FEATURE EXTRACTION AND CLASSIFICATION ALGORITHMS FOR ASSESSING MUSCLE CONDITION USING MULTI SENSORS

by

MOHAMED SARILLEE B. MOHAMED @ MOHD KASSIM

(1431311323)

A thesis submitted in fulfillment of the requirements for the degree of Master of Science in Biomedical Electronic Engineering

School of Mechatronic

UNIVERSITI MALAYSIA PERLIS

2016

UNIVERSITI MALAYSIA PERLIS

DECLARATION OF THESIS

Author's full name	:	Mohamed Sarillee B. Mohamed @ Mohd Kassim
Date of birth	:	29/12/1989
Title	:	Feature Extraction and Classification Algorithms for Assessing Muscle Condition using Multi Sensors
Academic Session	:	2014/2015

I hereby declare that the thesis becomes the property of Universiti Malaysia Perlis (UniMAP) and to be placed at the library of UniMAP. This thesis is classified as :

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*

☐ **RESTRICTED** (Contains restricted information as specified by the organization where research was done)*

☐ **OPEN ACCESS** I agree that my thesis is to be made immediately available as hard copy or on-line open access (full text)

I, the author, give permission to the UniMAP to reproduce this thesis in whole or in part for the purpose of research or academic exchange only (except during a period of _____ years, if so requested above).

Certified by:

______________________	______________________
SIGNATURE	**SIGNATURE OF SUPERVISOR**
891229-10-5567	**DR. HARIHARAN MUTHUSAMY**
(NEW IC NO. / PASSPORT NO.)	**NAME OF SUPERVISOR**
Date ______________________	Date : ______________________

NOTES: * If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization with period and reasons for confidentially or restriction.

ACKNOWLEDGEMENT

I wish to thank God the Almighty, without his consent, it would be impossible to achieve what had been done in this work. I would also like to thank my supervisors Dr.Hariharan Muthusamy and En.Anas Mohd Noor for their help and supervision in conducting this research. I greatly appreciate their care, dedication and constructive criticism. I have truly benefited and enjoyed working with them.

I would like to thank the management and authorities of Universiti Malaysia Perlis for the generous financial support, providing different administrative and facilities throughout School of Mechatronic Engineering.

I would like to express my sincere appreciation to my friends for their support during the progress of this work.

I also would like to extend my appreciation to family members for they understanding and patience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLE

LIST OF ABBREVIATIONS

AIF	Average Instantaneous Frequency
ALC	Anterior Cruciate Ligaments
AMG	Acoustic Myogram
ANOVA	Analysis of Variance
ARV	Average Voltage
bPSO	Binary Particles Swarm Optimization
CFS	Conventional Features Set
CWD	Choi-Williams Distribution
CWT	Continuous Wavelet Transform
DAP	Data Acquisition Protocol
Db	Daubechies
DWT	Discrete Wavelet Transform
EC	Excitation-Contraction
ECG	Electrocardiogram
EMG	Electromyogram
Flnsm-5	Spectral Indices
Fvar	Instantaneous Frequency Variance
GA	Genetic Algorithms
GDA	Gradient Descent
HF%	Relative Power Content
Hz	Hertz
ICA	Independent Component Analysis

IG	Information Gain
k-NN	K-Nearest Neighbor
LDA	Linear Discriminant Analysis
LLE	Locally Linear Embedding
L-M	Levenberg-Marquardt
MAVS	Mean Absolute Value
MDF	Median Power Frequency
MDW	Multisignal Denoising Using Wavelet
MF	Instantaneous Mean Frequency
MFS	M-Band Wavelet Transform Based Feature Set
MHW	Multiple Time Hamming Window
MLPNN	Multilayer Perceptron Neural Network
MMG	Mechanomyogram
MNF	Mean Power Frequency
MSW	Multiple Time Slepian Window
MTW	Multiple Time Trapezoidal Window
MU	Motor Unit
MVC	Maximum Voluntary Contraction
NB	Naïve Bayes
PC	Peak Count
PLI	Power Line Interference
Pps	Pulse-Per-Second
PSD	Power Spectral Density
PSO	Particle Swarm Optimization
RMS	Root Mean Square

SC	Spike Count
STFT	Short-Term Fourier Transform
SVM	Support Vector Machine
Var	Variance
WTP	Wavelet Packet Transform
WVD	Wigner Ville Distribution

LIST OF SYMBOLS

C_1	Self Confidence
C_2	Swarm Confidence
f_1	8Hz
f_{max}	Maximum Frequency
f_{med}	Median Frequency
f_{min}	Minimum Frequency
$P(f)$	Power Spectrum
P_g	Global Best
P_i	Particles Best
T	Current Iteration
t_{max}	Maximum Iteration
V_i	Velocity Vector
X	Signals
X_i	Particles Position
$\bar{x}$	Mean
ω	Inertial Weight

ALGORITAM KLASIFIKASI DAN PENGEKSTRAKAN CIRI BAGI MENILAI KEADAAN OTOT DENGANMENGGUNAKKAN PELBAGAI SENSOR

ABSTRAK

Kecederaan terikan otot hamstring adalah kecederaan paling biasa dalam persekitaran sukan khususnya dalam bola sepak. Selain sukan, tarian adalah satu lagi aktiviti fizikal yang mempunyai risiko tinggi untuk kecederaan terikan otot hamstring. Kelesuan otot adalah salah satu faktor risiko yang menyebabkan kecederaan otot. Kelesuan otot berlaku apabila otot gagal mengekalkan kekuatan otot yang diingini. Kerja ini bertujuan untuk menilai dan membezakan kelesuan otot dengan menggunakan tiga myogram yang berbeza (Electromyogram, Mechanomyogram dan Akustik Myogram). Untuk mencapai matlamat projek ini, 3 myogram berbeza direkod secara serentak dari otot hamstring (Biceps Femoris dan Semimembranosus) semasa pengecutan pelepaan dengan beban yang berbeza (5, 7.5 dan 10kg). 20 orang subjek lelaki sihat telah diambil untuk bekerja dalam kerja ini dengan umur 22.4 ±2.6 tahun. Myogram yang direkodkan ditapis dan dibahagikan. Korelasi Pearson dan regresi linear telah digunakan kepada 9 ciri-ciri domain masa dan frekuensi (diambil daripada setiap bingkai daripada setiap myogram) untuk mencari tingkah laku myogram dan hubungan antara myogram tersebut. Berdasarkan keputusan korelasi pearson terdapat hubungan yang kuat antara isyarat AMG dan EMG dan hubungan yang lemah antara isyarat MMG dan EMG. Walau bagaimanapun, hubungan MMG dan EMG isyarat menjadi kuat apabila muatan yang meningkat. Dua jenis ciri-ciri set telah diuntuk mengelaskan keadaan otot dan set ciri-ciri tersebut ialah ciri-ciri konvensional (CFS) dan ciri-ciri set berdasarkan M-band ubahan mengubah (MFS). Di CFS, 16 ciri-ciri telah diekstrak daripada setiap myogram, yang telah digunakan dalam kajian sebelum untuk menilai keadaan otot. MFS, myogram yang direkodkan telah diuraikan menggunakan 4-band ubahan mengubah dan 9 ciri-ciri diambil daripada setiap nod daripada setiap myogram. Oleh itu, 108 ciri-ciri telah dihasilkan bagi setiap myogram. Set ciri-ciri EMG, MMG and AMG daripada CFS dan MFS telah ditambahkan untuk membentuk satu set ciri-ciri yang baru. Dua peringkat ciri-ciri pemilihan telah digunakan untuk mengurangkan ciri-ciri dalam ruang kumpulan ciri-ciri baru. Dalam peringkat I dari ciri-ciri pemilihan, linear tempatan benaman (LLE) telah digunakan. Ruang reduktif ciri baru telah digunakan dalam peringkat II ciri-ciri pemilihan. Dalam peringkat itu, zarah binari swarm pengoptimuman (bPSO) telah digunakan. Pengelas k terdekat jiran (k-NN) telah digunakan untuk mengelaskan keadaan otot (bebas-keletihan atau kelesuan) dengan menggunakan myogram individu mempunyai set, set ciri-ciri digabungkan / diperkukuhkan dan selepas 2 peringkat ciri-ciri pemilihan ciri-ciri set. Dalam penilaian individu tersebut, set ciri EMG telah memperolehi ketepatan yang tinggi. Walaubagaimanapun, banding dengan set ciri-ciri digabungkan, set ciri-ciri digabungkan telah memberi ketepatan yang tinggi. Dua peringkat ciri-ciri pemilihan adalah diaplikasikan kepada set ciri-ciri digabungkan dan ciri-ciri berkurangan sebanyak 69% dan 89% untuk CFS dan MF dalam peringkat 1 masing-masing. Di peringkat 2 ciri tersebut berkurangan sebanyak 7% dan 40% untuk CFS dan MF masing-masing. Hasil pengelasan adalah meningkat kepada 75% dan 92% untuk CFS dan MSF masing-masing. Ia boleh menyimpulkan bahawa gabungan tiga myogram menyediakan maklumat yang lebih berguna daripada myogram yang tunggal.

FEATURE EXTRACTION AND CLASSIFICATION ALGORITHMS FOR ASSESSING MUSCLE CONDITION USING MULTI SENSORS

ABSTRACT

Hamstring muscle strain injury is the most common injuries in sport environment especially in football. Beside sport, dancing is another physical activity that has high risk for hamstring muscle strain injuries. Muscle fatigue is one of the risk factors that cause muscle injury. Muscle fatigue occurs when muscle fail to maintain desired muscle strength. The aim of this work is to assess and distinguish muscle fatigue using 3 different myograms (Electromyogram, Mechanomyogram and Acoustic myogram). To achieve the aim of this project, 3 different myograms were recorded simultaneously from hamstring muscle (Biceps Femoris and Semimembranosus) during isometric contraction with different loads (5, 7.5 and 10kg). 20 healthy male subjects were recruited in this work with aged 22.4 ±2.6 years. The recorded myograms were denoised and segmented. Pearson correlation and linear regression were applied on 9 time and frequency domain features (were extracted from each frame of each myograms) to find the behavior of the myograms and relationship between the myograms. Based on the pearson correlation results there are a strong relationship between AMG and EMG signals and a weak relationship between MMG and EMG signals. However, the relationship of MMG and EMG signals becomes significant when the load is increased. Two types feature sets were used to classify muscle condition and there are conventional features set (CFS) and M-band wavelet transform based feature set (MFS). In CFS, 16 features were extracted from each myograms, which were used in previous studies to assess muscle condition. In MFS, the recorded myograms were decomposed using 4-band wavelet transform and 9 features were extracted from each node of each myograms. Therefore, 108 features were generated for each myograms. The feature sets of EMG, MMG and AMG of CFS and MFS were augmented to form a new feature set. Two stage feature selection was used to reduce feature space in the new feature set. In stage I of features selection, linear locally embedding (LLE) was used. New reductive feature space was used in stage II for feature selection. In that stage, binary particle swarm optimization (bPSO) was used. k-nearest neighbor (k-NN) classifier was used to classify muscle condition (non-fatigue or fatigue) by using individual myogram feature sets, concatenated / augmented feature sets and after 2 stage feature selection feature sets. In the individual assessment, EMG feature sets have obtained high accuracy. However, compare with the concatenated feature sets, this feature sets have given high accuracy. Two stage feature selection was applied to the concatenated feature sets and the features were reduced by 69% and 89% for CFS and MFS in stage 1 respectively. In stage 2 the features were reduced by 7% and 40% for CFS and MFS respectively. The classification result was improved to above 75% and 92% for CFS and MSF respectively. It can be conclude that, the combination of three myogram provides more useful information than a single myogram.

CHAPTER 1

INTRODUCTION

1.1 Background

The hamstring muscle is composed of three distinct muscles – the Biceps Femoris, Semimembranosus and Semitendinosus – which involved in knee flexion and hip extension (Bobick & Balaban, 2008; Kaeding & Borchers, 2014). Hamstring muscle strain injury is the most common injuries in sport environment especially in football (Petroutsos; H. Liu, Garrett, Moorman, & Yu, 2012). According to epidemiology studies in Australian Rules Football, 6 to 29% of the injuries are hamstring muscle strain injuries. During the FELDA/FAM National Futsal League 2010 in Malaysia, the tournament medical team reported that 86 injuries in 141 matches and 33% of the injuries are upper leg injuries (Hamid, Jaafar, & Ali, 2014). Beside sport, dancing is another physical activity that has high risk for hamstring muscle strain injuries. 34% of the dancers have experienced acute hamstring muscle strain injuries while 17% had overuse injuries reported by Askling et al. (Askling, Lund, Saartok, & Thorstensson, 2002; H. Liu et al., 2012). The hamstring muscle strain injury also has high risk for re-injury with rate 12 to 31% of the players (H. Liu et al., 2012; Sullivan, Silvey, Button, & Behm, 2013). The consequence of the hamstring muscle strain injury is significant loss of training, competition time and significantly affects the quality of life of injured athletes (Petroutsos; H. Liu et al., 2012; Hamid et al., 2014). Understanding the cause of hamstring muscle strain injury helps to understand the way how to prevent the injury. There are few risk factors that cause the injury and it was categorized into non-modifiable risk factor and modifiable risk factor. Modifiable risk factors include shortened optimum muscle length, lack of muscle flexibility, strength imbalance, insufficient warm-up, fatigue, low back injury, and increased muscle neural tension. Non-modifiable risk factors include muscle compositions, age, race, and previous injuries.

Muscle fatigue is one of the risk factors that cause muscle injury. Based on the clinical observation, muscle strain injuries are occurred during the late portions of practices and competitions. According to (Mair, Seaber, Glisson, & Garrett, 1996), fatigued muscle significantly diminished the ability of muscle to absorb energy. It will reduce the muscle contractile strength and cause muscle failure or muscle injury (Small, McNaughton, Greig, & Lovell, 2010; H. Liu et al., 2012). Muscle fatigue also increases the knee flexion angle. This provides the knee to flex beyond the optimum degree and its cause the muscle to injury by strain the muscle (Duchateau, 2008; H. Liu et al., 2012). Therefore, investigating muscle fatigue was forces in this work.

1.2 Muscle Fatigue

Muscle fatigue is defined as a failure to maintain desired muscle strength (Place, Bruton, & Westerblad, 2009). Fatigue causes decline in force during sustained isometric and isotonic contraction (Binder-Macleod & Snyder-Mackler, 1993). Muscle fatigue is common symptom in our daily activities. It is also secondary outcome in many disease and health condition such as peripheral vascular disease, heart disease, chronic lung and kidney disease and anemia. Muscle fatigue occur when excitation-contraction (EC) coupling is impaired through *(i)* reducing the number of active crossbridges due to decrement in Calcium ion (Ca^{2+}) releasing; *(ii)* decreasing in sensitivity of the myofilaments to Ca^{2+}; and/or *(iii)* reducing in force produced by each active cross-bridge (Binder-Macleod & Snyder-Mackler, 1993).

1.3 Clinical Assessment of Muscle Fatigue

Two types of tests are used to assess muscle fatigue clinically. There are volitionally and electrically elicited fatigue tests. The volitional tests assess muscle fatigue by measuring the amount of fatigue, which known as fatigue index. Subjects perform repeated maximal knee extension efforts at 180°/s, and the torque output for each contraction was recorded on an isokinetic dynamometer. The subjects performed either 50 or 100 contractions (0.5 second contraction; 0.7 second relaxation). The average peak torque for the last three contractions was divided by the average peak torque of the first three contractions and used as a fatigue index.

Electrically elicited fatigue test is used as a clinical tool for assessing muscle performance. This test was modified to assess muscle performance in patient following surgical repair of anterior cruciate ligaments (ACL). The test consists of stimulating the muscle once per second with 330-millisecond, 40-pulse-per-second (pps) groups of electrical pulses (pulse trains) and measuring the percentage of decline in force production.

1.4 Problem Statement

Various techniques have been used to detected muscle fatigue. However, each technique was used individually to measure muscle fatigue. Therefore, this technique was utilized in this work to overcome these drawbacks:

- There is no clear understanding between the myograms. Different type statistical tests have been performed to find the difference, correlation and reliability of the features among the myograms.
- There are no clear view on the classification result of the myograms, individually or combined. Most of the presented work shown the interpretations of the myograms to due to muscle fatigue.
- There are tendencies to have limitations using a uni-modal system in assessing muscle fatigue. Electromyogram (EMG) has widely been used to assess muscle fatigue. However, there are limitations such as ability to monitor only a few site, crosstalk and electrode placement.

1.5 Objectives

- To investigate the relationship between the recorded myograms and muscle fatigue using Pearson Correlation and Linear regression.
- To propose and evaluate M-band wavelet transform based feature extraction for extracting salient features from 3 different myograms.
- To apply feature level fusion and two stage feature selection for improving the classification of muscle condition (non-fatigue and fatigue).

1.6 Scope of the Work

The main idea of this work is to distinguish the muscle condition (non-fatigue and fatigue) using three different myograms and also to investigate the relationship between the myograms due to muscle fatigue. To achieve aim of the work, EMG, MMG and AMG signals were recorded from healthy male volunteers of aged between 20 and 26. M-band wavelet transform is designed for feature extraction from the recorded myograms. Feature level fusion was applied by merging the myogram features. The dimension of concatenated feature sets was reduced by using two stage feature selection. The selected features were classified into fatigue and non-fatigue using k-NN classifier.

1.7 Thesis Organization

This thesis is composed of 5 chapters, including the present one.

Chapter 2 presents all the topics addressed in this research. First it explained what are the myograms, how the acquisition is usually made and the fields where these kind of signals were used. Next section explains the technologies involved in the processing of the myograms. Last section describes the relation between the myograms and fatigue. It also described the research observation and contributions.

Chapter 3 describes the proposed methodology and algorithms used in this work. This chapter also explains the experimental setup, which includes data acquisition system, data acquisition protocol, sensor placement and study participants and preprocessing techniques to denoise the myograms. This chapter also presents the statistical analysis and feature extraction technique for classification of muscle conditions, which include M-band wavelet transform, feature level fusion and feature selection.

Chapter 4 presents the statistical analysis and classification result and discussion. The first section describes the behavior of the myograms due to muscle fatigue and it also explains the relationship between the myograms. The next section shows the classification results obtained with different myograms. Followed by, classification result obtained after the feature level fusion and feature selection.

Chapter 5 presents the conclusions and future work.

CHAPTER 2

LITERATURE REVIEW

2.1 Overview

This chapter gives a brief introduction of the myograms (EMG, MMG and AMG). It also provides the signal processing (feature extraction) techniques that were implemented to analysis the myograms. This chapter also explains the relation between the myograms and the fatigue. At the end of the chapter, observations from previous works and the significance of multi-modal system in analysis of muscle fatigue are described.

2.2 Muscle Fatigue Analysis using Electromyogram

Electromyography (EMG) has widely been used to assess the muscle condition by recording electrical activity of the muscle which control by the nervous system. The muscle contains motor unit action potential trains the action potential to adhesion motor unit in the muscle (Andreassen & Arendt-Nielsen, 1987). The EMG is generated by summing up the action potential of the motor unit (Rash, 1999). It is complex signal which refer the muscle activity (Sadoyama & Miyano, 1981; Reaz, Hussain, & Mohd-Yasin, 2006).

2.2.1 Preprocessing

Pre-processing is a process to remove unwanted information such as crosstalk, motion artifacts and power line interference (PLI) from the raw EMG signal. The frequency range of raw surface EMG signal is around 10-600Hz (Rash, 1999). Different types of filter were applied to remove the unwanted information from the raw EMG signal and produce noise free EMG signal.

High pass filter was used to remove the motion artifacts in EMG signal. The frequency range of motion artifacts is 10-15 Hz, where the 10Hz represents walking artifacts while 15 Hz for the rapid movement artifacts (Stulen & De Luca, 1982; Rash, 1999). Therefore the cutoff frequency of the high pass filter depends on the activity was used. The low pass filter was applied to a raw EMG signal to optimize the frequency of the EMG signal and to remove high frequency components. The cutoff frequency depends on the type of sensors, for needle sensor is 1000Hz and for the surface electrode is 500Hz. Notch filter is used to remove the PLI with cut off frequency of 50/60 Hz.

2.2.2 Feature Extraction, Analysis and Classification

According to (Sadoyama & Miyano, 1981; Reaz et al., 2006), EMG is complex signal which refers to the muscle activity. Therefore, a suitable signal processing technique (which contain preprocessing and feature extraction) was needed to distinguish the muscle fatigue condition.

The root mean square (RMS) has widely been used to view muscle endurance with time. In some case, full rectifier RMS (RRMS) value was used to detect muscle fatigue (Kim, Ahad, Ferdjallah, & Harris, 2007; Oliveira & GonÃ§alves, 2009). The RMS and RRMS features are increasing due to muscle fatigue (Kroon & Naeije, 1991; Lalitharatne, Hayashi, Teramoto, & Kiguchi, 2012; Yung, Mathiassen, & Wells, 2012). In (Farina & Merletti, 2000), a study was conducted to view the effect of the amplitude estimation to the RMS value (Farina & Merletti, 2000). EMG amplitude estimation normally applied to raw EMG signal and the amplitude estimation process consist of 6 stages such as *(i)* noise and interference filtering, *(ii)* whitening, *(iii)* multiple-channel combination, *(iv)* demodulation, *(v)* smoothing, and *(vi)* re-linearization (Clancy, Bouchard, & Rancourt, 2001). It can be concluded that pre-whitening the EMG signals reduces considerably of the RMS values.

The average voltage (ARV) is another feature used to indicate the muscle fatigue. There is significant difference between before and after muscle fatigue (Arendt-Nielsen & Mills, 1988; Kim et al., 2007). Plus, this parameter increased due to muscle fatigue (Farina & Merletti, 2000).

A non-parameter feature was extracted from the EMG signal by counting the peak (PC) and spike (SC). The peak and spike, which is above the threshold was considered to count. The PC and SC also mimic the standard spectral indication of the muscle fatigue. The features were compared with median power frequency (MDF) which commonly used to view muscle fatigue. PC and SC had shown almost the same response which is decrement due to the fatigue (Dayan et al., 2012, November).