MINING SIGNIFICANT FEATURES SET FOR PREDICTING ISOTONIC MUSCLE FATIGUE USING SURFACE ELECTROMYOGRAPHY (sEMG) SIGNAL

KU MAN YI

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

BORANG PENGESAHAN STATUS TESIS

JUDUL: MINING SIGNIFICANT FEATURES SET FOR PREDICTING ISOTONIC MUSCLE FATIGUE USING SURFACE ELECTROMYOGRAPHY (sEMG) SIGNAL

SESI PENGAJIAN: 2014/2015

Saya KU MAN YI
(HURUF BESAR)

mengaku membenarkan tesis (PSM/~~Sarjana/Doktor Falsafah)~~ ini disimpan di Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dengan syarat-syarat kegunaan seperti berikut:

1. Tesis dan projek adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan Fakulti Teknologi Maklumat dan Komunikasi dibenarkan membuat salinan tesis ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. ** Sila tandakan (/)

________	SULIT	(Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)
________	TERHAD	(Mengandungi maklumat TERHAD yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)
________	TIDAK TERHAD	

____________________	____________________
(TANDATANGAN PENULIS)	(TANDATANGAN PENYELIA)
Alamat tetap: 9, Jalan Dato Senu 28, Taman Sentul Jaya, 51000, Kuala Lumpur.	PM. DR. CHOO YUN HUOY Nama Penyelia
Tarikh : ____________	Tarikh : ____________

CATATAN: *Tesis dimaksudkan sebagai Laporan Akhir Projek Sarjana Muda (PSM)
** Jika tesis ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa.

MINING SIGNIFICANT FEATURES SET FOR PREDICTING ISOTONIC MUSCLE FATIGUE USING SURFACE ELECTROMYOGRAPHY (sEMG) SIGNAL

KU MAN YI

This report is submitted in partial fulfilment of the requirements for the
Bachelor of Computer Science (Artificial Intelligence)

FACULTY OF INFORMATION AND COMMUNICATION TECHNOLOGY
UNIVERSITY TEKNIKAL MALAYSIA MELAKA
2015

DECLARATION

I hereby declare that this project report entitled

MINING SIGNIFICANT FEATURES SET FOR PREDICTING ISOTONIC MUSCLE FATIGUE USING SURFACE ELECTROMYOGRAPHY (sEMG) SIGNAL

is written by me and is my own effort and that no part has been plagiarized without citations.

STUDENT : ___________________________ Date: ________________

(KU MAN YI)

SUPERVISOR : ___________________________ Date: ________________

(PM. DR. CHOO YUN HUOY)

DEDICATION

To my beloved parents, friends, supervisor and evaluator.

ACKNOWLEDGEMENT

Firstly, I would like to express my special thanks of gratitude to my supervisor PM. Dr. Choo Yun Huoy for continuous support in this research, for her patience, motivation, enthusiasm and immense knowledge. I would like to thank you for inspiring and willingness to motivate me to contribute tremendously to this project. One simply could not wish for a better or friendlier supervisor.

My sincere thanks also goes to Dr. Zeratul Izzah Mohd. Yusoh for her evaluation and sincere comments that help in improving this research and assistance in writing report. I would also immensely grateful to my seniors Nur Shidah, Soh Jia Shan and Abu Bakar Yahya for provided me the valuable suggestions and encouragement throughout the project.

Honorable mention also goes to my university friends for their understandings and care, also for the sleepless nights we were working together and all the fun we have. Most importantly, I would like to thank my family for supporting me spiritually throughout my life.

ABSTRACT

Surface electromyography (sEMG) has been widely used in sport science as the muscular biofeedback in order to design the suitable training to increase their muscle endurance based on the signal feedback. However, the research of muscle fatigue caused by isotonic training receives less attention due to higher signal noise than isometric data. Good representative features are known to reduce signal noise. However, there exists no standard significant features set for isotonic training data. Therefore, this study aims to propose the significant features set for isotonic muscle fatigue. Correlation based feature subset selection (CFS) was used to select the significant features set. The commonly used features in sEMG analysis are Mean Absolute Value (MAV), Modified Mean Absolute Value 1 and 2 (MMAV1, MMAV2), Root Mean Square (RMS), Simple Square Integral (SSI), Variance of EMG (VAR), Waveform Length (WL), Mean Frequency (MF) and Median Frequency (MDF). Support Vector Machine was to measure the performance of features in prediction accuracy. Shapiro-Wilk normality test and ANOVA were performed to further verify the experimental results. Frequency domain features showed higher accuracy than time domain features, however features set with both domains showed better accuracy than a single domain. The MAV, RMS, MF and MDF features set is similar to full nine features set in performance. At the same time, CFS has selected two features which is MF and MDF. However, this deducts set is significantly worse than MAV, RMS, MF and MDF features set. Thus, features MAV, RMS, MDF and MF are proposed as the significant features set for isotonic muscle fatigue. For future improvement, research can be tested on different body muscle and more aggressive isotonic training.

ABSTRAK

Surface Electromyography (sEMG) telah digunakan secara meluas dalam bidang sains sukan untuk merangka latihan yang sesuai untuk meningkatkan daya tahan otot berdasarkan maklum balas isyarat. Walau bagaimanapun, penyelidikan di dalam keletihan otot yang disebabkan oleh latihan isotonik menerima kurang perhatian kerana terdapat gangguan pada isyarat yang lebih tinggi daripada data isometrik. Features yang baik dapat mengurangkan gangguan dalam isyarat. Walau bagaimanapun, masih tidak ada sebarang features set yang standard bagi data latihan isotonik. Oleh itu, kajian ini bertujuan untuk mencadangkan features set yang penting untuk keletihan otot isotonik. Correlation based feature subset selection (CFS) telah digunakan untuk memilih features yang penting. Features yang biasa digunakan dalam analisis sEMG adalah Mean Absolute Value (MAV), Modified Mean Absolute Value 1 and 2 (MMAV1, MMAV2), Root Mean Square (RMS), Simple Square Integral (SSI), Variance of EMG (VAR), Waveform Length (WL), Mean Frequency (MF) dan Median Frequency (MDF). Support Vector Machine adalah untuk mengukur prestasi ciri-ciri dalam ketepatan ramalan. Shapiro-Wilk normality test dan ANOVA telah dijalankan bagi mengesahkan lagi keputusan eksperimen ini. Frequency domain features menunjukkan ketepatan yang lebih tinggi daripada time domain features, namun set ciri daripada kedua-dua domain menunjukkan ketepatan yang lebih baik daripada satu domain. The MAV, RMS, MF dan MDF features set mempunyai persamaan di dalam prestatsi apabila dibandingkan dengan sembilan features yang yang lain. Pada masa yang sama, CFS telah memilih dua features iaitu MF dan MDF. Walau bagaimanapun, deduct set ini adalah lebih teruk apabila dibandingkan dengan daripada features set MAV, RMS, MF dan MDF. Oleh itu, features set MAV, RMS, MDF dan MF dicadangkan sebagai features penting yang ditetapkan untuk keletihan otot isotonik. Untuk penambahbaikan pada masa hadapan, penyelidikan boleh diuji pada otot badan yang berbeza serta berlainan dan bagi latihan isotonik yang lebih agresif.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AR	-	Autoregressive Coefficient
ARV	-	Average Rectified Value
ANOVA	-	Analysis of Variance
CfsSubsetEval	-	Correlation based Feature Subset Selection
EMG	-	Electromyography
FFT	-	Fast Fourier Transform
iEMG	-	Integrated Electromyography
MAV	-	Mean Absolute Value
MDF	-	Median Frequency
MF	-	Mean Frequency
MMAV	-	Modified Mean Absolute Value
RMS	-	Root Mean Square
sEMG	-	Surface Electromyography
SSC	-	Slope Sign Change
SSI	-	Simple Square Integral
STFT	-	Short Time Fourier Transform
SVM	-	Support Vector Machine
VAR	-	Variance of EMG
WAMP	-	Willison Amplitude
WL	-	Waveform Length
WPT	-	Wavelet Packet Transform
ZCR	-	Zero Crossing Rate

CHAPTER I

INTRODUCTION

1.1 Project Background

In the area of sport science, Electromyography (EMG) signal has been used in many field including biomechanics, movement analysis, athlete's strength training and sports rehabilitation (Komrad, 2006). Isotonic training is the muscle contraction that caused by dynamic movement, which is commonly used for athlete training. Therefore, it is essential to measure and analyse athlete muscle condition to arrange the suitable training specifically for each of them. When the muscle is continuously strengthening and contracting in a period of time, this will cause muscle fatigue followed by their motion will become slower and weaker. This happens because the person muscle unable to resists from the fatigue due to the continuously of muscle contraction. Therefore, it is essential to get the solution to measure and distinguish the isotonic muscle fatigue caused during training.

Electromyography (EMG) is the study of muscle function through the inquiry of the myoelectric activity which is the electrical impulses of muscle at rest or during contraction. Therefore, EMG can be established as an evaluation tool in many areas, such as sport science, medical research, rehabilitation and ergonomics. There are two types of EMG, which are surface EMG (sEMG) and integrated EMG (iEMG). sEMG records the muscle activity from the skin surface while iEMG records by placing needle electrode into the muscle. sEMG method is more safe and comfortable to the subject during the collection of the muscle signal. Besides, sEMG is more suitable for dynamic movement in training because the placement of electrode is on the skin surface only and subject is free to move without getting hurt. In this project, it will be focused on using sEMG signal as the muscular biofeedback.

sEMG is more common way of measurement, because it is non-invasive and can be conducted by personal, which cause no risk to the subject. Small electrical currents are generated by muscle fibres prior to the production of muscle force. These currents are generated by the exchange of ions across muscle fibre membranes, and the main part of the signalling process is obtain from the contraction of muscle fibres. Same as other electrophysiological signals, sEMG signal is small and contain noises so the signal recorded with electrode will not reflect the total and complete amount of motor units firing. As a solution, the sEMG signal first picked up by electrode have to undergo pre-processing, such as amplification, noise removal and others to get the more accurate result. Normally the current devices in collecting the sEMG signal is embedded with the pre-processing techniques in order to enable the sEMG signal is ideal for use.

In this research, it focus on introduce the significant feature extraction methods from the sEMG signal during isotonic training. Feature extraction transforms raw signals into more informative values (features vector) and remove the redundancy from an initial set of measured signals. The features can be used as the input data in classifier or modelling in order to predict the fatigue characteristics from the muscle signal. There are three types of EMG features in different domains, which are time domain, frequency domain and time-frequency domain features (Chowdhury et al., 2013). Therefore, the common used features within these domains used in sEMG such as Root Mean Square (RMS), Mean Absolute Value (MAV), and other features will be implemented and discussed in this research.

Besides, classification also needed to be implemented in order to evaluate the performance of the chosen features. In this research, multiclass Support Vector Machine (SVM) is used as the classifier and gives the recognition result from the extracted data. SVM is more generalized among sEMG classifier for the non-stationary and time-variable properties of sEMG signal. SVM classifier gave a very good accurate rate, which can up to 99% of accuracy (Alkan and Günay, 2012).

Following that, normality test and statistical test are used in result analysis. T-test and ANOVA are used for statistical test to identify the significant difference in the comparison of features or domains. These tests are important in proposing the significant features set at the end of the research. Nevertheless, feature selection also used to select the significant features set by eliminate the redundancy or less important features. Statistical test is tested again in this phase to ensure the accuracy after selection is maintained and not deviated too much.

At the end of the research, the significant features set for sEMG signal will be proposed through the evaluation. Therefore, the results can be used as the reference in the future for relevant research.

1.2 Problem Statements

EMG signal is commonly used in motion analysis and rehabilitation purpose, however it is not so common that the implementation of EMG in the area of sport science. Biofeedback are essential to be analyzed so that the suitable training can be rephrase to them, however there are no standard features used for dynamic training. EMG signal for the isotonic training which involve dynamic movement is more challenging because the instability and contains more noises. The raw sEMG signal is nearly meaningless, so feature extraction can be used in obtained the useful information within the raw sEMG signal. However, there are various kind of feature extraction methods, so the research will be study on the suitable feature extraction methods which are best fit with the problem.

1.3 Objectives

The project embarks on the following objectives:

1. To compare various feature extraction methods for dynamic movement data.
2. To propose a significant feature subset for dynamic movement in isotonic training.
3. To evaluate the outcome of the proposed feature methods.

1.4 Scope

In this project, it will be focused on several of scopes to fulfil my requirement to achieve the project objectives. Firstly, EMG can be categorized as sEMG and iEMG, the project will only focus on sEMG which is collecting the EMG signal from skin surface. Besides that, sEMG has been widely applied in many fields, therefore the field that focus in this research is sport science. Sport science involves various kind of training to enhance the sport men performance, and the training that selected in this project is isotonic training which involved the dynamic movement. The continuous dynamic movement causes a person decrease in performance due to the muscle do not able to resist to the fatigue, so the research aims to distinguish the fatigue and non-fatigue sEMG signal to design better training for sport men. In this studies, the commonly used existing feature extraction methods will be implemented and evaluated in order to propose a set of suitable features that best fit with the problem. All the feature extraction methods choose are basically from two domains, which are time domain and frequency domain and will be implemented in Matlab. Also, a classification technique, LibSVM which embedded in software WEKA is used to evaluate the performance during comparison of features and domains. Feature selection method, cfsSubsetEval which also embedded in WEKA also used. Last but not least, normality test and statistical test are carried out to make the analysis of the results acquired. For normality test, Shapiro-Wilk test is used because it is suitable for small samples in this research. Other than Shapiro-Wilk test, some ordinary methods like skewness and kurtosis, and graphical method also used for additional reference. T-test and ANOVA which are the statistical tests also used for identify the significant difference within the samples. At the end, a set of significant feature set is proposed.

1.5 Project Significance

The purpose of the project is to propose and evaluate the existing feature extraction methods toward the sEMG signal. The feature extraction methods will be applied into the sEMG fatigue dataset that been collected by a post-graduate UTeM student. The proposed set of features are specific for the condition during isotonic muscle fatigue happened. Therefore, the proposed features at the end of the research can be used as reference during analysing sEMG signal from isotonic training. Besides that, the proposed set of features can be beneficial to the future study in developing or improving some computational models which uses raw sEMG signal as input, such as Variable-load Intensity Model. In addition, other researches also able to use the research findings to make more contribution researches by using sEMG signal.

1.6 Expected Output

The project is expected to implement all the proposed features into the fatigue dataset that used for experiment and adapted to the problem. In order to implement the feature extraction methods, the methods coding had to be coded or modified in the Matlab software. Also, the pattern of features for fatigue and non-fatigue signal also had to be plotted and able to make the differentiation. Analysis had to be done to explain the reason of the changes of signal or features. At the end of the project, it is also expected to propose a set of features which specially used in sEMG isotonic muscle fatigue signal with the best accuracy with the least features required.

1.7 Summary

As a conclusion of this chapter, the overview of the project is discussed. The entire project will be carried on according to the scopes in order to achieve the project objectives. The previous studies on the problem and methods will be discussed in the next chapter.

CHAPTER II

LITERATURE REVIEW

2.1 Introduction

In this chapter, literature review on feature extraction methods for sEMG to prolong muscle endurance have been studied. This chapter contains seven sections. The second section will briefly introduce the sEMG and the signal it generated. The third section will narrow down the scope to the area of sport science which also explain the relationship of sEMG signals with muscle endurance and the trainings in sports. Then, the forth section provides the literature review on the fatigue happened during muscle contraction and how it affect the changes of the sEMG signal. Following that, the fifth section presents the literature review of the feature extraction techniques used in order to make the selection of the features for the project experimentation. The sixth section will discuss about classifier SVM to measure the features performance and it's method to use while the seventh section is discuss about the feature selection method. In addition, the eighth section explains the research for statisticak analysis. Lastly, the conclusion is summarized for the entire literature review chapter and determined the methods used in this project experimentation.

2.2 Surface Electromyography (sEMG)

EMG is the study of muscle function through analysis of electrical potential that emanate from muscle itself. Nowadays, EMG become common in acts as biofeedback in many biomedical and clinical application. EMG also established as an evaluation tool that applied in many areas, such as Medical Research, Ergonomics, Rehabilitation and Sport Science and others (Komrad, 2006). The project focus on the area of Sport Science. The study of EMG had initially discussed at early as 17th century (Basmajian and De Luca, 1985). They proposed that the electrical potential from muscle recorded for EMG is actually superposition of action potentials acting on fibre muscles. In shorts, EMG reflects the pattern of muscle contraction. Representation of electrical potential in form of time varying signal is called as EMG signal.

The EMG signal is based on the action potentials at the muscle fibre membrane during depolarization and repolarization processes. The ionic difference between the inner and outer spaces of muscle fibre membrane forms a resting potential which approximately -80 to -90 mV when without contraction. Resting potential is maintained by the ion pumping which negative intracellular charge at the inner space of muscle. During contraction, which conducts the excitation along the motor nerve makes the Na+ ions flow in which causes a membrane depolarization. Then, repolarization will happened immediately to restore by backward exchange of ions within the active ion pump mechanism. These are the reason EMG signals fluctuated between positive and negative values which shown in the red line at the figure below.

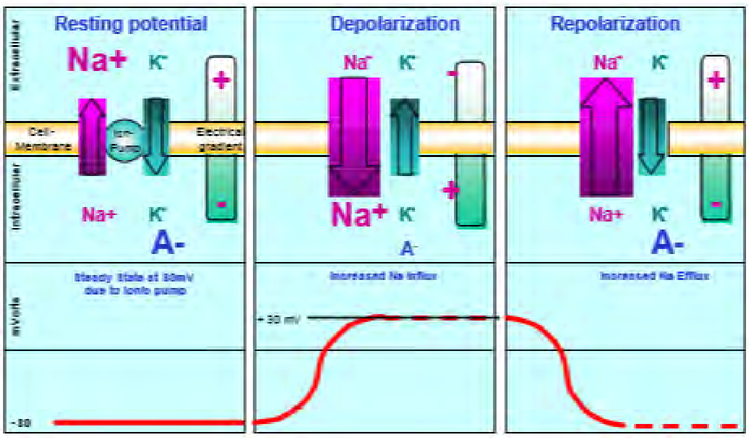

Figure 2 1 : The action potential acting on muscles.