that the presented framework have the potential to achieve a reliable classification of EMG signals, and to support the clinicians for making an accurate diagnosis of neuromuscular disorders.

Keywords: Electromyography (EMG), Motor unit action potentials (MUAPs), Multiscale Principle Component Analysis (MSPCA), ARburg (AR), Covariance (COV), Multivariate Covariance (MCOV), Eigen value, Multiple Signal Classification (MUSIC), discrete wavelet transform(DWT) and classifiers such as k-Nearest Neighbor (k-NN), Artificial neural network (ANN), Radial Basis Function (RBF), Support vector machine (SVM), Classification and Regression Trees (CART), C4.5 and Random Forest

ACKNOWLEDGEMENTS

There are many people who helped to make my years at the graduate school most valuable. First, I thank Prof. Dr. Abdulhamit SUBASI, who is my supervisor. Having the opportunity to work with her over the years was intellectually rewarding and fulfilling.

The last words of thanks go to my family. I thank my parents Azime and Ismail GOKGOZ and my brothers Kazim and Engin GOKGOZ for their patience and encouragement.

Lastly my precious thank goes to my wife Hatice GOKGOZ, for her extreme patience and endless support through this long and tough journey. Her support and help in this work cannot be expressed in words.

DECLARATION

I hereby declare that this PhD dissertation titled Effects of feature extraction and denoising methods in order to detect muscular diseases on EMG signals using classifiers is based on my original work except quotations and citations which have been duly acknowledged. I also declare that this dissertation has not been previously or concurrently submitted for the award of any degree, at International Burch University, any other University or Institution.

Ercan GOKGOZ

Date:

TABLE OF CONTENT

APPROVAL PAGE	iii
ABSTRACT	iii
ACKNOWLEDGEMENTS	v
DECLARATION	vi
TABLE OF CONTENTS.....	xi
LIST OF TABLES.....	xi
LIST OF FIGURES	xii
LIST OF APPENDICES	xiv
LIST OF ABBREVIATIONS.....	xv

CHAPTER

1. INTRODUCTION.....	1
1.1Background	1
1.2 Importance and Motivation.....	5
1.3 Objectives	7
1.4 Key Contributions	8
1.5 Overview of the Dissertation.....	8
2. ELECTROMYOGRAM (EMG).....	10
2.1 Introduction	10
2.2 The Human Muscle Physiology.....	13
2.2.1 Anatomy of the Muscle	13
2.2.2 The Structure and Function of Muscle	17

2.2.3 The Motor Unit Action Potential	24
2.3 Neuromuscular Disorders	29
2.3.1 Neuropathy Disorders: Mononeuropathies and Polyneuropathies ...	30
2.3.1.1 Mononeuropathies	31
2.3.1.2 Polyneuropathies	32
2.3.2 Motor Neuron Disease(MND)	37
2.3.3 Radiculopathy	39
2.3.4 Myopathy	40
2.3.5 Neuromuscular Junction Disorders (NMJ)	42
2.4 The Electromyograph (EMG)	43
2.4.1 The Electromyograph and Instrumentation	45
2.4.2 EMG Electrodes	45
2.4.3 Signal Acquisition	52
2.4.4 Signal Amplification and Filtering	53
2.4.5 Signal Digitization	54
2.4.6 Factors Influencing Electromyograph Readings	54
 3. BIOMEDICAL SIGNAL PROCESSING TECHNIQUES	 56
3.1 Fourier Series Analysis	56
3.1.1 Periodic Functions	56
3.1.2 Discrete Time Fourier Analysis	57
3.2 Spectral Analysis: Frequency Resolution	59
3.3 Truncated Fourier Analysis : Data Windowing	61
3.3.1 Power Spectrum	63
3.4 Ar - Spectral Analysis : Modern Techniques	66
3.4.1 Parametric Based Model Methods	66
3.4.1.1 Covariance	69
3.4.1.2 Multivariate Covariance	71
3.4.2 Non-Parametric Model Eigen Frequency Estimations	72
3.4.2.1 Burg's Method	74
3.4.2.2 Multiple Signal Classification (MUSIC) method	76
3.5 Time-Frequency Analysis	78
3.5.1 Short-Term Fourier Transform:The Spectrogram	78
3.5.2 Wigner-Ville Distribution : Special Case Of Cohen Class	79

3.6 Wavelet Analysis	82
3.6.1 The Continuous Wavelet Transform	83
3.6.2 Wavelet Time-Frequency Characteristics.....	84
3.6.3 Discrete Wavelet Transform.....	86
3.7 Multiscale Principal Component Analysis	91
4. MACHINE LEARNING TECHNIQUES	94
4.1 K - Nearest Neighbor (k-NN)	94
4.2 Artificial neural networks (ANN)	95
4.2.1 The Perceptron: An Artificial Neuron	97
4.2.2 Multilayer Perceptrons.....	99
4.2.3 The Back-Propagation Algorithm.....	100
4.4 Support vector machines (SVMs)	104
4.5 Classification And Regression Trees (CART)	106
4.6 C4.5	107
4.7 Random Forest	108
5. METHODOLOGY	109
5.1 Research Design	109
5.2 Data Collection	109
5.3 Procedure	110
5.4 Data Analysis Methods	111
5.4.1 Signal Denoising Method	111
5.4.2 Feature Extraction Methods.....	111
5.4.3 Classification Methods	111
6. RESULTS	112
6.1 Introduction	112
6.2 Performance Evaluation	112
6.3 Experimental Results	114
6.3.1 Results without MPSCA denoising	115
6.3.2 Results with MSPCA Denoising	116
6.3.3 Effect of denosing with MSPCA	118

7. DISCUSSIONS	123
7.1 Main Findings.....	123
7.2 Implications for Research.....	125
7.3 Implications for Practice	126
7.4 Limitations	126
7.5 Future Directions	127
 8. CONCLUSIONS	 128
 REFERENCES.....	 129
 APPENDICES	 139
 CURRICULUM VITAE.....	 162

LIST OF TABLES

TABLE 2.1 Table Shows Categorization of Neuromuscular Disorders and some Specific Diseases Associated with Each Group.....	30
TABLE 6.1 F-Measure results of classifiers for each feature extraction method without MSPCA denoising	116
TABLE 6.2 ROC (AUC) results of classifiers for each feature extraction method without MSPCA denoising	116
TABLE 6.3 Total Classification Accuracy (%) results of classifiers for each feature extraction method without MSPCA denoising	116
TABLE 6.4 F-Measure results of classifiers for each feature extraction method with MSPCA denoising	117
TABLE 6.5 ROC (AUC) results of classifiers for each feature extraction method with MSPCA denoising	118
TABLE 6.6 Total Classification Accuracy (%) results of classifiers for each feature extraction method with MSPCA denoising	118
TABLE 6.7 Total accuracies of all classifiers with each feature extraction method before and after MSPCA denoising.	119
TABLE 7.1 Effect of generalization parameter C in SVM method with different kernels	124
TABLE 7.2 Comparison of results and methods with other studies which used same dataset	126

LIST OF FIGURES

FIGURE 1.1 Schematic representation of feature extraction from EMG signal	7
FIGURE 2.1 Cross-section of skeletal muscle and connection to the bone via tendons	14
FIGURE 2.2 Diagram depicts the physiology of the human skeletal muscle, from the muscle fiber to the position of the sarcomeres.	16
FIGURE 2.3 Cross-section of the sarcoplasmic reticulum showing the T-tubes, myofibrils, and sarcolemma	21
FIGURE 2.4 Diagram depicts the thick myofilaments composed of myosin protein structures.....	21
FIGURE 2.5 Diagram depicts the thin myofilaments composed of actin, troponin, and tropomyosin protein structures	22
FIGURE 2.6 Structure of thin and thick myofilaments in cross-bridge motion	23
FIGURE 2.7 The generation of a MUAP of a single motor unit with four muscle fibers.	26
FIGURE 2.8 The radial nerve is derived from the posterior cord of the brachial plexus.....	33
FIGURE 2.9 In the upper arm, the radial nerve first derives several cutaneous nerves of the arm followed by muscular branches to the triceps brachii.	34
FIGURE 2.10 The peroneal nerve anatomy showing the nerve branches into the calf, feet, and toes.	35
FIGURE 2.11 The common, deep, and superficial peroneal nerves are shown.	36
FIGURE 2.12 The facial nerve and its major branches are depicted innervating some of the major face muscles.....	38
FIGURE 2.13 Block diagram describing the main steps in biomedical signal analysis.	45
FIGURE 2.14 Examples of concentric and monopolar EMG needle electrodes sold by Med Products.	47
FIGURE 2.15 Experimental setup using five channel sEMG for detecting limb motion.....	49
FIGURE 3.1 Schematic representation of model-based methods of spectral estimation	67
FIGURE 3.2 Schema of Selection of features in DWT approach.....	88
FIGURE 3.3 Plotting of each features after DWT feature extraction and MSPCA denoising for control signal	90
FIGURE 4.1 An artificial neuron or perceptron	97
FIGURE 4.2 Activation functions used in artificial neurons: (a) Step function (b) Gaussian (c) piecewise linear and (d) sigmoid.	98

FIGURE 4.3 A double-layer MLP network.	100
FIGURE 4.4 Two linearly separable classes by an infinite number of hyperplanes. ...	104
FIGURE 4.5 The nonlinear mapping makes inseparable points in input space separable in feature space.	105
FIGURE 4.6 The soft-margin support vector classifier.	106
FIGURE 5.1 Blok diagram of proposed system.	110
FIGURE 6.1 Graphical representation of Classifiers' performances after DWT feature extraction.	117
FIGURE 6.2 Graphical representation of results of both approaches after ARBURG Feature extraction	120
FIGURE 6.3 Graphical representation of results of both approaches after COV Feature extraction	120
FIGURE 6.4 Graphical representation of results of both approaches after MCOV Feature extraction	121
FIGURE 6.5 Graphical representation of results of both approaches after EIGEN Feature extraction	121
FIGURE 6.6 Graphical representation of results of both approaches after MUSIC Feature extraction	122
FIGURE 6.7 Graphical representation of results of both approaches after DWT Feature extraction	122

LIST OF APPENDICES

APPENDIX A Plotting of methods before and after denoising.....	140
APPENDIX B Table representation of results before and after denoising.....	156

LIST OF ABBREVIATIONS

EMG	Electromyography
MUAPs	Motor unit action potentials
ANN	Artificial neural network
SVM	Support vector machine
k-NN	k-Nearest Neighbor
MSPCA	Multiscale Principle Component Analysis
MUSIC	Multiple Signal Classification
CART	Classification and Regression Trees
RBF	Radial Basis Function
DWT	Discrete Wavelet Transform
COV	Covariance
MCOV	Multivariate Covariance
EDX	Electrodiagnostic
PSO	Particle Swarm Optimization
SNR	Signal Noise Ratio
MLPNN	Multi Layer Perceptron neural networks
DFNN	Dynamic fuzzy neural network
ANFIS	Adaptive Neuro-Fuzzy Inference System
ALS	Amyotrophic Lateral Sclerosis

CHAPTER 1

INTRODUCTION

1.1 Background

The nervous system controls the muscle activity during both contraction and relaxation processes. Electromyography (EMG) signal is a biomedical signal that measures electrical currents generated in muscles during both contraction and relaxation. Thus, ElectroMyoGraphy (EMG) has been investigated to analysis and evaluation of electrical activity in muscles not only biological but also for technical inferences.

Motor Unit Action Potentials (MUAPs) are smallest functional action units of the structure which provide an important source of information for the diagnosis of neuromuscular disorders. Although MUAPs have different shapes, almost the same for each discharge. Thus, MUAPs can be detected and investigated using pattern recognition and machine learning techniques to determine the disease, i.e. neuropathy or myopathy (McGill, 2002; Loudon, Jones, & Sehmi, 1992).

Recently signal processing techniques and machine learning methods have received extensive attention in EMG signal analysis and classification for diagnosis of neuromuscular disorders. Frequently used signal processing techniques are Fourier transform, autoregressive modeling, wavelet transform, time-frequency approaches (Subasi A. , 2005; Subasi & Kiymik, 2010; Begg, Lai, & Palaniswami, 2008). ANN used to classify motor unit action potentials (MUAP) of muscles (Schizas, Pattichis, & Middleton, 1990). Support vector machine (SVM) and artificial neural network were applied to analyze the classification of the electromyogram (EMG) signals which includes different subjects such as normal, neuropathy and myopathy.

AR coefficients were obtained from EMG signals by using autoregressive (AR) modeling. The AR coefficients were provided to SVM and ANN as input. Besides, these coefficients were tested both in ANN and SVM. The results show that SVM has high coherence level in the diagnosis of neuromuscular disorders (Güler & Koçer, 2005). SVM and ANN are also utilized together to diagnose neuromuscular disorders in another study (Akben, Subasi, & Tuncel, 2012). Another work investigates the performance of neuro-fuzzy system for analyzing and classifying EMG signals recorded from normal, neuropathy, and myopathy subjects in a similar approach. EMG signals were obtained from different disorder and subjects. Coefficients were obtained from the EMG signals using Autoregressive (AR) analysis and applied to neuro-fuzzy system for classification and definition (Koçer, 2010). By Adjusting of the kernel parameters and selecting features for SVM in training process better classification accuracies can be obtained. Genetic algorithm approach utilized for feature selection and parameters optimization to show how effective they are. Genetic algorithm based approach considerably improves the classification accuracy with having less input features for support vector machines (Huang & Wang, 2006). Tuning and upgrading of regularization and the kernel parameters increasing classification accuracy (Ho, Shu, & Chen, 2004). Machine learning techniques including Artificial Neural Networks (ANN), dynamic recurrent neural networks (DRNN), support vector machines (SVM) and fuzzy logic systems used for diagnosis of neuromuscular diagnosis. EMG signal decomposition has been done by wavelet spectrum matching and principle component analysis of wavelet coefficients with reasonable accuracy rates (Reaz, Hussain, & Mohd-Yasin, 2006). Multilayer perceptron neural networks (MLPNN), dynamic fuzzy neural network (DFNN), adaptive neuro-fuzzy inference system (ANFIS) and combined feature extraction methods (Autoregressive, discrete wavelet transform and wavelet packed energy) presented according to their effect on accuracy in the classification of EMG signals. Results of ANFIS is better than DFNN and MLPNN in terms of slightly higher recognition rate, insensitivity to overtraining and consistent outputs demonstrating higher reliability (Subasi A. , 2012). A novel Particle Swarm Optimization and SVM model has been combined to improve the EMG signal classification accuracy. In the same work, EMG signals were also decomposed into the frequency sub-bands using discrete wavelet transform (DWT). Statistical features were extracted from these sub-bands to represent the distribution of wavelet

coefficients. PSO combined with k-NN and the RBF classifiers too. But result shows great performance of proposed method with PSO. The PSO-SVM reached 97.41% overall accuracy on 1200 EMG signals (Subasi A. , 2013). Feed forward error back propagation artificial neural networks (FEBANN) and wavelet neural networks (WNN) based classifiers developed and compared according to their accuracy in classification of EMG signals (Subasi, Yilmaz, & Ozcalik, 2006). The success rate for the WNN technique was 90.7% and for the FEBANN technique 88% with autoregressive (AR) model feature extraction method. Principal component analysis (PCA) is in use for long time for detection by extracting relevant information from multivariate data. EEG signals which is another biomedical signal processed by using classical and model-based methods. Classical method (fast Fourier transform) and three model-based methods (Burg autoregressive, moving average, least-squares modified Yule-Walker autoregressive moving average methods) were utilized to process and classify EEG signals. Performance of Burg AR and least-squares modified Yule-Walker ARMA have been found significant (Subasi A. , 2005). Fatigue of biceps brachia muscle detected using time–frequency methods and independent component analysis (ICA). EMG signals were observed in time–frequency domain for determination of fatigue. Non-stationary and time varying time–frequency methods are used for the processing of signals. The proposed method for the detection of muscle fatigue is automated by using artificial neural networks (ANN). The results show that ANN with ICA separates EMG signals from fresh and fatigued muscles (Subasi & Kiymik, 2010). There is another approach which automatically detects the number of template MUAP clusters and classifies them into normal, neuropathic or myopathic. The cycle includes three steps: first one is preprocessing of electromyogram (EMG) recordings, second one is MUAP detection and clustering and third one is MUAP classification. The correct detection rate for MUAP clustering is 93% for normal, 95% for myopathic and 92% for neuropathic. 91% is the rate of superimposed MUAPs which were correctly classified. The obtained overall accuracy for MUAP classification is about 86% (Katsis, Goletis, Likas, Fodiatas, & Sarmas, 2006). There is another study which proposes another approach to EMG signals. The proposed method automatically decomposes MUAPs and classifies them into normal, neuropathic or myopathic using a two-stage feature-based classifier in four steps. First step is preprocessing of EMG recordings, second step is MUAP clustering and detection of superimposed

MUAPs, third step is feature extraction and fourth step is MUAP classification using a two-stage classifier. The proposed method employs Radial Basis Function Artificial Neural Networks and decision trees. Minimal tune of parameters is needed to support the classification decisions. The approach has been validated on real EMG recordings. The success rate for MUAP clustering is 96%, while the accuracy for MUAP classification is about 89% (Katsis, Exarchos, Papaloukas, Goletsis, Fodiat, & Sarma, 2007). Decomposition results can improve by detecting invalid Motor unit potential Trains (MUPTs). Numerous methods studied to validate an extracted MUPT using its motor unit potential (MUP) shape information. Some of these MUPT validation methods are newly developed adaptive methods. This method interprets the shape of the MUPs of MUPTs to recognize the activity of motor unit if it is a valid or invalid MUPT. The adaptive gap-based Duda and Hart (AGDH) method which is one of the new adaptive methods, is significantly fast and accurate. The adaptive gap-based Duda and Hart (AGDH) method can be used during or after the decomposition of EMG signals to improve accuracy according to 91.3% for simulated and 94.7% for real data classification total accuracies (Parsaei & Stashuk, 2010). Performance of two multi-channel feature extraction algorithms analyzed which are Montreal and Fuzzy Expert, by using both experimental and simulated data. Control data includes 12 subjects from tibialis anterior muscle which is collected by quadrifilar EMG needle. Performance evaluation of multiple automated algorithms is very important, because decomposing EMG data is difficult if they recorded under low signal to noise ratio (SNR). Thus, high agreement level is obtained between algorithms when more coherence is gained (Li, Dai, Clancy, Christie, & McGill, 2013). Two methods which are based on the gap statistic and jump algorithms are utilized to evaluate the validity of a given MUPT using its MUP shape information. Results show that the gap statistic method is more accurate than the jump method for both simulated and real data. The accuracy of the gap static method was 92.3% for simulated data and 93.8% for real data.

A system named EMGTools is developed to decompose electromyography (EMG) decomposition system that can extract the MUAPs and firing patterns (FPs) for quantitative analysis. The proposed system extracts and resolves MUAPs which are recorded by concentric needle electrodes to produce FPs. Critical parameters are avoided and replaced with adaptive solutions. The feature extraction algorithm has

three steps: segmentation, clustering, and resolution of compound segments. Segmentation completed by using time intervals. According to similarity of segments, clusters are created which are potential classes (PCL). A template is chosen to represent the PCL, from each PCL. The templates are used to identify the MUAPs which can be used as preprocessed data for classifiers (Nikolic & Krarup, 2011). In another study, multiple-level wavelet feature extraction applied to investigate usefulness of extraction of the EMG signal features. Different levels of mother wavelets were utilized to find meaningful resolution constituents from EMG signal. Optimal EMG sub-signal constituents investigated and then the reconstruction has done by the useful and meaningful information in sub-signal constituents. Noise and unnecessary components were filtered through this process. The investigated EMG signal which is base of reconstructed one is extracted with the popular feature extraction methods, i.e. mean absolute value and root mean square. Two criteria applied for the calculation of ratios which are Euclidean distance to a standard deviation and the scatter graph. It has been pointed that EMG features extracted from reconstructed from EMG signals of the first-level and the second-level coefficients. Optimal wavelet decomposition is obtained using the seventh order of Daubechies wavelet and the forth-level wavelet decomposition (Phinyomark, Limsakul, & Phukpattaranont, 2011).

Machine learning techniques such as Artificial Neural Networks (ANN), dynamic recurrent neural networks (DRNN), support vector machines (SVM) and fuzzy logic systems used for diagnosis of neuromuscular disorders (Reaz, Hussain, & Mohd-Yasin, 2006). Reasonable accuracy rates obtained by decomposing EMG signal with spectrum matching and principle component of wavelet coefficients (Begg, Lai, & Palaniswami, 2008). Adaptive Neuro-fuzzy inference system (ANFIS), multilayer perceptron neural networks (MLPNN), dynamic fuzzy neural network (DFNN) and combined feature extraction methods such as Autoregressive, discrete wavelet transform and wavelet packed energy, presented according to their effect on accuracy in the classification of EMG signals (Subasi A. , 2012). PSO optimized SVM classification technique improves accuracy of EMG signal classification (Subasi A. , 2013). Feed forward error back propagation approach developed with artificial neural networks (FEBANN) and wavelet neural networks (WNN) based classifiers and compared according to their accuracy in classification of EMG

signals. They used an autoregressive (AR) model feature extraction as an input to classification system. The success rate for the WNN technique was 90.7% and for the FEBANN technique 88% (Subasi, Yilmaz, & Ozcalik, 2006).

1.2 Importance and Motivation

Recent improvements and new approaches in technologies of signal processing and recording have made it practical and reliable to develop advanced EMG signal analysis (Katsis, Goletis, Likas, Fodiat, & Sarma, 2006; Loudon, Jones, & Sehmi, 1992; Pattichis & Pattichis, 1999). The success of wavelets and multi-scale methods in process monitoring and control has an interest in the investigation of wavelets based on detection methods. Popularity of EMG signal analysis is related to clinical diagnosis and biomedical applications usage field in real life. Analysis of EMG signals using powerful and advance methodologies is becoming new trend in biomedical signal processing with respect to improvements in the field. Because EMG signal analysis is crucial in clinical diagnosis of neuromuscular disorders. Although recording process been improved, actual activities in signal cannot be collected regarding to signal collection methods. There are two common types of data collection which are from surface by using stick tapes and from intramuscular by using needle.

Biologically, EMG signals are affected from environmental noise and signals while passing on different tissues. Thus, noise can seriously distort the EMG signal during the collection and recording process. Noise can be baseline wandering, motion artifact, power-line interference and electrode pop or contact noise. These are the reasons what makes EMG signal more complex and why need to be de-noised. The nature and characteristics of the signal can be understood and hardware integrations can be made for various EMG signal applications (Kimura, 2001).

Noise in the EMG data is removed with Multiscale Principal Component Analysis (MSPCA) technique. MSPCA is employed to generate a helpful representation of EMG signals that removes noise from the signal waveform. Besides denoising methods, the feature extraction methods are also very important for higher classification performance. The effect of the MSPCA denoising method discussed on

EMG signal classification. In addition, effect of Multiple Single Classification (MUSIC) feature extraction method presented and compared for the classification of EMG signals.

EMG signal need to be denoised since it can be distorted by environmental and neighbor constitutions during the collection and recording process. Because, if classifiers feed with more clear, relatively, understandable data, higher classification accuracies can be obtained. These are the reasons what makes EMG signal more complex and why need to be de-noised. Due to its multiscale manner, MSPCA is more convenient for modeling of biomedical signals whose behavior changes over time and frequency (Bakshi, 1998). Success of this approach is proved by results. To exemplify with some of results, after denoising with MSPCA, total classification accuracies upraised from 57 to 92.55 % for SVM, from 50.89 to 90.02 % for ANN, from 50.55 to 82.11 % for k-NN. Improvements are 35.55 % for SVM, 39.87% for ANN and 31.66% for k-NN.

After eliminate noise from EMG signal data it should be known how to work with which parameters. Since input to system is unknown, unknown parameters should be estimated to work with them. This estimation process is named as feature extraction in signal processing field. In addition, since one type of decomposed information is collecting, features in the signal should be extracted by decomposing which is also very important to increase level of performance of analysis. It can be said that EMG signal decomposition results can improve by using feature extraction methods which is depicted in Figure 1.1 (Parsaei & Stashuk, 2010).

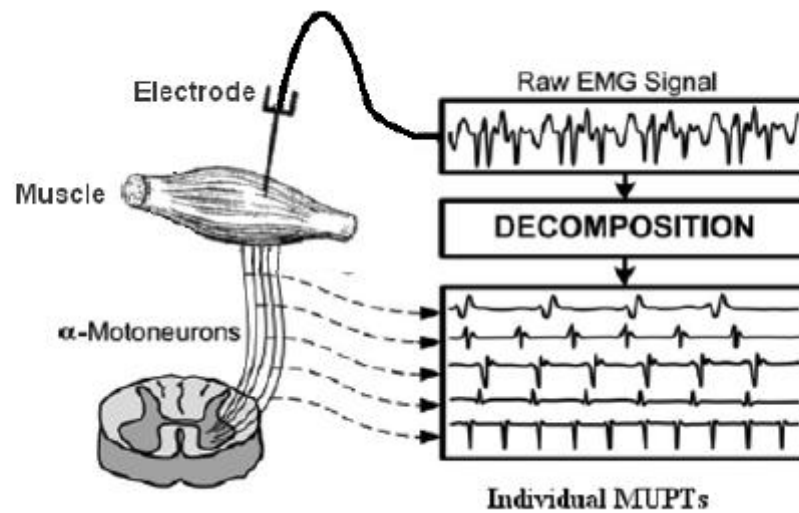


FIGURE 1.1 Schematic representation of feature extraction from EMG signal (Parsaei, 2011)

In this study numerous feature extraction methods are utilized such as discrete wavelet transform (DWT), Eigen Value Decomposition, Covariance (COV), Multivariate Covariance (MCOV), ARBurg (AR), and Multiple Signal Classification (MUSIC) by decomposing EMG signal. It is worth to note that, significantly different results obtained by using different feature extraction methods. For example, 40.55% total classification accuracy obtained by employing COV feature extraction method which is 73% with DWT. According to results coherence between methods can be obtained. It is obvious that each feature extraction methods have different characteristics.

1.3 Objectives

Aim of the study is defining coherence between methods, techniques and classifier algorithms in sight of investigated properties, features and specifications of biomedical signals. Because it has been approved in many studies, coherence between methods and techniques improved classification accuracy on EMG signal classification and analysis for diagnosis of neuromuscular disorders in numerous studies (Misra, Yue, Qin, & Ling, 2002; Bienvenu & Kopp, 1983; Subasi A. , 2013; Subasi A. , 2012; Begg, Lai, & Palaniswami, 2008).

This study investigates classification performance of different classification algorithms with feature extraction and MSPCA de-noising methods using intramuscular EMG signals for diagnosis of neuromuscular disorders. EMG signal classification accuracy improved by utilizing MSPCA de-noising and feature extraction methods. The effects of MSPCA de-noising and feature extraction methods are explained compared and discussed using different performance measures.

1.4 Key Contributions

Results are compared before and after denosing for all machine learning and decomposition techniques. Coherence between feature extraction and classification methods are obtained and discussed. Significant results were found by utilizing MSPCA denoising method. After denoising with MSPCA accuracy upraised notably. Beside that coherence and difference between feature extraction methods obtained. Novelty of this work can be listed as:

- It assumes noise in EMG Signals is an anomaly and needs to be eliminated.
- It evaluates effectiveness of de-noising and feature extraction methods to get higher classification accuracy.
- It examines which machine learning and feature extraction algorithms are more appropriate for classification of intramuscular EMG signal.

1.5 Overview of the Dissertation

The remainder of this study is organized as follows. Chapter 2 explains de-noising and feature extraction methods and how used in literature for biomedical signal classification. In addition, machine learning algorithms which are applied are introduced, such as Artificial Neural Networks, Support Vector Machines, K-Nearest Neighbors, Classification and Regression Trees, C4.5, and Random Forest algorithm. Chapter 3 introduces biomedical signal processing techniques with different analysis approaches. Applied machine learning techniques are introduced in chapter 4.

Methodology and approach explained with procedure in Chapter 5. Results are presented in Chapter 6 with respect to different classification accuracy measurements such as area under ROC curve, F-measure and classification accuracy with and without MSPCA denoising. In Chapter 7, Results are discussed and implications are explained in terms of practice and research. Chapter 8 is conclusion of dissertation with contributions.

CHAPTER 2

ELECTROMYOGRAM (EMG)

An electromyographic signal is reflection of the motor unit potential trains (MUPT) which created by the active background noise and motor units (MU). electromyography is technique for analyzing, evaluating and exploring EMG signals. Recent improvements in biomedical signal processing, computer technology and pattern recognition techniques have support the progress of new techniques for extracting valuable and useful information from the EMG signals. As a result, recently, electromyography placed in the middle of physiological investigations and clinical examinations for diagnosis of neuromuscular disorders, as a biomedical expert aid. The main and major approach is extraction of features by decomposing. EMG signal decomposition is a feature extraction process from actual MUPT which is produced by a MU. The functionality of EMG signal decomposition is estimating MUP template of each active MU. The decomposition process consists of digital signal processing and pattern recognition techniques. Numerous methods have been developed to decompose an EMG signal which are employed to estimate MU firing pattern statistics and MUP features (Parsaei, 2011).

2.1 Introduction

Main responsibility of the human skeletal muscular system is providing the required energy and information to perform different actions. This human skeletal muscular system contains two subsystems which are nervous system and muscular system. They are part of the neuromuscular system. The Central Nervous System provides energy, control and information via nerve signals. Nerves are kind of structured wires of electrical currents, where the nerve heads (nuclei) begin in the spinal column and widen axonal bodies and far and deep. The skeletal muscular

system includes muscle groups which are connected to bones by tendons and activity is achieved when nerve system is also interceptive to diseases like cardiovascular and any other signals cause contraction and relaxations of muscle that either pull or release the bone. The neuromuscular physiological system of the human body. Neuromuscular diseases include disorders which begins in the nervous system like spinal column conduction, in the muscle fibers, and neuromuscular junctions. All type of diseases have different levels from minor loss of strength to muscle or neuron death. On the other hand death is almost certain for some serious disorders such as amyotrophic lateral sclerosis (ALS) (Begg, Lai, & Palaniswami, 2008).

Electrical signals controls movement and position of limbs which traveling back and forward between muscles. Study of electromyograms which are electrical signals in muscles, can support realization, investigation and diagnosing abnormalities both for motor system and muscles. Electromyography (EMG) is representation and explanation of muscle actions. Until recently, aim of recording of electromyograms was investigation. After improvement of bioelectric and biomedical technologies, electromyograms have became very important tool to observe limb action. The investigation and improvement of clinical EMG depends on the technical informations and features used which are observed by electromyograms analysis. Second step after investigation is long paper EMG records which is evaluating by electromyographists with many years of experience. More quantitative approach applied such as potential interval, peak-to-peak amplitude, and number of phases by calculation of paper outputs with comparison to normal dataset which was collected from healthy ones of all ages. After computer placed in medical life which is Very recently, electromyograms are digitized, processed, saved on reusable devices and displayed on computer monitors with screen outputs (Henneberg K. , 2000).

Early and accurate diagnosis is an essential for improved and full recovery patient diagnosis systems. In most situations, clinical predictions and examinations are unsatisfactory for diagnose and localization of disorders because different disorders could cause different or similar symptoms (Preston & Shapiro, 2005). Neuromuscular diseases can be caused by many different muscle fibers or nerves, in some cases it is far from the regular symptoms. Accurate localization of the disorder is important because more focused treatment should be applied. Most of

electrodiagnostic (EDX) studies includes nerve conduction studies (NCS) and electromyography (EMG) for evaluating and diagnosing patients with neuromuscular disorders. EMG was introduced as a method for interpretation of neuromuscular states based on action potentials during contraction and relaxations of muscle. EMG waveform features are between 0.01-10mV and 10–2000 Hz on average. EMG pulse duration points the location and net metabolic state of the muscle fiber where myopathy signs can be found. However, EDX techniques helps to doctors in their diagnosis but not useful for determination of the diagnosis method. In addition, difficult cases are more impracticable methods such as muscle biopsies, sophisticated imaging techniques (Begg, Lai, & Palaniswami, 2008).

The analysis of EMG readings and interpretations is mostly performed by experienced and skilled neurologists. In addition to examining EMG waveforms, used techniques such as needle conduction studies and muscle acoustics is very important. Complexity of problem arise when there are few experts for decision. Therefore, it is becoming increasingly important to develop automated diagnostic systems based on EMG analysis. The requirement makes available a scope of Computational Intelligence techniques for the detection and classification of neuromuscular disorders based on EMG processing and analysis. These computational intelligent systems will assist and help medical experts in detecting of anomalies in the EMG signal outputs. Other applications of EMG signal processing will also benefit from improvements in Computational Intelligence technology such as motion control of prostheses, ergonomics, and intention detection (Begg, Lai, & Palaniswami, 2008).

The aim of intelligent diagnostic is process of the raw EMG signal and then extract meaningful information or features by using statistical and intelligent tools. Information and features of EMG data can be extracted by analyzing time and frequency domain information, Fourier coefficients, autoregressive coefficients, wavelet coefficients, and a wide range of statistical tools from other signal processing techniques. This information, features and results can be used as input data for classifiers, which can classify neuromuscular diseases. Designing a reliable, applicable and accurate automated diagnostic system, is very important and not as easy as a classification problem. Because, EMG readings should be patient-

dependent since their amplitude and duration can change regarding patient age, region and gender. Particularly Kids and babies have unstable EMG waveform morphologies regarding their growth speed which is making it difficult to design a general automated detection system based on readings from one age, gender and regional group (Begg, Lai, & Palaniswami, 2008).

Indeed, best signal-processing approach tends to be dependent on the particular EMG signal. For example, a particular neuromuscular disease should be detected by using the related EMG data and processing tools. We analyze the methods of processing EMG signals to extract meaningful and beneficial information. In addition, other hardest part of the area is to develop signal-processing techniques that analyzing and observing important information from EMG data to feed classifiers with good set of features to get higher accuracy. We will first describe the mechanics, anatomy and physiology of muscle and nerve connection occupied in actual limb which is an essential for understanding the MUAP. Importance of MUAP is being primary biosignal measured in EMG studies. The following section explains several neuromuscular disorders (Begg, Lai, & Palaniswami, 2008).

2.2 The Human Muscle Physiology

2.2.1 Anatomy of the Muscle

Muscles weight is about 40% of the normal human mass, ranging from the small muscles like extra ocular muscles that help turn the eyeball in its hollow to the large limb muscles that produce ability of turn and control stand up. The designs of muscles differ depending on the range of action and the forced. In the simplest composition, parallel fibers lengthen the full length of the muscle and connect to tendons at both ends. Muscles producing a large force to have a more complex organization in which many short muscle fibers connect to a flat tendon that lengthen over a large fraction of the muscle. This organization increases the cross-regional area and the contractile power of the muscle (Henneberg K. , 2000).

There are three types of human muscles, skeletal muscle, smooth muscle, and cardiac muscle, which have different functionalities. Muscles are activated by neurons both for contraction and relaxation. Flexible and expandable structure of muscles allows them to stretch when pulled and to contract when shorten regarding received information from neuron. In addition, muscles are elastic too. Advantage of being elastic is returning to their original position and length after a period of lengthening, relaxation, contraction or shortening which enables to achieve the required movements of daily life. Skeletal muscles provide us mechanical forces and power by moving bones.

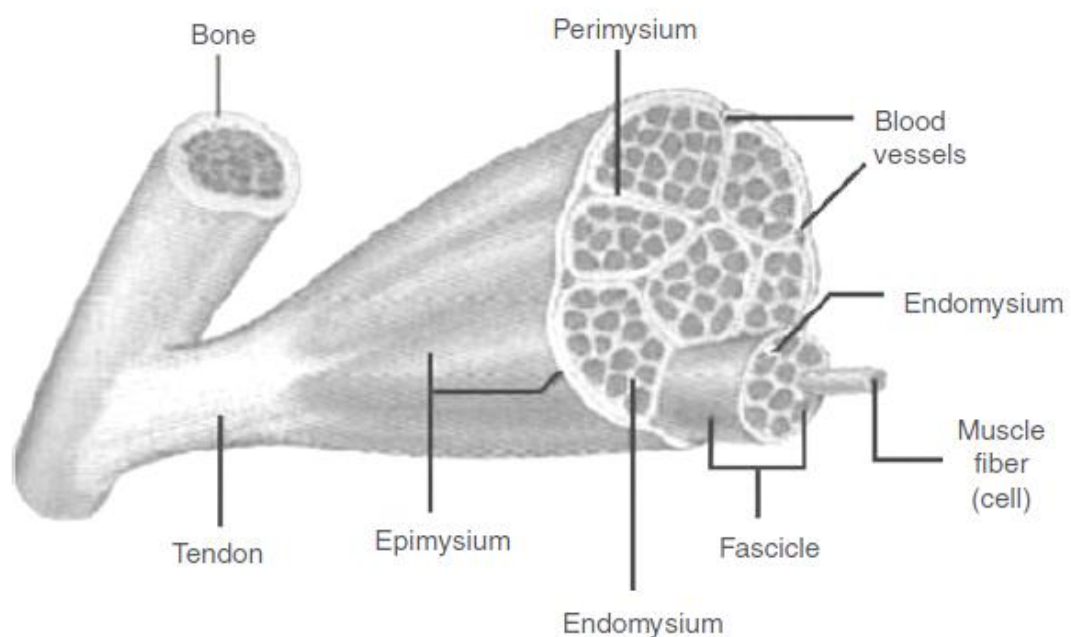


Figure 2.1 Cross-section of skeletal muscle and connection to the bone via tendons (Courtesy of <http://people.eku.edu/ritchisong/301notes3.htm>.)

On closer examination, which can be seen from Figure 2.1, muscle have a striated or striped structure, which is depends on the type of muscle. The appearance and structure includes numerous sarcomere filaments. Skeletal muscles are connected to bones by tendons and connective tissue *epimysium*. The muscle groups includes many *fascicles*, which are fiber groups covered by a coat called as *perimysium* as seen in Figure 2.1. *Sarcolemma* is membrane of cell of each fiber muscle, which works as a connector for impulses. The muscle cells are consist of myofibrils that are group of sarcomeres (see Figure 2.2). Sarcomere includes three type of multiprotein myofilaments which are called titin, myosin, and actin. Actin is

a monomer protein that places in thin filaments while myosin placed in thick filaments (Begg, Lai, & Palaniswami, 2008).

At the site of the neuromuscular connection, each motor neuron forms supplementary pullulates and surrounds several muscle fibers issued almost evenly within an elliptical or circular region ranging from 2 to 10 mm in diameter. The motor neuron and the muscle fibers it surrounds a functional unit, the motor unit. motor unit is called the motor unit territory (MUT) which is working to cross section of muscle. A typical muscle fiber is only surrounded at a single point, located within a crosssectional band according to as the end-plate zone. While the extensity of the end-plate zone is only a few millimeters, the zone itself may expand over a major part of the muscle (Henneberg K. , 2000).

The elastic structure of muscle consist of the large and wide protein structure titin. The sarcomere is between two Z-lines which are comprised by overlapped myofilaments and seems darker because of overlapping. The line between Z-lines and A-lines which is inner component, known as the isotropic band (I-band). The I-band is mostly includes actin filaments. Conversely, The A-band includes myosin filaments. The interaction between actin and myosin filaments is related with calcium ions which supplies contraction or relaxation properties to muscle. The structure and communication between components is similar to sliding cross-bridge system (Begg, Lai, & Palaniswami, 2008).

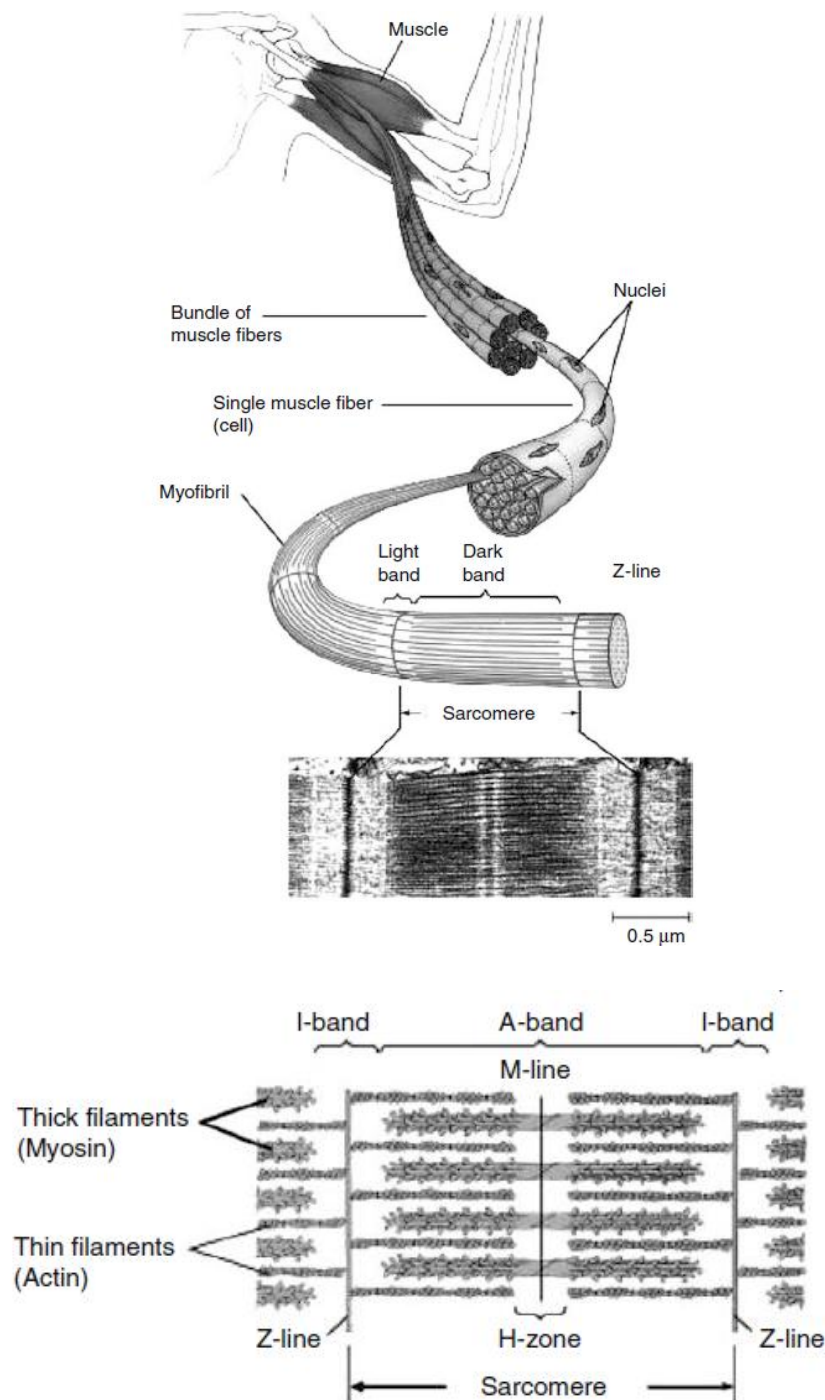


Figure 2.2 Diagram depicts the physiology of the human skeletal muscle, from the muscle fiber to the position of the sarcomeres. (Courtesy of <http://people.eku.edu/ritchisong/301notes3.htm>.)

Smooth muscles are can be placed around blood hollow structures of organs, intestines, and vessels. These are automated muscles, which have shaped muscle fibers. There are two structures of smooth muscle. The first is visceral generally muscle seems in blood vessel linings and the second one is multiunit muscle usually seen in large blood vessels. The muscle fibers consist of thick and thin myofilaments

but do not include sarcomeres. Cardiac muscle, which is special since it is involuntary, mononucleated, striated muscle and seems only in the heart. The cardiac muscle does not need electrical impulses from the nervous system for contraction and relaxation. The rate of contraction and relaxation, anyhow, is controlled by electrical nerve signals. Since one cardiac muscle fiber is placed next to another, contraction and relaxation of one muscle will stimulate next muscle fiber. This results in systematic propagation of the contraction action along the heart (Begg, Lai, & Palaniswami, 2008).

2.2.2 The Structure and Function of Muscle

The fibers of one motor unit are mixing with fibers of other motor units. Thus, several motor units reside within a given cross section. The fibers of the same motor unit are thought to be randomly or evenly distributed within the motor unit territory, anyhow, resurrounded fibers often results in the configuration of fiber clusters. The core of Electromyograms different the myocardium, skeletal muscles do not include motor cells from which excitations occur and extend. Electrical excitation of skeletal muscle is started and normalized by the central and peripheral nervous systems. Motor neurons carry nerve impulses from the anterior horn cells of the spinal cord to the nerve endings, where the axonal action triggers the happen of the neurotransmitter acetylcholine (Ach) into the slim cleft dividing the sarcolemma from the axon terminals. As Ach binds to the sarcolemma, Ach-sensitive sodium channels open, and miniature end-plate potentials occur in the sarcolemma. If enough Ach is released, the evaluation of minuscule end-plate potentials, the endplate potential, reaches the excitation limit, and sarcolemma action potentials circulate in opposite directions toward the tendons. As stimulation spreads down the fiber, it distributes into a highly branched transverse network of tubules (T system) which get in the myofibrils. The effective radial transference velocity (~ 4 cm/s) is about two orders of size slower than the lengthwise transference velocity (2 to 5 m/s). This is a matter of fact that the main piece of the total membrane capacity is located in the T system and that the lumen of the T system comprises a higher electrical protection than the myoplasm. The slower straight transference velocity implies an increasingly postponed onset of straight action potentials toward the core of the fiber relation to that of the sarcolemmal action potential. On the other hand,

measured with the time course of the consequent reduction, the spread of stimulation, along and within the muscle fiber is essentially sudden, by this means ensuring concurrent discharge of calcium from the sarcoplasmic reticulum during the whole volume of the muscle. If calcium discharge were limited to a small lengthwise section of the muscle fiber, only sarcomeres in this region would contract, and sarcomeres in the rest of the fiber would extend properly. Similarly, experiments in straightened muscle fibers, in which the persistence between the sarcolemmal and the straight membrane has been dislocated, have displayed that only a thin layer of superficial myofibrils contracts when straight action possibilities not activate to trigger calcium discharge deep in the muscle fiber. It is well known that the shape of the skeletal muscle action potential vary from that of fibre action potentials with refer to the repolarization phase. In skeletal muscle, the relaxation phase of the action potential is interrupted by a long, slowly reducing potential known as the after contraction. This late potential is caused by two opposing forces, the repolarization force due to the leak of potassium through the sarcolemma and a depolarization force due to an inflow of current from the interstitial space into the straight lumen. The latest current is drawn in through the straight openings by the repolarizing straight membrane. Large and long muscle fibers have a higher straight-sarcolemma field ratio. Therefore the inner flow of straight current increases with fiber size. The small fiber has only a small amount of straight membrane. Therefore the sarcolemmal potassium current has enough strength to fastly repolarize the membrane. For the large fiber, inner straight current actually depolarizes the sarcolemma slimly during repolarization of the straight membrane system, by this means producing a small bulge on the after contraction. Since the bulge on the after contraction is affected primarily by fiber length, this feature is more common for large fibers than for smaller human fibers. Findings have established that the sarcolemma action potential of straighten fibers hyperpolarizes in a way similar to that of a nerve action potential. The time course of the sarcolemmal existing density is compared with that of the current passing through the straight mouth during the time route of the sarcolemmal action potential. The positive (outer) peak of the straight current overlaps in time with the small capacitive sarcolemmal displacement current (initial positive peak) and the negative peak (inner sarcolemmal sodium current). As a result, the net current has a much larger positive peak than that of the sarcolemma alone, and the negative peak of the net current is only about half that of the sarcolemmal current.

The outer sarcolemmal potassium current (late positive phase) is almost completely opposed by an antisymmetric inner straight current, the current plotted into the straight lumen by the repolarizing T system. The joined effect of the sarcolemmal and straight interaction is a net current with an almost biphasic waveform and with similar amplitudes of the positive and negative peaks. As the net current source spreads toward the tendon, an extracellular potential area begins and follows the action potential complex. At a determined area point, this phenomenon is detected as a temporal potential waveform. On the other hand, a more comprehensive picture of the phenomenon is gained by marking the spatial distribution of potentials in the cross section of the motor unit. In a common use healthy motor unit, the mean inter fiber distance is on the order of a few hundred microns. It is obvious that single-fiber action potential (SFAP) overlapping arises between low-magnitude isopotential lines. In this case, higher-level isopotential lines overlap within the clusters, while regions with no fibers would appear as electrically silent. All fibers are supposed to be of the same size and with alike, perfectly axisymmetric potential distributions. The net single-fiber current source is a growing function of fiber size. Therefore the magnitude of the potential distribution will vary with varying fiber size. Fibers can to a good approximation be assumed as constant current sources. Thus, if the resistivity of the muscle tissue increases, due to increased fiber packing density, the potential variety between an observation point and a reference point also will raise. It follows that local variations in fiber packing density in the region of an active fiber will destroy the axisymmetric appearance of its potential distribution. Muscle fibers are not perfect cylinders, and angular deviation in the shape of the sarcolemma should be expected to create angular deviations in the potential distribution. Nevertheless, due to the relatively high conductivity of the volume conductor, it is plausible that such variations become more and more unimportant as the distance to the fiber is raised.

Single-fiber action potential (SFAP) is usually separated among the motor unit axis, and their potential allocations do not sum up in a simple way. SFAPs can't be arranged for numerous reasons. The end plates are not arranged. Therefore some SFAPs originate ahead of others. Variations in fiber size and packing density cause variations in transference velocity. Thus the spreading of SFAPs may actually raise with growing distance from the end-plate zone. Neuropathologic conditions may

affect the alignment of SFAPs. Complete or partial failure of collateral nerve sprouts to conduct nerve action potentials can abolish or delay the onset of action potentials, and immature sprouts formed during enervation may cause significant variability (jitter) in the onset of muscle action potentials. Enervated muscle fibers get smaller. Thus, newly reinnervated muscle fiber is expected to have a slower conduction velocity. An increased dispersion of SFAPs produces a very complex and irregular spatial potential distribution (Begg, Lai, & Palaniswami, 2008).

The sarcolemma or cell membrane of muscle fibers includes hollows which known as *transverse tubes* or “T-tubes”. T-tubes wrap around the myofibril and joins to others in the sarcolemma. Action potentials travelling from the surface of the sarcolemma to deep into the muscle center where the sarcoplasmic reticulum exists. The sarcoplasmic reticulum is a group of hollow to keep Calcium ions (Ca^{2+}), which is the main chemical firing element for contraction. The membrane performs using active transport energy which moves ions across the membrane. The membrane has ports to control ions flow. Calcium is constantly injected into sarcoplasmic reticulum from cytoplasm of the muscle fibers. During relaxation state, the sarcoplasmic reticulum has a high concentration of calcium ions compared to the muscle fibers or myofibrils.

The motion of muscle fibers is caused by interaction between thick and thin myofilaments. The thick myofilament has cross-bridge structure, which performs correlations with protein molecules.

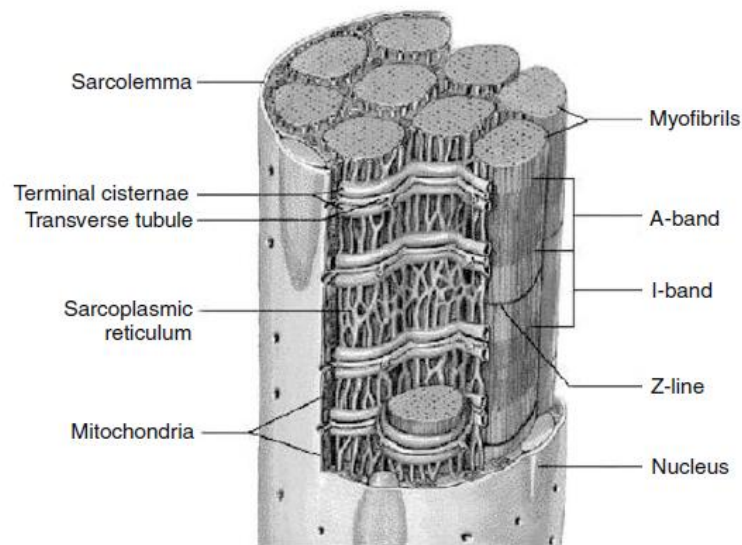


Figure 2.3 Cross-section of the sarcoplasmic reticulum showing the T-tubes, myofibrils, and sarcolemma. (Courtesy of <http://people.eku.edu/ritchisong/301notes3.htm>.)

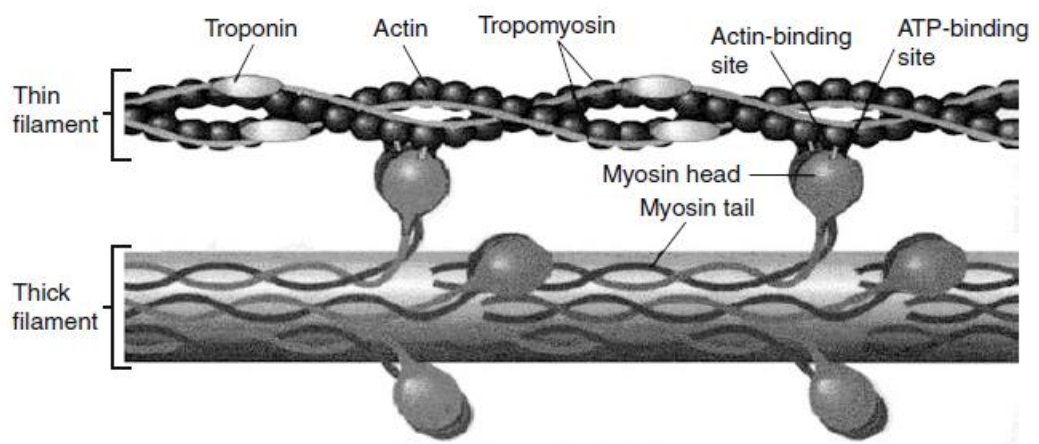


Figure 2.4 Diagram depicts the thick myofilaments composed of myosin protein structures (Courtesy of <http://people.eku.edu/ritchisong/301notes3.htm>.)

The cross-bridge structure includes adenosine triphosphate (ATP) molecules, which is source of energy. Thin myofilaments are includes actin, troponin, and tropomyosin protein molecules. Actin molecules moulds the long chain of the myofilament as seen in Figure 2.5.

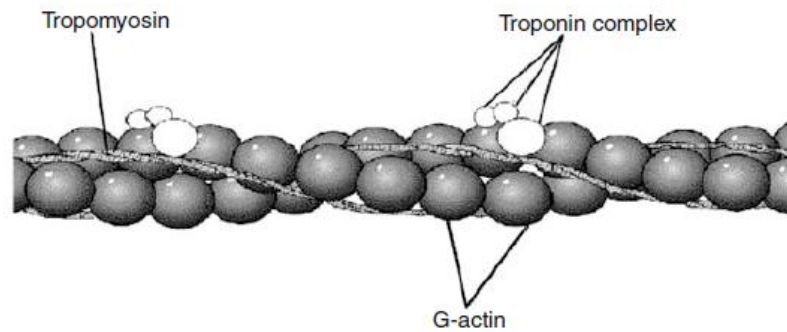


Figure 2.5 Diagram depicts the thin myofilaments composed of actin, troponin, and tropomyosin protein structures. (Courtesy of <http://people.eku.edu/ritchison/301notes3.htm>.)

Tropomyosin is single independent chain, which placed around actin chain with troponin molecule. During muscle's relaxation state, the tropomyosin chain in contact with the myosin heads. Troponin molecules have contact edges for calcium ions, which forms the molecule structure to reshape filaments while having calcium ions. At the contraction phase, a nerve impulse arrives to sarcolemma and the sarcoplasmic reticulum. the ion edges starts diffusion of the calcium ions into the myofibril filaments along the concentration phase when electrical impulse arrives to surface of the sarcoplasmic reticulum. During existence of calcium ions, troponin molecules changes shape and pull the tropomyosin chains along effecting the myosin head to contact with an actin molecule. The myosin head connects with an actin molecule in the existence of ATP when connection is performed. This chemical interaction transforms ATP into adenosine diphosphate (ADP), because of energy realization. The chemical transformation is represented by following equation where P is inorganic phosphate.



This impulsive release of energy effects the myosin edge to pull the actin chain or thin myofilament. A series of these interactions or movements in hundreds of thousands of thin myofilaments is performed as a single unit motion of the muscle fiber. ATP fills the binding edges on the myosin heads and disconnects the links between the myosin and actin molecules when the motion is completed. The process repeats as long as in existence of calcium ions. From this point of view, the thick myofilaments are pulling themselves along the thin myofilaments or sliding and the

muscle contracts. Position of the thin and thick myofilaments which allow motion via the cross-bridge mechanism is depicted in following figure (Begg, Lai, & Palaniswami, 2008).

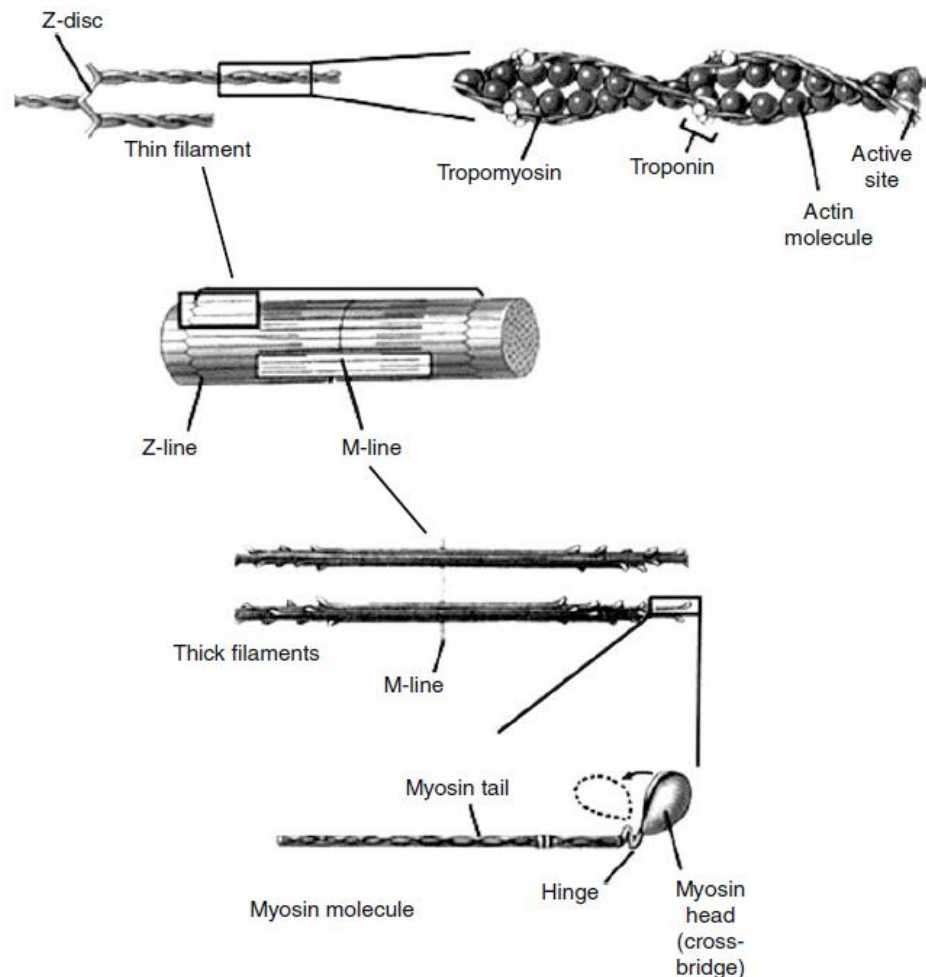


Figure 2.6 Structure of thin and thick myofilaments in cross-bridge motion. (Courtesy of <http://people.eku.edu/ritchisong/301notes3.htm>.)

Reduction of delivering ATP molecules is possible which is related with the capability of myosin heads to attach to new actin molecules. In this case, even in the existence of sufficient calcium ions, the duration of motion is reduced correlatively. This situation occurs naturally and is called as *muscle fatigue*. When the nerve impulse stops the ionic gates in the sarcoplasmic reticulum. During inexistence of calcium ions, the troponin molecules stops motion. The tropomyosin filaments contact with the myosin edges. Then, the muscle then turns into relaxation phase. There are two main types of contractions which are *isotonic* and *isometric* contraction. Either muscle shorten to generate force or muscle does not shorten and the force produced greater load (Begg, Lai, & Palaniswami, 2008).

Motor Unit Potentials (MUPs) with five or more peaks are classified like polyphasic. A satellite potential is an SFAP that is not so aligned from the main potential complex that its waveform is completely divided from the MUP. Electromyographic recordings amount of information referring the bioelectrical state of a muscle is stored in the timevarying spatial division of potentials in the muscle. Unfortunately, it is not clinically feasible to obtain high-resolution three-dimensional samples of the spatial potential distribution, since this would require the insertion of hundreds of electrodes into the muscles. For minimization the discomfort of the patient, routine EMG solutions commonly employ only a single electrode that is interleaved into different areas of the muscle. As the SFAPs of an active motor unit pass by the electrode, only their evaluation, the MUP, will be registered by the electrode. The electrode is sufficiently integrating out the spatial information hidden in the passing potential complex, leaving only a time-variant potential waveform to be recorded. It goes without saying that such a constraint on the recording process puts the electromyographist at a substantial disadvantage. To partially circumvent this setback, technical innovations and new procedures have continued to make better EMG examinations to such a level that some of the spatial information originally obscured by the electrode can be extracted intuitively from the temporal waveforms by an practiced electromyographist. With a detailed understanding of the bioelectric principles involved, the bioelectric sources, the volume conductor, and the recording properties of the electrode, the electromyographist can quickly identify and clarify the waveform characteristics associated with various neuromuscular abnormalities. To raise the amount of diagnostic information, some sets of EMG investigations may be evaluated using electrodes with vary recording characteristics. (Henneberg K. , 2000)

2.2.3 The Motor Unit Action Potential

Motor Unit Action Potential has anatomic and physiologic properties based on the integrity of neuromuscular transmission, fiber density, propagation velocity, and innervation ratio. These factors differ not only from one muscle group to another but also with age (Kimura, 2001).

The contraction and relaxation phases of muscles enables to achieve required movements of daily life by applying different parts of body such as the eyes, hands

and fingers etc. Muscles provide us mechanical forces and power by moving bones and other structures. According to function, muscles may be categorized as either skeletal, smooth, or cardiac. As it can be understood from its name, skeletal muscle is fastened to the body skeleton. Smooth muscle is found either around or in the smooth tissues such as intestines and blood vessels. Cardiac muscle places around the heart to build the heart walls and produces the contraction of the heart to create heartbeats. Contraction and relaxation of all types of muscles is controlled by electrical impulses, namely action potentials, which propagate between muscles and central or peripheral nervous systems. The action potentials are transmitted down the axons of the motor neurons to the muscle fibers which is originating in the brain or the spinal cord. Each motor neuron is connected to muscle fibers via specialized synapse which allows the action potentials to stimulate contraction or relaxation. According to function of the muscle, a single motor unit may includes a few or more than a thousand muscle fibers (Merletti & Parker, 2004; Basmajian & De Luca, 1985). Muscles have fewer muscle fibers than muscles, per motor unit, which control gross movements such as running and jumping. The contraction of a muscle fiber begins when neuronal action potentials reach the neuromuscular junction and fire action potentials which expand through the impulsive membranes of muscle fiber (Sörnmo & Laguna, 2005). A motor unit action potential (MUAP) results from spatial and temporal summation of individual action potentials as they spread through the different muscle fibers of a single motor unit, see Figure 2.7.

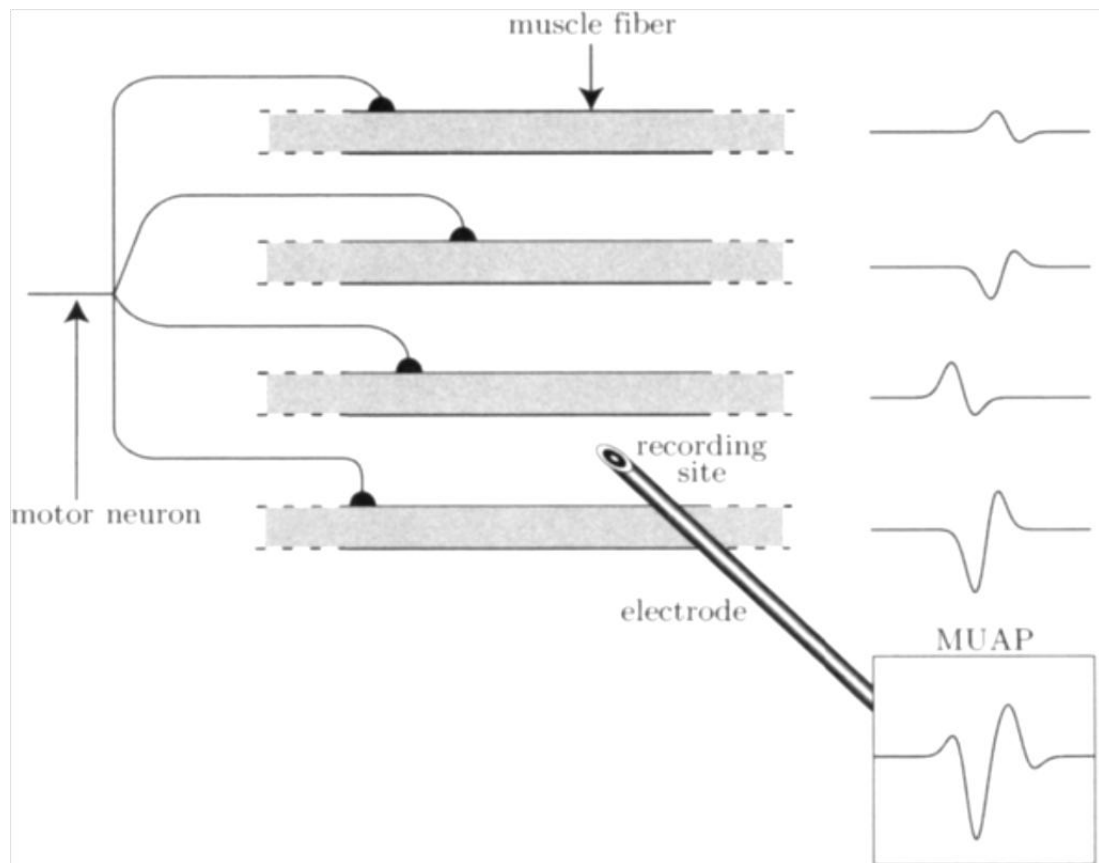


Figure 2.7 The generation of a MUAP of a single motor unit with four muscle fibers.

The number of MUAPs in the detection area related with the used type of electrode since electrodes sensitivities are not equal. A few square millimeters of a muscle may contain about 50 motor units. Motor unit labor is a fundamental muscular process force which is applied by muscle contraction and controlled by central nervous system via spatial and temporal motor units. Spatial recruitment means that force is increased by applying more motor units. Temporal recruitment means that force is increased by firing of action potentials, at faster rates (Sörnmo & Laguna, 2005).

MUAP morphology and firing rate of each motor unit is different than others. A high firing rate means that MUAP waveforms can be determined due to noise-like and random appearance. Motor units are generally employed according to their size. For example, larger units increasingly employed for increasing force of muscle contraction. As a matter of fact, if initial activation is tiny will be followed by a smooth increase. Analysis of motor unit labor at low levels of muscle contraction is important for diagnosing disorders related to the nervous system (neuropathy) and

muscle tissue (myopathy). The motor unit labor is usually analyzed with respect to firing rate and behavior of firings. The recruitment rate can be detected easily regarding presence of active motor units. For example, a neuropathic pattern can be recognized if the first motor unit exceeds 20 Hz before the second motor unit appears. On the other hand, a myopathic pattern can be detected if more than one motor units are active at minimal voluntary contraction phase to generate the required force. Thus, a myopathic pattern is connected with more than one MUAPs (Sörnmo & Laguna, 2005).

Intramuscular EMG signal can be separated into two groups which are morphology and firing pattern. Most useful models and processing algorithms developed for decomposition of the EMG signal. Morphologic analysis of the MUAP waveform is essential to perform clinical examination with observed features from the intramuscular EMG (Henneberg K. , 2000). The number of active muscle fibers determines the amplitude. In general, large amplitudes are associated with neuropathies and small amplitudes with myopathies (Sörnmo & Laguna, 2005).

The number of phases of a MUAP waveform is important information for action potential propagating in different single fibers. The recording electrode detects the action potentials of many fibers. Conversely, the action potential of a single fiber has only one or two phases, which are related to different degrees of waveform. On the other hand, polyphasic MUAPs have more than four phases and most of them for neuropathic and myopathic situations. Different number of phases and degrees can cause complexities (Sörnmo & Laguna, 2005). Commonly less than 12% of the total number of potentials are polyphasic. This percentage is a bit higher for deltoid and facial muscles with less than 25%. The number goes a little down for vastus lateralis and anterior tibial muscles with less than 20% (Thage, 1974; Buchthal, 1991). A MUAP may include a spike factor to separate from the main spike, which is also called as satellite potentials. The satellite potentials represent action potentials of a single or more muscle fibers. The satellite potentials separated from the main volume of the muscle fibers in the motor unit up to 3% of MUAPs of healthy muscle (Denny-Brown, 1949; Takahashi, 1966; Desmedt & Borenstein, 1973; Lang & Partanen, 1976; Nikolic M. , 2001)

Period length can be obtained after MUAP recording by the onset and end of the waveform. Commonly recording is taken as the first and last time directly. The period length related with the number of muscle fibers in the motor unit and increment of the number of the muscle fibers, since there is an interrelation between them. The regular MUAP duration is between 2 ms and 10 ms. Signals can be recorded with two main types which are intramuscular and surface. In addition, type of electrodes may affect MUAP morphology too. In a regular subject, MUAP morphology usually stable from discharge to discharge (Preston & Shapiro, 1998). Unstable morphology waveform may arise when the neuromuscular junction is weakened or damaged which can cause exact nerve action potentials (Sörnmo & Laguna, 2005).

In some situations, MUAP amplitude may double from one discharge phase to next or number of phases also can differentiate. Regular changes in MUAP morphology are obtained when the biochemical specifications of motor unit change, for example, during muscle fatigue. A muscle contraction is continued with recursive activation of the motor units. Each motor unit is associated with a specific train of MUAPs. Number of active motor units increases, unless there is more recording detected by electrodes. This will cause increasing number of concurrent MUAP trains, and separation of individual MUAP trains will become more difficult. Frequency of repeated MUAPs defines the firing pattern. By the way, the information helps to explain how the central nervous system controls the motor unit. There is an interrelation between firing patterns and consecutive interfiring intervals. For example, the delay between consecutive MUAPs and origin during contraction. A normal firing patterns is allows to determine the pattern by its average firing rate and average length of the interfiring intervals. A motor unit usually starts firing at a rate of 4-5 Hz, but the rate can increase with required force, directly proportional. The average firing frequency for normal muscle is between 7-12 Hz at the lowest level of muscle contraction, which can be demonstrated. This level of muscle contraction is usually calling as percentage of maximal voluntary contraction (MVC). If force increase, that means the average of firing rate increases too, directly proportional. Generally when assessing motor unit abnormalities, it is necessary to have their own

specific firing rates, because different muscles have different rates. It is applicable to characterize irregular firing patterns observed during dynamic contractions in force level, since the average firing rate is an important parameter for analysis of intramuscular EMGs. The same method can be applied for characterization of firing patterns with impaired neuromuscular transmission where the interfiring intervals more variable. Thus, possible parameters are guaranteed which effectively provide properties of a firing pattern (Sörnmo & Laguna, 2005).

90 percent of the variance of the data set can be covered by Principal components analysis employing three different features. These parameters can be the size of the motor units, variations in the arrival time at the recording electrode, and loss of muscle fibers within the motor unit territory (Hahn, Parkes, Bolton, & Stewart, 1991). Advanced techniques for spectrum analysis of motor unit action potentials enables successive measures of these aspects for quantitative investigations (Berardelli, Mercuri, Rona, Vacca, & Manfredi, 1996; Bertolasi, De Grandis, Bongiovanni, Zanette, & Gasperini, 1993; Crews, Kaiser, & Brooke, 1976; Kimura, 2001).

2.3 Neuromuscular Disorders

EMG studies applied to ease diagnostic problem by providing more precise localization. The neuromuscular disorders contains different diseases which states and effects the lower motor and sensory neurons, the motor and sensory roots, peripheral nerves, neuromuscular junctions, and muscles. Neuromuscular disorders are commonly related with peripheral nervous system. Type of disorders can be categorized regarding origin location. Myopathies are abnormalities of muscle usually caused by muscle degradation or muscle death and neuropathies are disorders of the nerves, which are main and most common types. We introduce many different disorders that affect the neuromuscular system by highlighting their clinical manifestations, etiologies and symptoms. Some of the usual disorders are summarized in Table 2.1. We describe components of the peripheral nervous system and mention underlying causes. More extensive groups and types of muscular

diseases can be found, at the Muscular Dystrophy Association (MDA) homepage (Begg, Lai, & Palaniswami, 2008).

Table 2.1 Table Shows Categorization of Neuromuscular Disorders and some Specific Diseases Associated with Each Group

Neuropathy (Mono and Poly)	Sensory neuronopathy
Entrapment	Paraneoplastic
Demyelinating	Autoimmune
Axonal	Infections
Mononeuritis multiplex	Toxins
Radiculopathy	Plexopathy
Disk herniation	Radiation induced
Spondylosis	Entrapment
Neoplastic	Neoplastic
Infarction	Diabetic
Infections	Hemorrhagic
Inflammatory	Inflammatory
Motor neuron disease	Myopathy
Amyotrophic lateral sclerosis	Muscular dystrophy
Polio	Congenital
Spinal muscular atrophy	Metabolic
Monomelic amyotrophy	Inflammatory
Neuromuscular junction disorders	Toxic
Myasthenia gravis	Endocrine
Lambert–Eaton myasthenic syndrome	
Congenital Myasthenic Syndrome	
Botulism	

Source: Adapted from Preston, D. and Shapiro, B. in *Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic Correlations*, Elsevier, Philadelphia, PA, 2005.

2.3.1 Neuropathy Disorders: Mononeuropathies and Polyneuropathies

Neuropathic disorders can be separated as mononeuropathies and polyneuropathies according to number of involved nerves. In mononeuropathies, a single nerve may

cause pain or some disability, while all nerves can be affected in a polyneuropath case. The nerve is affected either by disease or injury in both cases.

2.3.1.1 Mononeuropathies

A key point is that each mononeuropathy tends to have a specific types of pathophysiology. Contrarily, chronic mononeuropathies, either static or slowly progressive ones, typically occurs as transfer failure which generated by axons. Few of them interestingly due to demyelination, generally causes slower transmitting. Mostly, mononeuropathy in the arms and hands is namely median nerve entrapment. Entrapment is compressed or placed in an abnormal nerve which cause interruption. Entrapment syndromes occurs at the general sites of chronic or recurrent constriction of the radial, median, ulnar, common peroneal, and tibial nerves. The disorder generally arises in the carpal tunnel of wrist by causing pain and disabling the hand where the nerve has pressure. EMG is usually applied to separate the problem from brachial plexus, cervical nerve roots and lesions of the proximal median nerve (Encyclopedia of the Neurological Sciences, 2003).

EMG is the most feasible to the study with limbs. They can be employed for investigation of cranial nerves. Mononeuropathies may affects facial and trigeminal nerves which are the most expectable lesions. The facial nerve which also known as the cranial nerve is including different nerve fiber bundles, nerves to the facial muscles, parasympathetic motor fibers to the salivary and lacrimal glands, part of the external auditory meatus, somatic sensory fibers serving the ear surface and taste fibers to a major segment of the tongue (Begg, Lai, & Palaniswami, 2008).

The most common facial mononeuropathy is facial nerve palsy, which generally seems as palsy of idiopathic Bell. This facial nerve is related with several types of disorders such as stroke, lymphoma, leprosy, diabetes and herpes zoster. Symptoms of diseases depends on the location, pathophysiology, and severity of the lesion. For example, main lesion presents weakness in the lower facial region. Injury and solidity of the facial nerve may also cause hemifacial spasm which is characterized by involuntary contractions and spasms. Trigeminal neuropathy is another type of disorders which is responsible for transmitting nerve impulses to the face and motor

fibers and effects the trigeminal nerve. Tibial neuropathies are rarely seen and nearly always cause axon loss. Effect of the disorder is difficulty in opening the mouth or chewing and numbness over the ipsilateral face. Common peroneal neuropathies, most generalized to the fibular head, rarely causes to clinical weakness, so their underlying pathology is axon loss, demyelinating conduction block, or a combination. Trigeminal neuralgia is a rare type of disorders which cause strong pain in one or more branches of the trigeminal nerve and can be triggered by a light touch (Begg, Lai, & Palaniswami, 2008).

Injuries caused from acute or chronic recurring external pressure which may cause compressive neuropathy. In addition, chronic distortion of the nerve from an internal source causes entrapment neuropathy (Meyers, Gilden, Rinaldi, & Hansen, 1972; Zwarts, Van Weerden, Links, Haene, & Oosterhuis, 1988). Common types of peripheral nerve disorders may happen occupationally. To exemplify, patient may suffer from symptoms of cervical radiculopathies, thoracic outlet syndrome, and median, ulnar, and digital neuropathies (McManis, Lambert, & Daube, 1986). Numerous nerve lesions result from stretch, ischemia, compression, or laceration during a surgery (Cooper, Stokes, & Edwards, 1988). Abnormal parts of attachment can cause rare anomalies such as congenital ring constrictions of peripheral nerves (Nielsen, 1984; Kimura, 2001).

The diagnoses of a nerve lesion interrelated with clarification of weakness and weaken parts of muscles. Developments and improvements in electrodiagnostic studies help analyze, localize and characterize a focal lesion if kind of an extension of a physical examination. Electromyographic examination defines the exact distribution of proper muscles in localizing. The pattern of distribution helps understand the side of involvement (Kimura, 2001).

2.3.1.2 Polyneuropathies

The majority of polyneuropathies includes large myelinated fibers to a greater or lesser extent which are detectible in the electrodiagnostics laboratory. Polyneuropathies are categorized into two main types related with their underlying pathophysiology which are axon loss and demyelination. The nerve conduction

studies are the most important components of the electrodiagnostics examination with polyneuropathies (Begg, Lai, & Palaniswami, 2008).

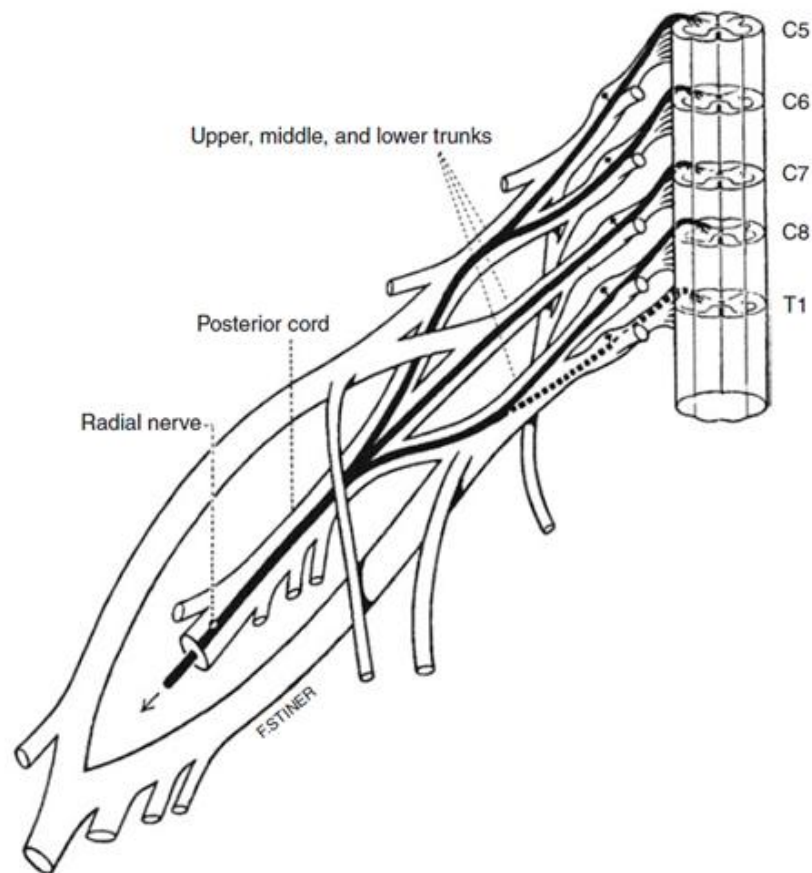


Figure 2.8 The radial nerve is derived from the posterior cord of the brachial plexus. It receives innervation from all three trunks of the brachial plexus and contributions from the C5–T1 nerve roots. (Reprinted with permission from Preston, D., and Shapiro, B., *Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic Correlations*, Elsevier, Philadelphia, PA, 2005.)

The next most common neuropathy is *ulnar neuropathy*, which is difficult to localize using either EMG or other electro diagnostic methods. Although it is commonly due to compression of the ulnar nerve at the elbow, it can be caused by entrapments of the nerve at the wrist making it difficult to pinpoint. Ulnar neuropathy at the wrist is rare and not easily detected during the early stages (Begg, Lai, & Palaniswami, 2008).

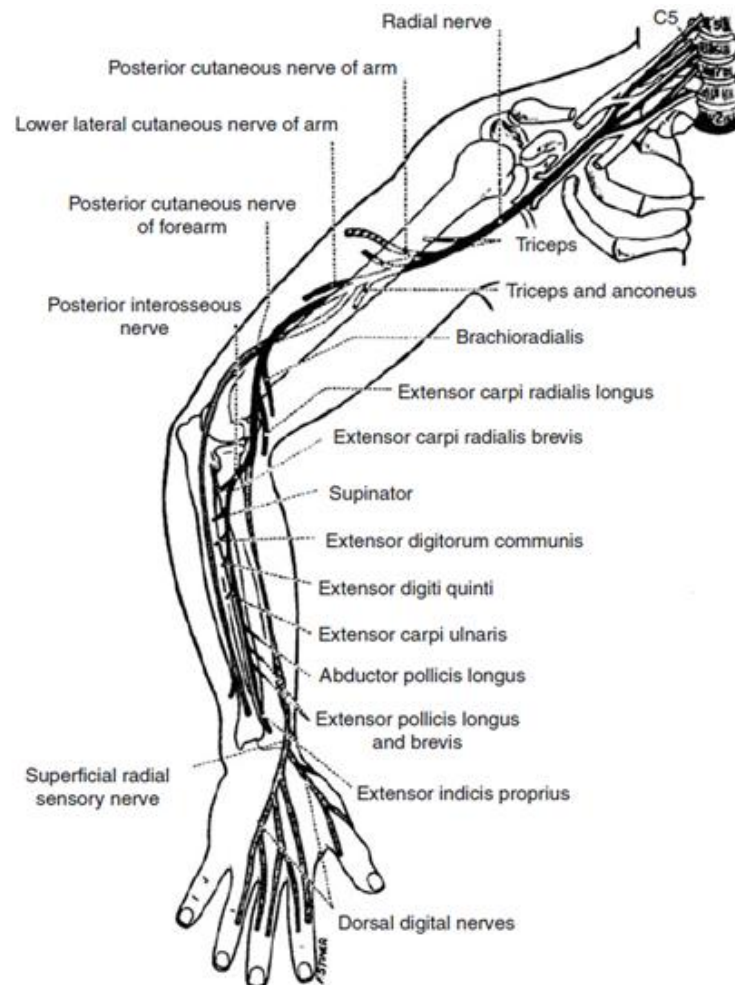


Figure 2.9 In the upper arm, the radial nerve first derives several cutaneous nerves of the arm followed by muscular branches to the triceps brachii. It then wraps around the humerus before descending down the elbow and deriving several smaller nerves, which branch into the hand and the fingers. (Reprinted with permission from Preston, D., and Shapiro, B., *Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic Correlations*, Elsevier, Philadelphia, PA, 2005.)

Nerve conduction studies may examine conduction abnormalities, which usually cause axonal differentiation. Stimulation above and below the abnormal part of lesion will show both slowing of conduction speed and changes in amplitude of the muscle. Signs of abnormalities makes a distinction an entrapment syndrome. This distinction may cause complexity in certain types of polyneuropathy in early stages (Kimura, 2001). There are several different disorders that include variants of chronic inflammatory demyelinating polyneuropathy and acute inflammatory demyelinating polyneuropathy (AIDP) localized by real sudden data. investigation of the nerve fibers and the fiber size can help understanding the etiology. However, EMG is

unlimitedly useful because patients have pure small fiber polyneuropathy and normal electrophysiological characteristics (Begg, Lai, & Palaniswami, 2008).

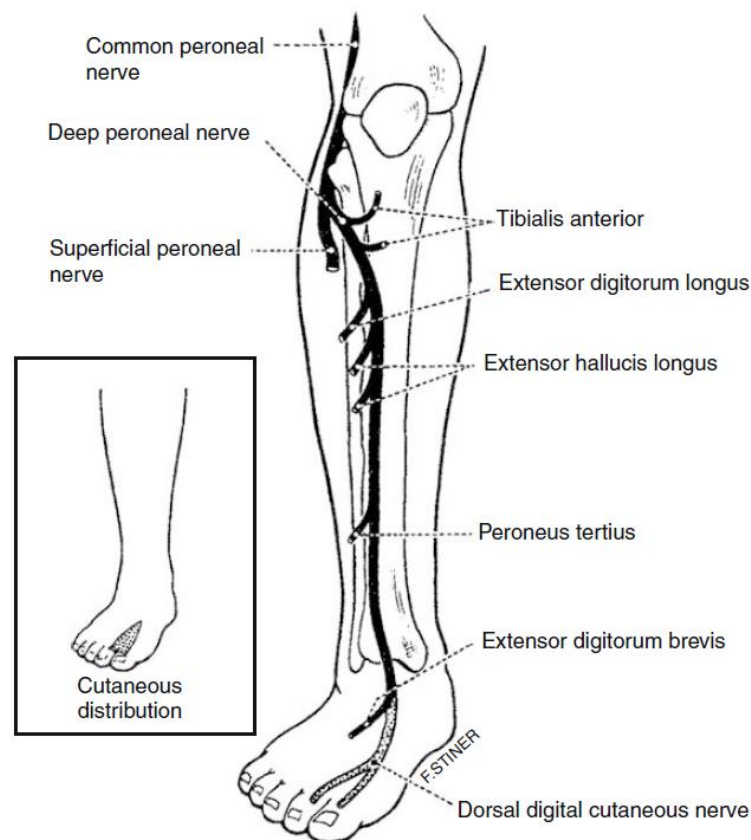


Figure 2.10 The peroneal nerve anatomy showing the nerve branches into the calf, feet, and toes. (Reprinted with permission from Preston, D. and Shapiro, B., *Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic Correlations*, Elsevier, Philadelphia, PA, 2005.)

Certain types of neuropathy may point wider sensory symptoms and others may start with more important approximal limitations. Generally, regular muscle stretch reflexes connects peripheral neuropathy. Acute pandysautonomic neuropathy achieves sparing of motor and sensory function, as usual (Lehmann-Horn, Kuther, Ricker, Grafe, Ballanyi, & Rudel, 1987). Milder autonomic dysfunction execute together with most peripheral neuropathies and it defines clinically detectable symptom in case of porphyria, amyloidosis, familial dysautonomia, diabetes and Guillain-Barré syndrome. Many autonomic disorders generally cause from acute demyelination or damage to small myelinated and unmyelinated fibers. Inflammatory neuropathies include Guillain-Barré syndrome and chronic inflammatory demyelinating neuropathy which is another type. The family history is

important for definition of inherited polyneuropathy conditions. When a patient doesn't have sufficient information, examination of family members is needed for definition and detection.

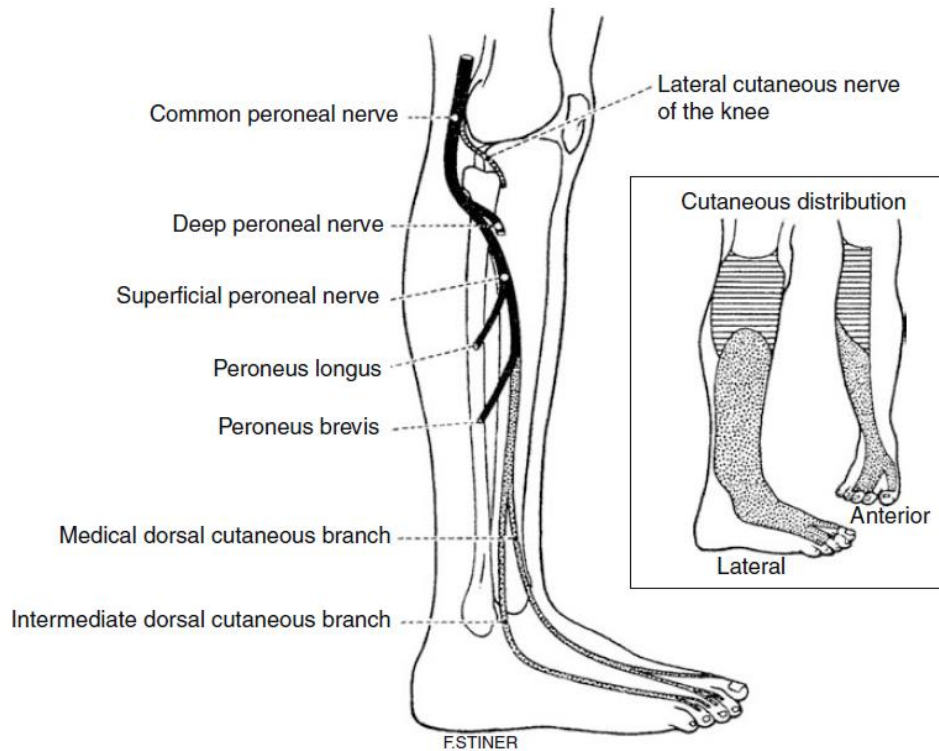


Figure 2.11 The common, deep, and superficial peroneal nerves are shown. (Reprinted with permission from Preston, D. and Shapiro, B., *Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic Correlations*, Elsevier, Philadelphia, PA, 2005.)

Anatomic diagnosis interrelated with clinical and electrodiagnostic interpretation, except few specific patterns of peripheral nerve involvement. Nerve transmission and electromyographic studies defines level and distribution of the lesions. Comparison with healthy one of electrophysiologic measures may give better estimation for diabetic polyneuropathy. Electrical studies infrequently differentiate clinical types of neuropathies. Resulting a specific diagnosis method and utilizing a set of therapy depend on clinical, electrophysiologic and histologic assessments (Kimura, 2001).

2.3.2 Motor Neuron Disease(MND)

Motor neuron disease (MND) characterized by differentiation of the motor neurons which results with death in some cases. In these cases longer than normal durations, irregular firing patterns and larger amplitudes can be seen in EMG readings (Preston & Shapiro, 2005; Hassoun, Wang, & Spitzer, 1994). These findings are results of spreading of action potentials. The general form of MND was presented in 1869 by French neurologist Jean-Martin Charcot. These type of diseases is also known as myotrophic lateral sclerosis derived from the Greek word *amyotrophic*.

Poor diagnosis and prediction of ALS becomes fatal in all cases. Approximately 10% of patients surviving longer. Although, the diseases usually affect group of people who are above 55 years, sometimes also defined on younger kids. Some types of ALS have been found with progressive bulbar palsy, primary lateral sclerosis and progressive muscular atrophy in both upper and lower motor neurons. Death of neurons generally cause from respiratory failure. Nerves controls the respiratory cycle and control Other signs are pulmonary embolus, sepsis, and pneumonia. On the other hand, real etiology of ALS is still unknown with few unconfirmed speculations ranging from genetics, viral infections to toxic poisoning (Begg, Lai, & Palaniswami, 2008).

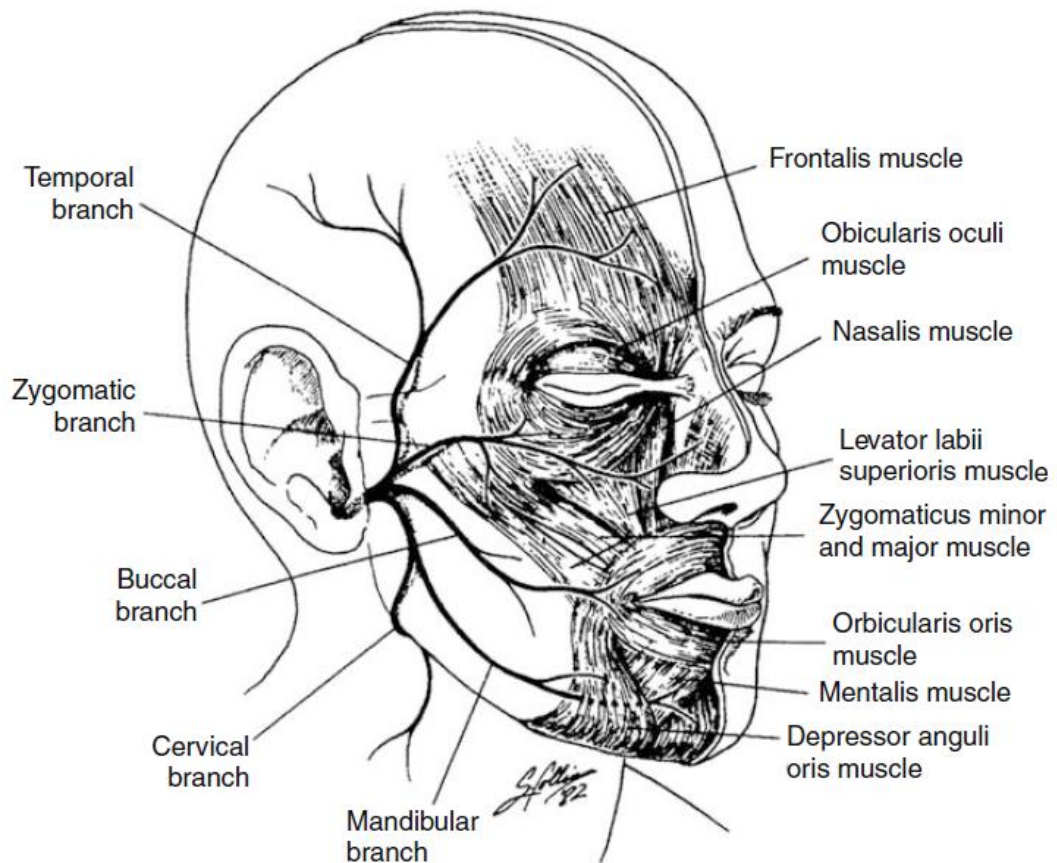


Figure 2.12 The facial nerve and its major branches are depicted innervating some of the major face muscles. (Reprinted with permission from Preston, D. and Shapiro, B., *Electromyography and Neuromuscular Disorders: Clinical-Electrophysiologic Correlations*, Elsevier, Philadelphia, PA, 2005.)

The chronic processes with marked collateral sprouting, such as spinal muscular atrophy which has the highest fiber density. Clinical studies focused on relationship between muscle strength and fiber density. The longer periods of the action potential can cause mixture of hypertrophic and atrophic fibers and slowly conducting, regenerated nerve sprouts, forming collateral reinnervation but not as much as in other neurogenic disorders (Deymeer, i dr., 1998). Conversely, quickly progressive diseases such as amyotrophic lateral sclerosis may cause increased jitter and blocking possibility. Clinically definitions of abnormalities are not possible with conventional electromyography. The frequency of force provided by each functional motor unit which helps to define recruitment method in disorders of the motor neuron, root or peripheral nerve. Increased effort to contract the muscle produces limited recruitment method information (Kimura, 2001).

Electromyography has quantitative information in assessing paresis which comes from upper motor neuron lesions. For instance, it may appear in the existence of surviving fibers that pass through the injured part of the spinal cord (Bertolasi, De Grandis, Bongiovanni, Zanette, & Gasperini, The influence of muscular lengthening on cramps, 1993). On the other hand, this can define changes in single motor unit firing characteristics (Eidelman, Nielsen, Moller, & Jannetta, 1985). Increased discharge rate in muscles can be sign of spinal cord injury (Garcia-Mullin & Daroff, 1973). In incomplete spinal cord injury can be defined by increment in the surface electromyography according to higher firing rates of activated motor units (Fabinyi & Adams, 1978; Kimura, 2001).

2.3.3 Radiculopathy

There is type of disorders which have been successfully involved into EMG studies with radioculopathy. Radiculopathy is generally caused by a differentiation of bone and disc. Disorders are specified by pain and paresthesia radiating from the nerve root. From etiological perspective, disorder is usually caused by structural lesions in the spinal column, herniated spinal discs, impingement of bone in the spinal column, and metastatic tumors of the spine. Known type of the disorder is spondylosis whose signs are acute pain in the neck and has difficulty sitting up, kneeling, or walking (Begg, Lai, & Palaniswami, Computational Intelligence in Biomedical Engineering, 2008).

Lesions generally cause partial relaxation of given muscle. Because most of muscles receive innervation from two or more roots and rarely affected by single radiculopathies. The diagnosis of a radiculopathy interrelated with abnormalities in the limbs which are innervated by the same root via different peripheral nerves. Defining same abnormalities in the correct paraspinal muscles can support the diagnosis. Another important factor for definition of radiculopathy is the time dependence and period. It can be done too early such as approximately in 21 days before fibrillation potentials or too late with static lesions approximately 6 months after onset. For these reasons, false and negative electrodiagnostic examinations are generally seen with radiculopathies (Encyclopedia of the Neurological Sciences, 2003).

2.3.4 Myopathy

Myopathies are diseases of nonfunctional skeletal muscles. Myopathies can be differentiated from other disorders. Additionally, disorders of muscles can be categorized so that it is possible to characterize, specify and categorize a particular myopathy based on its characteristic features. Myopathy disorders are related with injury of muscle group. The frequent one is muscular dystrophy which caused by degenerative muscle states. EMG readings are specified by MUAPs with shorter durations and reduced amplitudes (Begg, Lai, & Palaniswami, 2008).

Congenital muscular dystrophy (CMD) is a disorder which begins before birth related with genetic mutations by affecting the production of necessary proteins for muscles. Duchenne muscular dystrophy (DMD) or pseudohypertrophic disorder affects involuntary muscles such as the cardiac muscle. Myopathy is a genetic origin disorder and generally detected in kids between 2 and 6 years. The disorder is very rare and not fatal after 30 years old. Becker muscular dystrophy (BMD) spreads through the heart muscle which can cause heart failure and occurs similarly to DMD. Myotonic muscular dystrophy (MMD) or Steinert's disease can appear in early adulthood which is not rare. The disease begins slowly beginning with the weakness in the facial muscles and spreading throughout the body. Distal muscular dystrophy (DD) is also interrelated with gene mutation and appears young to middle-aged adults. Emery–Dreifuss muscular dystrophy (EDMD) is a slowly proceeding one which caused by gene mutations. Mostly appears on kids and cause weakness and wasting of the upper body. Another gene mutation disorder is limb–girdle muscular dystrophy (LGMD) which affects the waist and lower limbs (Begg, Lai, & Palaniswami, 2008).

The most important step of evaluating a patient with a myopathy is history information. After obtaining the history, the findings on the physical examination and weakness pattern helps for definition of the diagnosis. The results of the laboratory studies with patient history and physical examination such as blood tests, electromyogram, muscle biopsy, and molecular studies utilized to confirm the preliminary diagnosis. Symptoms history of muscle disease can be classified as negative regarding to loss of function. Conversely, can be classified as positive

regarding to problems of new functions. Negative Symptoms of a patient with muscle disease commonly can be weakness, a loss of muscle strength. For instance, when the weakness is in the legs, patients will complain of difficulty climbing. Similar type of symptoms in the arms and legs points to muscle weakness, which is most possible weakness in a myopathic disorder. It is important to define the frequency of the felt weakness. Patients tend to ask if the weakness is show all of the time. Myopathies may show either constant or episodic periods of weakness in case of muscular dystrophies and inflammatory myopathies. Muscle disorders can have one acute in less than 4 weeks, subacute in 4–8 weeks, or chronic ones in more than 8 weeks. The frequency of constant weakness can be caused from:

- 1) Acute or subacute in some inflammatory myopathies
- 2) chronic slow progression over years
- 3) fixed weakness with little change over decades.

In conclusion, both constant and episodic myopathic disorders can have monophasic or relapsing symptoms. To exemplify, a myositis can have an acute monophasic course and returns to normal strength within weeks or months. Weakness may be the most reliable negative symptom of a patient with a myopathy. Many patients who complain of diffuse common weakness or fatigue do not have a disorder.

A specific category of pain is the involuntary muscle cramp. Cramps are generally placed in a specific muscle region and last from seconds to minutes, according to how strong it is. Cramps can happen with dehydration, hyponatremia, azotemia, myxedema, and in disorders of the motor neuron or nerve. Muscle contractures are usually last longer than cramps. They can be defined clearly with electromyography needle. However, contractures are electrically silent where cramps are combined with rapid-firing motor units. Patients can complain about muscle solidness in almost any muscle group. They will have difficulty while releasing their handgrip after a handshake. The myotonia caused disabilities can be fixed with repeated exercises. Paramyotonia is the paradoxical occurrence of exercise making the myotonia worse. Myotonia is related with cyclical depolarization of the muscle. It should be asked to patient if urine has ever turned dark or red during or after these episodes, indicating myoglobinuria which follows too much myoglobin from muscle during periods. Other major points in the patient history is related with the time and period first symptoms and the age can provide important approaches to the diagnosis. Disorders

such as myotonic dystrophy and oculopharyngeal dystrophy may not become symptomatic until middle age or later. The family history is has great importance in correctly diagnosing the hereditary myopathies. It is necessary to have a detailed family tree to search for autosomal dominant, recessive, and X-linked patterns of transmission. Defining a correct pattern can not only help in correctly diagnosing the disorder but also important for genetic observation (Encyclopedia of the Neurological Sciences, 2003).

Different diagnosis approaches depends on the pattern of inheritance. The distribution of muscle weakness, and the time course of progression. Recessively inherited disorders mostly contains unfunctional units. Homozygous or hemizygous patients have only imperfect copies of imperfect gene. On the contrary, dominantly inherited disorders mostly contains functional differences. Heterozygous patients have not only normal but also mutant copies of the gene.

Beneficial monitoring and definition tests includes identification of creatine kinase level and erythrocyte sedimentation rate. Muscle biopsy samples provide histologic and histo-chemical implication. Clinical myopathic disorders' biopsy appears well-known myopathic features, although myopathy of young people will be affected neurogenic changes (Kimura, 2001).

2.3.5 Neuromuscular Junction Disorders (NMJ)

The Neuromuscular Junction Disorders (NMJ) is a middle layer of transportation of electrical nerve impulse to muscle. NMJ diseases are generally interrupt receptor interaction and spreading. EMG readings show movements of the MUAP waveform where we can examine the characteristic if damaged or healthy. NMJ can be detected by single-fiber electromyography (SFEMG) efficiently. There are types of NMJ disorders which are different. Myasthenia gravis (MG) is an autoimmune disorder where the body's immune system attacks the NMJ area. The disease affects women more than men regarding weakness or tiredness in voluntary muscles. Lambert–Eaton myasthenic syndrome (LES) is similar to MG which is combined with weakness in the lower limbs about 60%. Congenital myasthenic syndrome (CMS) anomaly of producing necessary proteins of NMJ. Symptoms and diagnosis methods

may differ according to source of the disorder because problems could occur with the presynaptic and postsynaptic part (Begg, Lai, & Palaniswami, 2008).

Typically, early symptoms include lid droop and double vision, difficulty of chewing, swallowing, and speaking and possibly common limb weakness and tiredness. Small percentage of patients' muscle weakness can improve. The clinical findings can cover ptosis, diplopia, jaw weakness, dysarthria, dysphagia, and limb muscle weakness. In addition, some types of bilateral median mononeuropathies at the wrists (Encyclopedia of the Neurological Sciences, 2003).

2.4 The Electromyograph (EMG)

The EMG shows the electrical activity of skeletal muscles and contains information about the structure and function of muscles which moves different parts of the body. EMG studies are more reliable and accurate than clinical examinations for definition of muscle fiber types which are involved in a disorder. Although, EMG studies can examine abnormal sensory nerve conduction, the patient is in ordinary poses. This approaches to sensory neuropathy narrows the possible disorders. Investigation of pathophysiological problems related to axonal loss or demyelination is also possible with EMG which will increase the clinician response time and allow some disorders to be treatable by applying different approaches. Furthermore, EMG studies can assist a clinician while deciding a diagnosis method without requiring a muscle biopsy (Sörnmo & Laguna, 2005).

Although it has been known that muscles generate electricity since the 1960s, electromyography has come into daily usage. Recently, the accuracy of EMG signal classification has been significantly improved due to progress in recording technology and computer-based signal analysis systems. Different types of electrodes are available in these days. One can record localized activity of individual muscle fibers while noninvasive electrodes providing two-dimensional spatial characterization (Sörnmo & Laguna, 2005).

It is possible to look deeper into the electrophysiology of muscles with the development of progressive signal processing techniques. For instance, by

decomposing and denoising EMG signal into different features, the activity of individual muscle units can be understood. Improvements in especially medical engineering have produced electromyography further than the traditional diagnostic applications to diverse areas such as ergonomics, exercise physiology, rehabilitation, movement analysis, biofeedback, and myoelectric control of prosthesis (Sörnmo & Laguna, 2005).

2.4.1 The Electromyograph and Instrumentation

Electromyograph is the primary recorder of MUAPs with an instrument used to record the action potential of the motor unit. There are two general types of EMGs, needle fine wire EMG and surface EMG (sEMG). Surface electrodes records total electrical activity from many muscle or nerve fibers, but needle electrodes separate individual motor unit potentials within a narrow radius from the recording tip. EMG instrumentations are electrodes, both low-pass and bandpass filters, audio and visual displays and amplifiers to increase frequency of signal. EMG system should accept minimum or zero noise and artifacts. Generally EMG instruments are having standard settings. It should be readjusted according to signal characteristics (Begg, Lai, & Palaniswami, 2008; Kimura, 2001).

Clinical approaches based on bioelectrical signals are widely using in hospitals all around the world. In some cases computer oriented systems are using as an efficient and cost-effective solution for signal analysis. This kind a system employs one or several sensors, external hardware for patient insulation card with analog/digital (A/D) conversion and software for signal analysis (Webster, 1998).

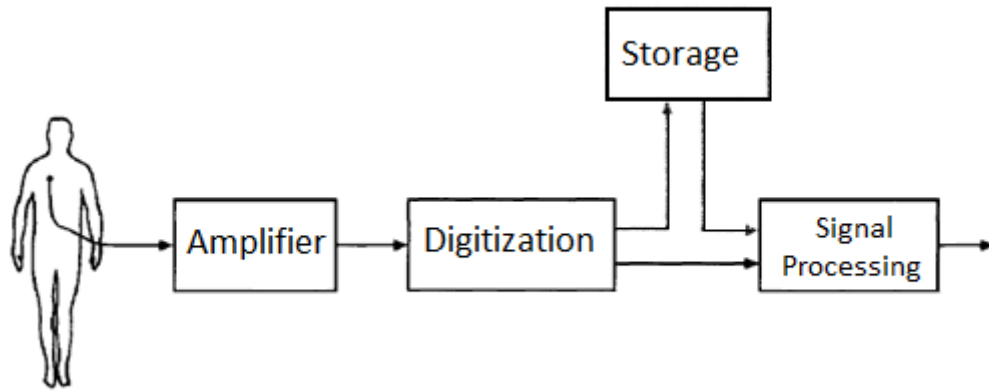


Figure 2.13 Block diagram describing the main steps in biomedical signal analysis.

In a computer oriented system, signal analysis is usually done locally by employing digital signal processor (DSP) card. The signal can be processed at the time of its achievement or can be stored on a local hard disk to use later. It is also possible to record the signal at one physical location or clinic using the computer oriented system, then can be processed in another location. Since the recorded signal mostly stored in a database, it can be better to process the signal on the server since it may have more computational ability (Sörnmo & Laguna, 2005).

2.4.2 EMG Electrodes

There are many types of electrodes ranging according to their pickup range. The concentric and monopolar electrodes have an intermediate pickup range and are used in conservative recordings. The single-fiber electrode is a more recent innovation. It has a very small pickup range and is used to get recordings from only one or two muscle fibers. The macro electrode, which is the cannula of either the concentric or single-fiber electrode in mixture with a remote reference electrode, picks up potentials throughout the motor unit territory. Adrian and Bronk developed the concentric electrode in order to get a pickup range that is smaller than that of wire electrodes. The modern version of the concentric electrode include a platinum or stainless steel wire located in the lumen of a stainless steel cannula with an outer diameter of about 0.5 mm. The tip is beveled at 15 to 20 degrees, thus exposing the central wire as an oblique elliptical area of about $150 \times 580 \mu\text{m}$. The central wire is insulated from the cannula with araldite or epoxy (Henneberg K. , 2000).

The concentric electrode is connected to a differential amplifier. Thus, common signals are effectively rejected, and a relatively stable baseline is achieved. The cannula cannot be regarded as an indifferent reference electrode because it is located within the potential distribution and thus will pick up potentials from active fibers. Simulation studies have shown that the cannula shields the central wire from picking up potentials from fibers located behind the tip (Henneberg & Plonsey, 1993). The sensitivity of the concentric electrode is for that reason largest in the hemisphere facing the oblique elliptical area. Due to the asymmetric sensitivity function, the waveshape of the obtained potentials will vary if the electrode is rotated about its axis. This problem is not observed with the axisymmetric monopolar electrode, however, this electrode has a more unreliable baseline due to the remote location of the reference electrode. Both the concentric and monopolar electrodes are used in conventional EMG. Owing to the differences in recording characteristics, however, concentric and monopolar recordings cannot be analyzed easily. During the concentric needle examination, the investigator searches for abnormal insertional activity, spontaneous activity in relaxed muscles, and motor unit potentials with abnormal shaped. The waveshape of motor unit potentials is assessed on the basis of the quantitative waveform features (Henneberg K. , 2000).

In daily usage of clinics same electrodes using for different patients. That is why need to sterilize to prevent the transmission of infection. Sterilization of needle electrodes in boiling water for at least 20 minutes prior to use. Gas sterilization also enough to prevent the transmission of infection. The chemicals and boiling water usage may damage the other parts of electrodes. Electrode manufacturers delivers instructions and manuals for optimal sterilization methods (Begg, Lai, & Palaniswami, 2008).

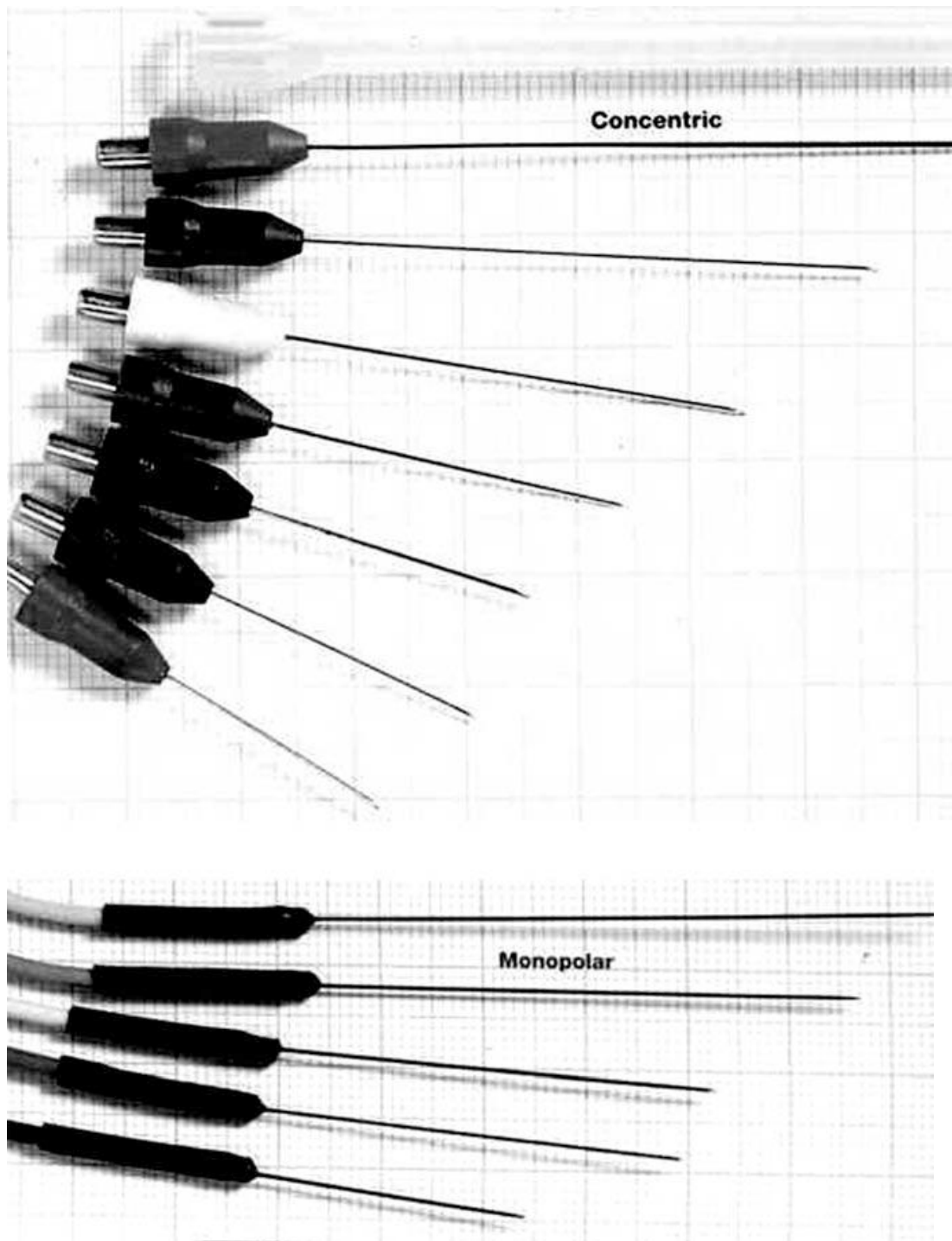


Figure 2.14 Examples of concentric and monopolar EMG needle electrodes sold by Med Products. (From <http://www.medproducts.com>. With permission.)

Surface electrodes located on the muscle to record activities from many motor units. Individual motor unit potentials during mild muscle contraction can be recorded by using needle electrode. Motor units can be affected by other neighbor activities which makes harder to analyze, classify and identify. Electromyographers use standard,

bipolar concentric or monopolar needles, for daily clinical uses. Single-fiber electrodes allow recording of potentials derived from single muscle fibers. Rarely used electrodes are multielectrode, flexible wire electrode, and the microelectrode which can be placed intracellular (Sörnmo & Laguna, 2005).

Another approach for sterilization is discarding needle electrodes after use in each patient which is less costly and practical. The American Association of Electromyography and Electrodiagnosis recommends this way of use to prevent possible transmission of diseases, especially after studying a patient with AIDS, hepatitis, or any other contagious disorder. More safer application before disposal of used needles and blood-contaminated materials is autoclaving them for one hour at 120° C and 15 PSI (Begg, Lai, & Palaniswami, 2008).

There is many needle electrodes due to electrical properties. Electrodes should be checked periodically if they are functioning correctly and properly. If recording surface is increased, the amplitude and area of the recorded motor unit potentials should be reduced, relatively. Electrodes should be maintained and tested periodically to be sure about reliability. One battery can be connected to the lead of a needle and the other battery to an ammeter with a small exploring metal hook to test needle insulation. The electricity can flow if the exploring hook touches the exposed tip of the needle. An ammeter should record no current if connected to the battery through the two leads of standard or bipolar concentric needles (Begg, Lai, & Palaniswami, 2008)

Surface electrodes employs square or round metal plates made of platinum or silver. These metal plates may come in different sizes around 1 X 1 cm. Transpiration can cause short circuit from surface electrodes. On the other hand, surface EMG electrodes are a more practical alternative to needle and wire methods. These electrodes are generally discs or bars shaped to gather signal effectively due to applied part of body. Gel or paste can be used to obtain a good interface for signal acquisition instead of using dry or isolated electrodes, to prevent noise from environment. Once alcohol or medical napkins should be applied to cleanse the skin and the area is shaved to avoid motion artifacts before attaching these electrodes

(Begg, Lai, & Palaniswami, 2008). A surface electrode is shown in Figure 2.17 with five channel sEMG experimental setup.

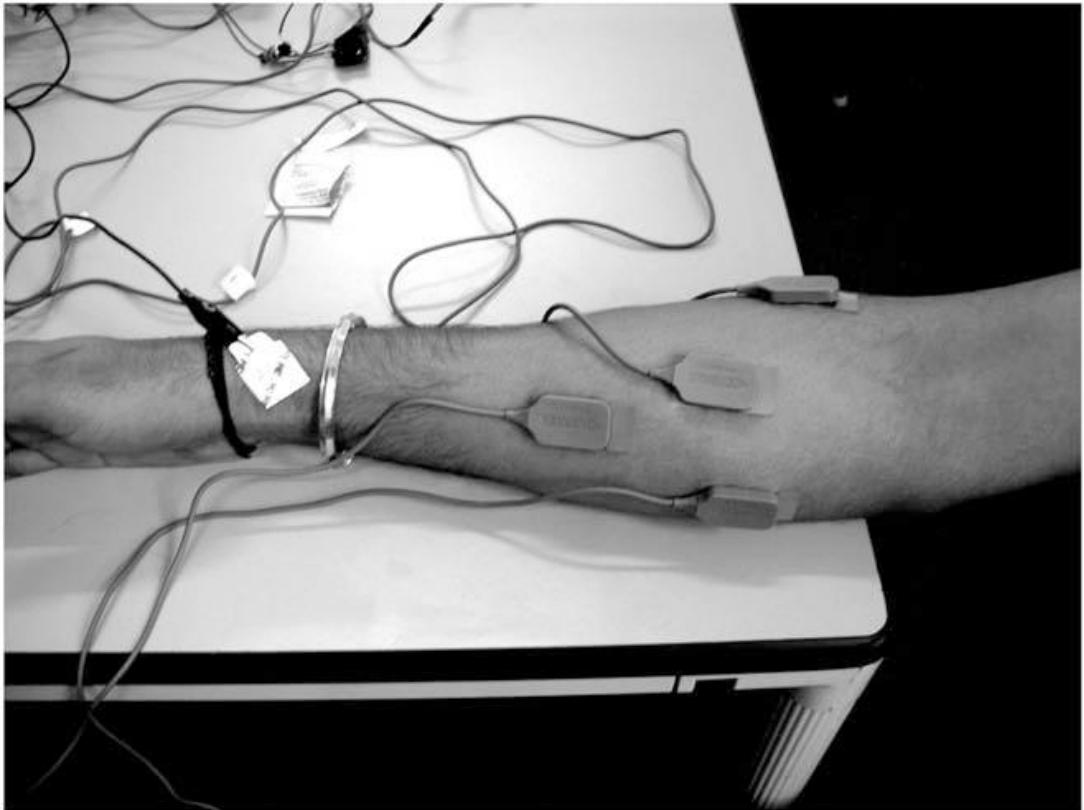


Figure 2.15 Experimental setup using five channel sEMG for detecting limb motion. Experiments performed in the biosignals lab of RMIT, Australia.

Some surface electrodes allow movement between electrolyte and skin to reduce possibility of having extra noise. Surface electrode is more coherent for monitoring voluntary muscle contraction during kinesiological studies and recording of muscle action potentials. Surface electrode records electrical activities from a wider region. In addition, recordings with surface electrode covers the recording radius of some 20 mm selective pickup from a 500 μm radius by a needle electrode. The size of recorded muscle action potentials decreases with increasing electrode size. (Kimura, 2001).

Standard Or Coaxial Concentric Needle electrode, presented by Adrian and Bronk in 1929 which has a stainless-steel cannula similar to hypodermic needles, with a wire in the center of the shaft. The wire, generally made of chrome, silver, or platinum. This type of the needle has an oval shape with an area which is about 150 μm x 600 μm . It records the potential difference between the wire and the shaft when the needle near source of electrical activity in a defined recording area. A separate

surface electrode applied with a suction cup which performs as ground. Non reusable concentric needles usually can be compared with reusable electrodes. Electric or physical testing of the leads may not be predicted due their recording characteristics (Kimura, 2001).

Bipolar concentric needle contains two fine stainless steel or platinum wires. As a result, it has a larger diameter than the standard concentric needle for the same size of wires. The electrode records the changes of potentials between the two inside wires. Thus, the bipolar electrode detects potentials which has much smaller volume than the standard needle. The three terminals connects the cable whose two are active leads and a ground. A small number of single muscle fibers utilized as the source for electrical activity by this type of recording from a restrictly defined area. Concentric electrodes may detect more spontaneous potentials than monopolar needles probably because of increased tissue injury (Kimura, 2001).

Monopolar needle electrode made of stainless steel and it is well insulated except at the distal 0.2 to 0.4 mm. The wire coated by a Teflon which has an around 0.8 mm diameter. Its sharp tip is painy during insertion which is less stable electrically and noisier than the concentric electrode. Presoaking improves the resolution of low amplitude signals. A monopolar needle records voltage changes between the tip of the electrode and the reference. The spatial recording characteristics can change from one type of needle to another. In general, a monopolar needle records a potential from the same source (Kimura, 2001).

Single-fiber electromyography uses an electrode with a smaller leading edge to record from individual muscle fibers instead of motor units. A wire with 25 μm length is mounted on the side of a needle which gives the maximal amplitude between near and distant muscle fiber potentials. Single-fiber needles may contain two or more wires along the shaft as in concentric electrodes. The mostly used type has one inserted wire into a cannula which is few millimeters at the back of the tip. Single-fiber needles has greater potential with radial distance method compared to concentric or monopolar electrodes (Kimura, 2001).

Multielectrodes includes three or more insulated wires which are generally around 1 x 1 mm size and placed through cannula. The difference between the leads and the side of multielectrode verifies the recording radius. The generally applied distances for the motor unit territory is around 0.5 mm for myopathy and 1.0 mm for neuropathy. The Single-fiber needle may include multiple wires exposed along the shaft. Multi-lead electrodes may utilize for intraoperative nerve recording (Kimura, 2001).

A flexible wire and hypodermic needle are tightly interrelated which allows freedom of movement in kinesiologic examination and provides getting less noise. Some researchers prefer to use a bipolar electrode which made of nylon-coated Evanohm alloy wire and has 25 μm in diameter. Mostly used ones has insulated platinum wires with 50–100 μm diameter and bare tip. A small hole created in the insulation of the wire which may provide smaller lead-off surfaces. These electrodes cannot be fitted for quantitative studies of action potentials due to lack of rigid standardization (Kimura, 2001).

Glass microelectrodes can be used for intracellular recording. It consists of fine glass tubing filled with potassium chloride solution. Due to its fragile structure, cannula should be used as a carrier to attach the electrode through the skin and a micromanipulator for muscle connection. The electrode has a very good tip which is less than 1 μm in diameter. Therefore, recording from a glass microelectrode need amplifier to balance impedance and diameter size (Kimura, 2001).

Electrodiagnostic examinations range in amplitude from microvolts to millivolts. The amplifier includes some stages to achieve in the defined range of sensitivity. Amplifier is a must for EMG analysis due to its complex structure to have a high input impedance, a low noise level, and a large dynamic range. This arrangements and processes raise signal-to-noise ratio by allowing major amplification of the signal (Kimura, 2001).

2.4.3 Signal Acquisition

EMG data generally recording in a hospital or in a lab by an electromyographer using either noninvasive surface or needle electrodes due to situation and requirements. Once the location should be determined for EMG recording, like biceps brachii on the upper arm. Then the EMG begins to copy for a predetermined time by applying predefined methods about amplification, bandpass filters and digitization. Many EMG acquisition instrumentations supports these basic facilities. For example, the needle EMG is a standard clinical tool which can be used for diagnostic purposes, while the surface EMG can be analyzed and processed in different applications such as prosthesis control, ergonomics, movement and gait analysis and sports medicine (Sörnmo & Laguna, 2005).

All type of EMG recordings either from surface EMG or needle EMG differentiates due to voltage potentials and intracellular activities even if nerve or muscle (Preston & Shapiro, 2005). The transmission process from intracellular electric potential to the extracellular fluid and tissue is known as *volume conduction* which can be modeled as near-field or farfield potentials. Recent studies in the area has focused on modeling surface EMG based on volume-conduction (Mesin & Farina, 2005; Mesin, Joubert, Hanekom, Merletti, & Farina, 2006). This approach is successful and useful for predefinition and decomposition of the surface EMG waveform into the individual MUAP waveforms (Begg, Lai, & Palaniswami, 2008).

Recordings should be made as close as to the source of the action potential not to be affected by near-field potentials. Because every other signal or effect from environment or near-field potentials will damage desired recording. The amplitude of the recording is due to the electrode proximity and distance between the recording electrodes. The recording can be involved not only by near-field potentials but also compound MUAPs. The single MUAP waveform which is triphasic, can be employed as action potential measuring method (Begg, Lai, & Palaniswami, 2008).

A smaller positive phase of EMG can be observed as the potential. If the electrode is attached directly to the source of the action potential, the initial positive may be canceled and a negative spike can be observed once. If the muscle action potential

passes through the electrode, one generally gets a small positive effect. This circumstance can be happen when recording a muscle which is connected by the median nerve. In this case, MUAPs never directly reach the target muscle and its action potentials. Far-field motor unit action potentials can easily propagate to a wider area. In this case two electrodes can be employed to record the same potential which will be combined (Begg, Lai, & Palaniswami, 2008).

Various information can be obtained by the electromyographer due to distance between surface and far-field potentials such as:

a. *Insertional activity*. Initial insertion and movement of the needle electrode cause a short peak potentials in normal muscles and nerves which is calling as *insertional activity*. These peak potentials have been lasting in 1 to 3 ms and amplitudes of 0.3 mV, most probably due to physical bump to the electrode.

b. *Spontaneous activity*. Occasionally abnormal or spontaneous depolarization of muscle can be happen even without impulse activation. The reason can be due to spontaneous depolarization of muscle cell, fasciculation potentials, myokymic discharges, and other factors (Preedy & Peters, 1994).

c. *MUAP morphology and stability*. The MUAP waveform can be processed and analyzed when the initial activity finished. The waveform and shape of the individual MUAPs informs about etiology of the disorder (Begg, Lai, & Palaniswami, 2008).

2.4.4 Signal Amplification and Filtering

Once the signal band-pass filtered and sampled then low-pass filtered again, after amplification. Filters are generally employed in the instrumentation and there is almost no need to specify the frequencies. Conversely, some international journals strict about the type and application of filtering. This is essential since an important part in the EMG signal is known to be contained between 5 and 500 Hz frequency. This frequency range preferred between 10–350 Hz band for surface EMG

recordings while between 10–150 Hz or 50–350 Hz for filtering due to structure of waveform. Application of one filter to all type of EMG signals is not favorable because signal's power may be above 150 Hz and below 50 Hz from time to time which should not be ignored. Intramuscular recording and filtering should be made according to higher frequencies, e.g, 10–450 Hz bandpass filter. These are probable ranges, which can be adjusted according to justification. Bandpass filter with cutoffs can be applied at 3–10 kHz, and with sampling rate of 20 kHz. The filtered signal is had a simplification for the analysis with eliminated signal samples (Begg, Lai, & Palaniswami, 2008).

2.4.5 Signal Digitization

After the signal is collected from muscle it needs be amplified then filtered, to analyze it better and deeply. To apply these steps, signal has to be converted from analog to digital. If signal could be saved, it can be used for sometime after. For these requirements and technological improvement, signal digitized and saved in digital format after conversion and sampling. Range of bandpass filter of sampling is double of normal cases during digitization to save all details. For example, if a bandpass filter is applied between 10–350 Hz, the minimum sampling rate will be 700 Hz (350×2). In some cases, a low-pass filter is used first to smooth the EMG signal. The raw unfiltered EMG signal can also be stored in the computer using a sampling rate greater than 2.5 kHz which should be preprocessed before analyzing, to cancel faults and extraordinary situations. The conversion of the analog EMG signal to digital signal can be done by an A/D convertor with a resolution of 8–12 bits.

2.4.6 Factors Influencing Electromyograph Readings

In all EMG studies, locations of electrodes are very important because wrong location can cause incorrect EMG signal observation. Positional factors also affect the collected EMG signal and are important to prevent miscalculations. Difficult part of recording correct EMG readings is deep muscles and their small and tiny amplitudes. This makes harder to read signal which can easily corrupted and affected by environmental and neighbor noise. Concentric needle EMG is produced to obtain

more clear signal which is harder to use due to physiological reasons. To exemplify, EMG signal may distort by a patient who is uncomfortable during the reading. In addition environment of the place where EMG is collecting is also important. It should be silent and no one should not enter the room while recording to increase comfort of patient.

The EMG instrumentations like the type of electrode and the filter properties can affect MUAP waveforms. Waveform and amplitude of EMG recordings also related with the distance between the muscle fiber and the recording location. In addition, action potentials may route to the electrode site can be mixed because of crosstalk. Deep muscles propagate through the surface by passing layers of muscle and fat which have different properties and structures. Waveform speed, and the waveform can be distorted or dispersed because of these layers. In addition to intermuscular effects, crosstalk can also arise from cables and switches between the electrodes and locations. choosing appropriate electrode size, interelectrode distance, and location is also important. Extra attention should be paid when working on areas where many muscles located (e.g., forearm) or when working with superficial/thin muscles (e.g., trapezius).

CHAPTER 3

BIOMEDICAL SIGNAL PROCESSING TECHNIQUES

3.1 Fourier Series Analysis

3.1.1 Periodic Functions

Advantage of the Fourier transform approach is energy contain sinusoids at one frequency. The amplitude of these sinusoids must be interrelated with frequency component in the waveform. It's been known that any periodic waveform can be represented by a series of sinusoids which are at the same frequency. The set of sinusoids can be expressed either as sines or cosines. Assume that sines and cosines are used to find the appropriate amplitude of these components. It is required to correlate the waveform with the sine and cosine set, and average over the complete waveform. This procedure can be formulized as follows.

$$\begin{aligned} a(m) \\ = \frac{1}{T} \int_0^T x(t) \cos(2\pi m f_T t) dt \end{aligned} \quad (3.1)$$

$$\begin{aligned} b(m) \\ = \frac{1}{T} \int_0^T x(t) \sin(2\pi m f_T t) dt \end{aligned} \quad (3.2)$$

T is the period or length of the waveform, $f_T = 1/T$, and m is set of integers which defines the family member. This becomes set of sines and cosines by having harmonically related frequencies, $m f_T$ (Semmlow, 2004).

The Fourier series analysis uses an investigation function in the set of interrelated sinusoids. In the set, the sines and cosines have valid frequencies only at values of m/T . The same frequency or higher multiples (when $m > 1$) are termed as *harmonics*. The approach is called as *harmonic decomposition* since represents

waveforms by harmonically related sinusoids. The Fourier transform and Fourier series form a mutual transformation. The Fourier transform can be applied to a waveform to get the sinusoidal components. Sine and cosine components of Fourier series can be processed to reconstruct the original waveform by following equation:

$$x(t) = a(0/2 + \sum_{m=0}^{\infty} a(k) \cos(2\pi m f_T t) + \sum_{m=0}^{\infty} b(k) \sin(2\pi m f_T t) \quad (3.3)$$

Mostly the number of sine and cosine components is limited which is fairly short. The reconstructions are inaccurate particularly for the triangular wave by using 3 sine waves. Spectral information is generally processed as a frequency plot, a plot of sine and cosine amplitude. $f = m/T$ is conversion formula where m is component number, f is frequency, and T is the period. The sampling frequency may be used to define the spectral frequency in digitized signals. It is more applicable to plot the amplitude and phase angle of a sinusoidal wave rather than plot sine and cosine amplitudes by using the rectangular-to-polar transformation:

$$a \cos(x) + b \sin(x) = C \sin(x + \Theta) \quad (3.4)$$

where

$$C = (a^2 + b^2)^{1/2} \text{ and } \Theta = \tan^{-1}(b/a) \quad (3.5)$$

The magnitude of the sinusoidal component becomes quite small after the first few components which explains why the triangle function can be so accurately represented by only 3 sine waves (Semmlow, 2004).

3.1.2 Discrete Time Fourier Analysis

The discrete-time Fourier series analysis modified by two operations: sampling and windowing which is an extension of the continuous analysis. The sampling process makes the spectra recurring. As a matter of fact, the discrete Fourier series of waveform is theoretically infinite. Since it is periodic and symmetric, all of the information is enclosed in the frequency range. The sampling theorem is the original analog waveform to be limited on its highest frequency. The digitized data is representation of the analog waveform.

The digitized waveform must be cropped due to environmental and physical limitations. The process is calling as *windowing*. The windowing process can be multiplication of the data by a window shape. If the waveform is basically truncated and no shaping is applied on the digitized waveform which is often the case, then the rectangular is the default window shape. Different shapes can be applied on the data by multiplying the digitized waveform. The effectiveness of windowing processes is preparing the data which can be more applicable. The computation equations of Fourier series are same as summation of continuous data except the integration. Generally these type of equations are utilizing complex variables. For that reason both sine and cosine terms can be utilized by a single exponential term using Euler's identity:

$$e^{jx} = \cos x + j \sin x \quad (3.6)$$

The equation for the discrete Fourier transform becomes as follows by utilizing complex variables:

$$X(m) = \sum_{n=0}^{N-1} x(n) e^{(-j2\pi mn/N)} \quad (3.7)$$

where N is the total number of points and m represents the group member. Result of the equation can be both positive and negative when used in complex notation like $m = -N/2, \dots, N/2-1$. In Eq. (3.7), $f_m(n)$ is replaced by $e^{-j2\pi mn/N}$. The inverse Fourier transform can be calculated as:

$$x(n) = \frac{1}{N} \sum_{m=0}^{N-1} X(m) e^{(j2\pi n f_m T_s)} \quad (3.8)$$

$|X(m)|$ gives the magnitude for the sinusoidal representation of the Fourier series while the angle of $X(m)$ gives the phase angle for this representation, since $X(m)$ can also be written as:

$$X(m) = \sum_{n=0}^{N-1} x(n) \cos\left(\frac{2\pi mn}{N}\right) - j \sum_{n=0}^{N-1} x(n) \sin\left(\frac{2\pi mn}{N}\right) - j \quad (3.9)$$

For computational reasons, $X(m)$ can be both positive and negative values for m ; negative values points negative frequencies. But these are only computational requirements and don't means anything physically. In some versions of the Fourier series equations which is shown in Eq. (3.7) is multiplied by T_s (the sampling time) while in Eq. (3.8) is divided by T_s . It can be said that the sampling period is integrated clearly into the Fourier series coefficients. Different versions of fourier equations can be found in the literature. The discrete Fourier transform produces a function of m . To convert this to frequency notes that:

$$f_m = mf_1 = m/T_P = m/NT_s = mf_s/N \quad (3.10)$$

where $f_1 \equiv f_T$ is the fundamental frequency, T_s is the sample interval; f_s is the sample frequency; N is the number of points in the waveform; and $TP = NT_s$ is the period of the waveform. The equation for the discrete Fourier transform can also be written after swapping as follows:

$$X(f) = x(n) = \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j2\pi n f_m T_s} \quad (3.11)$$

which may be more easier and applicable for manual calculations. If the waveform is truly periodic, then the approach produces an accurate spectrum of the waveform. Such analysis be termed as Fourier series analysis, which also known as Fourier transform analysis. The second term using for aperiodic or truncated waveforms. In all cases, used algorithms are the same, thus the Fourier transform is widely applied to all spectral analyses, especially based on decomposing a waveform. Currently, the upgraded version of Fourier transform is fast Fourier transform (FFT) which is more computationally efficient algorithm that crops the number of computations from N^2 to $2 \log N$, where N is the length of the digital data (Semmlow, 2004).

3.2 Spectral Analysis: Frequency Resolution

The number of points produced by the operation is N is the number of points in the data set which is from the discrete Fourier series equation above (Eq. (3.7)). The spectrum generated symmetrically around the middle point which is $N/2$ (or $f_s/2$ in

frequency). Each point in the spectra represents a frequency point of $1/(NT_s)$ when sampling time is T_s . The frequency resolution of the spectra will be similar to frequency spacing with $1/(NT_s)$. Frequency resolution can be influenced by the type of applied windowing. Frequency of the spectrum produced by the Fourier transform can be decreased or increased due to the length of the data, N . Frequency resolution can be improved by increasing sample interval, since decrease in f_s means that the maximum frequency in the spectra is reduced by limiting the spectral range. Application of zero padding is acceptable since the undigitized part of the waveform is always assumed as zero. Zero padding basically adds more unsampled waveform regarding this assumption. Since the frequency interval is smaller it seems improved because of zero padding process. Zero padding does not increase the underlying resolution of the transform since the number of points provide information which is same with different frequencies. On the other hand, zero padding make available interpolated transform with a smoother appearance of waveform. Additionally, it may behave as a filter in practice when a signal has a center frequency that lies between the $1/NT_s$ which makes easier estimation of peaks in the spectra (Semmlow, 2004).

Deterministic signal model also can be adopted with mathematical tools for spectral analysis. It should be pointed that application of these tools requires that the aperture from $-\infty$ to $+\infty$. Once, we should approach with the thought of every sample function of a stochastic process is a deterministic function. Accordingly we can utilize Fourier series and transforms to compute a spectrum for stationary processes. Deterministic functions and sequences are employed in many studies of stationary processes for example, the autocorrelation sequence, which is a deterministic sequence. Spectra can be defined by deterministic signals which can be used to summarize important features of stationary processes. Continuous-time signals are most practical applications of spectrum estimation. For example, in electroencephalography, spectrum estimation is employed to study on sleep disorders and its effect of medication on the brain (Duffy, Iyer, & Surwillo, 1989). There are three steps for numerical computation of the spectrum of a continuous-time signal:

1. Sampling the continuous-time signal to obtain a sequence of samples.

2. Collecting number of nearby samples to use for the computation of the spectrum which is known as *windowing*, or *tapering*.
3. Computing the values of the spectrum at predefined frequencies.

Once, the continuous-time signal is processed after applying a low-pass filter then should be sampled to obtain a discrete-time signal. Finally, a proper length of the windowed data should be used for spectrum estimation, which is then analyzed (Manolakis, Ingle, & Kogon, 2005).

3.3 Truncated Fourier Analysis : Data Windowing

Biomedical engineering examples and analysis can be found in EEG, EMG and ECG. Waveforms of these signals being analyzed with real subjects with actual data. But, only some part of such waveforms can be represented and processed in the finite memory of computers, and some attention must be paid and proper methods should be applied while truncating the waveform. A segment is can be basically cropped from the an overall waveform and stored in computer which will be part of the actual waveform. This is approach is calling as rectangular *windowing* over the waveform. The analysis is limited with the *windowed* portion of the waveform. Windowing process is similar to the sampling process which is explained above. Different window shapes can be applied other than rectangular by multiplying the waveform. It should be assumed, points which are outside of the window are zero. Data set should be windowed if the data set is larger than the memory storage. Thus, the frequency characteristics of the window become part of the spectral result. By the way, all windows produce artifact. The absence of a window function is a rectangular window, as a default. All type of windows produces two types of artifact which are *mainlobe and sidelobe*. The actual spectrum is widened by an artifact called as *mainlobe*, and additional peaks called *sidelobes*. Most alternatives of rectangular window reduce the sidelobes, but produces wider mainlobes. Figures 3.1 and 3.2 depicts the shape and frequency spectra of two popular windows which is *triangular and hamming window* (Manolakis, Ingle, & Kogon, 2005). The algorithms for these windows are as follows:

Triangular window:

for odd n :

$$w(k) = \begin{cases} 2k/(n-1) & 1 \leq k \leq (n+1)/2 \\ 2(n-k-1)/(n+1) & (n+1)/2 \leq k \leq n \end{cases} \quad (3.12)$$

for even n :

$$w(k) = \begin{cases} (2k-1)/n & 1 \leq k \leq n/2 \\ 2(n-k+1)/n & (n/2) + 1 \leq k \leq n \end{cases} \quad (3.13)$$

Hamming window:

$$w(k+1) = 0.54 - 0.46(2\pi k/(n-1))k = 0, 1, \dots, n-1 \quad (3.14)$$

These and others windowing can easily be implemented in MATLAB, particularly with the Signal Processing Toolbox. Selecting the proper window, like other aspects of signal analysis, depends on what spectral features. We can work out with the spectrum of a signal by using a finite-duration segment. For the reason that:

1. The spectral composition of the signal may changes in time.
2. We have only a finite set of data.
3. We desired to save the computational complexity to an acceptable level.

These are the reason why it is required to partition $x(n)$ into blocks of data before processing. This set of processes is called as *frame blocking* and it is employing two parameters which are *length* of frame, N , and the *overlap* between frames, N_0 . The main problem for frequency analysis is determination of the spectrum of a signal. Its values will be in a finite interval $0 \leq n \leq N-1$, from a finite-duration segment. If we consider that the signal is periodic, we can easily determine its spectrum by calculating its fourier series. On the other hand, in practical applications we cannot make this consideration since the available block of data could be either part of the period of a periodic signal or a segment from an aperiodic signal (Manolakis, Ingle, & Kogon, 2005). There are three ways to deal with this issue:

1. *Periodic extension.* We consider that $x(n)$ is periodic with period N
2. *Windowing.* We consider that the signal is zero outer the interval of observation. This is very similar to applying the rectangular windowing

$$\omega R(n) \triangleq \begin{cases} 1 & 0 \leq n \leq N-1 \\ 0 & \text{elsewhere} \end{cases} \quad (3.15)$$

3. *Extrapolation.* Periodic extension and windowing can be thought as simplest forms of extrapolation. A successful extrapolation can be done with better spectrum estimates. Periodic extension is a essential application of the DFT. Extrapolation needs to employ some form of a sophisticated signal model (Manolakis, Ingle, & Kogon, 2005).

3.3.1 Power Spectrum

Fourier transform of the autocorrelation function is generally defined as power spectrum. Equation of the power spectrum is described below, both for continuous and discrete.

$$PS(f) = \int_0^T r_{xx}(\tau) e^{-2\pi f_T \tau} d\tau \quad (3.16)$$

$$PS(f) = \sum_{n=0}^{N-1} r_{xx}(n) e^{-j2\pi n f_T T_s} \quad (3.17)$$

$r_{xx}(n)$ is the autocorrelation function. Since the autocorrelation function has odd symmetry, the sine terms will all be zero and Eq. (3.17) can be simplified by applying real cosine.

$$PS(f) = \int_0^T r_{xx}(\tau) \cos(2\pi m f_T t) d\tau \quad (3.18)$$

$$PS(f) = \sum_{n=0}^{N-1} r_{xx}(n) \cos(2\pi n f_T T_s) \quad (3.19)$$

These equations both in continuous and discrete form are named as *cosine transform*. This computation approach to power spectrum is supporting *direct approach*, which makes fast fourier transform more efficient. On the other hand, a variation of this approach is used in certain time–frequency methods. The direct approach is formed by the energy analog signal which is placed in $x(t)$.

$$E = \int_{-\infty}^{\infty} |x(t)|^2 dt \quad (3.20)$$

By an extension of Parseval's theorem it is easy to show that:

$$\int_{-\infty}^{\infty} |x(t)|^2 dt = \int_{-\infty}^{\infty} |X(f)|^2 df \quad (3.21)$$

Consequently $|X(f)|^2$ equals to the energy density function which is known also energy spectral density, the *power spectral density*, or simply the *power spectrum*. The power spectrum is calculated as squared of the Fourier transform of the waveform.

$$P_s(f) = |X(f)|^2 \quad (3.22)$$

Generally, power spectral analysis is applied to truncated data, when the data has noise inside. Averaging is used while the power spectrum is calculating by FFT. Power spectrum evaluation is an estimation process by using statistical properties of waveform.

The selection of data window and averaging method is generally based on experimentation of the actual data by using statistical inferences. Data windowing tend to be described and applied for analysis of power spectral. Averaging is generally calculated by dividing the waveform into a number of segments. The final spectrum is produced from an average of the Fourier transforms which is obtained from the various segments. Segmentation is reducing the number of data samples. Reliability of the spectral estimation will increase by applying shorter segment length. As a result, frequency resolution will decrease due to segment length. A Fourier transform is calculated from each segment, and then averaged. The periodogram produced from the segmented and averaged data which is much smoother (Semmlow, 2004).

Overlapping segments are used and a window is applied to each segment In this approach. More segments can be averaged for a given segment and data length, by overlapping segments. More overlapping has been recommended in other

applications. Maximum overlap has been recommended when computing time is not an issue. The windowing for sidelobe control is not very important when the spectra expected to be relatively flat. In addition, windows should be applied whenever the spectra are expected have large amplitude differences (Semmlow, 2004).

Real-time spectrum analyzers mostly analyze the spectrum which is computed by using the DFT. If the number of DFT samples is greater than or equal to frame length N , the exact continuous spectrum can be obtained by reconstructing the frequency domain (Oppenheim & Schaffer, 1989; Proakis & Manolakis, 1996). This reconstruction requires a periodic sine function which is not a practical function to implement, especially in real-time applications. Therefore a simple linear interpolation is used for plotting. Linear interpolation can lead to unaccepted results even the computed DFT sample values are correct. DFT equation of periodic N -point period is formulated as follows.

$$\tilde{X}(k) = \sum_{n=0}^{N-1} x(n)e^{-j(2\pi/N)kn} \quad (3.23)$$

There should be an interrelation between the DFT and the Fourier transform or the Fourier series of discrete-time signals. It can be seen that $x(n)$ is also periodic with period N By a diverse approach to inverse DFT.

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} \tilde{X}(k)e^{j(2\pi/N)kn} \quad (3.24)$$

This is an interesting result since there is no information about the signal $x(n)$ except the interval $0 \leq n \leq N - 1$. The period in the time domain can be easily applied by utilizing sampling in the frequency domain. Continuous-time sinusoidal signal can be sampled to understand effects of spectrum sampling. The discrete time fourier transform plots has mainlobes and sidelobes, but the discrete fourier transform samples sidelobes at nonzero values. For that reason length of the window is an important variable in spectrum estimation (Manolakis, Ingle, & Kogon, 2005).

3.4 Ar - Spectral Analysis : Modern Techniques

3.4.1 Parametric Based Model Methods

Every discrete-time periodic signal $x(n)$ can be expressed by the following Fourier series, with period N .

$$\begin{aligned} x(n) &= \sum_{k=0}^{N-1} X_k e^{j(2\pi/N)kn} \end{aligned} \quad (3.25)$$

Where

$$\begin{aligned} X_k &= \frac{1}{N} \sum_{n=0}^{N-1} x(n) e^{-j(2\pi/N)kn} \end{aligned} \quad (3.26)$$

The sequence X_k , $k = 0, \pm 1, \pm 2, \dots$, is called as *spectrum* of the periodic signal $x(n)$. Equation $X_{k+N} = X_k$ is the spectrum of a discrete-time periodic signal. The power of the periodic signal $x(n)$ can be determined by Parseval's relation as follows

$$P_x = \frac{1}{N} \sum_{n=0}^{N-1} |x(n)|^2 = \sum_{k=0}^{N-1} |X_k|^2 \quad (3.27)$$

The sequence $|X_k|^2$ is known as the *power spectrum* of the periodic sequence $x(n)$ (Manolakis, Ingle, & Kogon, 2005). These methods are the most powerful spectral estimators. Although some information of the data can be helpful for window selection, no need to plenty of features for estimation, due to source of the data.

In that kind classical approaches, the waveform which outside of the data window is processed as zero. Since the approach is not applicable for all situations, it can cause distortion while predicting (Marple, 1987). Modern approaches of spectral analysis can deal with some of distortions which are based on the classical approach and are rarely performs well when data segments are short. Modern techniques can be

classified into two main classes which are parametric, model-based or eigen decomposition, and nonparametric. These techniques trying to deal with the restriction and weakness of traditional methods.

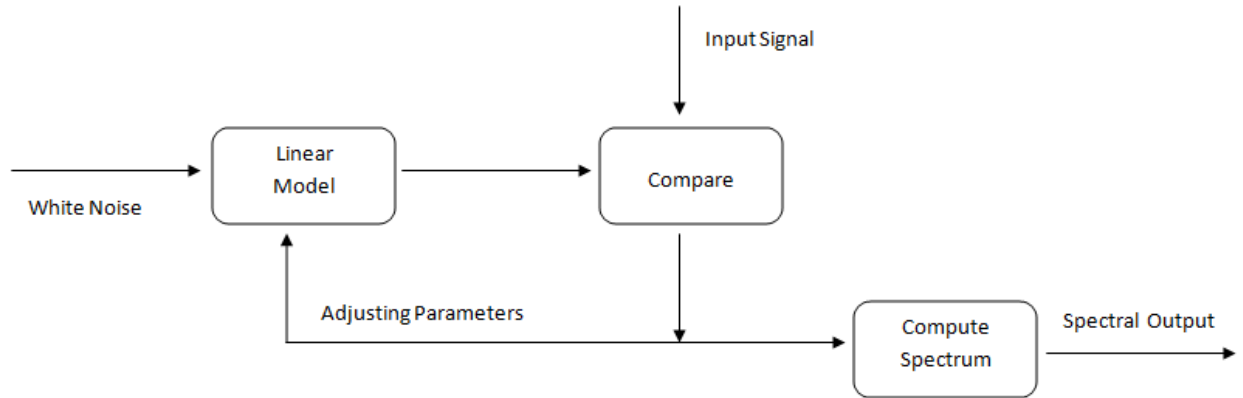


Figure 3.1 Schematic representation of model-based methods of spectral estimation (Semmlow, 2004).

Methods can make prediction if any component of the waveform is known. When the waveform includes too much noise, spectral resolution may improve by filtering data for windowing. Improvement in resolution and fidelity will depend strongly on selection of proper model (Marple, 1987). Consequently, modern approaches, especially parametric spectral methods, applies more ruling than classical methods. One of the better known criteria for selecting the model order has been proposed by Akaike, called as Akaike information criterion (AIC) (Akaike, 1974). Parametric methods generally used for estimation of power spectrum. The output of this model is compared with the input waveform. Parameters of the model can be adjusted to find the best coherence between model output and the waveform. Generally it is better not to use default parameters. The model outputs the best estimation when the best match is obtained (Semmlow, 2004).

Numerous different model types are used in this approach. There is three most popular models which are *autoregressive* (AR), *moving average* (MA), and *autoregressive moving average* (ARMA). Selection of the proper model can be done by the shape information of the spectrum. Generally AR model is results better for

estimating spectra which have sharp peaks. Transfer function of AR model has a constant variable in the formula. Therefore, this model is also named as *all-pole*.

$$y(n) = - \sum_{k=1}^p a(k)y(n-k) + u(n) \quad (3.28)$$

The output is calculated by the weight function, $a(k)$ in eq. (3.28). If there is no sharp peaks it is better to use *moving average(MA)* model for evaluating spectra. The transfer function of the model is also named as *all-zero*.

$$y(n) = - \sum_{k=1}^q b(k)u(n-k) \quad (3.29)$$

Input function is $x(n)$ and q is the model order. If spectrum includes both sharp peaks and the valleys, then AR and MA can be combined to cover entire spectrum. In that case, the transfer function of an ARMA model will include both numerator and denominator polynomials, so it can be called as a pole-zero model.

$$y(n) = - \sum_{k=1}^p a(k)y(n-k) + \sum_{k=1}^q b(k)u(n-k) \quad (3.30)$$

It is required to select the model order, p or q , while selecting the type of model. Any information can be helpful while generating the process model. There are few developed schemas to assist in model selection. The common concept is that model spectrum should be fit the signal spectrum. In many cases, trial-and-error basis models are utilizing. There is many techniques for evaluating the parameters of AR model and performance of MA and ARMA are not as good as AR. MA and ARMA model based power spectral analysis could not be popular on autoregressive spectral estimation since only the AR model is placed in the MATLAB Signal Processing Toolbox (Semmlow, 2004).

AR spectral estimation techniques have two main approaches which are algorithms for processing block data and sequential data. There is more block-processing

algorithms for biomedical engineering which are placed in the MATLAB Signal Processing Toolbox. By utilizing AR spectral approach better results are obtained for short data segments. There are a numerous different approaches for estimating the AR model coefficients. The four approaches which are the most popular ones are Yule-Walker, Burg, covariance, and modified covariance methods which are placed in the MATLAB Signal Processing Toolbox (Semmlow, 2004).

The Yule-Walker method produced the spectra with the least resolution and the most smoothing among the four. Modified covariance method produces the sharpest peaks, which can be used to study with sinusoidal components (Marple, 1987). Spectra of Burg and covariance methods are similar. The Yule-Walker method and low-order model produces a smooth spectrum, especially in the low frequency range. The highest order model estimates with less accuracy on the low frequency spectral feature. The produced spectrum by the covariance method is seems similar to the one which is produced by Yule-Walker method with the same model order. A low-order AR model, smoothes the noise effectively. Higher order model improves the frequency resolution by producing sharper peaks. To sum up, by using AR model-based methods effective spectral estimations can be done, if the nature of the signal is known. On the other hand, selection of model order and model type has a major role (Semmlow, 2004).

3.4.1.1 Covariance

Mathematically, correlations calculations can be done by multiplying and normalizing. Multiplication is using for both *covariance* and *correlation* to evaluate the linear interrelation between two variables. Additionally, the coefficients are normalized between zero and one in correlation to make the correlation coefficients insensitive.

In many signal processing applications covariance is applicable since the variable scales are similar. Correlation and covariance can be applied to two or more waveforms or to multiple segments of the same waveform. Correlation and covariance operations are not only make comparisons over at range of times but also used to compare different waveforms. The *crosscorrelation function* is an example

of this process. The correlation function is produced by two waveforms which is given here in both continuous and discrete form.

$$r_{xx}(t) = \int_0^T y(t)x(t + \tau)d\tau \quad (3.31)$$

$$r_{xx} = \sum_{k=1}^M y(k + n)x(k) \quad (3.32)$$

Equations (3.31) and (3.32) show that the difference between crosscorrelation and convolution is the direction of the change. The waveforms are shifted in opposite directions for convolution. Since there is no normalization process, the outputs of Eq. (3.31) and (3.32) can be named as *raw correlation*. An unusual case of the correlation function arise when the comparison is between two different waveforms. In that case *autocorrelation function* will be needed which outputs how similarity a waveform (Semmlow, 2004).

The autocorrelation function is interrelated with the bandwidth of the waveform, if it has sharper the peaks or broader bandwidth. In case of having white noise which is infinite, adjacent points will be uncorrelated. When data lengths are finite like in signal processing, the crosscorrelation and autocorrelation will be needed (Semmlow, 2004).

The crosscovariance function is applying same approach like crosscorrelation function. Only difference is means of data is removed before calculation. For that reason, the equation has a small update on Eq. (3.32), as shown below:

$$\text{Cov}(n) = \sum_{k=1}^M [y(k + n) - \bar{y}] [x(k) - \bar{x}] \quad (3.33)$$

Correlation and covariance utilizes similar operations to Equation (3.31), (3.32) and (3.33) when used alone. For example, the covariance between two functions is given as follows (Semmlow, 2004).

$$\text{Cov} = \sigma_{x,y} = \sum_{k=1}^M [y(k) - \bar{y}] [x(k) - \bar{x}] \quad (3.34)$$

Plotting of method presented in appendix before and after denoising process.

3.4.1.2 Multivariate Covariance

Multivariate covariance can be applied to *multivariate* data where multiple responses or observations needs to be processed. An example from biosignal processing is with EEG where the signal has more than one waveforms which is observed in different positions. The covariance and correlation matrices suppose that for multivariate data, the matrix of different variables and the rows provides different *observations*. The rows are the waveform of time samples, in signal processing. Variance can be calculated by the covariance matrix by using data matrix (Semmlow, 2004).

$$S = \begin{bmatrix} \sigma_{1,1} & \sigma_{1,2} & \cdots & \sigma_{1,N} \\ \sigma_{2,1} & \sigma_{2,2} & \cdots & \sigma_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{N,1} & \sigma_{N,2} & \cdots & \sigma_{N,N} \end{bmatrix} \quad (3.35)$$

Normalized version of the correlation matrix is utilizing in an ordinary signal processing. Particularly, there is a interrelation between the correlation matrix and the covariance matrix by the following equation.

$$C(i, j) = \frac{C(i, j)}{\sqrt{C(i, j)C(j, j)}} \quad (3.36)$$

Correlation coefficients produce the correlation matrix as a set of waveform observations. Similar positional relationship in the covariance matrix can be represented as follows (Semmlow, 2004):

$$R_{xx} = \begin{bmatrix} r_{xx}(0) & r_{xx}(1) & \cdots & r_{xx}(L) \\ r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(L-1) \\ \vdots & \vdots & \ddots & \vdots \\ r_{xx}(L) & r_{xx}(L-1) & \cdots & r_{xx}(0) \end{bmatrix} \quad (3.37)$$

Plotting of method presented in appendix before and after denoising process. Difference can be observed easily by comparing other methods' plotting.

3.4.2 Non-Parametric Model Eigen Frequency Estimations

Better resolution and better frequency estimation can be obtained by utilizing Eigenanalysis spectral methods. The method has a significant performance at high noise levels if the noise is not white which does not mean. The starring feature of eigenvector approaches is division of the information in the data waveform which can be replaced with autocorrelation function. The eigen decomposition produces eigenvalues which is one of the most important variable of eigenvectors. Functions can be computed based on either signal or noise subspace in sight of eigenvalues and eigenvectors. This approach is not a true power spectral estimator. Better frequency estimators can provide spectra. Selection of the proper signal or noise subspace is the most important step of applying eigenvector spectral analysis. The signal subspace can be defined according to features of eigenvalues. But, especially with short data segments, the method is not working in practice (Marple, 1987). The definition of the signal subspace generally due to trial-and-error approach which can be done by applying an AR model (Semmlow, 2004).

Aperiodic signal frequency analysis of a continuous-time can be done by using the Fourier transform by using following formulation.

$$X_c(F) = \int_{-\infty}^{\infty} x_c(t) e^{-j2\pi Ft} dt \quad (3.38)$$

Dirichlet conditions should be satisfied with some requirements, which are

- (1) having a finite number of maxima or minima within any finite interval,
- (2) having a finite number of interruptions within any finite interval, and
- (3) be absolutely integrable.

With the usage of inverse Fourier transform, the signal $x_c(t)$ can be analyzed from its *spectrum* $X_c(F)$ as following formula

$$x_c(t) = \int_{-\infty}^{\infty} X_c(F) e^{j2\pi Ft} dF \quad (3.39)$$

With the usage of Parseval's theorem, the energy of $x_c(t)$ can be computed in either the time or frequency domain.

$$E_x = \int_{-\infty}^{\infty} |x_c(t)|^2 dt = \int_{-\infty}^{\infty} |X_c(F)|^2 dF \quad (3.40)$$

If result of the function which is $|X_c(F)|^2$ equal or bigger than zero, as a function of frequency, it can show the distribution of energy of $x_c(t)$. Therefore, it is called as *energy spectrum* of $x_c(t)$. It can be said that continuous-time aperiodic signals have *continuous spectra*.

$$\sum_{n=-\infty}^{\infty} |x(n)| < \infty \quad (3.41)$$

Any summable discrete-time signal, can be described by the discrete-time Fourier transform (DTFT).

$$X(e^{j\omega}) \triangleq F[x(n)] = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \quad (3.42)$$

$\omega = 2\pi f$ is the frequency variable in *radians per sampling interval* and f is the frequency variable in *cycles per sample*. By usage of the inverse Fourier transform, the signal $x(n)$ can be analyzed from its *spectrum* (Manolakis, Ingle, & Kogon, 2005).

$$x(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(e^{j\omega}) e^{j\omega n} d\omega \quad (3.43)$$

The function $X(e^{j\omega})$ is periodic with fundamental period 2π , if $x(n)$ is real-valued. By using Parseval's theorem, the energy of the signal can be evaluated either in time or frequency domain.

$$E_x = \sum_{n=-\infty}^{\infty} |x(n)|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |X(e^{j\omega})|^2 d\omega = \int_{-\pi}^{\pi} \frac{|X(e^{j\omega})|^2}{2\pi} d\omega \quad (3.44)$$

The function which is also called as *energy spectrum* calculates the distribution of the energy at various frequencies (Manolakis, Ingle, & Kogon, 2005).

3.4.2.1 Burg's Method

Burg's method which is for estimation of the AR parameters. Burg's method is based on minimization of the forward and backward. The Levinson-Durbin approach utilized in the minimization process contrarily to the modified covariance method. The method estimates the parameters of a prediction error filter with lattice structure which is named after its inventor John Burg. The parameters transformed as predictor parameters. Consequently, by using these parameters, the power spectrum will be able to calculate. Derivation started by recalling and calculating the forward and backward prediction errors. Combination of the relations with the Levinson-Durbin recursion described in following equation

$$a_p = \begin{bmatrix} a_{p-1} \\ 0 \end{bmatrix} - \gamma_p \begin{bmatrix} 0 \\ \tilde{a}_{p-1} \end{bmatrix} \quad (3.45)$$

The forward prediction error can be computed recursively with proper distribution of $\tilde{x}_p(n)$ which described as follows

$$\begin{aligned} e_p^+(n) &= \begin{bmatrix} a_{p-1}^T & 0 \end{bmatrix} \begin{bmatrix} \tilde{x}_{p-1}(n) \\ x(n-p) \end{bmatrix} - \gamma_p \begin{bmatrix} 0 & \tilde{a}_{p-1}^T \end{bmatrix} \begin{bmatrix} x(n) \\ \tilde{x}_{p-1}(n-1) \end{bmatrix} \\ &= e_{p-1}^+(n) - \gamma_p e_{p-1}^-(n-1) \end{aligned} \quad (3.46)$$

A recursion approach can be used for the backward prediction error, in the same way

$$\begin{aligned} e_p^-(n) &= \begin{bmatrix} a_{p-1}^T & 0 \end{bmatrix} \begin{bmatrix} x_{p-1}(n-1) \\ x(n) \end{bmatrix} - \gamma_p \begin{bmatrix} 0 & \tilde{a}_{p-1}^T \end{bmatrix} \begin{bmatrix} x(n-p) \\ x_{p-1}(n) \end{bmatrix} \\ &= e_{p-1}^-(n-1) - \gamma_p e_{p-1}^+(n). \end{aligned} \quad (3.47)$$

Both recursions are initialized by the input signal $x(n)$,

$$e_0^+(n) = e_0^-(n) = x(n). \quad (3.48)$$

An important component of the recursive equations of $e_p^+(n)$ and $e_p^-(n)$ is definition of lattice filter in one stage. The entire prediction error filter is formulated with different lattice stages. The lattice formulization for error prediction can be defined by the parameters $\gamma_1, \dots, \gamma_p$, which are commonly called as *reflection coefficients* (Sörnmo & Laguna, 2005).

Error can be minimized by using the predicted errors which are produced by lattice equations as follows.

$$\sigma_e^2 = E[|e_p^+(n)|^2] + E[|e_p^-(n)|^2] \quad (3.49)$$

This error calculation can be presented with following equation, after substitution of the lattice recursions in (3.153) and (3.154),

$$\begin{aligned} \sigma_e^2 = & (1 + \gamma_p^2) (E[|e_{p-1}^+(n)|^2] + E[|e_{p-1}^-(n-1)|^2]) \\ & - 4\gamma_p E[e_{p-1}^+(n)e_{p-1}^-(n-1)] \end{aligned} \quad (3.50)$$

The optimal p^{th} reflection coefficient can be calculated by

$$\gamma_p = \frac{2E[e_{p-1}^+(n)e_{p-1}^-(n-1)]}{E[|e_{p-1}^+(n)|^2] + E[|e_{p-1}^-(n-1)|^2]} \quad (3.51)$$

Priority in determination of parameter prediction is coefficients which will be used for computing the prediction errors to use in spectral analysis. These parameters can be easily computed by using Levinson-Durbin recursion. From a perspective of Burg's method, it is better to involve estimation and parameter definition methods (Sörnmo & Laguna, 2005). Plotting of method presented in appendix before and after denoising process. Difference can be observed easily by comparing other methods' plotting.

3.4.2.2 Multiple Signal Classification (MUSIC) method

EMG signal includes its own signal components. Raw EMG data need to be extracted to define the composed features. Signal should be preprocessed in order to feed classification algorithms to classify efficiently (Reaz, Hussain, & Mohd-Yasin, 2006). The frequency content of the signal extracted from raw EMG signal by MUSIC method.

The multiple signal classification is a frequency estimation method as a development of Pisarenko harmonic decomposition which is both first frequency estimation method and based on eigen decomposition (Pisarenko, 1973; Schmidt, 1986; Bienvenu & Kopp, 1983). The dimensional space is divided into signal and noise components using the eigenvectors of the correlation matrix (Manolakis, Ingle, & Kogon, 2005).

$$\mathbf{Q}_s = [\mathbf{q}_1 \mathbf{q}_2 \cdots \mathbf{q}_P] , \mathbf{Q}_w = [\mathbf{q}_{P+1} \cdots \mathbf{q}_M] \quad (3.52)$$

Length of the time window $M = P + 1$ that means number of complex exponentials are 1 greater than the number of complex exponentials to allow the size of the time window like $M > P + 1$. Hence, dimension of noise subspace is greater than 1. Advantage of larger dimension is procuring an improved and stronger frequency estimation method. Due to the orthogonality between the noise and signal subspaces, all time window frequency vectors of the complex exponentials are orthogonal to the noise subspace from

$$\mathbf{P}_s \mathbf{v}(f_p) = \mathbf{v}(f_p) \quad \mathbf{P}_w \mathbf{v}(f_p) = \mathbf{0}. \quad (3.53)$$

Thus, for each eigenvector ($P < m \leq M$)

$$\mathbf{v}^H(f_p) \mathbf{q}_m = \sum_{k=1}^M q_m(k) e^{-j2\pi f_p(k-1)} = 0 \quad (3.54)$$

for all the P frequencies f_p of the complex exponentials. Therefore, if we compute a pseudospectrum for each noise eigenvector as

$$\bar{R}_m(e^{j2\pi f}) = \frac{1}{|v^H(f)q_m|^2} = \frac{1}{|Q_m(e^{j2\pi f})|^2} \quad (3.55)$$

The polynomial Q_m has $M-1$ roots. These roots produce P peaks in the pseudospectrum from Equation (3.55) (Manolakis, Ingle, & Kogon, 2005).

The pseudospectra of $M - P$ noise eigenvectors shares its roots with the signal subspace. The remaining roots of the noise eigenvectors procure at different frequencies. There are no restrictions on the range of roots. If there is some near the group circle, can produce extra peaks in the pseudospectrum.

$$\bar{R}_{music}(e^{j2\pi f}) = \frac{1}{\sum_{m=p+1}^M |v^H(f)q_m|^2} = \frac{1}{\sum_{m=p+1}^M |Q_m(e^{j2\pi f})|^2} \quad (3.56)$$

The P complex exponentials' frequency estimation taken as the P peaks in pseudospectrum. Since in the Equation (3.56) does not include information about the powers of the complex exponentials or the background noise level, pseudospectrum term is used (Manolakis, Ingle, & Kogon, 2005).

The presumption in the MUSIC pseudospectrum is the noise eigenvalues equal to $\lambda_m = \sigma^2 w$, which means the noise is white. But in practice the noise eigenvalues will not be equal when prediction method used instead of correlation matrix. The variations become more obvious when the correlation matrix is predicted by small samples of data. Therefore, a little variation on the MUSIC algorithm, known as the *eigenvector (ev) method*, was intended to define potentially different noise eigenvalues (Johnson & DeGraaf, 1982). The pseudospectrum is

$$\bar{R}_{ev}(e^{j\omega}) = \frac{1}{\sum_{m=p+1}^M \frac{1}{\lambda_m} |v^H(f)q_m|^2} = \frac{1}{\sum_{m=p+1}^M \frac{1}{\lambda_m} |Q_m(e^{j2\pi f})|^2} \quad (3.57)$$

where λ_m is the eigenvalue corresponding to the eigenvector $\mathbf{q}_m$, for the method. The pseudospectrum is normalized by its corresponding eigenvalue for each eigenvector. The eigenvector and MUSIC methods are identical in case of equal noise eigenvalues ($\lambda_m = \sigma^2 w$) for $P + 1 \leq m \leq M$ (Manolakis, Ingle, & Kogon, 2005).

The peaks in the MUSIC pseudospectrum corresponds approaches to zero in equation (7). As a matter of fact, we may prefer z-transform instead of this denominator which is sum of the z-transforms and formulated as (Manolakis, Ingle, & Kogon, 2005).

$$\bar{P}_{music}(z) = \sum_{m=p+1}^M Q_m(z) Q_m^* \left(\frac{1}{z^*} \right) \quad (3.58)$$

Plotting of method presented in appendix before and after denoising process. Difference can be observed easily by comparing other methods' plotting.

3.5 Time-Frequency Analysis

3.5.1 Short-Term Fourier Transform:The Spectrogram

The first time–frequency methods were based on slicing waveform approach. Standard Fourier transform applied for segment analysis which has similar functionality like windowing function. Fourier transform is applied to a segment of data which is isolated from the entire waveform. Name of the method is *spectrogram* or “short-term Fourier transform” (STFT) since the Fourier Transform is applied and is shorter than the entire waveform. Selection of the most proper window length is important since shorten data segments are used. This approach has been widely used in a numerous biomedical applications (Semmlow, 2004). In continuous domain, the basic equation for the spectrogram is:

$$X(t,f) = \int_{-\infty}^{\infty} x(\tau) w(t - \tau) e^{-j\pi f \tau} d\tau \quad (3.59)$$

$w(t-\tau)$ is the window function and τ is the variable for cutting the waveform. The discrete version of Eq. (3.59) is similar to general probing function which is applied with sinusoids in complex format (i.e., $e^{-jnm/N}$):

$$X(m,k) = \sum_{n=1}^N x(n) [W(n - k) e^{-jnm/N}] \quad (3.60)$$

There are two main problems with the spectrogram which are selection of optimal window length for data segments and measurement of shortening the data length, N . Shortening the data segment may cause the loss of low frequencies. For this reason, when the window is small enough to improve the time resolution, the frequency resolution is decreased. The time–frequency tradeoff has been formulized as follows where B is bandwidth and T is time.

$$BT \geq 1/4\pi \quad (3.61)$$

The STFT has been applied widely to numerous problems. Frequency resolution is not a problem for high frequency components. Especially speech processing field has applied STFT application regarding coherence between method and speech waveform. The STFT is a simple and classical theory since it is based on the Fourier transform (Semmlow, 2004).

3.5.2 Wigner-Ville Distribution : Special Case Of Cohen Class

Different approaches have been developed to deal with spectrogram problems. *Wigner-Ville distribution* which is one of the widely applied time–frequency method. Once, the method was developed by Wigner to use in physics formulations, then applied to signal processing problems by Ville, for this reason the method is dual named. Wigner-Ville approach is an extraordinary case of transformations which known as subtopic of *Cohen's class* approach.

For calculation of the power spectrum, the Wigner-Ville distribution and Cohen's class, use an approach for the early use of the autocorrelation function. The waveform is compared for all possible shifts or lags to build the autocorrelation function. The equation is similarly repeated here for both continuous and discrete approach:

$$r_{xx}(\tau) = \int_{-\infty}^{\infty} x(t) x(t + \tau) dt \quad (3.62)$$

and

$$r_{xx}(n) = \int_{k=1}^M x(k) x(k + n) \quad (3.63)$$

In the standard autocorrelation function, time is summed out of the result. The Wigner-Ville and other Cohen's class distributions use a variation of autocorrelation function with time remains result. This is obtained by comparison of the waveform. The comparison is done for all possible values of time instead of integrating over time. This comparison supports the defining of equation.

$$R_{xx}(t, \tau) = x(t + \tau/2)x^*(t - \tau/2) \quad (3.64)$$

$$R_{xx}(n, k) = x(k + n)x^*(k - n) \quad (3.65)$$

These equations can be applied to either real or analytic signal which is more complex version. There are also advantages of using the analytic signal methods which are autocorrelation function keeps both lags and time and consequently a two-dimensional function. Output of a sinusoid ordinary autocorrelation function would be a sinusoid at the same frequency. These *cross products* are result from the multiplication in the instantaneous autocorrelation equation, Eq. (3.65) which are source of problems for all instantaneous autocorrelation based methods. The Wigner-Ville approach applies the Fourier transform of the instantaneous autocorrelation function.

The approach outputs a function of both frequency and time. It is expected to filter the autocorrelation function before application of fourier transform to improve features of the resulting power spectrum when one dimensional power spectrum is computed. All other approaches apply a filter to instantaneous autocorrelation function before applying the Fourier transform while there is no filtering application for Wigner-Ville distribution.

The general equation for determining a time–frequency distribution from Cohen's class is:

$$\rho(t, f) = \iiint g(v, \tau) e^{j2\pi v(u - \tau)} x(u + 1/2\tau)x^*(u - 1/2\tau) e^{-j2\pi f\tau} dv du d\tau \quad (3.66)$$

$g(v,\tau)$ makes applicable the two-dimensional filtering of the instantaneous autocorrelation which is also known as *kernel* and rest of the equation is the Fourier transform of the instantaneous autocorrelation function (Semmlow, 2004).

Main advantage of Wigner-Ville over the STFT is better marginals and conditional moments and greatest strength that produce “a remarkably good picture of the time-frequency structure” (Cohen, 1992). The marginals are related with the summation over time or frequency of to the signal energy. For instance, if the Wigner-Ville distribution is summed over frequency at a fixed time, the equal energy at that point in time will be obtained, which is also same for frequency. The conditional cases of the Wigner-Ville distribution also has significant formulation where $p(t)$ is the marginal in time.

$$f_{\text{inst}} = 1/p(t) \int_{-\infty}^{\infty} f p(f,t) df \quad (3.67)$$

The Wigner-Ville distribution has a number of other properties which can be important for some applications. It is possible to reconstruct the original signal from the distribution, except constant one. The transformation is stationary while shifting in time and frequency. To exemplify, shifting the signal in time with delay of T seconds, the same distribution will be shifted T times on the time axis. Accordingly, the same could be said for frequency shifting similarly to STFT. Finite support in time and frequency is a feature of Wigner-Ville distribution which is not applicable for STFT. In time, finite support means that the distribution is zero before the signal starts and after it ends, which means the distribution does not contain frequencies, in frequency. The Wigner-Ville includes some nonexistent energy which comes from cross products in the time and frequency boundaries of the original signal. For the reason, the Wigner-Ville distribution is not zero when the signal is zero, which is called *strong finite support* (Semmlow, 2004).

The Wigner-Ville distribution also poor noise properties. The noise is producing all time and frequency including noise of cross products. In some cases, the cross products and noise level can be reduced by windowing. The desired window function can be applied to lag dimension of the instantaneous autocorrelation

function similar to application of the time function. Windowing will reduce frequency resolution as in Fourier transform analysis. In practice, a cooperation is required between cross products and loss of frequency resolution. (Semmlow, 2004)

3.6 Wavelet Analysis

A transform is a kind of remapping of a signal which presents more information. For example, the Fourier transform provides new approaching over the frequency information. Although, Fourier transform upgraded in terms of both time and frequency characteristics of the waveform with different approaches, none of them is not able to completely solve the time–frequency problem. The wavelet transform can be used to investigate the properties of a waveform which can change over time by dividing segments of scale instead of sections of time.

The waveform was compared to a sine function in the Fourier transform. This approach updated by multiplication of the waveform with the sinusoidal functions as follows

$$X(\omega_m) = \int_{-\infty}^{\infty} x(t)e^{-j\omega_m t} dt \quad (3.68)$$

Eq. (3.68) is used to define the discrete Fourier transform. Sinusoidal functions are mostly used one because of their frequency characteristics. They have the energy at only one specific frequency. Functions can be used to evaluate some specific behavior or characteristic of the waveform. It would be proper to translate probing function over the waveform $x(t)$, as described in short-term Fourier transform (STFT) in following equation (Semmlow, 2004).

$$\text{STFT}(t,f) = \int_{-\infty}^{\infty} x(\tau)(w(t - \tau)e^{-2j\pi f\tau})d\tau \quad (3.69)$$

F is the frequency and $w(t - \tau)$ is sliding window function where t translating the window over x . More common translated probing function is:

$$X(t,m) = \int_{-\infty}^{\infty} x(\tau)f(t - \tau)_m d\tau \quad (3.70)$$

If family functions $f(t)m$ is large enough, then it should be able to represent all properties of the waveform $x(t)$ for reconstruction. The family of base functions is so large which includes unnecessary set of descriptions. Capabilities of The Fourier transform is not enough, STFT and other distributions which applies two dimensions, should be employed (Semmlow, 2004).

3.6.1 The Continuous Wavelet Transform

In wavelet analysis, numerous different probing functions may be used. The family of functions includes upgraded or compressed versions of the basic function. This approach provides equation definitions for the continuous wavelet transform (CWT):

$$W(a,b) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{|a|}} \Psi^*((t-b)/a) dt \quad (3.71)$$

The probing function Ψ could be any number of different functions. The operation points of combination and the normalizing factor $1/\sqrt{a}$ shapes the energy is the same for all values of a . If $b=0$, and $a=1$ and the wavelet is in its natural form, this is calling as the *mother* wavelet. One of the popular wavelet is *Morlet* wavelet and defined by the following equation:

$$\Psi(t) = e^{-t^2} \cos(\pi \sqrt{2/\ln 2} t) \quad (3.72)$$

$W(a,b)$ which is the wavelet coefficients, describe the correlation between the waveform and the wavelet at various translations and scales. The coefficients presents the amplitudes of a series of wavelets, over a range of scales and translations. In addition, coefficients will be employed while reconstructing the original signal. Thus, wavelet analysis can be thought as an investigation over the waveform. Wavelet coefficients effects the changes in the waveform, mostly type of changes which are on the same scale as the wavelet. It is easier to analyze or recognize patterns using CWT. If the wavelet function $x(t)$ is properly chosen, then it is possible to reconstruct the original waveform from the wavelet coefficients. A double summation or integration is need to be applied to reconstruct the original

signal since the CWT decomposes the waveform into a and b coefficients, which is formulized as follows (Semmlow, 2004):

$$x(t) = 1/C \int_{a=-\infty}^{\infty} \int_{b=-\infty}^{\infty} W(a, b) \Psi_{a,b} da db \quad (3.73)$$

where:

$$C = \int_{b=-\infty}^{\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega$$

and $0 < C < \infty$ (the so-called *admissibility condition*) for recovery using Eq.(3.73).

Reconstruction of the original waveform is rarely performed using the CWT coefficients. Discrete wavelet transform should be employed when reconstruction of the original waveform is required due to its appropriate structure (Semmlow, 2004).

3.6.2 Wavelet Time-Frequency Characteristics

Wavelets do not exist at a specific time or a specific frequency. In fact, wavelets provide an interconnection between time and frequency localization. They are well localized and defined in both time and frequency. Δt_{ψ} which is measure of a specific wavelet in a time range, can be specified by the square root of a given wavelet about its central time (Akansu & Haddad, 1992):

$$\Delta t_{\psi} = \sqrt{\frac{\int_{-\infty}^{\infty} (t-t_0)^2 |\Psi(t/a)|^2 dt}{\int_{-\infty}^{\infty} |\Psi(t/a)|^2 dt}} \quad (3.74)$$

where t_0 is the central time, or first moment of the wavelet, and is given by:

$$t_0 = \frac{\int_{-\infty}^{\infty} t |\Psi(t/a)|^2 dt}{\int_{-\infty}^{\infty} |\Psi(t/a)|^2 dt} \quad (3.75)$$

Similarly the frequency range, $\Delta \omega_{\psi}$, is given by:

$$\Delta\omega t_\psi = \sqrt{\frac{\int_{-\infty}^{\infty} (\omega - \omega_0)^2 |\Psi(\omega)|^2 d\omega}{\int_{-\infty}^{\infty} |\Psi(\omega)|^2 d\omega}} \quad (3.76)$$

$\Psi(\omega)$ is the frequency domain of $\psi(t/a)$, and ω_0 is the center frequency of $\Psi(\omega)$. The center frequency is given by an equation similar to Eq. (3.75):

$$\omega_0 = \frac{\int_{-\infty}^{\infty} \omega |\Psi(t/a)|^2 d\omega}{\int_{-\infty}^{\infty} |\Psi(t/a)|^2 d\omega} \quad (3.77)$$

The time and frequency ranges can be obtained from the mother wavelet using Eqs. (3.74) and (3.76). The variable a changes the time range by multiplying Δt_ψ . Accordingly, the time range of $\psi(a,0)$ is defined as $\Delta t_\psi(a) = |a|\Delta t_\psi$. Another relationship between time and frequency can be obtained by applying *Mexican hat wavelet* which is given by following equation:

$$\psi(t) = (1 - 2t^2)e^{-t^2} \quad (3.78)$$

The frequency range can be the change of mother Wavelet divided by a : $\Delta\omega_\psi(a) = \Delta\omega_\psi / |a|$. If we multiply the frequency range by the time range, constant will be produced by following equation:

$$\Delta\omega_\psi(a)\Delta t_\psi(a) = \Delta\omega_\psi\Delta t_\psi = \text{constant}_\psi \quad (3.79)$$

Equation (3.79) shows that the ranges are invariant and inversely related. When the frequency range increasing, the time range decreases. These ranges interrelated with the time and frequency resolution of the CWT. There is a time–frequency trade-off between behind these relations. For example, decreasing the wavelet time range provides more accurate features of time characteristics in terms of frequency resolution. (Semmlow, 2004)

The CWT will present better frequency resolution when a is large and the length of the wavelet is long. Conversely, the wavelet is short and the time resolution is maximum when a is small. Since a is just a variable, there is an inner trade-off

between time and frequency resolution which is key and core to the CWT and makes it well suited for signal analyzing (Semmlow, 2004).

3.6.3 Discrete Wavelet Transform

In applications better transformations are required. For this reason, transforms, which produces the minimum number of coefficients is preferred to reconstruct the original signal accurately. The *discrete wavelet transform* (DWT) achieves manner by limiting the variation in translation and scaling. The DWT may won't need to perform a transform if the wavelet is chosen carefully (Semmlow, 2004).

Briefly explanation of a signal transform can be rearranging of a signal that provides more information and feature than the original one. To exemplify, the Fourier transform supports the description quite strongly. The wavelet transform which is one of transformation methods can be used as another way to describe and extract features of a waveform which changes over time. To handle this situation the waveform is divided into segments of scale instead sections of time (Semmlow, 2004). In wavelet transforms, the capability of the transform to a meaningful information extraction over time and frequency characteristics of the signal, depends strongly on the selection of the mother wavelet $\psi(t)$ which is certain function and generally like following equation (Begg, Lai, & Palaniswami, 2008).

$$\Psi(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \quad (3.80)$$

s used as a scale and u as a translation parameter. Because of flexibilities of these parameters, functions with various frequencies and midpoint localities can be produced in time, which are called as *baby wavelets* or *wavelet atoms*. Handling of correlations at different frequencies are gained by a form of the signal $x(t)$ with following wavelets as

$$W_x(u,s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi^*\left(\frac{t-u}{s}\right) dt \quad (3.81)$$

This formula is a WT of $x(t)$ at scale s and position u , the coefficient $1/\sqrt{s}$ is for normalization to keep the energy in the transform (Begg, Lai, & Palaniswami, 2008). The equation $Wx(u, s)$ is also known as the WT coefficient. Largeness of coefficients effects correlations in a positive way. By using the position parameter u , we can also define the strongest or weakest correlation position in the signal. WT coefficients also can be used to define the frequency in the signal, thus localization information of the signal $x(t)$ in time and frequency can be detected. The transform (Equation 3.81) is known as the continuous wavelet transform (CWT) because the signal $x(t)$ is assumed to be continuous on the time domain t . In the case of discrete or finite signals the discrete wavelet transform (DWT) is used with some modifications as follows

$$(s_j, u_k) = (2^j, k 2^{-j} : j, k \in \mathbb{Z}) \quad (3.82)$$

In wavelet transforms, the capability of the transform to a meaningful information extraction over time and frequency characteristics of the signal, depends strongly on the selection of the mother wavelet $\psi(t)$. Wavelet transforms widely have used by biosignals using DWT in most biomedical applications. It is worthy to note that using discrete wavelet bases improves accuracy of border problems (Begg, Lai, & Palaniswami, 2008).

DWT utilizes two different sets of functions which are scaling and wavelet functions. These functions are interrelated to low-pass and high-pass filters. Each step combines two filters and two down-samplers. These high-pass and low-pass filters observe the detail, D1 and approximation, A1 for the first one, by using the down-sampled outputs. There are set of processes for each approximation, from A1 to A6, which are decomposition, processing and reconstruction, orderly. Details are given in (Akay, 1997; Subasi A. , 2007; Subasi A. , 2012).

Entropy of the EMG signal can be obtained clearly by using extracted discrete wavelet coefficients in time and frequency with some mathematical techniques. Reducing dimensionality of statistics and extracted feature vectors is one of the mathematical techniques which is used in numerous studies (Kandaswamy , Kumar ,

Ramanathan , Jayaraman , & Malmurugan , 2004; Subasi A. , 2007; Subasi A. , 2012). EMG signals represented by using following statistical features:

- (1) Mean of the coefficients for each sub-band.
- (2) Average power of the wavelet coefficients in each sub-band.
- (3) Standard deviation of the coefficients in each sub-band.
- (4) Ratio of mean values of neighbor sub-bands.

1st and 2nd features represent the frequency distribution of the signal. 3rd and 4th features represents the changes in frequency distribution. Seven different values for features (1), (2) and (3); six different values for feature (4), hence total 27 feature was extracted. Selection of features and schema of DWT depicted in Figure3.9. These feature vectors, calculated for the frequency bands A6 and D1–D6 and were used as input to classifiers as depicted in figure 3.9. (Qi, Hui, Yongji, & Jian, 2009). The “db4” wavelet filter gave higher classification accuracy and lower computation cost than the others in (Qi, Hui, Yongji, & Jian, 2009). As a matter of fact that we used “db4” wavelet filter in this study.

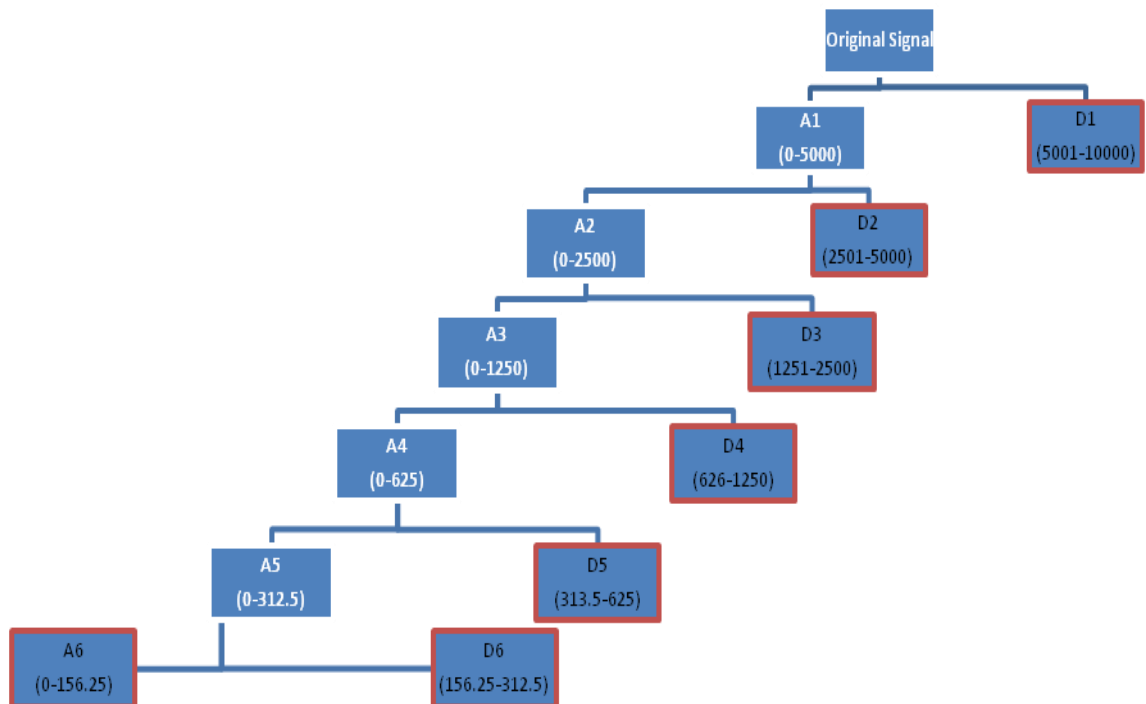


FIGURE 3.2 Schema of Selection of features in DWT approach

All the extracted features of DWT method is plotted to observe differences between features. Each feature behaving dissimilarly which means each of them contains

different information, as it can be seen from following figure. These unlikely extracted features are improving quality of features, consequently total classification accuracy. Plotting of method presented in appendix before and after denoising process. Difference can be observed easily by comparing other methods' plotting in appendix.

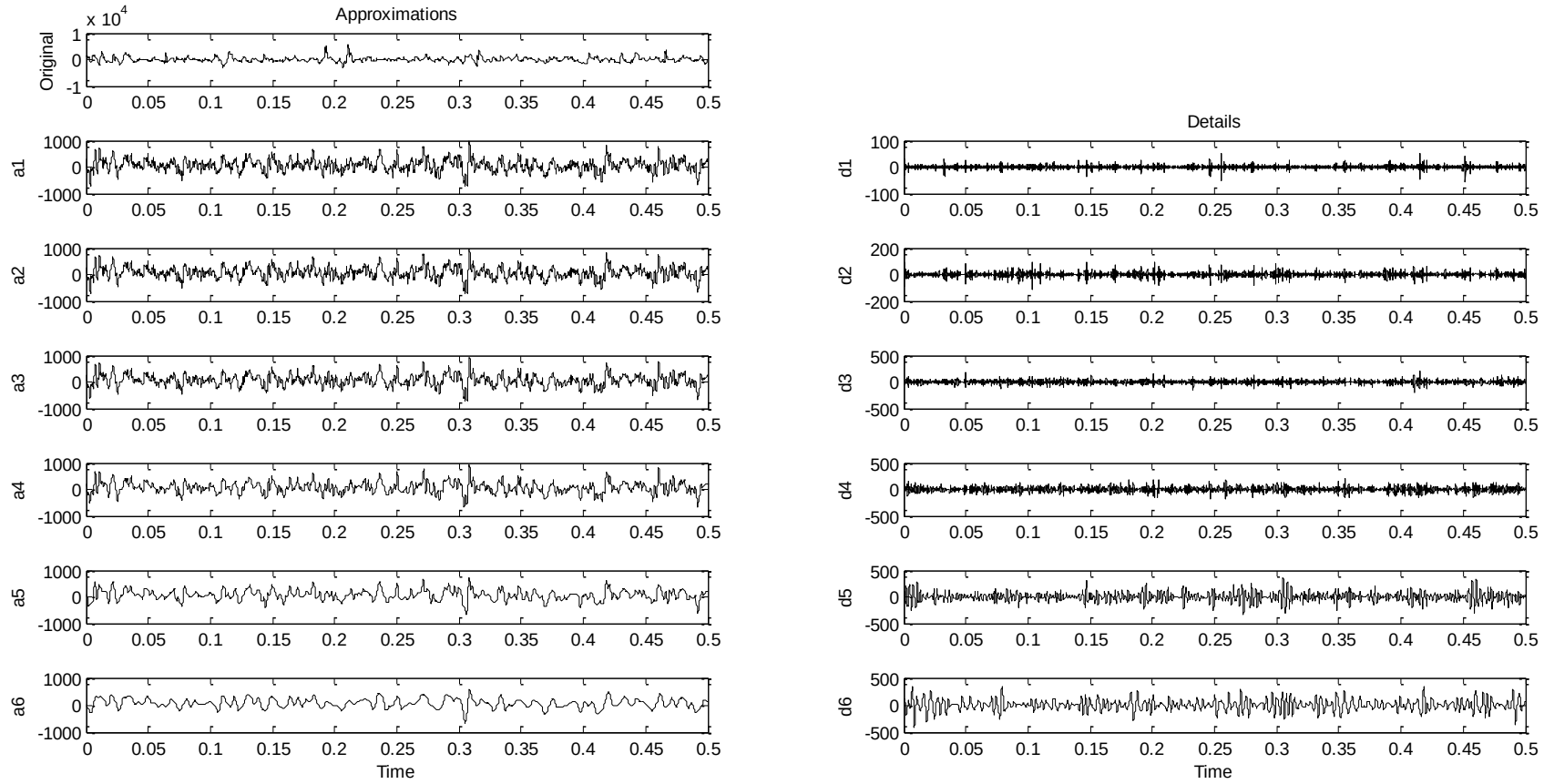


FIGURE 3.3 Plotting of each features after DWT feature extraction and MSPCA denoising for control signal

3.7 Multiscale Principal Component Analysis

Multiscale Principal Component Analysis (MSPCA) combines the features and capability of Principal Components Analysis (PCA) to decorrelate the variables by obtaining a linear relationship. MSPCA decorrelate autocorrelated measurements by using wavelet analysis to get significant features. MSPCA evaluates the PCA of the wavelet coefficients at each scale and merge the results at related scales. Due to its multiscale manner, MSPCA is convenient for modeling of data for signals whose behavior changes over time and frequency (Bakshi, 1998).

Signal processing by PCA is extensively used as a classical multiscale signal processing tool (Chawla, 2009; Hotelling, 1933; Joliffe, 2002). PCA has been applied in different fields of science and engineering to better utilize its ability (Bakshi, 1998; Bakshi, 1999; Misra, Yue, Qin, & Ling, 2002). For biomedical signals like EMG, a robust extension of classical PCA by analyzing shorter signal segments is suggested (Kotas, 2006). It may be used in data reduction, beat detection, classification, signal separation and feature extraction (Castells, Laguna, Sornmo, Bollmann, & Roig, 2007). The desired quality of processed signals is achieved by selecting the principal components (PC) based on energy features in selected wavelet subset matrices. The number of PC selection is based on cumulative percentage of total variation of variances to control the quality of denoised signals. PCA is performed on these matrices for signal denoising. PCA denoising follows the idea that retaining only the principal components with highest variance to reconstruct the decomposed signal. The choice of multiscale matrices and selection of eigenvalues preserve the desired energy in the processed signals (Bakshi, 1998; Bakshi, 1999; Misra, Yue, Qin, & Ling, 2002). PCA transforms the data matrix statistically diagonalizing method by the covariance matrix. The process is obtaining the correlation between variables in the data. If the measured and evaluated variables are related, the first few entities present the relationship between the variables. The MSPCA methodology decomposes each variable on a selected type of wavelets. The PCA model is determined one by one for the coefficients at each scale (Bakshi, 1998).

PCA transforms $n \times p$ data matrix as $\mathbf{X}$. The variables calculated as a linear weighted sum obtained as,

$$\mathbf{X} = \mathbf{T}\mathbf{P}^T \quad (3.84)$$

where, $\mathbf{P}$ represents the principal component loadings, $\mathbf{T}$ is the principal component scores, n and p are represents number of dimensions and variables. The principal component loadings represent direction and position of the hyperplane which evaluate the highest probability of residual variance referred measured variables. Multiscale PCA (MSPCA) combines the capability of PCA to extract the crosscorrelation or relationship between the variables with orthonormal wavelets. The measurements for each variable (column) are decomposed to its wavelet coefficients using the same orthonormal wavelet for each variable in order to combine the benefits of PCA and wavelets. This results in transformation of the data matrix, $\mathbf{X}$ into a matrix, $\mathbf{WX}$, where $\mathbf{W}$ is an $n \times n$ orthonormal matrix representing the orthonormal wavelet transformation operator containing the filter coefficients. The matrix, $\mathbf{WX}$ is as big as the original data matrix size, $\mathbf{X}$, but related to the wavelet decomposition, the deterministic component for each variable in $\mathbf{X}$ is compressed in a relatively small number of coefficients in $\mathbf{WX}$, when the stochastic supplement in each variable is approximately decorrelated in $\mathbf{WX}$, and propagate over all components according to its power spectrum. To make use the multiscale components of the data, PCA of the covariance matrix coefficients at every scale is calculated separately from the other scales. The outcoming scores at every scale are not cross correlated because of PCA, and their autocorrelation is roughly decorrelated because of the wavelet decomposition. Being dependent on the character of the application, a lesser subset of the principal component scores and wavelet coefficients can be chosen at every scale. The amount of principal components, to be absorbed at every scale is not altered by the wavelet decomposition because it does not influence the core relationship connecting the variables at any scale. As a result, existing techniques like, cross validation, screen test, or parallel analysis may be used for the data matrix in the time domain or to every wavelet coefficient to choose the appropriate amount of components. For choosing the relevant wavelet coefficients, distinct methods may be applied according to the application. Using split thresholds at every scale permits MSPCA to

be more responsive to scale-varying signal characteristics like autocorrelated measurements. Limiting of the coefficients at every scale points the area of the time-frequency space. Details are given in (Bakshi, 1998). Different feature extraction methods are applied before and after MSPCA denoising. It is observed that capability of methods are increased after denoising when compared to before denoising. Plotting of all feature extraction methods presented in appendix which are clearly showing great performance of MSPCA denoising. In addition effect of MSPCA denoising is compared in experimental results section with figures and graphs.

CHAPTER 4

MACHINE LEARNING TECHNIQUES

4.1 K - Nearest Neighbor (k-NN)

K-nearest neighbor (k-NN) is a popular algorithm due to its speed, because there is almost no learning process. Theoretical background of k-NN is pretty simple. There are K training inputs with a specified volume (Yang, 2010).

We first point an input by $\mathbf{x}$ and the number of inputs in one class as N_k , where $k \in \{1, 2\}$. We use two simple probabilities to quantify (Yang, 2010). The simple probability equations are described below

$$p(\mathbf{x} | k) = \frac{N_k}{N} \quad (4.1)$$

where $N = N_1 + N_2$. The background probability is evaluated by the following equation

$$p(k|\mathbf{x}) = \frac{p(\mathbf{x}|k)p(k)}{p(\mathbf{x}|1)p(1) + p(\mathbf{x}|2)p(2)} \quad (4.2)$$

where $p(k)$ is the a prior probability of the k^{th} class. If $p(1) = p(2)$ we have

$$p(k | \mathbf{x}) = \frac{N_k}{N} \quad (4.3)$$

The query vector is tagged as class 1 when $N_1 > N_2$. If the prior information is updated to $p(1) > p(2)$ then we have

$$p(k | \mathbf{x}) = \frac{p(k)Nk}{p(1)N_1 + p(2)N_2} \quad (4.4)$$

There is also another way to find same result. The probability of the k^{th} class is defined for given a volume V

$$p(\mathbf{x} | k) = \frac{Nk}{N} \quad (4.5)$$

The background probability of the k^{th} class is described as

$$p(k | \mathbf{x}) = \frac{\frac{N_k}{V} p(k)}{\frac{N_1}{V} p(1) + \frac{N_2}{V} p(2)} \quad (4.6)$$

It seems like Bayes rule which is the base for deriving a k-nearest neighbor classification. k-NN has a model selection problem like other machine learning algorithms. k number must be carefully selected based on model specifically. Distance measure formulas should be selected because different formulas may cause different outputs such as Manhattan and Euclidean distance formulas (Yang, 2010).

4.2 Artificial neural networks (ANN)

Artificial neural networks (ANNs) are computing systems with many simple, highly interconnected processing elements which are called nodes or artificial neurons to abstractly simulate the organization, structure and operation of the biological nervous system. A biological neuron collects input signals from near neurons via the dendrites, conjugates the signals, and produces an output response then forward to other neurons via the axon. In addition that, a neuron makes a connection to other neurons by axons with the dendrite tree of a neighboring neuron. The neurons are characterized as mathematical function with numerical inputs and a single output (Rosenblatt, 1958). It is obvious that learning can be modeled mathematically like classification, if the first perceptron has a number of inputs. Learning in ANNs is accomplished through special training algorithms developed based on learning rules presumed to imitate the learning mechanisms of biological systems. There are many architectures and types of neural networks especially for learning functions in signal

processing field. The details are well documented in (Subasi & Erelebi, 2005; Haykin, 1999).

The perceptron and the biological neuron are comparable in the following ways (Begg, Lai, & Palaniswami, 2008). The input of the perceptron has different features, thus becomes similar to real dendrite branches of the biological neuron. The synapse connection w_j is modeled as weights which is very important for summation and activation function. The processing of the weighted inputs to calculate the output of the perceptron is done by summation node using following mathematical function which acts like nucleus of the biological neuron:

$$c_i = f(\sum_j w_j x_{ij}) = f(u_i) \quad (4.7)$$

The activation function f effects neuron's behavior and it is generally has a limited increasing and nonlinear function like sigmoid function:

$$f(u_i) = \frac{1}{e^{-u_i}} \quad (4.8)$$

The output c_i is result of calculation which is spreading along the axon. In the perceptron, learning quality and rate is equivalent to optimum adjustments of weights to compare the real network output and the expected network output for a particular training input. The perceptron learning algorithm is not able to classify when the data is not separable. But, perceptron could produce much more separating limitations if connected like a network that has links in each other (Minsky & Papert, 1988). This neural network design is known as the multilayer perceptron (MLP) which is one of the most common neural network designs. MLP design has a single input layer attached to output layer via a hidden layer. In the network, every node, in each layer, is linked to every node in the next layer. An N -layer MLP network generally has an input layer, $N - 1$ hidden layers and an output layer. The complexity of the separation boundary depends on the number of hidden layers and nodes per layer. The activation functions create separator boundaries instead of linear function (Begg, Lai, & Palaniswami, 2008).

4.2.1 The Perceptron: An Artificial Neuron

An artificial neuron can be modeled as the system. An artificial neuron usually denoted as perceptron, which is the simplest and basic unit of networks. An artificial neuron or perceptron is gathering inputs to penetrate from inputs to outputs.

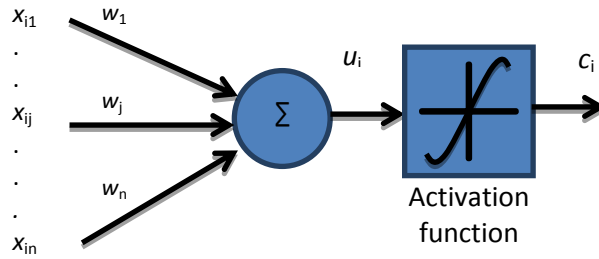


FIGURE 4.1 An artificial neuron or perceptron

The perceptron and the biological neuron are comparable in the following ways (Begg, Lai, & Palaniswami, Computational Intelligence in Biomedical Engineering, 2008):

i. Inputs. The x_{ij} is inputs to the neuron. The input of the perceptron has different features, thus becomes similar to real dendrite branches of the biological neuron. The synapse connection w_j is modeled as weights which is very important for summation and activation function.

ii. Summation and activation function. The processing of the weighted inputs to calculate the output of the perceptron is done by summation node using following mathematical function which acts like nucleus of the biological neuron:

$$c_i = f\left(\sum_j w_j x_{ij}\right) = f(u_i)$$

(4.9)

The activation function f effects neuron's behavior and it is generally has a limited increasing and nonlinear function like sigmoid function:

$$f(u_i) = \frac{1}{1 + e^{-u_i}} \quad (4.10)$$

iii. Output. The output c_i is result of calculation which is spreading along the axon.

iv. Feedback. Several ANN types receive the output back as input to make the neural network adaptive.

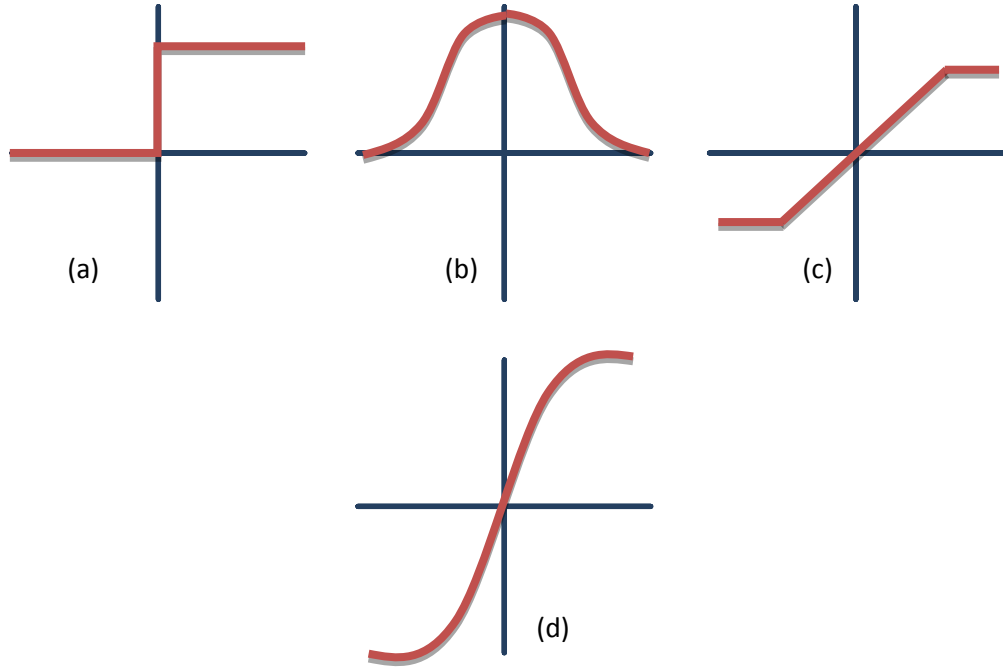


FIGURE 4.2 Activation functions used in artificial neurons: (a) Step function (b) Gaussian (c) piecewise linear and (d) sigmoid.

In the perceptrons, learning quality and rate is equivalent to optimum adjustments of weights to compare the real network output and the expected network output for a particular training input. The perceptron needs to learn the conversion from training inputs x_i to expected outputs y_i . The error between real output c_i and the expected output y_i is formulized as,

$$E_i = y_i - c_i \quad (4.11)$$

For one perceptron, all errors of all training patterns are formulized as:

$$E_{SN} = \sum_i E_i = \sum_i y_i - c_i \quad (4.12)$$

The sum of errors is,

$$E_{SN}^{sq} = \sum_i (y_i - c_i)^2 \quad (4.13)$$

Minimization of summed errors is necessary for learning of perceptron. This process is empirical risk minimization (ERM) to represent error by risks (Vapnik V. , 2000).

The learning process of perceptron can be shown as:

Feed the neurons with weights $0.2 < w_j < 2.1$

the training pattern x_i evaluate the real output c_i ,

weights updated by using

$$w_j^{t+1} = w_j^t - \lambda(y_i - c_i)x_{ij} \quad (4.14)$$

λ represents the learning rate. The value is usually between 0.05 and 0.75.

Low values assigned to weights in order to prevent large initial error. Thus we can make learning time shorter. Bigger λ generally means shorter and faster training times (Rosenblatt, 1958).

4.2.2 Multilayer Perceptrons

The perceptron learning algorithm is not able to classify when the data is not separable. But, perceptrons could produce much more separating limitations if connected like a network that has links in each other's (Minsky & Papert, 1988). This neural network design is known as the multilayer perceptron (MLP) which is one of the most common neural network design (Begg, Lai, & Palaniswami, Computational Intelligence in Biomedical Engineering, 2008).

MLP design has a single input layer attached to output layer via a hidden layer. In the network, every node, in each layer, is linked to every node in the next layer. An N-layer MLP network generally has an input layer, N - 1 hidden layers and an output layer.

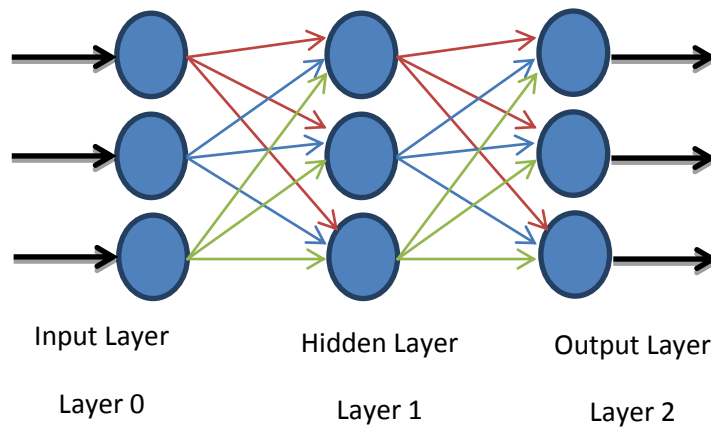


FIGURE 4.3 A double-layer MLP network.

The complexity of the separation boundary depends on the number of hidden layers and nodes per layer. The activation functions create separator boundaries instead of linear function. (Begg, Lai, & Palaniswami, 2008).

In this study, neural network application (i.e., classification of EMG data) employed for designing classifiers, namely the MLPNN. This type of networks is a statistical satisfactory approximator for linear problems. Additional hidden processing layers are created to handle the problem's complexity, capacity and nonlinearity. Maybe just one hidden layer can be sufficient to map an arbitrary function to any level of accuracy. Hence, MLP architecture of ANNs was adopted for the present study.

4.2.3 The Back-Propagation Algorithm

Back-propagation algorithm is invented while perceptron learning algorithm was fitting to MLP. Back-propagation algorithm can be applied to any feedforward neural network. The process contains two steps:

Calculation of the derivatives of the network training error to feed the weights. To adjust the weights gradient descent technique should be applied in order to decrease the output error.

To exemplify, w_{ij} represent the weight from node j to node i . If there is no connection from node j , then $w_{ij} = 0$. Since only forward connections are allowed in

feedforward network, $w_{ji} = 0$. First of all, the outputs are evaluated by training pattern x_k as input at each node c_{ki} . Then, total of squared errors E_{MLP}^{sq} , which represents total of all individual squared errors of each training pattern E_k , is calculated as:

$$E_k = \frac{1}{2} \sum_i (y_{ki} - c_{ki})^2 \quad (4.15)$$

$$E_{MLP}^{sq} = \sum_k E_k = \frac{1}{2} \sum_k \sum_i (y_{ki} - c_{ki})^2 \quad (4.16)$$

where k sorts the training patterns out and i indexes the output and y_{ki} is expected output for k -th training pattern. The coefficient $1/2$ is for sufficiency of evaluating the derivative part of this error in order to use for weights, which bring about the equation below:

$$\frac{\partial E_{MLP}^{sq}}{\partial w_{ij}} = \sum_k \frac{\partial E_k}{\partial w_{ij}} \quad (4.17)$$

After applying derivative chain rule for each training pattern, next equation becomes:

$$\frac{\partial E_k}{\partial w_{ij}} = \frac{\partial E_k}{\partial c_i} \frac{\partial c_i}{\partial u_i} \frac{\partial u_i}{\partial w_{ij}} = \delta_i \frac{\partial u_i}{\partial w_{ij}} \quad (4.18)$$

For output nodes we have

$$\delta_i = -(y_i - c_i) f_i' \quad (4.19)$$

where f_i' represents the derivative part of the activation function. The sigmoid activation function has a derivative like $f_i' = c_i(1 - c_i)$, which is also match up with $\tanh$ activation function $f_i' = 1 - c_i^2$.

The last derivative on the right side of next equation can be found as follows:

$$\frac{\partial u_i}{\partial w_{ij}} = \frac{\partial \left(\sum_i w_{ij} c_j \right)}{\partial w_{ij}} = c_j \quad (4.20)$$

For a hidden node or layer j which connects directly to the output node i, we have

$$\delta_j = \frac{\partial E_k}{\partial u_j} = \sum_i \frac{\partial E_k}{\partial u_i} \frac{\partial u_i}{\partial u_j} = \sum_i \delta_i \frac{\partial u_i}{\partial u_j} \quad (4.21)$$

If we update u_i in terms of findings about u_j , we have

$$\delta_j = \sum_i \delta_i \frac{\partial \left(\sum_j f(u_j) w_{ij} \right)}{\partial u_j} = f'_j \sum_i w_{ij} \delta_i \quad (4.22)$$

where δ_j represents the processed sum of δ_i , which must be evaluated first. The procedure starts at the output layer and propagates backward to the hidden, that is why the method is named as backpropagation.

With calculated derivatives, the weights are updated using a gradient descent technique:

$$w_{ij}^{t+1} = w_{ij}^t - \lambda \frac{\partial E_{MLP}^{sq}}{\partial w_{ij}} \quad (4.23)$$

The backpropagation algorithm is updating weights in two ways such as the batch and online learning types. The batch type is evaluating the total error of derivatives by summing the error derivatives for all training patterns. The process run for entire training set. MLP operating needs many processes in the batch type to being saturatedly trained, which is the drawback of this type (Begg, Lai, & Palaniswami, 2008).

The online learning type is a growing type for the weights which are updated referring to single training pattern. The batch type updates the weights, while online

learning type updates K times which are number of samples in training data set (Begg, Lai, & Palaniswami, 2008).

4.3 Radial Basis Function (RBF)

A radial basis function is a real-time function which uses distance from origin. RBFs a powerful tool for working with untidy data. They are used extensively for interpolating scattered data (Meinguet, 1979). Radial basis function approaches are usually finite linear combinations of translated radial symmetric basis functions by applying Euclidean norm (Buhmann, 2003). It is based on distance between two given points.

In the case function of multilayer perceptron, X_{n+1} will be equal to 1 which is called as bias. The bias weights vector can simply be interrelated with weights of matrix. After such process, it can be simplified to use in RBF networks. The hidden layer activation functions representing that this particular network represents a multilayer perceptron. The structures of multilayer perceptrons, RBF network and fuzzy logic models are same or very similar. The basic difference between sigmoidal and radial basis function networks is how processed each particular neuron (Kecman, 2001).

The basic computation is very simple for neural networks. After an input vector is inserted to network, the input spreads to all hidden layer neurons as scalar values either using dot or inner products of weights and vectors. A radial basis is calculated by two sets of parameters: the center $\mathbf{c}$, which defines the position and shape parameters that determines the shape. In case of a multivariate input vector, the parameters defines the shape of the hyper-Gaussian function of a covariance matrix. The covariance matrix contains the shape parameters. The output from each neuron depends on the type of used activation function. The most common activation functions in multilayer perceptrons are the squashing sigmoidal functions. All nonlinear modeling tools have the same structure. Thus, their representational capabilities are the same or very similar after successful training. All three models learn from a set of training data and try to be as good as possible in modeling typically sparse and noisy data pairs in high-dimensional space. Accuracy should be measured while trying the best model in terms of goodness, performance, or quality.

Each has different error, risk or cost function. This is where the main difference between the models. The application of different functions also leads to different learning and optimization procedures (Kecman, 2001).

Management of classifier examined using two criteria in some studies. The first criterion is application with improved interpretation of classifier output. The second one is reducing the classification error of the unknown ventricular and supraventricular arrhythmic beats. Ventricular beats was maintained close to 0% as average and which is about 35% for supraventricular beats. This approach improved the usefulness of the NN applications with the advantage of minimizing uncertainty and reducing global error (Begg, Lai, & Palaniswami, 2008).

4.4 Support vector machines (SVMs)

Vapnik established Support vector machines (SVMs) in 1992 as a Supervised Learning Method (Vapnik V. , 2006). SVMs are set of supervised machine learning methods which is using in computer science for classification. Although basic SVM is type of tree algorithm, SVMs are capable to classify via different kernel approaches. The main advantage of SVM is having the best generalization capability on statistical learning theory field (Dobrowolski, Wierzbowski, & Tomczykiewicz, 2011). Base of SVM is a linear separator by a hyperplane with a given margin sizes. The hyperplane divides into two classes which are linearly dividable groups of points in a multidimensional space.

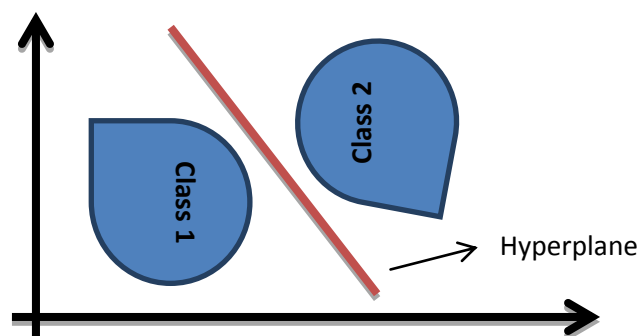


FIGURE 4.4 Two linearly separable classes by an infinite number of hyperplanes.

In theory this situation is possible but in daily life both classes may not linearly separable since they mixed up. This case confuses the classifier which is using linear hyperplane. Due to fact that, there is a requirement to apply a nonlinear transformation from input space to feature space (Begg, Lai, & Palaniswami, 2008). The hyperplane shows evaluated sum of all training samples x_i . Algorithm searching for the next hyperplane to make possible the separation linearly in feature space,

$$f(x) = \sum_{i=1}^n w_i \phi(x_i) + b \geq 1 \quad \text{if } y_i = +1 \quad (4.24)$$

$$f(x) = \sum_{i=1}^n w_i \phi(x_i) + b \leq -1 \quad \text{if } y_i = -1 \quad (4.25)$$

Equation (4.24) and (4.25) can be used together for a strict linear discriminant function

$$y_i f(x) \geq 1 \quad (4.26)$$

The closest points to the hyperplane have an important role since they detect the size of class margins and boundaries.

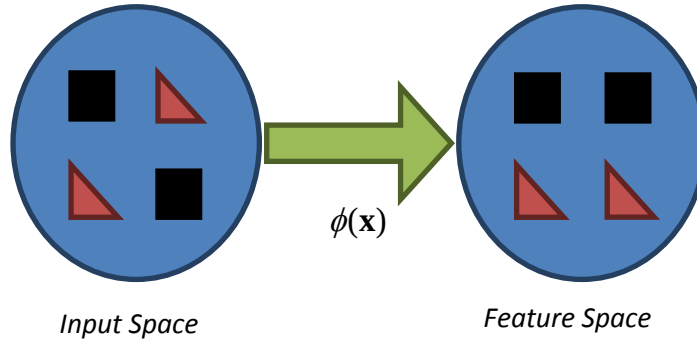


FIGURE 4.5 The nonlinear mapping makes inseparable points in input space separable in feature space.

If linear separation not applied to both two classes in feature space, strict discriminant function cannot be weakened which is represented in Equation (19). To make weaker, the soft-margin SVM classifier is defined as (Cristianini & Shawe-Taylor, 2000):

$$\tilde{S}(X, \tilde{S}) = \frac{1}{2} \sum_{i=1}^n w_i^2 + C \sum_{i=1}^n S_i \quad (4.27)$$

subject to

$$y_i f(x) \geq 1 - S_i \quad S_i > 0 \quad (4.28)$$

The parameter ‘C’ is an adjuster for the relationship between the training errors and generalization capability (Begg, Lai, & Palaniswami, 2008).

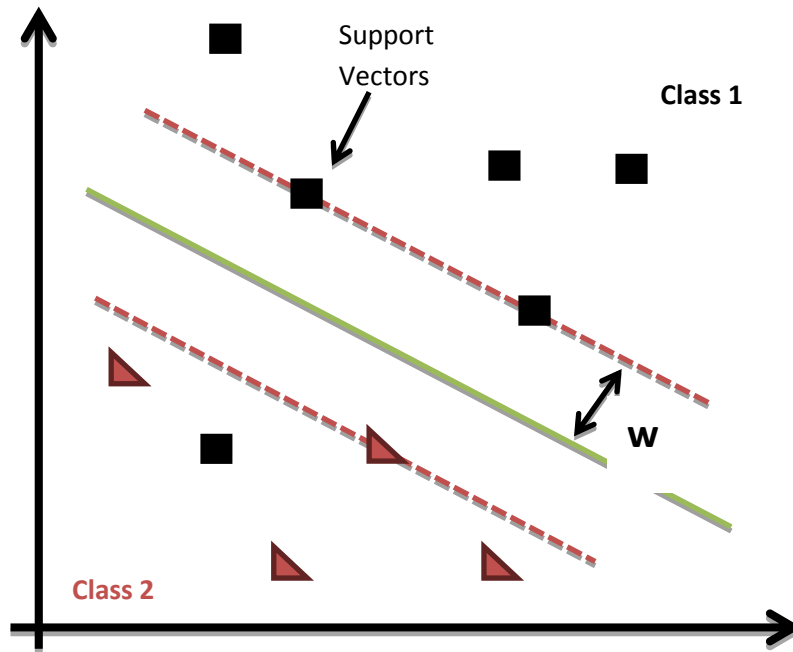


FIGURE 4.6 The soft-margin support vector classifier.

4.5 Classification And Regression Trees (CART)

The Classification and Regression Tree (CART) is a Decision Tree algorithm developed by (Breiman, Friedman, Ohlsen, & Stone, 1984). CART uses a recursive partitioning technique which has splitting criteria to create nodes. Tree is building by using created and split nodes referring splitting criteria and function. Before applying the split criteria we need to know what the best split point is. The quality of the splitting criteria is measured by a function which is obtained by processing variance function. Generated function is applying to each split point to calculate best point for splitting (Shien-Ping & Chih-Chou, 2009). Some criteria can use for defining the splits such as Gini which is commonly used, towing and ordered towing (Breiman, Friedman, Ohlsen, & Stone, 1984).

$$\text{Gini}(t) = 1 - \sum_i P_i^2 \quad (4.29)$$

Gini's minimum value is zero when all the data belongs to one of categories. Reversely, if all data belongs to other category in the target variable Gini reaches the maximum value (Soman, Diwakar, & Ajay, 2006). Advantages of this method are applicable to numerical data and categorical data either ordinal or non-ordinal structure and dealing with missing variables. In addition, it is relatively automatic "machine learning" method (Lewis, 2000).

CART Method has four process steps. First one is building the tree by using recursive splitting of nodes where reached to splitting criteria. Second one is stopping tree building process after fitted and formed according learning dataset. Third one is pruning of tree, through cutting off of important nodes to produce simpler trees. The last one is selection of optimal tree from the sequence of pruned trees, which are not over fitting with the information when fitting with the learning dataset (Lewis, 2000).

4.6 C4.5

C4.5 is a popular decision tree algorithm which is an upgraded version of ID3 decision tree algorithm that is developed also by Quinlan, too (Quinlan R. J., 1993). C4.5 has additional features such as handling missing values, categorization of continuous attributes, pruning of decision trees, rule derivation. For pruning and splitting, it applies heuristic approach which is based on the statistical information to increase quality of splitting process (Ali & Wasimi, 2007).

C4.5 algorithm starts by selection of attribute to test root of tree. Each attribute is calculated statistically to check feasibility by comparing training examples. The most proper attribute is used to test root node. Branch nodes of root node are created after examining each potential attribute values in respect to training examples. This process is repeated for each branch node to select the most proper attribute by testing referring training examples. Essence of C 4.5 algorithm is choosing correct attribute according to information gain to test nodes (Mitchell, 1997; Quinlan J. R., 1996). Steps of C4.5 decision tree algorithm are.

- (1) The root node consider that all training examples and selected attributes are most proper ones.
- (2) The information of training examples passed to branch nodes which keeps in leaf nodes to give decisions.
- (3) Rules are produced by using the path from the root node to leaf node (Ali & Wasimi, 2007; Mazid, Shawkat Ali, & Stickle, 2010).

4.7 Random Forest

Machine learning techniques supported improvement of modern biology for sizable and complex biological data analysis (Breiman L. , 2001). In the area of Bioinformatics and Biomedical Engineering, the Random Forest (RF) classifier, which is set of decision trees became a popular option in the machine learning processes. Increased use of random forest can be seen in computational biology field, owing to its advantages in dealing with limited sample size, complex data structures and multidimensional feature space (Qi Y. , 2012). The reasons that we used random forest are possible lower error rates for large data sets and usage of random inputs and features to have better results in classification.

Random forest uses numerous independent decision trees which are created by randomly chosen variables. The independent trees are built by an algorithm. After trees are created they vote to find out the most popular class (Breiman L. , 2001). The algorithm guarantees that all trees in forest are different. Randomness is applied in two steps. First step is using different bootstrap sample data to build each tree. Second step is choosing subset of predictors randomly then splitting each node of trees with the best subset instead of all predictors (Yi & Pan, 2010). There are two reasons to have bootstrap step. First one is accuracy increment when random features are used and the second one is reducing generalization error (Breiman L. , 2001). Random selection of splitting does better than bagging in terms of generalization error (Dietterich, 2000). Total accuracy of a random forest depends on the strength of the individual tree classifiers. Sometimes, this algorithm works very well compared to other classifiers such as support vector machines, neural networks and discriminant analysis (Liaw & Wiener, 2002). Even though random forest was presented for decision trees, can work with other classifiers.

CHAPTER 5

METHODOLOGY

5.1 Research Design

In this study 2 types of neuromuscular disorders, namely Myopathy and ALS, and normal EMG signals by using 6 different feature extraction methods (AR, COV, MCOV, EIGEN, MUSIC, DWT) 7 classification methods (k-NN, ANN, RBF, SVM, CART, C4.5, Random Forest) processed as independent variables to classify into 3 dependent variables which are Myopathy, ALS and Control. Results are discussed and interpreted in terms of series of experiments.

5.2 Data Collection

The EMG signals were recorded and analyzed under usual conditions for MUAP analysis. The recordings were obtained at low (just above threshold) voluntary and constant level of contraction. Visual and audio feedback was used to monitor the signal quality. A standard concentric needle electrode was used. The EMG signals were recorded from five places in the muscle at three levels of insertion (deep, medium, low). As boundaries, the high and low pass filters of the EMG amplifier were set at 2 Hz and 10 kHz. The recorded signal data consisted of a normal control group, a group of patients with myopathy and a group of patients with ALS. The control group consisted of 10 normal subjects aged 21-37 years, 4 females and 6 males. 6 out of 10 were in very good physical shape, and the remaining except one was in general good shape. None in the control group had signs or history of neuromuscular disorders. The myopathy group includes 7 patients; 2 females and 5 males aged 19-63 years. All 7 had clinical and electrophysiological signs of myopathy¹⁵.

The ALS group consisted of 8 patients; 4 females and 4 males between 35-67 years old. In the meantime, clinical and electrophysiological signs compatible with ALS, 5 of 8 patients died in a couple years after beginning of the disorder, which supporting the diagnosis of ALS. In addition, no one in the control group had signs or history of neuromuscular disorders. The brachial biceps and medial vastus muscles were used in this study since they were the most frequently investigated in the two patient groups (Nikolic M. , 2001; Gokgoz & Subasi, 2014).

5.3 Procedure

EMG signal need to be denoised since it can be distorted by environmental and neighbor constitutions during the collection and recording process. Because, higher classification accuracies can be obtained by using clearer input. Denoised signal cause easier extraction of features which is next step of the design. Higher classification rates can be reachable by extracting meaningful and efficient features from signal. We applied 6 feature extraction method to investigate which one is more appropriate and why. Lastly, the denoised and feature extracted data given to classifiers as input. Then performance of classifiers are compared by using cross validation accuracy, F-measure and area under (ROC) curve performance evaluation methods to prove reliability of the design and approach.

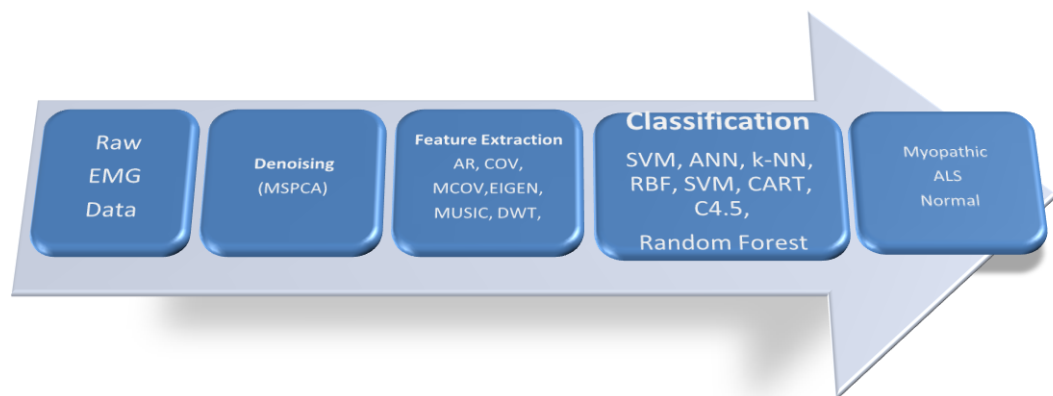


FIGURE 5.1 Blok diagram of proposed system.

5.4 Data Analysis Methods

5.4.1 Signal Denoising Method

Multi-scale principal component analysis (MSPCA) is used for detection and diagnosis. MSPCA extracts both PCA approach and wavelet approach at the same time. The individual signals are decomposed into approximations by using wavelets and details at each scales. Findings from each scale are collected in different matrices and a PCA model is constructed to extract correlation at each scale. The multi-scale nature of MSPCA formulation makes it suitable to work with process data at different scales (Misra, Yue, Qin, & Ling, 2002).

5.4.2 Feature Extraction Methods

In this study numerous feature extraction methods are utilized such as discrete wavelet transform (DWT), Eigen Value Decomposition, Covariance (COV), Multivariate Covariance (MCOV), ARBurg (AR), Multiple Signal Classification (MUSIC) by decomposing EMG signal. It is worth to note that, significantly different results obtained by using different feature extraction methods. For example, 40.55% total classification accuracy obtained by employing COV feature extraction method which is 73% with DWT. According to results coherence between methods can be obtained. It is obvious that each feature extraction techniques have different characteristics.

5.4.3 Classification Methods

All extracted features are given as input to each machine learning techniques to with the coherence between methods before and after denoising process. Machine Learning Techniques such as k-Nearest Neighbor (k-NN), Artificial neural network (ANN), Radial Basis Function (RBF), Support vector machine (SVM), Regression Trees (CART), C4.5 and Random Forest classifiers are utilized in this study, since they are popular ones in biomedical signal processing field. Different techniques applied which has different structures and capabilities.

CHAPTER 6

RESULTS

6.1 Introduction

The decomposition of EMG signals into integrated MUAPs and their classification into groups of similar characteristics is a common machine learning problem (Subasi, Yilmaz, & Ozcalik, 2006). Complexity of the problem is growing since MUAPs waveforms have unknown and dissimilar shapes. Main goals of analyzing signals are to get better accuracy in shorter time with flexible and useful methods whose reliabilities are accepted. In this study, EMG data are collected from 25 patients. The control group consisted of 10 normal subjects aged 21-37 years, 4 females and 6 males. 6 out of 10 were in very good physical shape, and the remaining, except one, were in general good shape. In control group none of them have any signs or history of neuromuscular disorders. The group with myopathy consisted of 7 patients; 2 females and 5 males aged 19-63 years. All 7 had clinical and electrophysiological signs of myopathy¹⁵. The ALS group consisted of 8 patients; 4 females and 4 males aged 35-67 years (Nikolic M. , 2001).

6.2 Performance Evaluation

Some of combinations resulted very high accuracy on EMG signal data. In order to analyze the coherence between method combinations of our approach, the whole EMG data is separated as training and test data since sets must be selected independently from each other. The classification model built via training set and verified by the test set is used for validation.

K-fold cross-validation method is a well-known and reliable method for predicting the error rate of a classification technique. We use it as many researchers do to decrease bias which is related with random sampling of the produced data sets (Ambroise & McLachlan, 2002; Sandvig, Mobasher, & Burke, 2008; Subasi A. , 2012). K-fold cross validation randomly divide the data into k subsets that are calling as folds which have almost same size. The cross-validation accuracy (CVA) is the average of the k individual accuracy measures which is formulated as follows

$$CVA = \frac{1}{k} \sum_{j=1}^k A_i \quad (6.30)$$

where k is the number of used folds, and A_i is the accuracy measure of each fold (Zhang, McAllister, Scotney, McClean, & Houston, 2006). The training data is used for building of the classification model and the test data is used for validation.

The number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) are used to evaluate performance of classifiers. Different definitions are used for different domains. The *sensitivity* and *specificity* are extensively used in diagnostic and detection tests. Sensitivity shows the amount of people with disease, who have a positive test result,

$$Sensitivity = \frac{TP}{TP+FN} \times 100\% \quad (6.31)$$

Specificity shows the amount of people without disease who have a negative test result, which is $1 - FP$ and defined as:

$$Specificity = \frac{TN}{TN+FP} \times 100\% \quad (6.32)$$

Accuracy is used as a measure, which is:

$$Accuracy = \frac{TP+TN}{TP+FN+TN+FP} \times 100\% \quad (6.33)$$

F-measure is another measurement which is used to evaluate performance and formulated as follows (Subasi A. , 2013).

$$F - measure = \frac{2 TP}{2 TP + FN + FP} \times 100\% \quad (6.34)$$

In addition to F-measure, performances of classifiers are evaluated with ROC curve too. A ROC curve can be observed by drawing of all sensitivity namely true positive values, on the y axis, contrary to equivalent values which is 1-specificity namely false-positive thresholds. The reliability of the approach depends on the number of thresholds tested. The area under the ROC curve (AUC) is generally thought as key performance since it provides a single quantity of overall accuracy with following equation (Witten & Frank, 2005).

$$M = (FP/FN) \times ((1-P)/P) \quad (6.35)$$

Accuracy is calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (6.36)$$

We calculated performances of different machine learning methods. Classification accuracy is the number of correctly classified instances divided by the full number of instances.

6.3 Experimental Results

Features of raw EMG signals extracted with AR, EIGEN, MUSIC, DWT, COV and MCOV decomposition methods. Performances and results compared using different classification algorithms such as CART, C4.5, ANN, k-NN, SVM, RBF and Random Forest before and after denosing with MSPCA method. Matlab R2011b is used for biomedical signal processing with signal processing toolbox. The performances of classifiers are tested in Weka, regarding its reliability and ease of use.

F-measure is another performance measurement which supports our results. As shown from Table 6.7 best total accuracy obtained by Random Forest. ALS result of all classifiers are higher than total results which shows that ALS is more applicable for feature extraction and denoising.

The performances of classifiers are also measured via receiver operating characteristic (ROC) curve which is a graphical technique. It is produced by plotting true positives as percentage of all positives in the sample versus negative ones using cross-validation. For each fold of cross-validation it counts the true and false positives in the test set and plot the results to ROC axes (Witten & Frank, 2005). Then, the classification performance can be measured by the mean area, which is the area under the ROC curve (AUC). The mean of area which is under curve as an average gives result using drawn and calculated axis points to evaluate how reliable the estimated results. Correspondingly, the quality of the estimation of a curve depends on the number tested thresholds (Handley & McNeil, 1982; Obuchowski N. A., 2003; Swets, 1979; Deleo, 1993; Devos, i dr., 2004; Zweig & Campbell, 1993). The AUC is generally taken as the index of performance because it generates one accuracy output which not depends on any defined threshold (Hanley & McNeil, 1982; Obuchowski N. A., 2003; Swets, 1979).

According to fact, Random Forest is better than the others for EMG patterns with MSPCA denoising and DWT feature extraction method. Advantage of using DWT and Random Forest methods is having high dimensional feature space which is proved by results. Numerous studies prove that random forest and the other tree algorithms pretty reliable and accurate since they are widely using as classification algorithm (Liaw & Wiener, 2002; Díaz-Uriarte & Alvarez de Andrés, 2006; Svetnik, Liaw, Tong, Culbertson, Sheridan, & Feuston, 2003). It can be said that, in some cases Random forest better than SVMs, referring to Liaw and Wiener (Liaw & Wiener, 2002). Consequently, we can say that EMG data is more coherent to classify with Random Forest classifier either not denoised or denoised. Further figures and Tables can be found in appendix.

6.3.1 Results without MPSCA denoising

Raw EMG signals are affected from environmental noise and signals while passing through other tissues which seriously distort the EMG signal during the collection and recording process. Performances of classifiers measure before denoising process to compare results after denoising. Appendix A and B can be seen for further graphical and table representations of results based on feature extraction methods.

TABLE 6.1 F-Measure results of classifiers for each feature extraction method without MSPCA denoising

Feature Extraction Method	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
AR	0.672	0.695	0.676	0.476	0.623	0.34	0.39
COV	0.646	0.679	0.721	0.482	0.613	0.363	0.27
MCOV	0.658	0.682	0.685	0.408	0.599	0.367	0.34
EIGEN	0.649	0.631	0.676	0.506	0.568	0.453	0.39
MUSIC	0.592	0.532	0.577	0.541	0.506	0.19	0.4
DWT	0.688	0.699	0.731	0.733	0.66	0.733	0.62

TABLE 6.2 ROC (AUC) results of classifiers for each feature extraction method without MSPCA denoising

Feature Extraction Method	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
AR	0.822	0.785	0.831	0.661	0.716	0.56	0.61
COV	0.8	0.781	0.865	0.667	0.708	0.568	0.55
MCOV	0.798	0.772	0.84	0.636	0.699	0.277	0.55
EIGEN	0.782	0.725	0.835	0.705	0.674	0.484	0.61
MUSIC	0.729	0.653	0.757	0.701	0.629	0.523	0.6
DWT	0.817	0.773	0.886	0.887	0.744	0.831	0.78

TABLE 6.3 Total Classification Accuracy (%) results of classifiers for each feature extraction method without MSPCA denoising

Feature Extraction Method	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
AR	67.2	69.67	67.89	48.44	62.22	40	42
COV	65	68	72.33	48.55	61.22	40.55	38.1
MCOV	65.78	68.33	68.67	42.11	59.89	36.66	37
EIGEN	64.89	63	67.67	51.89	56.66	48.44	43.9
MUSIC	59.11	53.3	58	53.89	50.55	52.33	44.1
DWT	68.67	69.89	73.33	73.22	65.89	73	61.9

6.3.2 Results with MSPCA Denoising

The improvement on total classification accuracy for all three classifiers is significant with applying MSPCA Denoising method. To exemplify, MSPCA + Random Forest perform much better than just Random Forest. The accuracy for Random Forest without MSPCA is 73.33%, with improved 96.67%. The 23.34% improvement proves the great performance and advantage of MSPCA denoising for

EMG data. Appendix A and B can be seen for further graphical and table representations of results based on feature extraction methods.

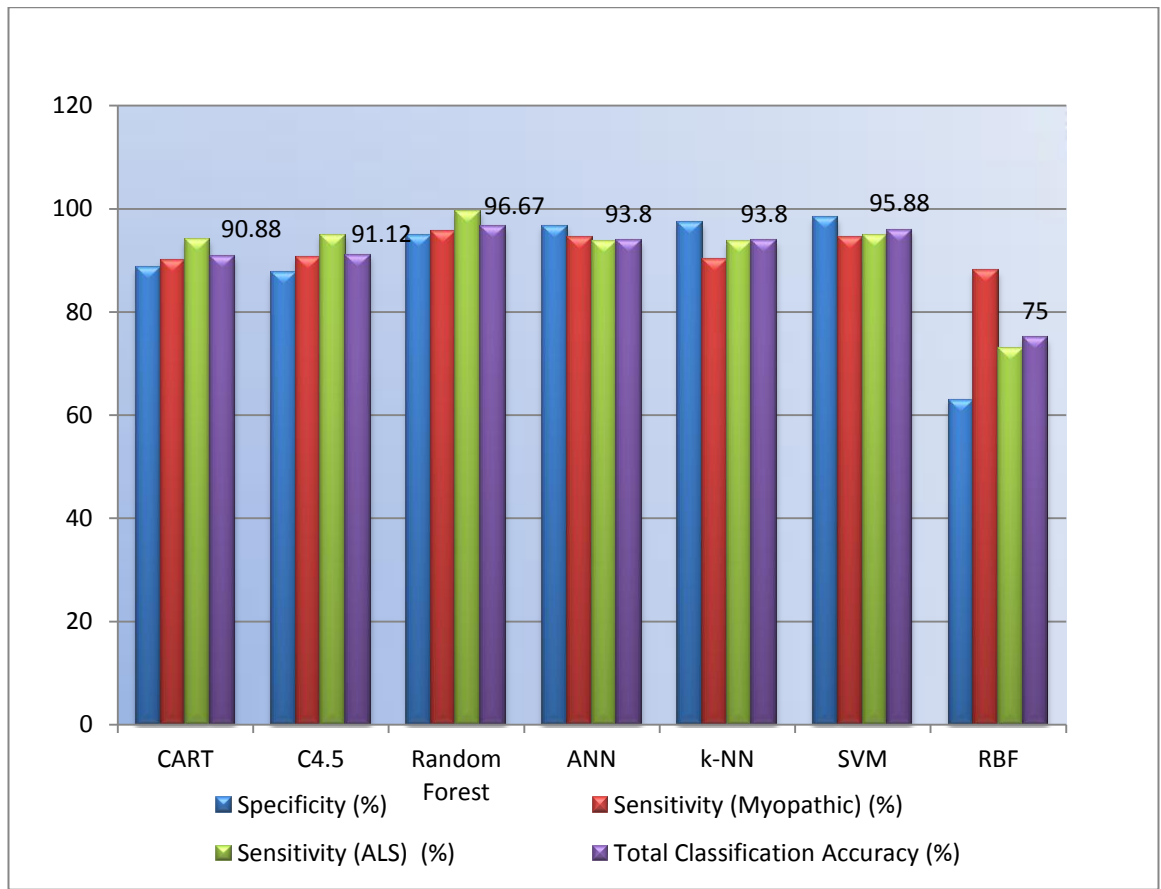


FIGURE 6.1 Graphical representation of Classifiers' performances after DWT feature extraction.

TABLE 6.4 F-Measure results of classifiers for each feature extraction method with MSPCA denoising

Feature Extraction Method	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
AR	0.728	0.739	0.799	0.363	0.748	0.30	0.27
COV	0.784	0.787	0.855	0.379	0.806	0.259	0.33
MCOV	0.76	0.795	0.842	0.353	0.811	0.368	0.31
EIGEN	0.784	0.781	0.849	0.408	0.835	0.294	0.29
MUSIC	0.816	0.821	0.884	0.888	0.813	0.841	0.55
DWT	0.909	0.896	0.94	0.947	0.938	0.924	0.75

TABLE 6.5 ROC (AUC) results of classifiers for each feature extraction method with MSPCA denoising

Feature Extraction Method	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
AR	0.813	0.831	0.923	0.587	0.81	0.377	0.51
COV	0.877	0.853	0.954	0.588	0.854	0.287	0.58
MCOV	0.833	0.858	0.949	0.557	0.858	0.257	0.54
EIGEN	0.857	0.854	0.946	0.615	0.877	0.362	0.57
MUSIC	0.875	0.859	0.969	0.963	0.859	0.841	0.71
DWT	0.94	0.932	0.987	0.986	0.953	0.949	0.9

TABLE 6.6 Total Classification Accuracy (%) results of classifiers for each feature extraction method with MSPCA denoising

Feature Extraction Method	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
AR	72.89	74	80	38.77	74.77	37.33	35.1
COV	78.33	78.55	85.55	39	80.55	35.89	40.9
MCOV	76	79.44	84.33	35.66	81	36.67	38.1
EIGEN	78.55	78.2	85.11	41.56	83.55	36.22	38.4
MUSIC	81.56	82.1	88.44	90	81.22	91	54.9
DWT	90.89	89.55	96.67	94.9	93.78	95.88	74.6

6.3.3 Effect of denosing with MSPCA

Same feature extraction and classification methods applied to same data before and after MSPCA denoising, to observe how effective denoising EMG data with MSPCA method. For example, the total accuracy for SMV with MSPCA denoising is 92.55%, conversely, 57% without MSPCA. This proves the great advantage of applying MSPCA denoising. To exemplify, effect of MSPCA denoising on classification accuracy for SVM is 35.55%, for ANN is 39.87%, for k-NN is 31.66%. Results for other classifiers are presented in Table 6.7 and following figures.

TABLE 6.7 Total accuracies of all classifiers with each feature extraction method before and after MSPCA denoising.

	AR		EIGEN		COV		MCOV		MUSIC		DWT	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
Random Forest	67.89	81	67.67	88	72.33	87	68.67	86	58	91.5	73.33	96.67
SVM	40	60	48.44	65	40.55	63.5	36.66	56	52.33	91	73	95.88
ANN	48.44	44	51.89	41	48.55	39	42.11	36	53.89	90	74.22	94.9
k-NN	62.22	75	56.66	88	61.22	82	59.89	81	50.55	82	65.89	93.8
C4.5	69.67	74.2	63	78	68	80	68.33	79	53.3	83	69.89	91.11
CART	67.2	73	64.89	79	65	79	65.78	77	59.11	81	68.67	90.88
RBF	42	35	43.9	38	38.1	38	37	38	44.1	73	61.9	89

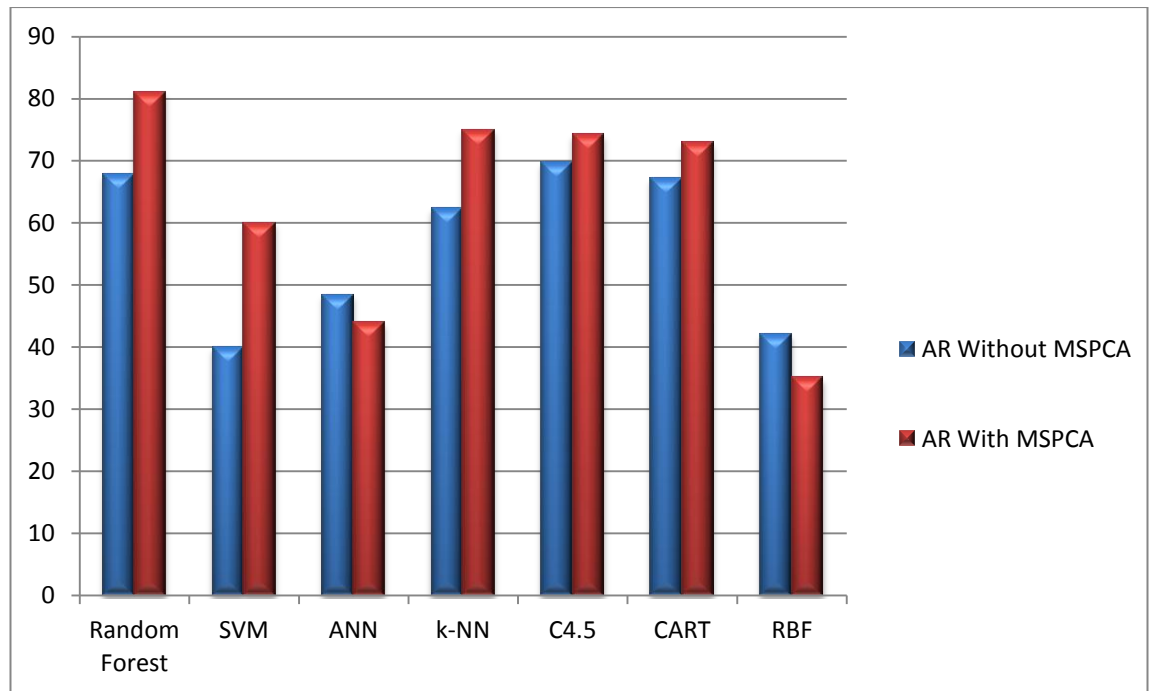


FIGURE 6.2 Graphical representation of results of both approaches after ARBURG Feature extraction

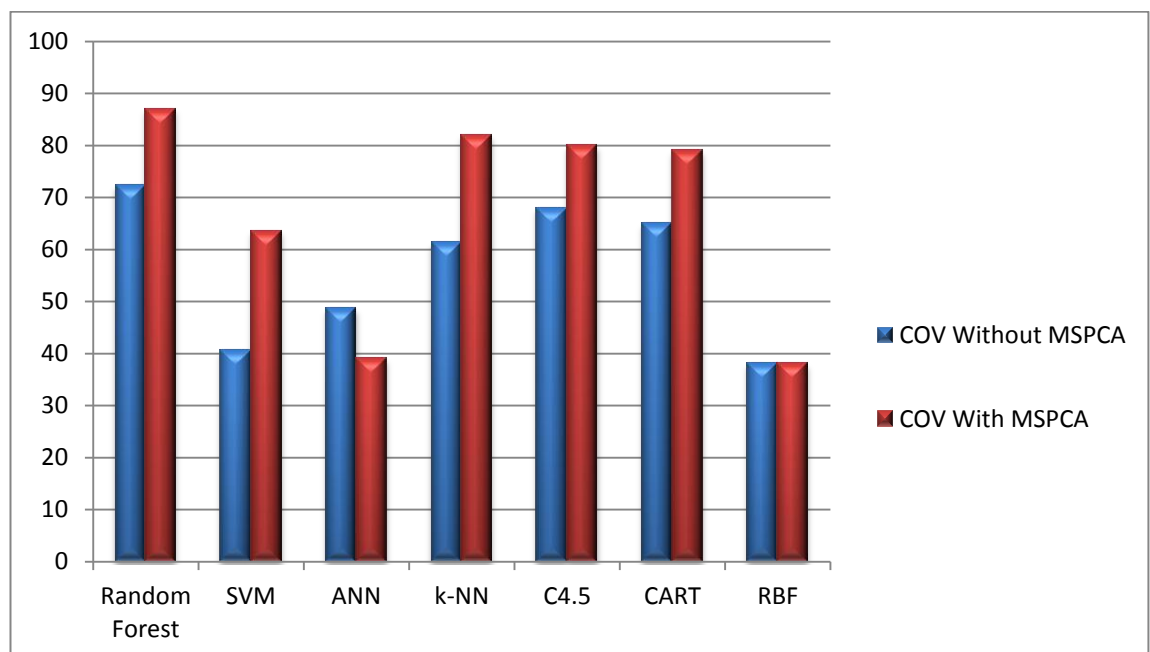


FIGURE 6.3 Graphical representation of results of both approaches after COV Feature extraction

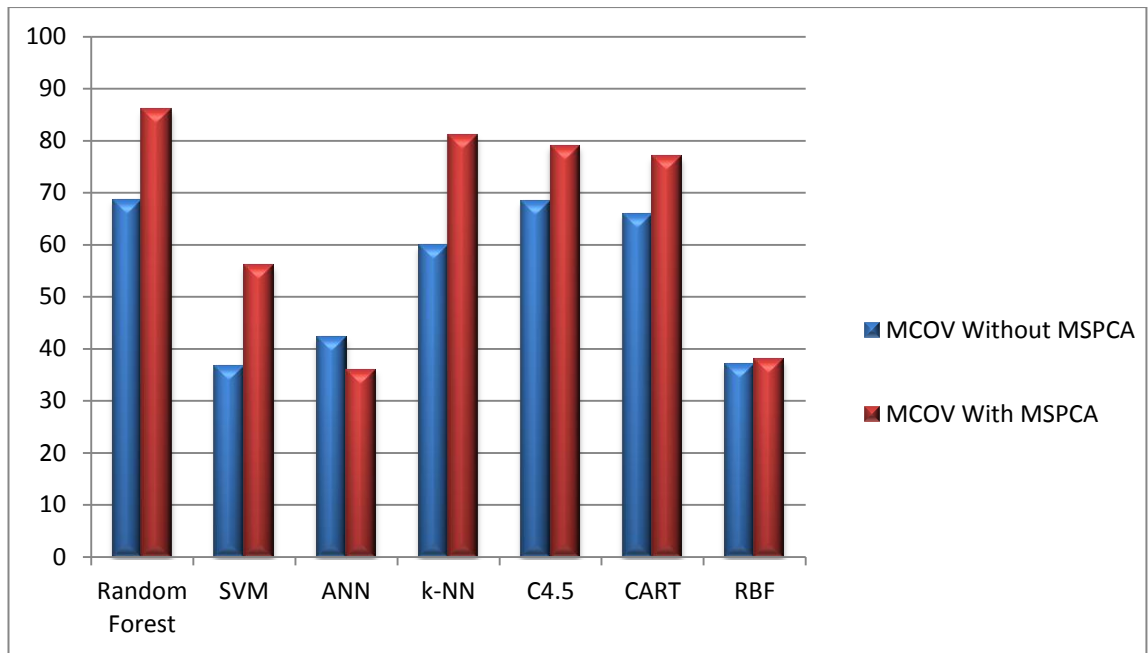


FIGURE 6.4 Graphical representation of results of both approaches after MCOV Feature extraction

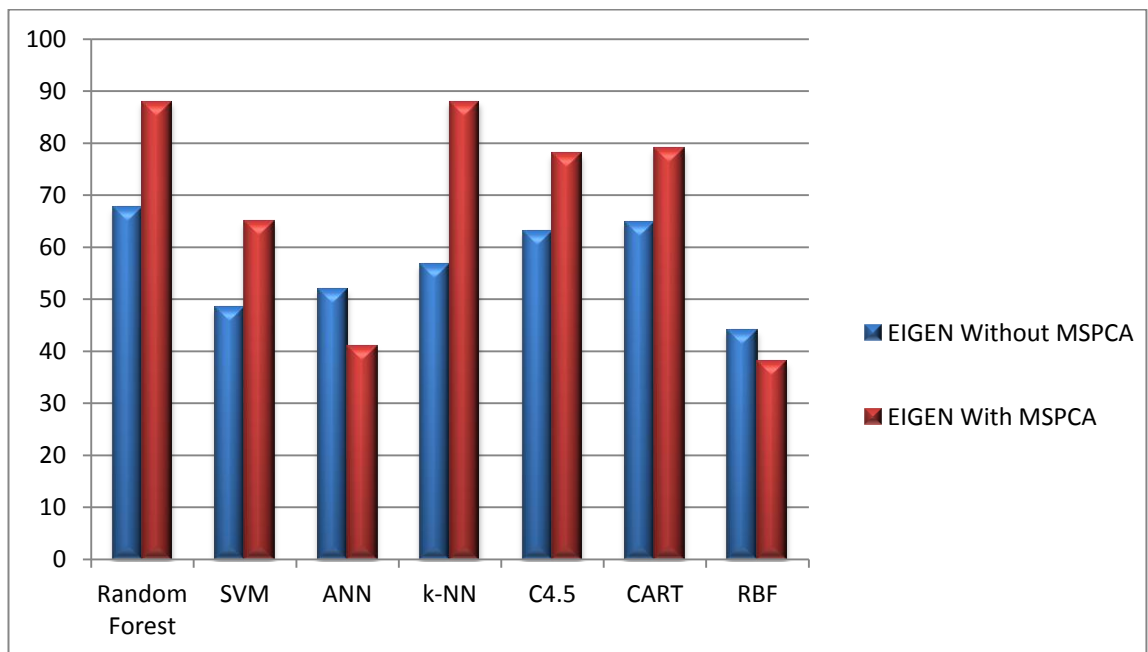


FIGURE 6.5 Graphical representation of results of both approaches after EIGEN Feature extraction

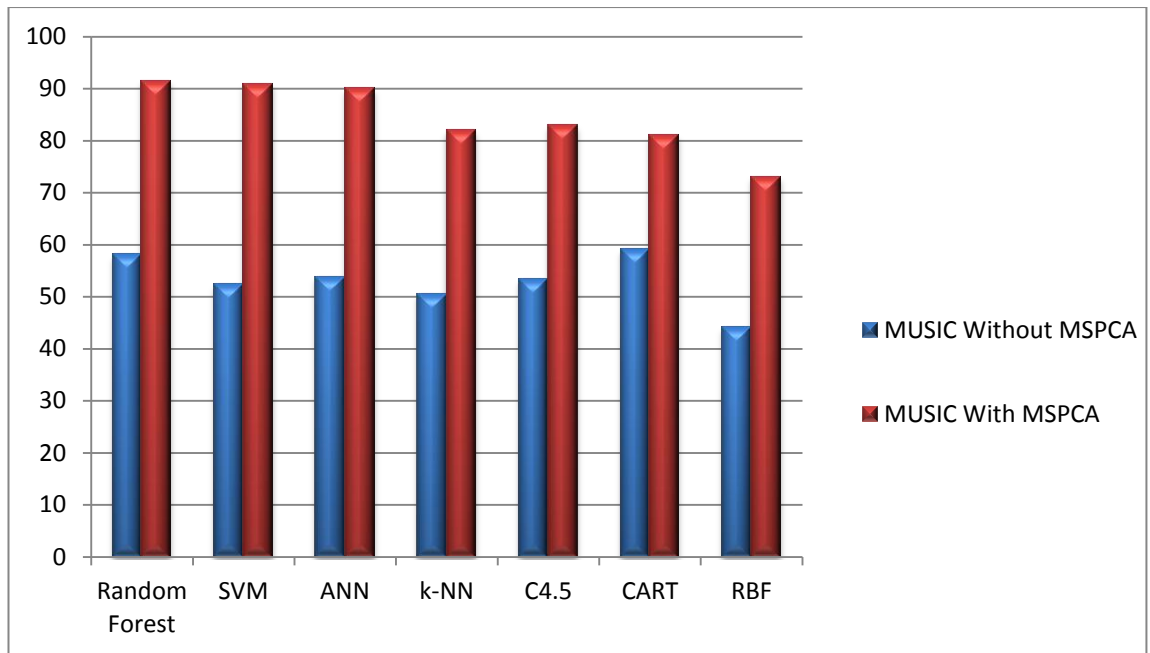


FIGURE 6.6 Graphical representation of results of both approaches after MUSIC Feature extraction

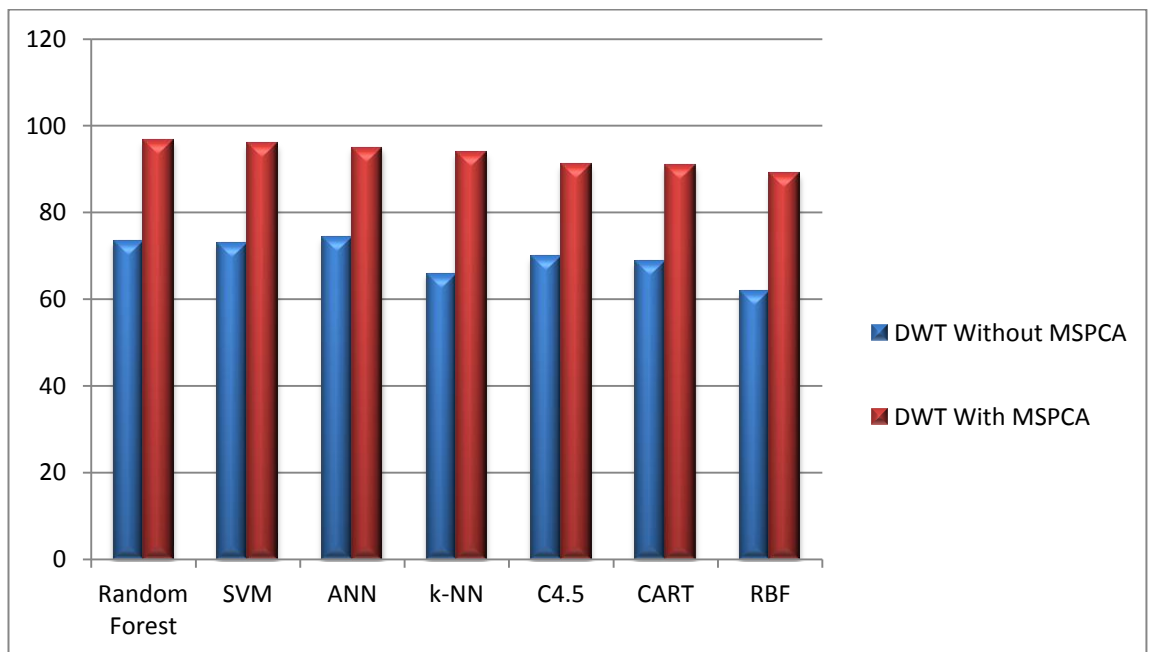


FIGURE 6.7 Graphical representation of results of both approaches after DWT Feature extraction

CHAPTER 7

DISCUSSIONS

7.1 Main Findings

Other significant EMG classification studies can be found in literature such as (Katsis, Goletis, Likas, Fodiat, & Sarmas, 2006; Pino, Stashuk, Boe, & Doherty, 2008; Yan, Wang, & Xie, 2008; Kamayuko, Farina, Yoshida, & Jensen, 2009; Dobrowolski, Wierzbowski, & Tomczykiewicz, 2011; Subasi A. , 2012). Data collecting method is important to compare EMG classification results delicately. In some studies EMG signal data obtained from surface and classified using some other techniques while some of them using simulated EMG datasets. We used intramuscular EMG signal data which gathered using needle from muscle. Since collection method is different, it cannot be said both of the data is EMG data and by using same methods, same accuracy rate can be achieved.

As said in literature SVMs have one of the best generalization capabilities. In some cases SVM can classify effectively but in our case which is classification of EMG data, SVMs could not achieve efficient performance. Since, EMG has a complex structure, SVMs and other classifiers need to be feed with more clear data. In spite of the fact that features of EMG data needs to be extracted to obtain better performance from classifiers (Dobrowolski, Wierzbowski, & Tomczykiewicz, 2011). Since SVMs have an adjuster parameter for the relationship between the training errors and generalization capability, error rate is not increasing when capability increased (Begg, Lai, & Palaniswami, 2008). It is obtained that to confirm the optimum number of the parameter within a logical range is higher accuracies obtained by using higher values.

Many classifiers applied to classify EMG signals due to their wide generalization capability, feasibility and flexibility compared to traditional classification techniques. While some of them works well with limited sample size, complex data structures, some others more coherent for multidimensional feature space. Advantage of the capability provides higher classification accuracies and a lower sensitivity. Effects of the parameters of kernel functions are affecting results between %1 and %15 depending method and data. To exemplify, SVM has a cost parameter C which organizes the range of support vector to support decision function. As a result, by using low values of C, big numbers of support vectors were obtained and more computations required determining the decision function. Hence, low number of support vectors can be used as parameter while choosing C. It is obtained that to confirm the optimum number of support vectors applying sufficiently high values of C, number of support vectors does not decrease anymore. Higher accuracies obtained by using higher C values. As an experiment, it is found that for lower values of C good generalization ability is achieved. There is another variable of SVM named as gamma which is affecting margin size and generalization capability (Cristianini & Shawe-Taylor, 2000; Massart, Vandeginste, Buydens, De Jong, Lewi, & Smeyers-Verbeke, 1997; Belousov, Verzakov, & Frese, 2002). Different kernels, C and gamma values were tried to find the best accuracy results. The best accuracy obtained by using poly kernel with degree two, C value 650 and gamma value 0.001. Effect of C parameter with different kernels can be seen from following table.

TABLE 7.1 Effect of generalization parameter C in SVM method with different kernels

	MUSIC		EIGEN		AR	
	Puk kernel	RBF kernel	Puk kernel	RBF kernel	Puk kernel	RBF kernel
C=1	85.2	85.2	42.1	33.4	46.7	35.0
C=10	89.6	89.6	44.0	34.5	53.9	35.4
C=100	90.4	90.4	57.0	37.7	61.1	37.1
C=200	90.4	90.4	59.2	37.5	61.8	38.3
C=300	90.4	90.4	60.4	38.4	62.3	38.9
C=400	90.4	90.4	90.4	89.0	63.9	39.1
C=500	90.4	90.4	90.4	89.3	64.3	40.6
C=1000	90.4	90.4	90.4	90.4	65.1	49.0

As it can be seen from Table 6.2 and 6.5, AUC results are both close to previous results and in the same order. But there is an interesting investigation for ALS data. According to fact, it can be said that, accuracy results for ALS are significantly higher than Normal and Myopathic data. This situation is same not only for SVM and Random Forest but also for all classifiers.

Since multi-space principal component analysis is utilized, number of trees also increased to enhance the dimensionality and number of variations. Obtained results proved great advantage of multidimensional and multi-spaced approach. Random Forest and MSPCA combination obtained the best classification rate with 96,67%.

7.2 Implications for Research

Total classification accuracy result of my approach is the best when compared to other studies which used same EMG data set. The reason of being best among others is using MSPCA denoising. While some of others used partial denoising approach, some of them not. In one study multiscale approach applied as parser but without decomposition method. According to comparison results it can be said that appropriate decomposition and denoising methods performs best on EMG signal. On the other hand simulated used in some studies since it is hard to find reliable EMG dataset. But, different classification results outcomes by application same set of methods to simulated and real data. Higher total classification accuracies can be obtained by using simulated data. But it can be said that, since simulated data is not real one, results of the study is not reliable for clinical usage or as a decision support tool. Even if they are not reliable for clinical usage, it can be used for interpretation and inference.

TABLE 7.2 Comparison of results and methods with other studies which used same dataset

Studies	Approach	Accuracy (%)
(Parsaei & Stashuk, 2011)	Adaptive gap-based Duda and Hart (AGDH) method used for component analysis on motor unit potential trains	91.3%
(Filho, da Silva, & de Carvalho, 2008)	Multiscale Multidimensional Parser applied and results evaluated percent root mean for encoding EMG signal	91.1% of the False Positives, 95% correctly identified
(Parsaei & Stashuk, 2010)	Gap statistics and Motor unit potential trains utilized together for decomposition of EMG signal	93.8%
(Yejin, Chenyun, & Clancy, 2013)	Montreal and Fuzzy Expert decomposition techniques applied	95.7% (at %20 contraction level)
(Marateb, Muceli, McGill, Merletti, & Farina, 2011)	A novel decomposition method applied on Motor unit potential trains	>90%
(Nikolic & Krarup, 2011)	A novel EMGtools Decomposition method applied on constituent MUAPs and firing patterns	95% correctly identified
My Proposal	Features of denoised EMG signal extracted by using Discrete Wavelet Transform then classified with Random Forest	96.67%

7.3 Implications for Practice

Popularity of analysis of EMG signal is growing related to clinical diagnosis, clinician support tool and biomedical applications usage field in real life. Analysis of EMG signals using recent, powerful and advance methodologies is becoming new trend in biomedical signal processing with respect to developments and investigations in computer sciences. Because EMG signal analysis is crucial in clinical diagnosis of neuromuscular disorders.

7.4 Limitations

As every study has we also have some limitations and boundaries by covering these limitations better results can be obtained. Reliability level also can be improved after eliminating some limitations. Limitations in this study can be listed as follows.

- Number of patients is 25 and the age distribution between 19 and 67 years old. More subjects should be included in future studies.
- The brachial biceps and medial vastus muscles were used in this study since they were the most frequent ones. Other muscle groups also should be analyzed.

7.5 Future Directions

Every study can be updated and upgraded in sight of newer or other approaches. Our results made us think about other approaches. In short time, it is planned to combine some methods because when one method results better for one disease, other can result better than for other disease. To exemplify, although total classification accuracy of ANN better than Random Forest with MUSIC feature extraction, ALS sensitivity of Random forest is considerably higher than ANN. If different classification methods can be used for different type of data, better results can be observed.

Another study can be about classification and investigation of ALS disease in terms of biomedical signal processing. Because, According to obtained results accuracy of ALS almost everytime higher than myopathy and normal data. Deeper and wider investigation can be done just on ALS disease regarding its coherence to work with machine learning techniques.

Next and the last future study can be comparison of denoising methods on EMG signal. We applied just one denoising method which has very impressive effect on total classification accuracy. The aim can be investigating best EMG denosing algorithm.

CHAPTER 8

CONCLUSIONS

In this study, we find out considerable accompanies between classifiers, denoising and feature extraction methods, which proved by the results. Coherence between feature extraction and classification methods are obtained and discussed. Significant results were found by utilizing MSPCA denoising method. After denoising with MSPCA accuracy upraised notably. By the way effectiveness is investigated. Beside that coherence and difference between feature extraction methods obtained. Proposed approach assumes that noise on the EMG signal is anomaly which needs to be eliminated to make it more pure.

Novelty of this study is achieving significantly better performance by denoising with MSPCA over the three EMG signal patterns: normal, myopathic and ALS. This was concluded using classifiers and feature extraction methods with MSPCA denoising. While searching for coherence between methods, we realized how to get higher accuracy by rearranging kernel parameters of classifiers referring to their characteristics and effects. Generally it is better not to use default kernel parameters of classifiers. The classification accuracy with MSPCA improved to 96.67% which was 73.33% without MSPCA using same DWT feature extraction + Random Forest combination. The 39.87% improvement proves the great performance and advantage of MSPCA denoising. Consequently, classification accuracy of other classifiers also increased. To exemplify, improvement of classification accuracy for SVM is 35.55%, for ANN is 39.87%, for k-NN is 31.66%. MSPCA denoising can become helpful for diagnosing of neuromuscular disorders regarding these results and methods.

REFERENCES

- Akansu, A. N., & Haddad, R. A. (1992). *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets*. Boston, MA: Academic Press.
- Akay, M. (1997). Wavelet applications in medicine. *IEEE spectrum* , 34, 50-56.
- Akben, S. B., Subasi, A., & Tuncel, D. (2012). Analysis of Repetitive Flash Stimulation Frequencies and Record Periods to Detect Migraine Using Artificial Neural Network. *Journal of Medical Systems* , 36 (2), 925-931.
- Ali, A. B., & Wasimi, S. A. (2007). *Data Mining : Methods and Techniques*. Victoria: Thomson, Australia.
- Ambroise, C., & McLachlan, G. J. (2002). Selection bias in gene extraction on the basis of microarray gene-expression data. *Proceedings of the National Academy of Science of USA* , 6562–6566.
- Bakshi, B. R. (1999). Multiscale analysis and modeling using wavelets. *Journal of Chemometrics* , 13 (3-4), 415 - 434.
- Bakshi, B. R. (1998). Multiscale PCA with application to MSPC monitoring. *AIChE Journal* , 44 (77), 1596 - 1610.
- Basmajian, J. V., & De Luca, C. J. (1985). *Muscles Alive. Their Functions Revealed by Electromyography*. Baltimore: Williams & Wilkins.
- Begg, R., Lai, D. T., & Palaniswami, M. (2008). *Computational Intelligence in Biomedical Engineering*. Boca Raton: CRC Press.
- Belousov, A. I., Verzakov, S. A., & Frese, J. (2002). A flexible classification approach with optimal generalisation performance:support vector machines. *Chemometrics and Intelligent Laboratory Systems* , 64 (1), 15-25.
- Berardelli, A., Mercuri, B., Rona, S., Vacca, L., & Manfredi, M. (1996). Pathophysiological mechanisms of dystonia. *Recent Advances in Clinical Neurophysiology* , 933-938.

- Bertolasi, L., De Grandis, D., Bongiovanni, L. G., Zanette, G. P., & Gasperini, M. (1993). The influence of muscular lengthening on cramps. *Ann Neurol* , 33, 176-180.
- Bienvenu, G., & Kopp, L. (1983). Optimality of high resolution array processing using the eigensystem approach. *IEEE Transactions on Acoustics, Speech and Signal Processing* , 31 (5), 1235-1248.
- Breiman, L. (2001). Random forests. *Machine Learning* , 45, 5-32.
- Breiman, L., Friedman, J. H., Ohlsen, R. A., & Stone, C. I. (1984). Classification and. *Wadsworth International Group: Belmont, California* , 43-49.
- Buchthal, F. (1991). Electromyography in the Evaluation of Muscle Diseases. *Methods in Clinical Neurophysiology* , 2, 22-45.
- Buhmann, M. (2003). *Radial Basis Functions: Theory and Implementations*. Cambridge: Cambridge University Press.
- Castells, F., Laguna, P., Sornmo, L., Bollmann, A., & Roig, J. (2007). Principal Component Analysis in ECG Signal Processing. *EURASIP Journal on Advances in Signal Processing* , 2007, 98-119.
- Chawla, M. P. (2009). A comparative analysis of principal component and independent component techniques for electrocardiograms. *Neural Comput and Applic, Springer-Verlag London Limited* , 18 (6), 539-556.
- Cohen, L. (1992). Instantaneous scale and the short-time scale transform. *Time-Frequency and Time-Scale Analysis, 1992., Proceedings of the IEEE-SP International Symposium* (pp. 383-386). Victoria, BC: IEEE.
- Cooper, R. G., Stokes, M. J., & Edwards, R. H. (1988). Physiological characterisation of the “warm up” effect of activity in patients with myotonic dystrophy. *J Neurol Neurosurg Psychiatry* , 51, 1134-1141.
- Crews, J., Kaiser, K. K., & Brooke, M. H. (1976). Muscle pathology of myotonia congenita. *J Neurol Sci* , 28, 449-457.
- Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
- Deleo, J. M. (1993). Receiver operating characteristic laboratory (ROCLAB): software for developing decision strategies that account for uncertainty. *Proceedings of the Second International Symposium on Uncertainty Modelling and Analysis* , 318-325.

- Denny-Brown, D. (1949). Interpretation of the electromyogram. *Arch Neural Psychiatry* , 61, 99-128.
- Desmedt, J. E., & Borenstein, S. (1973). Collateral reinnervation of muscle fibre by motor axons of dystrophic motor units. *Nature* , 246, 500-501.
- Devos, A., Lukas, L., Suykens, J. A., Vanhamme, L., Tate, A. R., Howe, F. A., et al. (2004). Classification of brain tumours using short echo time 1H MR spectra. *Journal of Magnetic Resonance* 170 , 164-175.
- Deymeer, F., Cakirkaya, S., Serdaroglu, P., Schleithoff, L., Lehmann-Horn, F., Rudel, R., et al. (1998). Transient weakness and compound muscle action potential decrement in myotonia congenita. *Muscle Nerve* , 21, 1334-1337.
- Díaz-Uriarte, R., & Alvarez de Andrés, S. (2006). Gene selection and classification of microarray data using random forest. *BMC Bioinformatics* , 7 (3).
- Dietterich, T. G. (2000). An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization. *Machine Learning* , 40 (2), 1-22.
- Dobrowolski, A. P., Wierzbowski, M., & Tomczykiewicz, K. (2011). Multiresolution MUAPs decomposition and SVM-based analysis in the classification of neuromuscular disorders. *Comput. Methods Programs Biomed* , 107 (3), 393-403.
- Duffy, F., Iyer, V. G., & Surwillo, W. W. (1989). *Clinical Electroencephalography and Topographic Brain mapping: Technology and Practice*. Michigan: Springer-Verlag.
- Eidelman, B. H., Nielsen, V. K., Moller, M., & Jannetta, P. J. (1985). Vascular compression, hemifacial spasm and multiple cranial neuropathy. *Neurology* , 35, 712-716.
- Encyclopedia of the Neurological Sciences*. (2003). USA: Elsevier Science.
- Fabinyi, G. C., & Adams, C. B. (1978). Hemifacial spasm: Treatment by posterior fossa surgery. *J Neurol Neurosurg Psychiatry* , 41, 829-833.
- Garcia-Mullin, R., & Daroff, R. B. (1973). Electrophysiological investigations of cephalic tetanus. *J Neurol Neurosurg Psychiatry* , 36, 296-301.
- Gokgoz, E., & Subasi, A. (2014). Effect of Multiscale PCA de-noising on EMG signal classification for Diagnosis of Neuromuscular Disorders. *Journal of Medical Systems* , In Press.

- Güler, N. F., & Koçer, S. (2005). Use of Support Vector Machines and Neural Network in Diagnosis of Neuromuscular Disorders. *Journal of Medical Systems* , 29 (3), 271-284.
- Hahn, A. F., Parkes, A. W., Bolton, C. F., & Stewart, S. A. (1991). Neuromyotonia in hereditary motor neuropathy. *J Neurol Neurosurg Psychiatry* , 54, 230-235.
- Handley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143 , 29-36.
- Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic(ROC). *Radiology* 143 , 29-36.
- Hassoun, M., Wang, C., & Spitzer, A. (1994). NNERVE: neural network extraction of repetitive vectors for electromyography. *IEEE Transactions on Biomedical Engineering* , 41 (11), 1953-1061.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice-Hall.
- Henneberg, K. A., & Plonsey, R. (1993). Boundary element analysis of the directional sensitivity of the concentric EMG electrode. *IEEE Transactions on Biomedical Engineering* , 40, 621-631.
- Henneberg, K. (2000). Principles of Electromyography. In J. D. Bronzino, *The Biomedical Engineering Handbok:Second edition*. Boca Raton: CRC Press LLC.
- Ho, S. Y., Shu, L. S., & Chen, J. H. (2004). Intelligent evolutionary algorithms for large parameter optimization problems. *IEEE Trans. Evolut.* , 8 (6), 522-541.
- Hotelling, H. (1933). Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology* , 24 (6), 417-441.
- Huang, C. L., & Wang, C. J. (2006). A GA-based attribute selection and parameter optimization for support vector machine. *Expert Syst. Appl.* , 31 (2), 231-240.
- Johnson, D. H., & DeGraaf, S. (1982). Improving the resolution of bearing in passive sonar arrays by eigenvalue analysis. *IEEE Transactions, Acoustics, Speech, Signal Processing* , 30 (4), 638–647.
- Joliffe, I. T. (2002). *Principal Component Analysis, second edition*. New York, NY, USA: Springer.
- Kamayuako, E. N., Farina, D., Yoshida, K., & Jensen, W. (2009). Relationship between grasping force and features of sgnle channel intramuscular EMG signals. *Journal of Neuroscience Methods* , 185 (1), 143-150.
- Kandaswamy , A., Kumar , C. S., Ramanathan , R. P., Jayaraman , S., & Malmurugan , N. (2004). Neural classification of lung sounds using wavelet coefficients. *Computers in Biology and Medicine* , 34 (6), 523-537.

- Katsis, C. D., Exarchos, T. P., Papaloukas, C., Goletsis, Y., Fodiatis, D. I., & Sarmas, I. (2007). A two-stage method for MUAP classification based on EMG decomposition. *Computers in Biology and Medicine* , 1232-1240.
- Katsis, C. D., Goletis, Y., Likas, A., Fodiatis, D. I., & Sarmas, I. (2006). A novel method for automated EMG decomposition and MUAP classification. *Artificial Intelligence in Medicine* , 37 (1), 55-64.
- Kecman, V. (2001). *Learning and Soft Computing*. Cambridge,Massachusetts: MIT Press.
- Kimura, J. (2001). *Electrodiagnosis in Diseases of Nerve and Muscle: Principle and Practice* (3rd ed.). Newyork: Oxford University Press.
- Koçer, S. (2010). Classification of Emg Signals Using Neuro-Fuzzy System and Diagnosis of Neuromuscular Diseases. *Journal of Medical Systems* , 34 (3), 321-329.
- Kotas, M. (2006). Application of projection pursuit based robust principal component analysis to ECG enhancement. *Biomedical Signal Processing and Control* , 1 (4), 289-298.
- Lang, A. H., & Partanen, I. (1976). Satellite potentials and the duration of potentials in normal, neuropathic and myopathic muscles. *J Neurol Sci* , 513-525.
- Lehmann-Horn, F., Kuther, G., Ricker, K., Grafe, P., Ballanyi, K., & Rudel, R. (1987). Adynamia episodica hereditaria with myotonia: A noninactivating sodium current and the effect of extracellular pH. *Muscle Nerve* , 10, 363-374.
- Lewis, R. J. (2000). An Introduction to Classification and Regression Tree (CART) Analysis. *Annual Meeting of the Society for Academic Emergency Medicine*. San Francisco, California.
- Li, Y., Dai, C., Clancy, E. A., Christie, A., & McGill, K. C. (2013). Cross-Comparison Between Two Multi-Channel EMG Decomposition Algorithms. *39th Annual Northeast Bioengineering Conference*, (pp. 191-192).
- Liaw, A., & Wiener, M. (2002). Classification and regression by Random Forest. *R News* , 2, 18-22.
- Loudon, G. H., Jones, N. B., & Sehmi, A. S. (1992). New signal processing techniques for the decomposition of EMG signals. *Medical & Biological Engineering & Computing* , 30 (6), 591-599.
- Manolakis, D. G., Ingle, V. K., & Kogon, S. M. (2005). *Statistical and Adaptive Signal Processing*. London: Artech House.

- Marple, L. S. (1987). *Digital Spectral Analysis: With Applications*. NJ: Prentice Hall.
- Massart, D. L., Vandeginste, B. G., Buydens, L. M., De Jong, S., Lewi, P. J., & Smeyers-Verbeke, J. (1997). *Handbook of Chemometrics and Qualimetrics*. Amsterdam: Elsevier.
- Mazid, M. M., Shawkat Ali, A. B., & Stickle, K. (2010). Improved C4.5 Algorithm for Rule Based Classification. *AIKED'10 Proceedings of the 9th WSEAS international conference on Artificial intelligence, knowledge engineering and data bases*, (pp. 296-301). Wisconsin.
- McGill, K. C. (2002). Optimal resolution of superimposed action potentials. *IEEE Transactions on Bio-medical Engineering*, 49 (7), 640-650.
- McManis, P. G., Lambert, E. H., & Daube, J. R. (1986). The exercise test in periodic paralysis. *Muscle Nerve*, 9, 704-710.
- Meinguet, J. (1979). Multivariate interpolation at arbitrary points made. *Zeitschrift Angewandte Math. Phys.*, 292–304.
- Merletti, R., & Parker, P. A. (2004). *Electromyography. Physiology, Engineering, and Noninvasive Applications*. NJ: IEEE Press Series in Biomedical Engineering.
- Mesin, L., & Farina, D. (2005). Estimation of M-wave scale factor during sustained contractions at high stimulation rate. *IEEE Transactions on Biomedical Engineering*, 52 (5), 869-877.
- Mesin, L., Joubert, M., Hanekom, T., Merletti, R., & Farina, D. (2006). A finite element model for describing the effect of muscle shortening on surface EMG. *IEEE Transactions on Biomedical Engineering*, 53 (4), 593-600.
- Meyers, K. R., Gilden, D. H., Rinaldi, C. F., & Hansen, J. L. (1972). Periodic muscle weakness, normokalemia, and tubular aggregates. *Neurology (Minneapolis)*, 22, 269-279.
- Minsky, M., & Papert, S. (1988). *Perceptrons*. Cambridge: MIT Press.
- Misra, M., Yue, H. H., Qin, S. J., & Ling, C. (2002). Multivariate process monitoring and fault diagnosis by multi-scale PCA. *Computers and Chemical Engineering*, 26 (9), 1281 – 1293.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Nielsen, V. K. (1984). Pathophysiology of hemifacial spasm: I. Ephaptic transmission and ectopic excitation. *Neurology*, 34, 418-426.

- Nikolic, M. (2001). *Findings and Firing Pattern Analysis in Controls and Patients with Myopathy and Amyotrophic Lateral Sclerosis*. Copenhagen: Faculty of Health Science, University of Copenhagen.
- Nikolic, M., & Krarup, C. (2011). EMGTools, an Adaptive and Versatile Tool. *IEEE Transactions On Biomedical Engineering* , 58 (10), 2707-2718.
- Obuchowski, N. A. (2003). Receiver operating characteristic curves and their use in radiology. *Radiology* 229 , 3-8.
- Oppenheim, A. V., & Schafer, R. W. (1989). *Discrete-time signal processing*. NJ: Prentice Hall.
- Parsaei, H. (2011). *EMG Signal Decomposition using Motor Unit Potential Train Validity*. Systems Design Engineering. Ontario: University of Waterloo.
- Parsaei, H., & Stashuk, D. W. (2010). Evaluating a motor unit potential train using cluster validation methods. *Canadian Student Conference on Biomedical Computing and* , 31-35.
- Pattichis, C. S., & Pattichis, M. S. (1999). Time-scale analysis of motor unit action potentials. *IEEE Transactions on Bio-medical Engineering* , 46 (11), 1320-1329.
- Phinyomark, A., Limsakul, C., & Phukpattaranont, P. (2011). Application of Wavelet Analysis in EMG Feature Extraction for Pattern Classification. *Measurement Science Review* , 11 (2), 45-52.
- Pino, L. J., Stashuk, D. W., Boe, S. G., & Doherty, T. J. (2008). Motor unit potential characterization using pattern discovery. *Medical Engineering & Physics* , 30 (5), 563-573.
- Pisarenko, V. F. (1973). The retrieval of harmonics from a covariance function Geophysics. *Journal of the Royal Astronomical Society* , 33 (3), 347-366.
- Preedy, V. R., & Peters, T. J. (1994). Alcohol and muscle disease. *Journal of the Royal Society of Medicine* , 87 (4), 188-190.
- Preston, D. C., & Shapiro, B. C. (1998). *Electromyography and Neuromuscular Disorders Clinical-Electrophysiologic Correlations*. Boston: Butterworth-Heinemann.
- Preston, D., & Shapiro, B. (2005). *Electromyography and Neuromuscular Disorders: Clinical-electrophysiologic Correlations*. Philadelphia: Elsevier.
- Proakis, J. G., & Manolakis, D. G. (1996). *Digital signal processing (3rd ed.): principles, algorithms, and applications*. NJ: Prentice Hall.

- Qi, X., Hui, Z., Yongji, W., & Jian, H. (2009). Fuzzy support vector machine for classification of EEG signals using. *Medical Engineering & Physics* , 31, 858–865.
- Qi, Y. (2012). Random Forest for Bioinformatics. In Z. Cua, & Y. Ma, *Ensemble Machine Learning Methods and Applications* (pp. 307-323). NJ: Springer.
- Quinlan, J. R. (1996). Improved use of continuous attributes in C4.5. *Journal of Artificial Intelligence Reserach* , 4, 77-90.
- Quinlan, R. J. (1993). *C4.5 : Programs for Machine Learning*. USA: Morgan Kaufmann.
- Reaz, M. B., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: detection, processing, classification and Applications. *Biol. Proced. Online* , 8, 11-35.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* (65), 386-408.
- Sandvig, J. J., Mobasher, B., & Burke, R. (2008). A Survey of Collaborative Recommendation and the Robustness of Model-Based Algorithms. *IEEE Computer Society Technical Committee on Data Engineering* , 31 (2), 3-13.
- Schizas, C. N., Pattichis, C. S., & Middleton, L. (1990). Artificial neural nets in computer-aided macro motor unit potential classification. *IEEE Engineering Min Medicine Biology Magazine* , 9 (3), 31-38.
- Schmidt, R. O. (1986). Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propagation* , 34 (3), 276-280.
- Semmlow, J. L. (2004). *Biosignal and Biomedical Image Processing*. New York: Marcel Dekker.
- Shien-Ping, H., & Chih-Chou, C. (2009). Process Monitoring with ICA-based Signal extraction Technique and CART Approach. *Quality and Reliability Engineering International* , 25, 631-643.
- Soman, K. P., Diwakar, S., & Ajay, V. (2006). *Insight into Data Mining: Theory and Practice*. India: Prentice-Hall.
- Sörnmo, L., & Laguna, P. (2005). *Bioelectrical Signal Processing in Cardiac and Neurological Applications*. London: Elsevier Academic Press.
- Subasi , A. (2012). Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines. *Computers in Biology and Medicine* , 42 (8), 806–815.

- Subasi, A. (2005). Application of Classical and Model-Based Spectral Methods to Describe the State of Alertness in EEG. *Journal of Medical Systems* , 29 (5), 473-486.
- Subasi, A. (2012). Classification of EMG signals using combined features and soft computing. *Applied Soft Computing* , 12 (8), 2188–2198.
- Subasi, A. (2013). Classification of EMG signals using PSO optimized SVM for diagnosis. *Computers in Biology and Medicine* , 43 (5), 576–586.
- Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Systems with Applications* , 32, 1084-1093.
- Subasi, A., & Erçelebi, E. (2005). Classification of EEG signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine* , 78 (2), 87-99.
- Subasi, A., & Kiymik, M. K. (2010). Muscle Fatigue Detection in EMG Using Time–Frequency Methods, ICA and Neural Networks. *Journal of Medical Systems* , 34 (4), 777-785.
- Subasi, A., Yilmaz, M., & Ozcalik, H. R. (2006). Classification of EMG signals using wavelet neural network. *Journal of Neuroscience Methods* , 156 (1-2), 360–367.
- Svetnik, V., Liaw, A., Tong, C., Culbertson, J. C., Sheridan, R. P., & Feuston, B. P. (2003). Random Forest: A Classification and Regression Tool for Compound Classification and QSAR Modeling. *Journal of chemical information and modeling* , 43 (6), 1947-1958.
- Swets, J. A. (1979). ROC analysis applied to the evaluation of medical imaging techniques. *Invest. Radiol.* 14 , 109-121.
- Takahashi, T. (1966). The coupling discharge in neurogenic muscular atrophy. *Arch Neurol* , 14, 617-623.
- Thage, O. (1974). *Quadriceps weakness and wasting*. Copenhagen: FADLs Forlag.
- Vapnik, V. (2006). *Estimation of Dependencies Based on Empirical Data*. springer.
- Vapnik, V. (2000). *The Nature of Statistical Learning Theory*. Springer.
- Webster, J. G. (1998). *Medical Instrumentation. Application and Design*. Boston: John Wiley and Sons.
- Witten, I., & Frank, E. (2005). *Data Mining, Practical Machine Learning Tools and*. San Francisco, CA: Morgan Kaufmann Publishers (Elsevier).

- Yan, Z., Wang, Z., & Xie, H. (2008). The application of mutual information-based feature selection and fuzzy LS-SVM-based classifier in motion classification. *Computer Methods and Programs in Biomedicine* , 90 (3), 275-284.
- Yang, Z. R. (2010). *Machine Learning Approaches to Bioinformatics*. Singapore: World Scientific Publishing Co. Pte. Ltd.
- Yi, Z., & Pan, J. (2010). Application of Random Forest to Stellar Spectral Classification. *3rd International Congress on Image and Signal Processing* (pp. 3129-3133). Yantai: School of Computer Science, Yantai University.
- Zhang, R., McAllister, G., Scotney, B., McClean, S., & Houston, G. (2006). Combining Wavelet Analysis and Bayesian Networks for the Classification of Auditory Brainstem Response. *IEEE Trans. Info. Tech. Biomed* , 10 (3), 458-467.
- Zwarts, M. J., Van Weerden, T. W., Links, T. P., Haene, H. T., & Oosterhuis, H. J. (1988). The muscle fiber conduction velocity and power spectra in familial hypokalemic periodic paralysis. *Muscle Nerve* , 11, 166-173.
- Zweig, M. H., & Campbell, G. (1993). Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clinical Chemistry* , 39 (8), 561-577.

APPENDICES

Appendix A. Plotting of methods before and after denoising

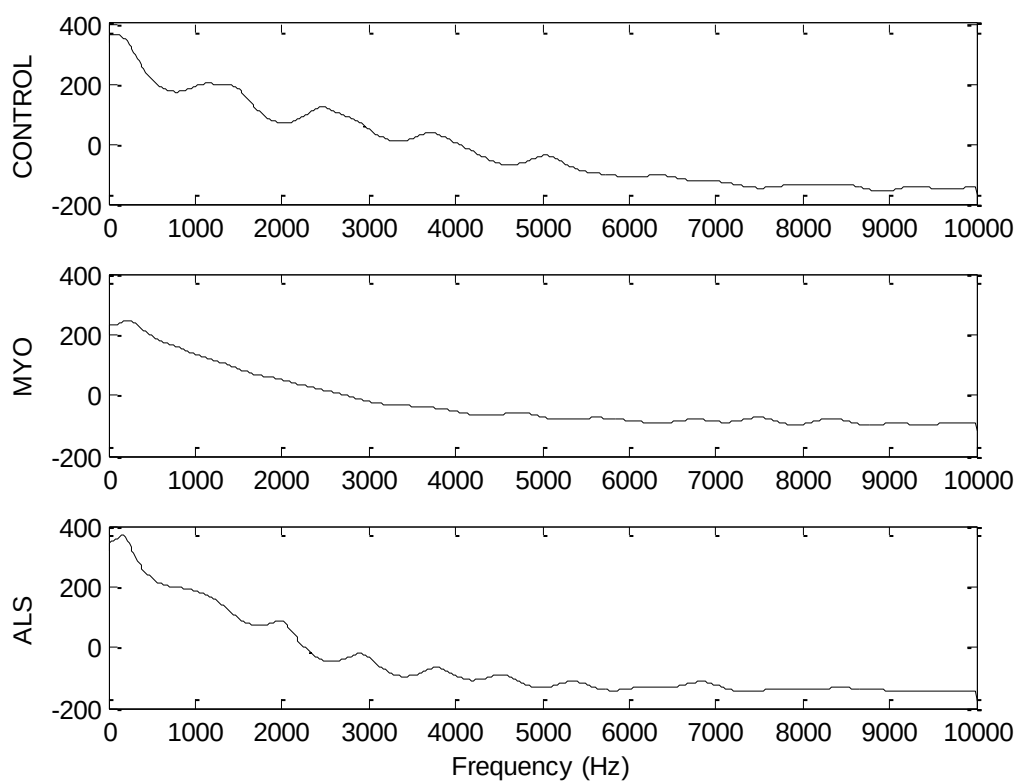


FIGURE A.1 Plotting of signal after ARBurg feature extraction before MSPCA denoising

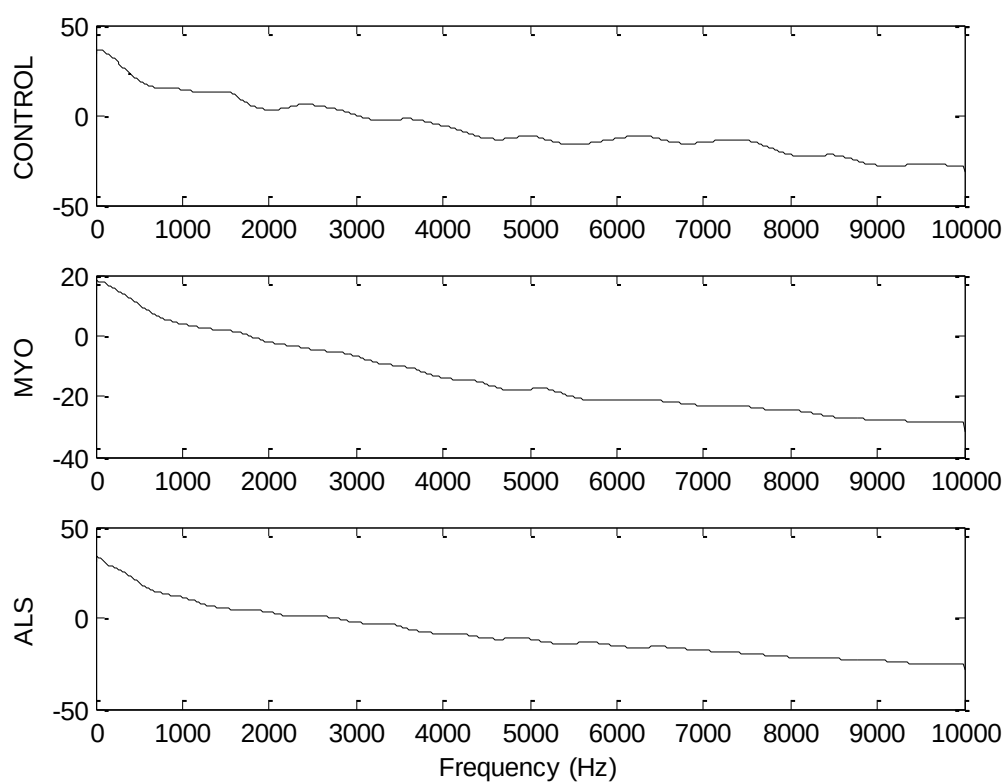


FIGURE A.2 Plotting of signal after ARBurg feature extraction and MSPCA denoising

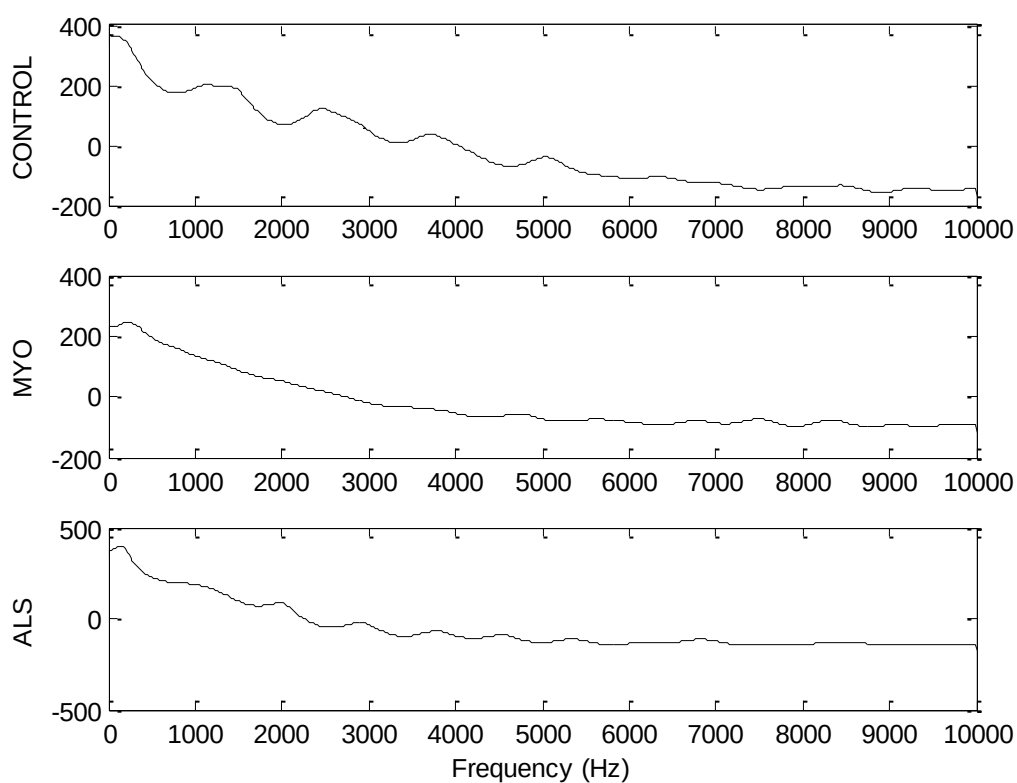


FIGURE A.3 Plotting of signal after COV feature extraction before MSPCA denoising

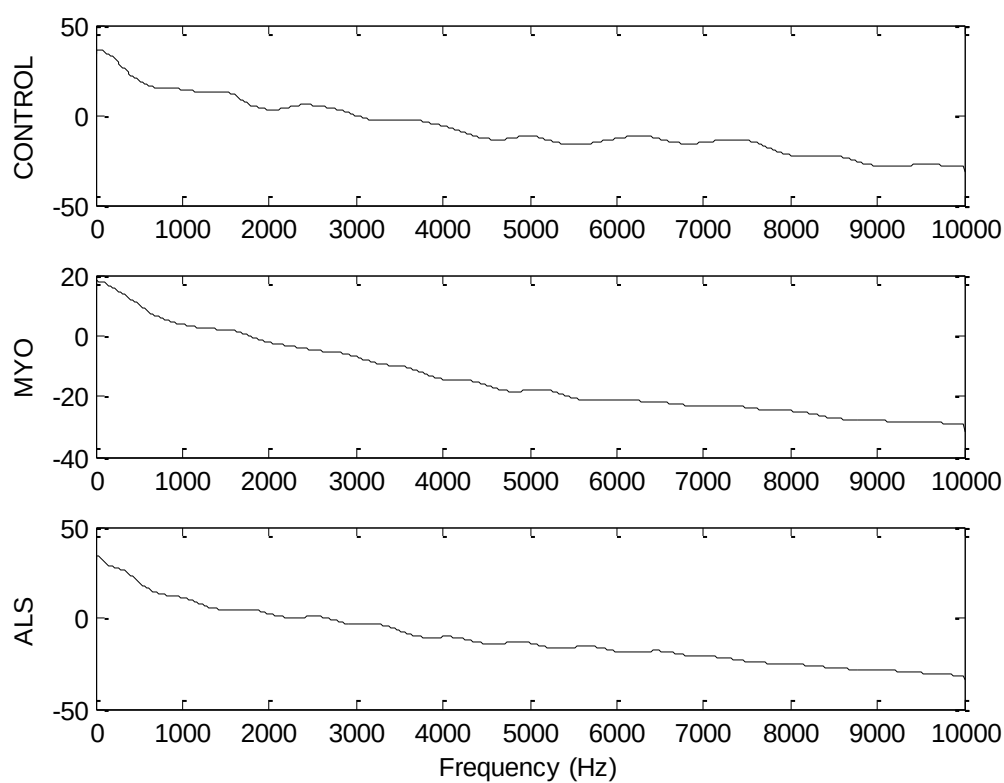


FIGURE A.4 Plotting of signal after COV feature extraction and MSPCA denoising

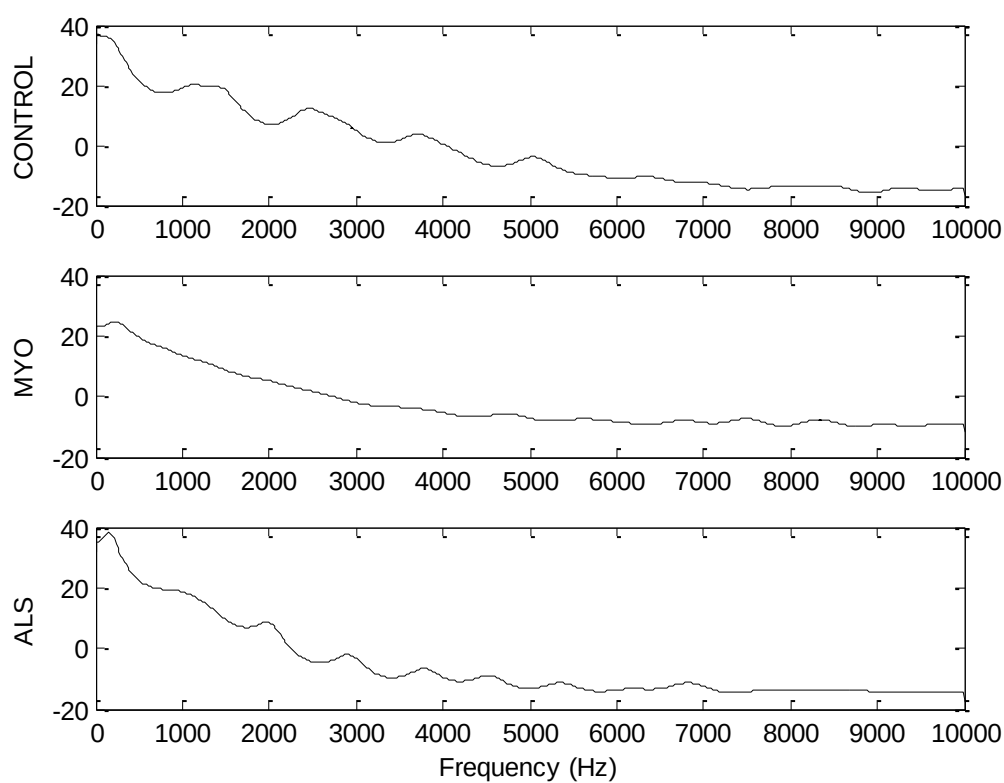


FIGURE A.5 Plotting of signal after MCOV feature extraction before MSPCA denoising

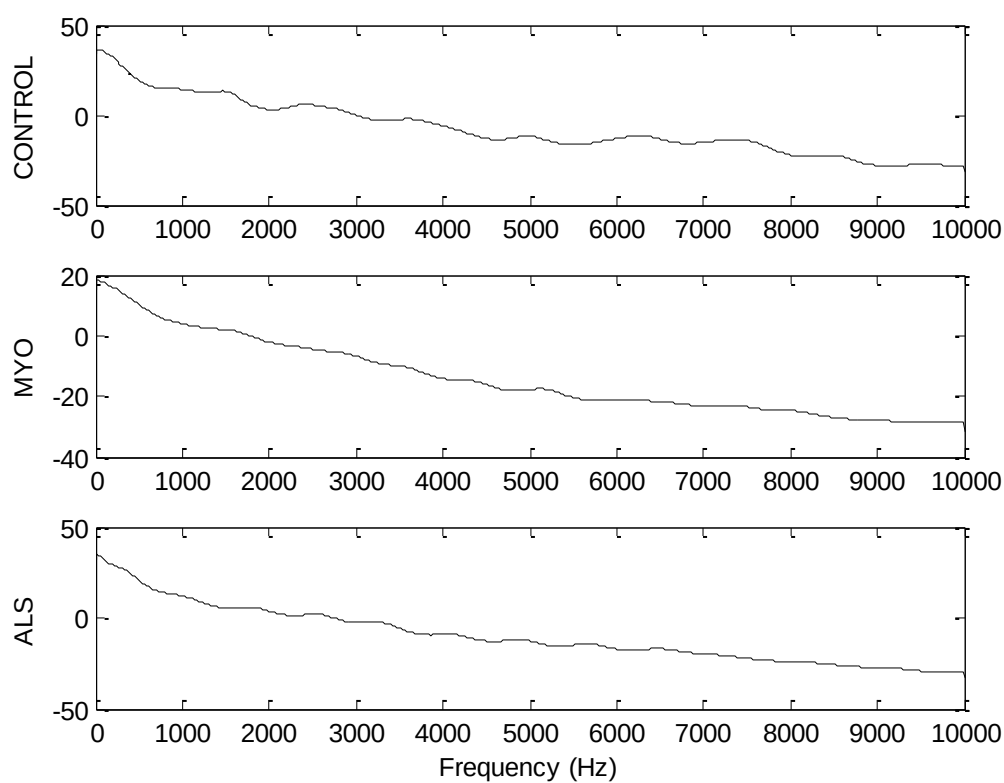


Figure A.6 Plotting of signal after MCOV feature extraction and MSPCA denoising

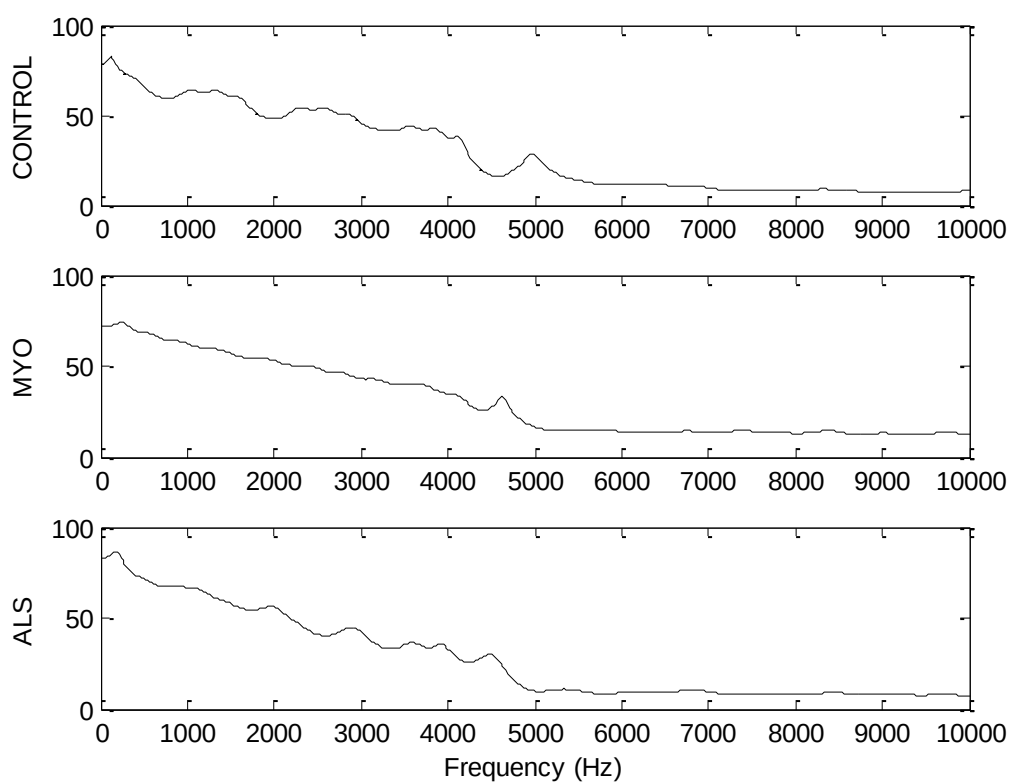


FIGURE A.7 Plotting of signal after EIGEN feature extraction before MSPCA denoising

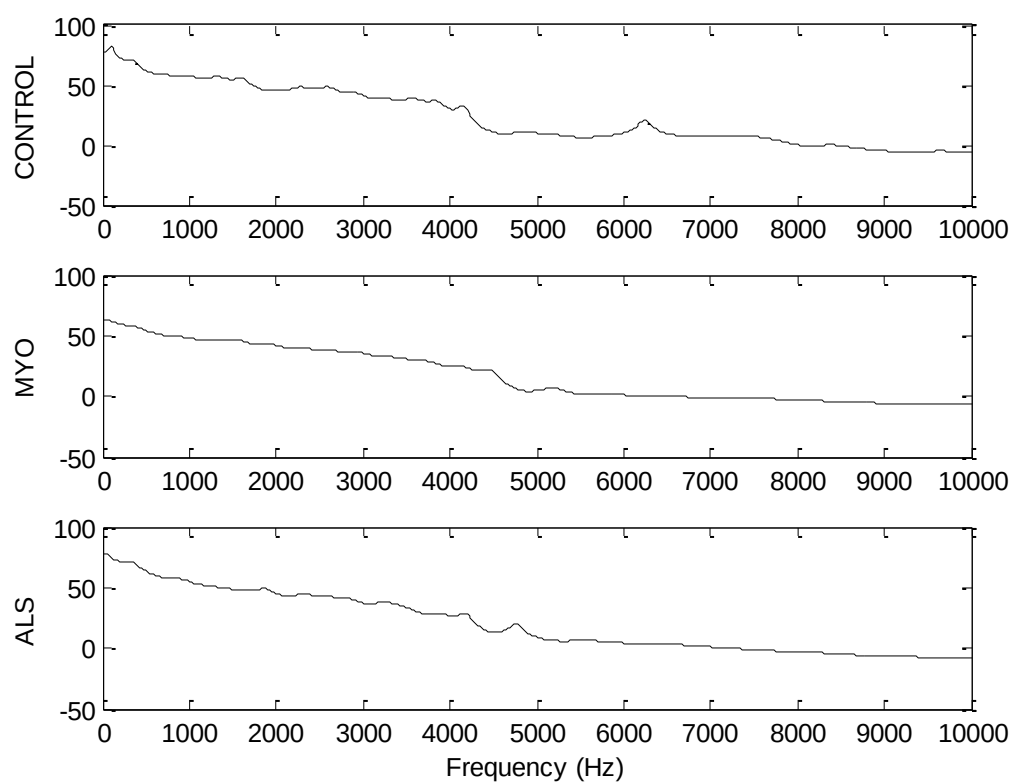


FIGURE A.8 Plotting of signal after EIGEN feature extraction and MSPCA denoising

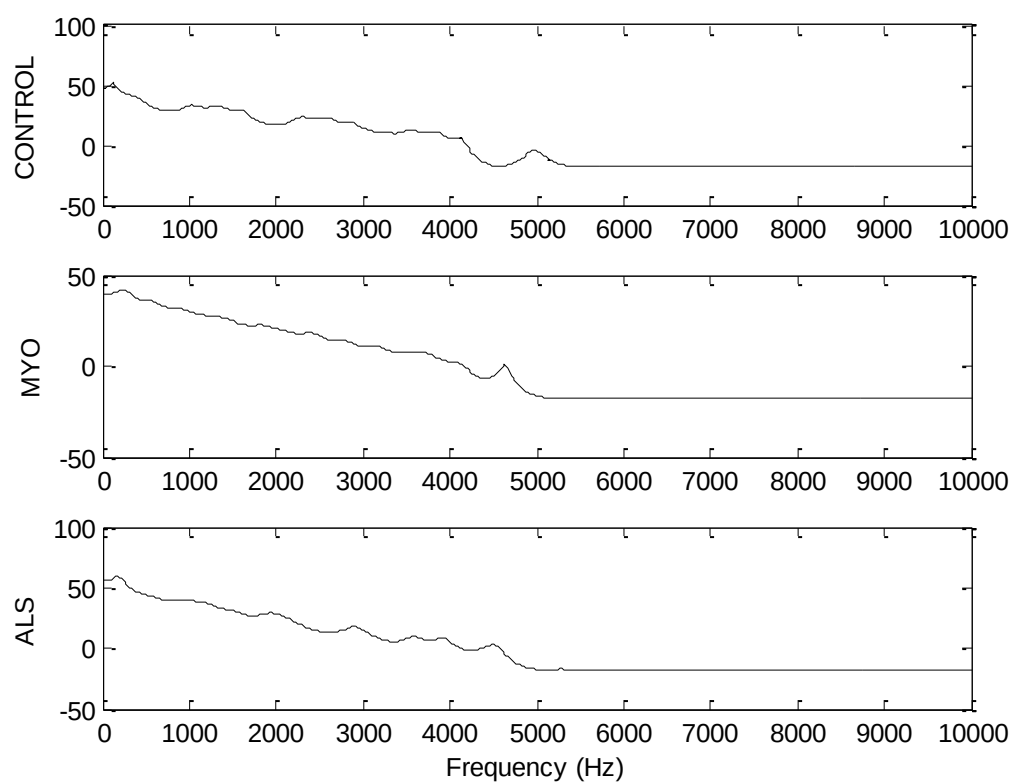


FIGURE A.9 Plotting of signal after MUSIC feature extraction before MSPCA denoising

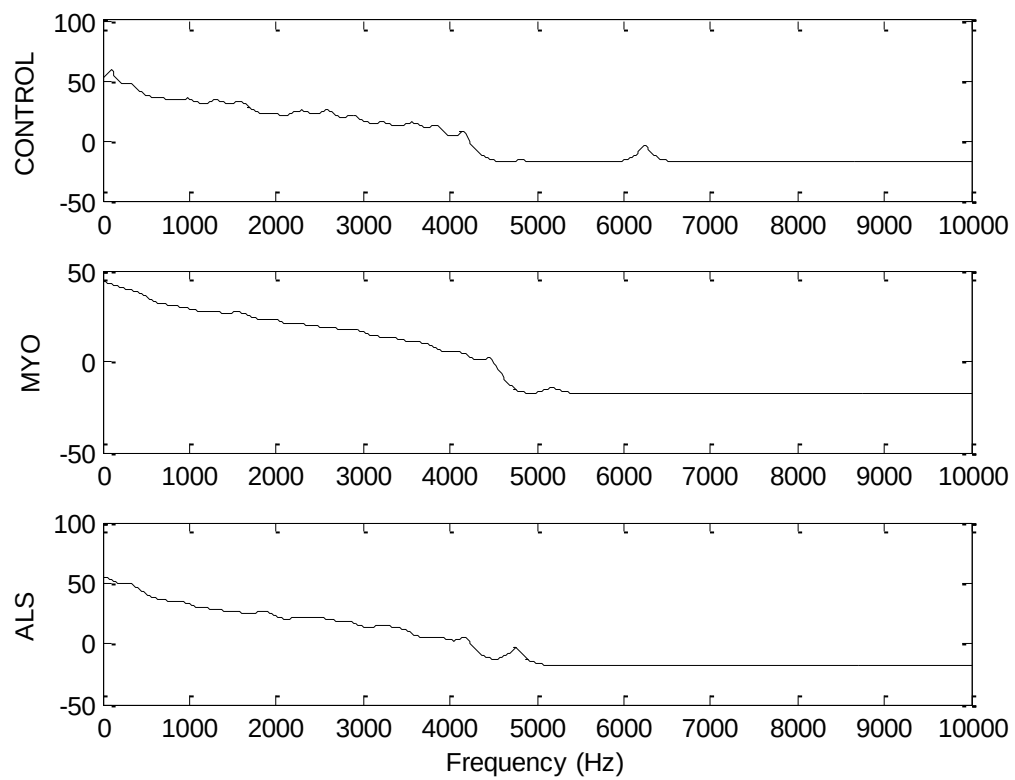


FIGURE A.10 Plotting of signal after MUSIC feature extraction and MSPCA denoising

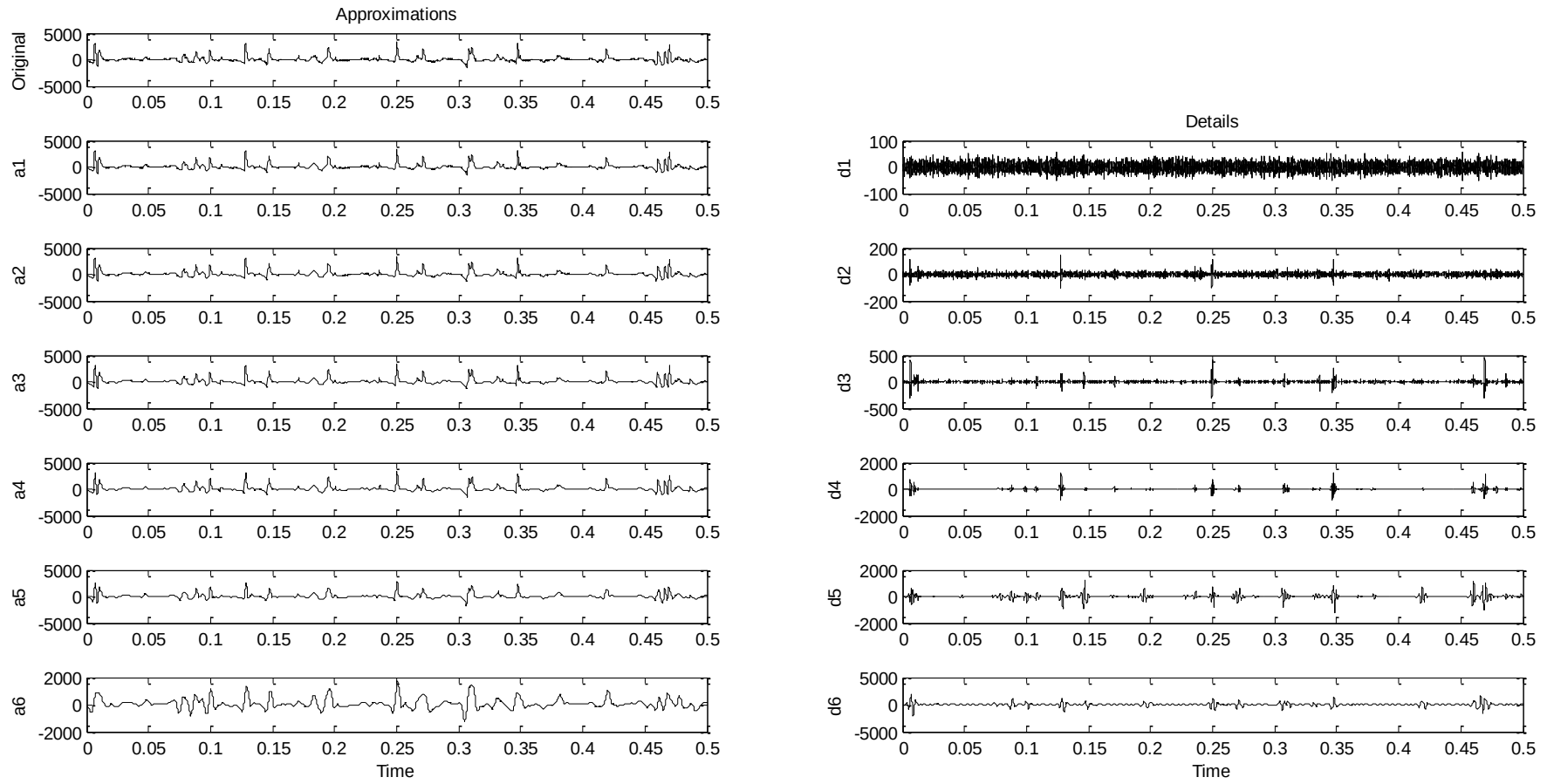


FIGURE A.11 Plotting of signal after DWT feature extraction before MSPCA denoising for control signal

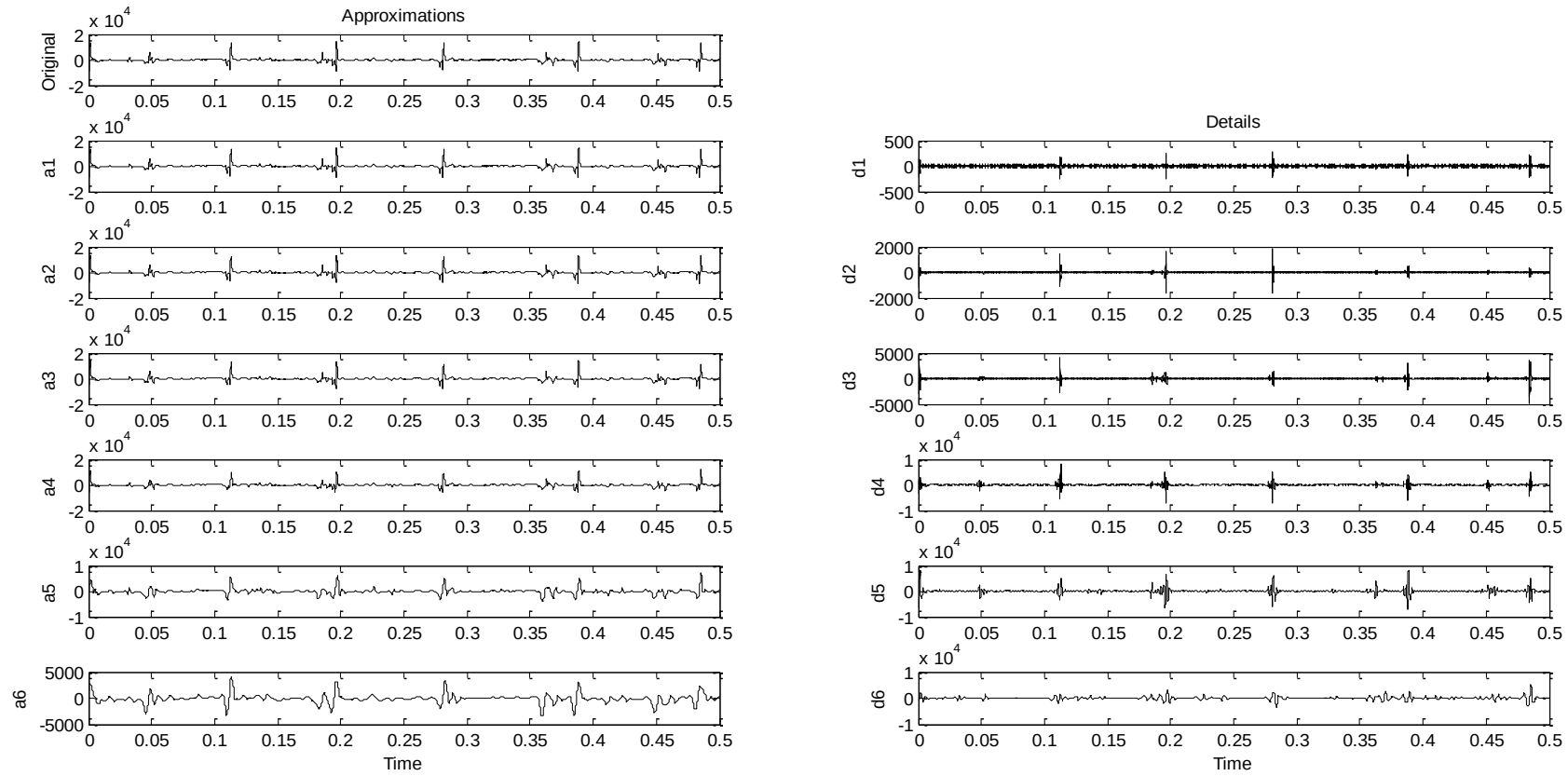


FIGURE A.12 Plotting of signal after DWT feature extraction before MSPCA denoising for als signal

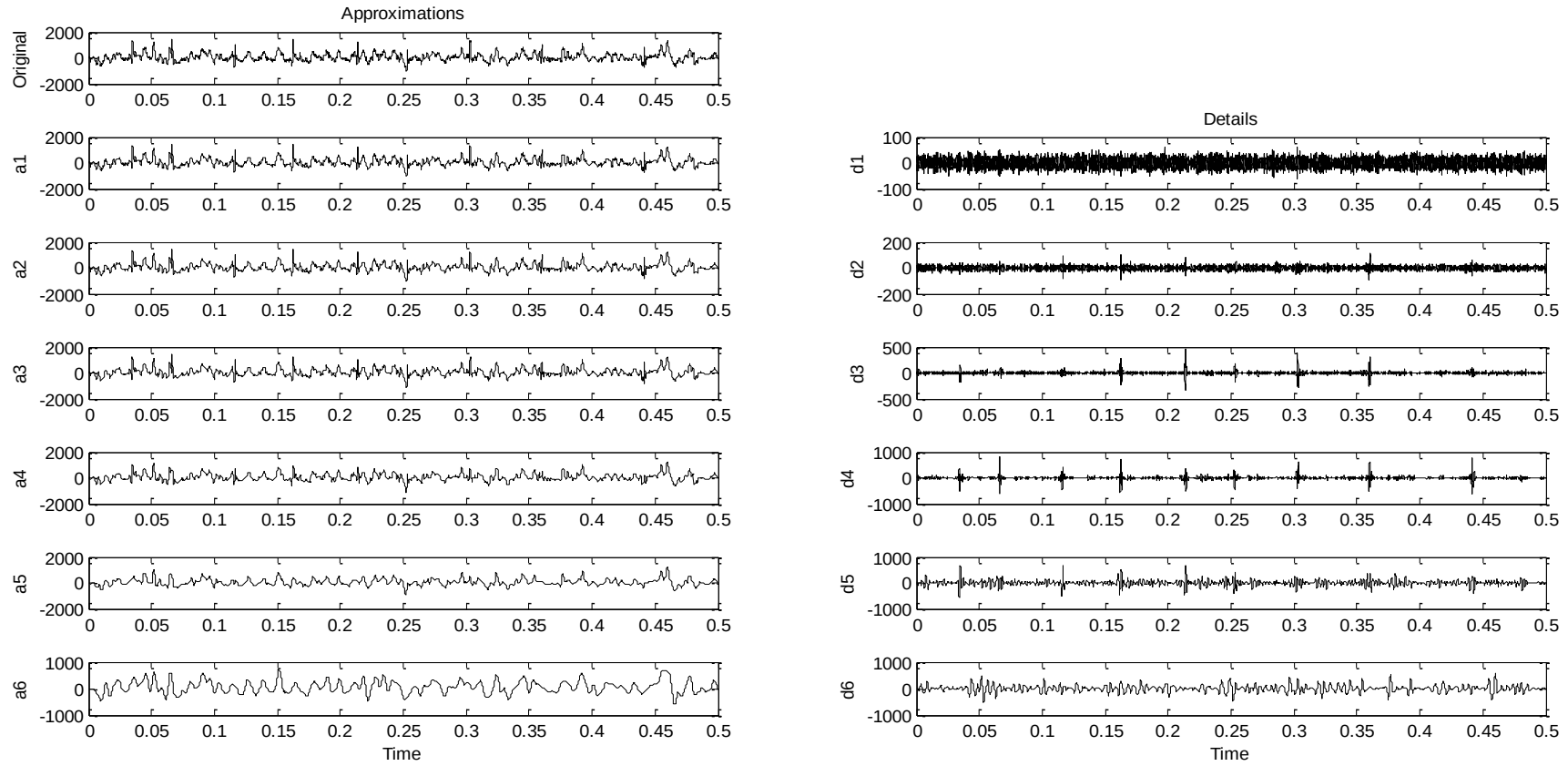


FIGURE A.13 Plotting of signal after DWT feature extraction before MSPCA denoising for myo signal

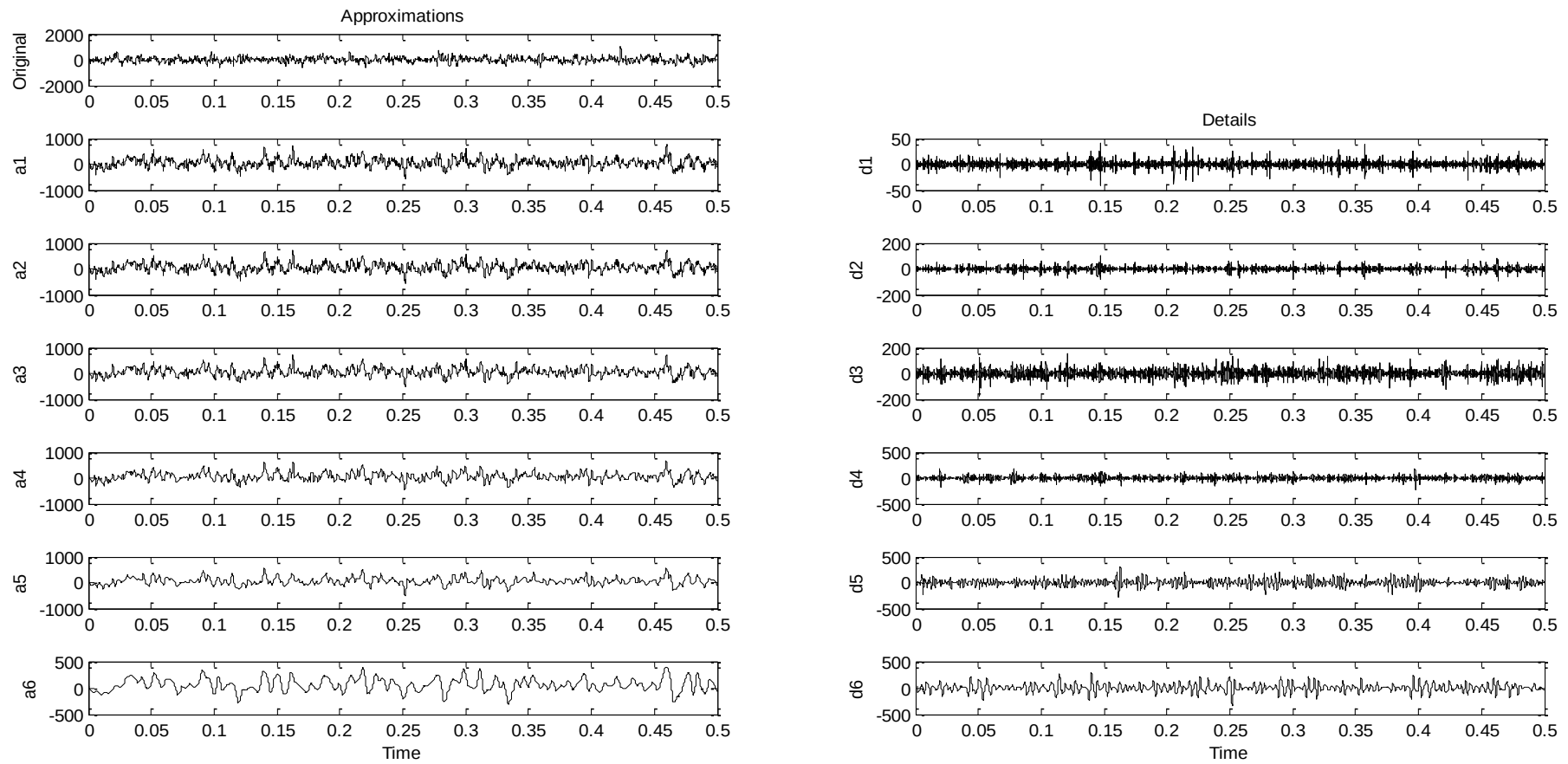


FIGURE A.14 Plotting of signal after DWT feature extraction and MSPCA denoising for myo signal

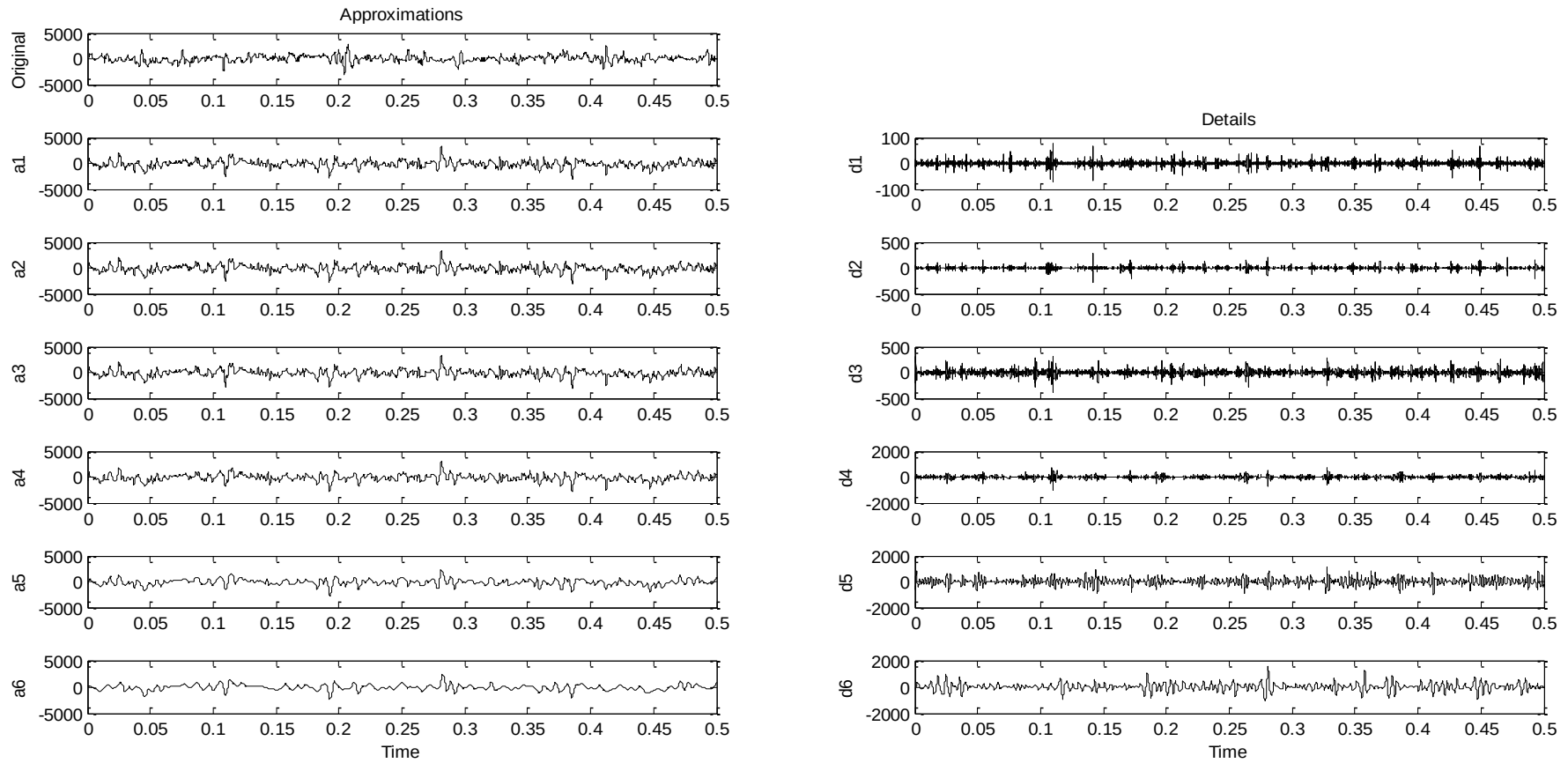


FIGURE A.15 Plotting of signal after DWT feature extraction and MSPCA denoising for als signal

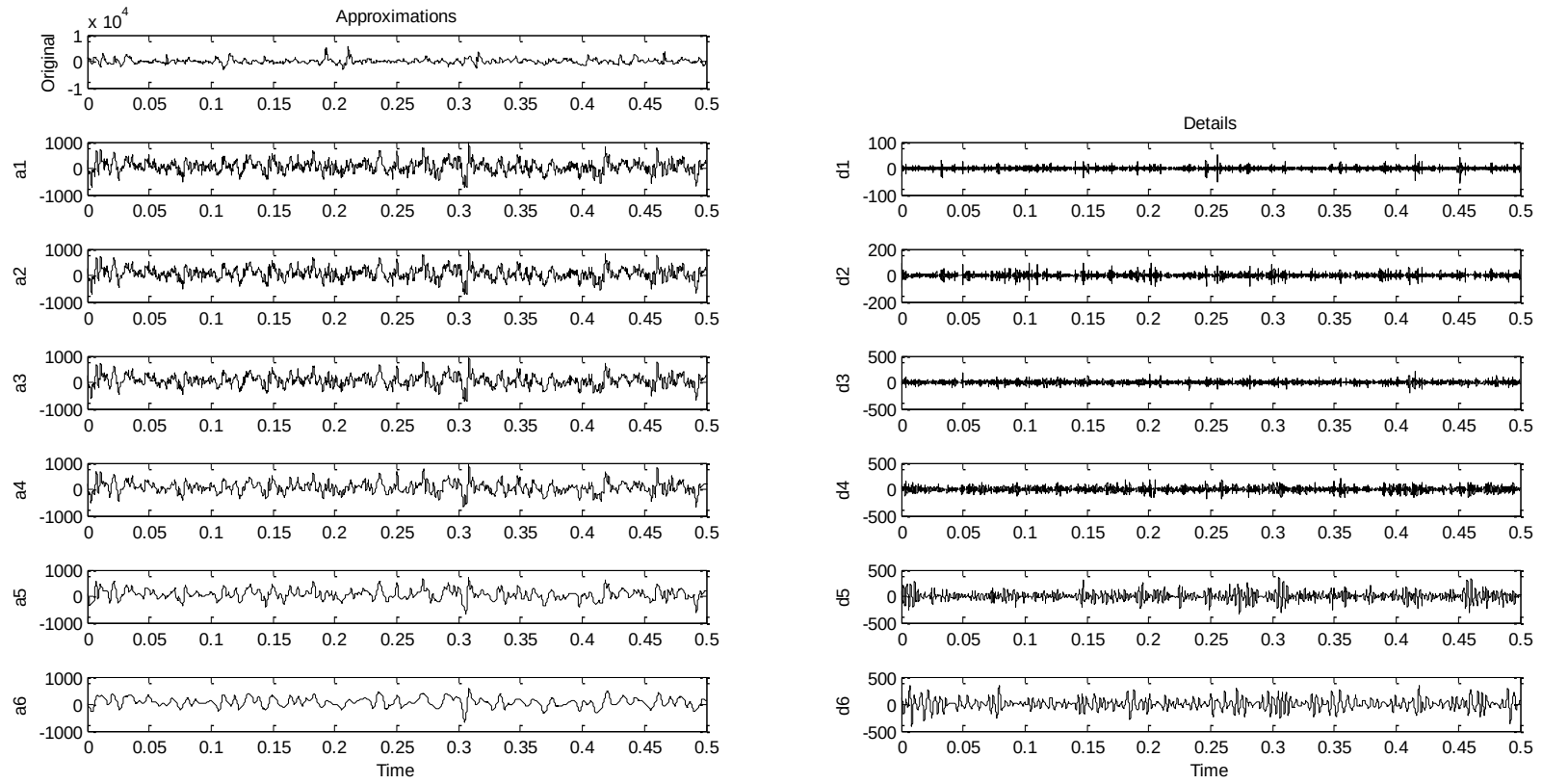


FIGURE A.16 Plotting of signal after DWT feature extraction and MSPCA denoising for control signal

Appendix B. Results and comparison of classifiers for each feature extraction methods before and after denoising

TABLE B.1 Comparison of performance of Classifiers without MSPCA denoising and AR feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	57	58	54.66	27.33	59	14.33	72.6
Sensitivity (Myopathic) (%)	73.33	74.33	70.66	63	60	87.33	17
Sensitivity (ALS) (%)	71.33	76.66	78.33	55	67.66	18	36.6
Total Classification Accuracy (%)	67.22	69.66	67.89	38.77	62.22	40	42

TABLE B.2 Comparison of performance of Classifiers without MSPCA denoising and COV feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	48.33	59	60.66	31.33	57	79.66	91.6
Sensitivity (Myopathic) (%)	73.66	68.66	76.66	56.33	58.33	26.66	6.6
Sensitivity (ALS) (%)	73	76.33	79.66	58	68.33	15.33	24.3
Total Classification Accuracy (%)	65	68	72.33	48.55	61.22	40.55	40.8

TABLE B.3 Comparison of performance of Classifiers without MSPCA denoising and MCOV feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	56.66	57	57.33	21.66	54	5.66	66.3
Sensitivity (Myopathic) (%)	66.33	70.66	73.33	60.33	59.33	21.33	28.3
Sensitivity (ALS) (%)	74.33	77.33	75.33	44.33	68.33	80.66	16.3
Total Classification Accuracy (%)	65.78	68.33	68.67	42.11	59.89	35.88	37

TABLE B.4 Comparison of performance of Classifiers without MSPCA denoising and EIGEN feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	55	55.33	59	26	55	88.66	86.6
Sensitivity (Myopathic) (%)	70.66	64	70.33	64.66	52.66	18.66	12
Sensitivity (ALS) (%)	69	69.66	73.66	65	62.33	38	33
Total Classification Accuracy (%)	64.89	63	67.67	51.88	56.66	48.44	43.8

TABLE B.5 Comparison of performance of Classifiers without MSPCA denoising and MUSIC feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	49	44.67	45.33	59.33	51.33	71.67	50.6
Sensitivity (Myopathic) (%)	64.33	52	59	50.67	51.33	47	37.3
Sensitivity (ALS) (%)	64	63.33	69.67	51.67	49	38.3	44.3
Total Classification Accuracy (%)	59.11	53.3	58	53.89	50.55	52.33	44.1

TABLE B.6 Comparison of performance of Classifiers without MSPCA denoising and DWT feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	61.33	62.66	61	68	64	73.33	46.3
Sensitivity (Myopathic) (%)	68	70.66	77.66	72	63	72	77
Sensitivity (ALS) (%)	76.66	76.33	81.33	79.66	70.66	73.66	62.3
Total Classification Accuracy (%)	68.66	69.89	73.33	73.22	65.89	73	61.8

TABLE B.7 Comparison of performance of Classifiers with MSPCA denoising and AR feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	61.66	63.33	71.33	68	69.33	15	5
Sensitivity (Myopathic) (%)	76.33	79.33	83.33	69.33	74	88	83.3
Sensitivity (ALS) (%)	80.66	79.33	85.33	27	81	10	17
Total Classification Accuracy (%)	72.89	74	80	38.77	74.78	37.66	35.1

TABLE B.8 Comparison of performance of Classifiers with MSPCA denoising and COV feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	71	72.66	77.33	29	74.66	80.66	95.3
Sensitivity (Myopathic) (%)	80.66	83.66	88.33	60.66	80.33	21.33	2.6
Sensitivity (ALS) (%)	83.33	79.33	91	27.33	86.66	5.66	16.3
Total Classification Accuracy (%)	78.33	78.56	85.55	39	80.55	35.89	38.1

TABLE B.9 Comparison of performance of Classifiers with MSPCA denoising and MCOV feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	68	57	74.33	40	77.66	10	66.3
Sensitivity (Myopathic) (%)	81	70.66	88	41	80	90.33	28.3
Sensitivity (ALS) (%)	79	77.33	90.66	16	85.33	9.66	16.3
Total Classification Accuracy (%)	76	68.33	84.33	35.66	81	36.67	37

TABLE B.10 Comparison of performance of Classifiers with MSPCA denoising and EIGEN feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	67.33	67.33	73.66	26.66	76.33	81.66	93
Sensitivity (Myopathic) (%)	82	82	88.66	60.33	86.33	18.33	2.6
Sensitivity (ALS) (%)	86.33	85.33	93	37.66	88	28.66	19.6
Total Classification Accuracy (%)	78.55	78.22	85.11	41.55	83.55	36.22	38.4

TABLE B.11 Comparison of performance of Classifiers with MSPCA denoising and MUSIC feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	88.7	87.7	91.66	96.7	97.3	98.3	63
Sensitivity (Myopathic) (%)	90	90.7	93	94.3	90.3	94.3	88
Sensitivity (ALS) (%)	94	95	97.33	93.7	93.7	94.7	73
Total Classification Accuracy (%)	90.88	91.11	94	94.9	93.8	95.88	75

TABLE B.12 Comparison of performance of Classifiers with MSPCA denoising and DWT feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
Specificity (%)	88.66	86	94.75	93.66	93.66	88	72.6
Sensitivity (Myopathic) (%)	90	89.33	95.66	94	90.33	91.66	88.3
Sensitivity (ALS) (%)	94	93.33	99.58	96.33	97.33	97.66	62.6
Total Classification Accuracy (%)	90.88	89.55	96.67	94.66	93.78	92.44	74.5

TABLE B.13 Comparison of performance measurement of Classifiers with MSPCA denoising and AR feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
F-Measure	0.728	0.739	0.799	0.363	0.748	0.30	0.27
ROC(AUC)	0.813	0.831	0.923	0.587	0.81	0.377	0.51
Total Classification Accuracy (%)	72.89	74	80	38.77	74.77	37.33	35.1

TABLE B.14 Comparison of performance measurement of Classifiers with MSPCA denoising and EIGEN feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
F-Measure	0.784	0.781	0.849	0.408	0.835	0.294	0.29
ROC (AUC)	0.857	0.854	0.946	0.615	0.877	0.362	0.57
Total Classification Accuracy (%)	78.55	78.2	85.11	41.56	83.55	36.22	38.4

TABLE B.15 Comparison of performance measurement of Classifiers with MSPCA denoising and MUSIC feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
F-Measure	0.816	0.821	0.884	0.888	0.813	0.841	0.55
ROC(AUC)	0.875	0.859	0.969	0.963	0.859	0.841	0.71
Total Classification Accuracy (%)	81.56	82.1	88.44	88.89	81.22	84.11	54.9

TABLE B.16 Comparison of performance measurement of Classifiers with MSPCA denoising and DWT feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
F-Measure	0.909	0.896	0.94	0.947	0.938	0.924	0.75
ROC(AUC)	0.94	0.932	0.987	0.986	0.953	0.949	0.9
Total Classification Accuracy (%)	90.89	89.55	94	94.66	93.78	92.44	74.6

TABLE B.17 Comparison of performance measurement of Classifiers with MSPCA denoising and COV feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
F-Measure	0.784	0.787	0.855	0.379	0.806	0.259	0.33
ROC(AUC)	0.877	0.853	0.954	0.588	0.854	0.287	0.58
Total Classification Accuracy (%)	78.33	78.55	85.55	39	80.55	35.89	40.9

TABLE B.18 Comparison of performance measurement of Classifiers with MSPCA denoising and MCOV feature extraction

Statistical Parameters	CART	C4.5	Random Forest	ANN	k-NN	SVM	RBF
F-Measure	0.76	0.795	0.842	0.353	0.811	0.368	0.31
ROC(AUC)	0.833	0.858	0.949	0.557	0.858	0.257	0.54
Total Classification Accuracy (%)	76	79.44	84.33	35.66	81	36.67	38.1

TABLE B.19 Confusion matrix of CART classifier with MUSIC feature extraction before and after MSPCA denoising

	CART					
	ALS		MYO		NORMAL	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
ALS	192	264	53	10	55	26
MYO	20	9	193	239	87	52
NORMAL	51	24	102	45	147	231

TABLE B.20 Confusion matrix of ANN classifier with MUSIC feature extraction before and after MSPCA denoising

	ANN					
	ALS		MYO		NORMAL	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
ALS	155	290	64	4	81	6
MYO	42	5	152	255	106	40
NORMAL	36	4	86	32	178	255

TABLE B.21 Confusion matrix of k-NN classifier with MUSIC feature extraction before and after MSPCA denoising

	k-NN					
	ALS		MYO		NORMAL	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
ALS	147	268	66	9	87	23
MYO	63	10	154	227	83	63
NORMAL	55	16	91	48	154	236

TABLE B.22 Confusion matrix of C4.5 classifier with MUSIC feature extraction before and after MSPCA denoising

	C4.5					
	ALS		MYO		NORMAL	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
ALS	190	269	55	7	55	24
MYO	48	11	156	239	96	50
NORMAL	68	24	98	45	134	231

TABLE B.23 Confusion matrix of Random Forest classifier with MUSIC feature extraction before and after MSPCA denoising

	Random Forest					
	ALS		MYO		NORMAL	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
ALS	209	284	43	2	48	14
MYO	45	4	177	269	78	27
NORMAL	73	19	91	38	136	243

TABLE B.24 Confusion matrix of RBF classifier with MUSIC feature extraction before and after MSPCA denoising

	RBF					
	ALS		MYO		NORMAL	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
ALS	133	166	76	30	91	104
MYO	79	41	112	156	109	103
NORMAL	64	56	84	72	152	172

TABLE B.25 Confusion matrix of SVM classifier with MUSIC feature extraction before and after MSPCA denoising

	SVM					
	ALS		MYO		NORMAL	
	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA	Without MSPCA	With MSPCA
ALS	115	272	68	6	117	22
MYO	15	9	141	244	144	47
NORMAL	14	16	71	43	215	241

TABLE B.26 AUC results for different classifiers with MSPCA denoising and ARBURG feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.719	0.761	0.877	0.763	0.529	0.493	0.514
Myopathic	0.856	0.854	0.94	0.81	0.593	0.564	0.51
ALS	0.863	0.878	0.953	0.859	0.638	0.539	0.516
Total	0.813	0.831	0.923	0.81	0.587	0.532	0.514

TABLE B.27 AUC results for different classifiers with MSPCA denoising and COV feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.815	0.786	0.921	0.805	0.536	0.499	0.539
Myopathic	0.906	0.898	0.967	0.852	0.599	0.543	0.526
ALS	0.911	0.876	0.974	0.906	0.631	0.524	0.597
Total	0.877	0.853	0.954	0.854	0.588	0.522	0.554

TABLE B.28 AUC results for different classifiers with MSPCA denoising and MCOV feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.763	0.804	0.914	0.823	0.536	0.511	0.526
Myopathic	0.867	0.889	0.965	0.854	0.55	0.526	0.532
ALS	0.87	0.88	0.968	0.896	0.586	0.531	0.573
Total	0.833	0.858	0.949	0.858	0.557	0.523	0.544

TABLE B.29 AUC results for different classifiers with MSPCA denoising and EIGEN feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.776	0.794	0.908	0.833	0.555	0.506	0.536
Myopathic	0.891	0.871	0.95	0.882	0.614	0.542	0.559
ALS	0.903	0.897	0.98	0.916	0.675	0.546	0.611
Total	0.857	0.854	0.946	0.877	0.615	0.531	0.568

TABLE B.30 AUC results for different classifiers with MSPCA denoising and MUSIC feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.923	0.837	0.995	0.822	0.838	0.849	0.65
Myopathic	0.962	0.864	0.996	0.831	0.853	0.887	0.725
ALS	0.984	0.936	1	0.925	0.934	0.96	0.773
Total	0.956	0.879	0.997	0.859	0.875	0.899	0.716

TABLE B.31 AUC results for different classifiers with MSPCA denoising and DWT feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.917	0.9	0.981	0.94	0.983	0.921	0.834
Myopathic	0.937	0.941	0.984	0.942	0.98	0.958	0.96
ALS	0.967	0.954	0.994	0.978	0.996	0.967	0.902
Total	0.94	0.932	0.987	0.953	0.986	0.949	0.898

TABLE B.32 F-measure of Different Classifiers with MSPCA denoising and ARBURG feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.627	0.652	0.733	0.685	0.25	0.203	0.09
Myopathic	0.766	0.771	0.82	0.746	0.477	0.52	0.486
ALS	0.791	0.793	0.845	0.813	0.362	0.176	0.233
Total	0.728	0.739	0.799	0.748	0.363	0.3	0.27

TABLE B.33 F-measure of Different Classifiers with MSPCA denoising and COV feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.706	0.708	0.808	0.739	0.315	0.471	0.513
Myopathic	0.82	0.841	0.865	0.802	0.457	0.282	0.051
ALS	0.824	0.811	0.891	0.877	0.364	0.107	0.226
Total	0.784	0.787	0.855	0.806	0.379	0.287	0.276

TABLE B.34 F-measure of Different Classifiers with MSPCA denoising and MCOV feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.679	0.739	0.787	0.761	0.371	0.058	0.269
Myopathic	0.811	0.834	0.857	0.807	0.377	0.507	0.504
ALS	0.79	0.811	0.882	0.863	0.312	0.206	0.269
Total	0.76	0.795	0.842	0.811	0.353	0.257	0.311

TABLE B.35 F-measure of Different Classifiers with MSPCA denoising and EIGEN feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.695	0.694	0.782	0.779	0.305	0.477	0.506
Myopathic	0.808	0.803	0.854	0.837	0.464	0.249	0.051
ALS	0.849	0.846	0.912	0.89	0.457	0.158	0.308
Total	0.784	0.781	0.849	0.835	0.408	0.294	0.288

TABLE B.36 F-measure of Different Classifiers with MSPCA denoising and MUSIC feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.878	0.875	0.952	0.75	0.849	0.9	0.507
Myopathic	0.908	0.916	0.958	0.799	0.863	0.925	0.559
ALS	0.942	0.942	0.99	0.902	0.954	0.951	0.59
Total	0.909	0.911	0.967	0.817	0.888	0.925	0.552

TABLE B.37 F-measure of Different Classifiers with MSPCA denoising and DWT feature extraction methods.

	CART	C4.5	Random Forest	k-NN	ANN	SVM	RBF
Normal	0.878	0.856	0.914	0.914	0.927	0.899	0.664
Myopathic	0.908	0.901	0.939	0.93	0.946	0.94	0.885
ALS	0.942	0.93	0.967	0.97	0.967	0.933	0.691
Total	0.909	0.896	0.94	0.938	0.947	0.924	0.747

CURRICULUM VITAE

Ercan GOKGOZ was born on 04 October 1987, in Samsun, Turkey. He received his BS and M.Sc. degree in 2011 from Computer Engineering Department of Fatih University, Turkey. He is currently working as a research assistant at the Information Technologies Department of International Burch University since 2011. His research interests include machine learning, biomedical signal processing and classification.