VOL. 11, NO. 7, APRIL 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

FEATURE REDUCTION USING LOCALLY LINEAR EMBEDDING FOR CLASSIFICATION MUSCLE FATIGUE

Mohamed Sarillee, M. Hariharan, Anas M. N., Omar M. I., Aishah M. N. and Q. W. Oung
Biomedical Electronic Engineering, School of Mechatronic Engineering, Universiti Malaysia Perlis Campus Pauh Putra, Arau, Perlis, Malaysia
E-Mail: sarilleelee@gmail.com

ABSTRACT

The aim of this work was to classify muscle condition (non-fatigue and fatigue) using a mutil-modal system. In order to realize this aim, electromyogram (EMG), mechanomyogram (MMG) and acoustic myogram (AMG) signals were recorded from activated muscle during isometric contraction from 20 healthy volunteers. Sixteen features were extracted from each recorded myograms (EMG, MMG and AMG) and concatenated to form a feature set with 48 features. Feature reduction using Locally Linear Embedding (LLE) was proposed to select best discriminative features to enhance the classification of muscle condition. k-nearest neighbor (k-NN) classifier was used and obtained highest accuracy of 93.50% after applying LLE.

Keywords: muscle fatigue, EMG, MMG, AMG.

INTRODUCTION

Hamstring injuries are the most common injury occur to athletes in many sports. The hamstring muscles run down the back of the thigh and consist of 3 types of muscles known as Semitendinosus, Semimembranosus and Biceps femoris. Hamstring injured athletes are suffer in terms of decrease in performance, severe long-term pain and physical disability (Hamilton, 2012; Woolf and Pfleger, 2003). Although, several techniques have been used to treat muscle injuries but most of the treatments are failing to optimize the muscle healing process (Hamilton, 2012; Lim *et al*., 2014). Therefore, it is crucial for the athletes to avoid this injury before it occurs. The risk factors of muscle injuries are pervious injured, age of the person and muscle fatigue. Muscle fatigue is a common muscle condition caused by declining force during isometric and isokinetic contraction.

Assessment of muscle fatigue is very common in sports and clinical environment. Surface electromyogram (sEMG) has frequently been used to assess muscle fatigue by recording electrical activity of the muscle. Recently, mechanomyogram (MMG) and acoustic myogram (AMG) have been used to assess muscle fatigue. MMG and AMG signals present mechanical activity of the muscle through vibration signal and sound signal respectively. There are many advantages have been found when using MMG and AMG, which are easy and simple to implement, do not contain power line interference and have a higher signal to noise ratio (Islam *et al.*, 2013; Mohamed Irfan *et al*., 2011).

The aim of this work is to detect muscle fatigue using a multimodal system which consists of 3 sensors to record the EMG, MMG and AMG. Each myogram has its own advantages to represent a muscle condition. EMG represents the firing rate of the motor unit and recruitment of motor unit (Basmajian, 1978; Rash, 1999) while the MMG represent force production and muscle stiffness and intramuscular fluid pressure (Barry *et al*., 1992; Beck, 2010b; Sarillee *et al.,* 2014). AMG signals present the degree of muscle contraction and quantify the development of the forces (Barry *et al*., 1985; Grass, 2004).

This paper is organized into 5 different sections. Section I is the introduction, which explains about the work, and aim of the work. Follow up by summary of works that has been conducted in Section II. Section III explains about the methodology, which consists of integration of hardware, data acquisition protocol, details of subject, sensor placement, signal processing, feature selection and classification. Section IV discusses the classification result, discussion of the results and comparison with previous works. Final section concludes the work.

PREVIOUS WORKS

MMG is vibration signals generated by the muscle, while contracting. The amplitude of the MMG is correlated with force production, even a small change in force which reflected in the MMG amplitude. In (Beck, 2010b), the authors also suggested that the MMG frequency range (0-100 Hz) should be used. The sound or acoustic signal called as (AMG) is audible produced when the muscle is contracted. The AMG is a low frequency signal, in 1810; William Hyde Wollaston discovered the frequency of the sound is 25 Hz. The power spectrum of acoustic signal is fairly broad, ranging from about 10 to 50 Hz (Barry *et al.*, 1985).

Over the past decade, most research in muscle fatigue has emphasized the use of EMG or MMG or AMG. In some cases, EMG and MMG/AMG have been used to assess muscle fatigue. However, the EMG was used to map the relationships between MMG/AMG and muscle fatigue. As discussed in the introduction, each myograms represent different muscle activity such as the

ARPN Journal of Engineering and Applied Sciences

EMG recorded electrical activity, while MMG and AMG reflect the mechanical activity of the muscle. Therefore, combining these myograms may provide information in detecting muscle fatigue. In (Beck, 2010a; Orizio *et al*., 1992), the authors suggested that combination of EMG and MMG may be useful for investigating the neural and peripheral mechanisms underlying muscle fatigue.

However, the features play big role in investigating muscle fatigue. Therefore, time, frequency and time-frequency domain features have been extracted to assess muscle fatigue. The most common features that have been used are the Root Mean square (RMS), median power frequency (MDF) and mean power frequency (MNF). The RMS mimics firing rate of the motor unit and force of EMG, MMG and AMG respectively. The MNF feature correlated with the muscle oxygenation (EMG). In most of the previous work, the researchers have used the statistical approach and visual representation to assess the muscle fatigue.

METHODOLOGY

The different stages for detecting muscle fatigue (as in Figure-1) which include; data acquisitions (recording EMG, MMG and AMG of hamstring muscles), feature extraction from time, frequency and time-frequency domain, feature reduction using locally linear embedding and classification of muscle condition using k-nearest neighbors.

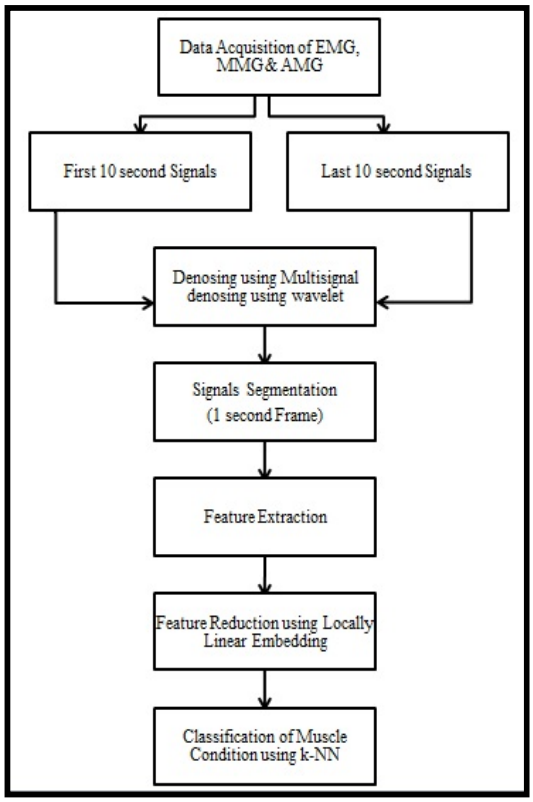

Figure-1. The block diagram of methodology.

Data acquisition

The data acquisition system consists of 3 sensors; there are surface electrode (to record EMG), accelerometer (to record MMG) and cardio microphone (to record AMG). Surface electrode and cardio microphone were connected to PowerLab 4/25T and accelerometer was connected to Data Translation (DT 9837A). PowerLab 4/25T and Data Translation were used to digitize with sample rate of 1 kHz. The interface software was developed using LABVIEW software to collect MMG through Data Translation, whereas EMG and AMG were collected using LABCHART.

The data acquisition protocol (DAP) started with oral explanation followed by answering the questionnaire (Consent Form), to complete the personal details and health condition of the subject. After the briefing, each subject was directed to perform lower limb exercise as a warm up and then followed by the sensor placement. The bipolar surface EMG electrode, accelerometers and cardio microphone were used to measure the muscle activation at the hamstring muscle (biceps femoris and semimembranosus) during the isometric contraction. Two surface EMG electrodes (Meditrace Pellet Ag/AgCl discs and 10 mm in diameter, Graphic Controls Ltd., Buffalo, NY) were placed at 2-4 cm apart over the midpoint of the muscle belly between the gluteal fold and the popliteal fold (Sullivan *et al*., 2013) (as in Figure-2). Accelerometer and cardio-microphone were set at the centre of the muscle belly between the gluteal fold and the popliteal fold (Lee *et al*., 1992)(as in Figure-2). To improve the electrical conductivity, the skin surfaces, where the electrodes were placed were shaved, abraded, and cleansed with alcohol.

The experiment was begun after the sensor placement. The subject was laid on the lying leg curl machine to perform the isometric contraction. The subjects were instructed to bend/curl the leg by lifting 5kg load (as in Figure-3). They were requested to maintain the contraction / forces until exhaust.

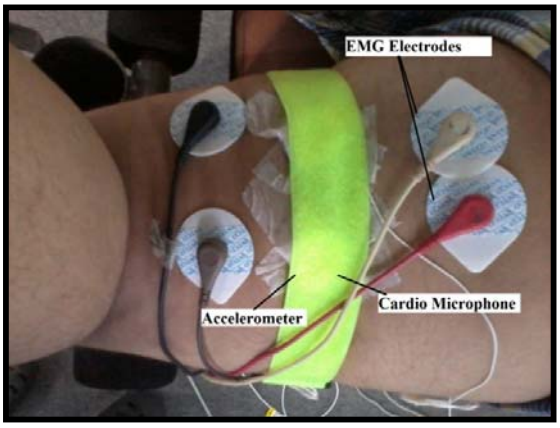

Figure-2. Sensor placement.

ARPN Journal of Engineering and Applied Sciences

Figure-3. Experimental setup for hamstring curl exercise.

Signal processing

The myograms were recorded averagely 6.06 minutes. The first and last 2 seconds data were removed to obtain a stable myograms. The first and last 10 seconds of the myograms were segmented. The first 10 seconds signals were labeled as non-fatigue signals. The last 10 seconds signals were denoted as fatigue signals (Wang *et al*., 2007).

Wavelet denoising

Multisignal denoising using wavelet technique was applied in this work to denoise the myograms. This technique has widely been used to remove artifact and unwanted signals. Daubechies 4 (db4) was used as mother wavelet with 4 levels of decomposition (Gradolewski *et al.*, 2015). Principle of Stein's Unbiased Risk threshold rule method was used to determine the threshold. Before and after the denoised signal is shown Figure-4 and Figure-5.

Feature extraction

The denoised signals were segmented into 1 second frame without overlap for feature extraction. Statistical features were extracted from segmented myograms.

Time domain

Time domain refers to the variation of the amplitude of signal with time, which is important because the muscle is endurance with time. 5 features were extracted from each recorded myograms and described as follows.

Mean ($\bar{x}$): Mean was extracted from the myograms to view the changes of the amplitude. So, mean can be expressed as:

$$\bar{x} = \frac{\sum_{i=1}^{N} x_i}{N} \tag{1}$$

Variance (Var): Var captures the power of EMG signal and magnitude of the MMG and AMG. Variance can be defined as:

$$Var = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N-1} \tag{2}$$

Root Mean Square (RMS): RMS is related to constant force contraction. Generally, it is similar to standard deviation, which can be expressed as:

$$RMS = \sqrt{\frac{\sum_{1=i}^{N} x_i^2}{N}} \tag{3}$$

Peak (PC) and spike (SC) were extracted from the myograms by counting them (Dayan *et al.*). The peak and spike, which are above the semi-amplitude, were considered as count. PC and SC mimic the standard spectral indication of the muscle fatigue condition.

Frequency domain

A total of 4 frequency domain features were extracted from each myograms. Fast Fourier Transform (FFT) was applied to transform the signals from time domain to frequency domain. Four different

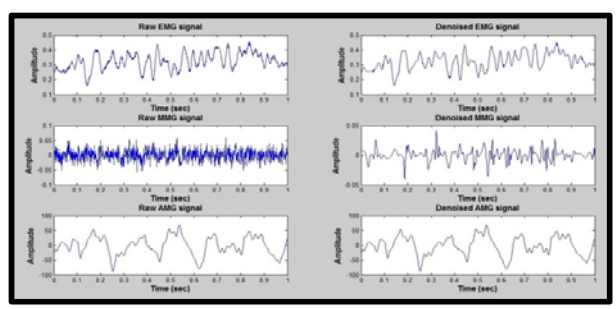

Figure-4. Before and after denoised signals for non-fatigue muscle condition.

ARPN Journal of Engineering and Applied Sciences

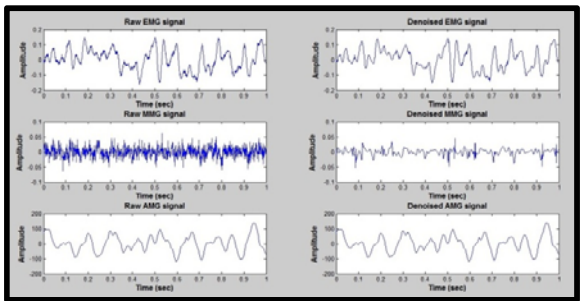

Figure-5. Before and after denoised signals for fatigue muscle condition.

features were extracted in this work are described as follows:

Mean Power Frequency (MNF) and Median Power Frequency (MDF) are the most common feature that has been used to present the musculoskeletal conditions. These features were included in this work and they can be expressed as:

$$MNF = \frac{\int_{f_{\min}}^{f_{\max}} f.P(f)}{\int_{f_{\min}}^{f_{\max}} P(f)} \tag{4}$$

$$MDF = \int_{f_{\min}}^{f_{med}} P(f) = \int_{f_{med}}^{f_{\max}} P(f) = \frac{1}{2}\int_{f_{\min}}^{f_{\max}} P(f) \tag{5}$$

where f_{min}, f_{max} and f_{med} are the minimum, maximum and median of frequency respectively and $P(f)$ is the power spectrum of the signal.

An average instantaneous frequency (AIF) was used to indicate muscle fatigue condition and it can be defined as:

$$AIF = \frac{\int_{f_1}^{f_{\max}} P(f)}{f_{\max} - f_1} \tag{6}$$

where f_1 is 8Hz, f_{max} is maximum of frequency and $P(f)$ is the power spectrum of the signal.

A feature knows as spectral indices (FInsm-5) were extracted from the myograms. It was the ratio between the spectral moment order -1 and order 5 and it is represented as:

$$FInsm-5 = \frac{\int_{f_1}^{f_{\max}} f^{-1}P(f)df}{\int_{f_1}^{f_{\max}} f^{5}P(f)df} \tag{7}$$

Time-Frequency domain

Two different time-frequency algorithm were used to extract 7 features from each myograms. They are Choi-Williams distribution (CWD) and Discrete Wavelet Transform (DWT). The CWD is a Cohen's generalized classes, which consist of Wigner-Ville distribution and the spectrogram. The characteristic of Cohen's classes is exponential kernel which used to take down the magnitude of cross terms in CWD. Two features such as Instantaneous Mean Frequency (MF) and Instantaneous Frequency Variance (F_{var}) were extracted using CWD (Dimitrov *et al.*, 2006; M.González-Izal; *et al.*, 2009).

$$MF = \frac{\int_{f_1}^{f_{\max}} f.PS_{cw}(f,t)df}{\int_{f_1}^{f_{\max}} PS_{cw}(f,t)df} \tag{8}$$

$$F_{\text{var}}(t) = \frac{\int_{f_1}^{f_{\max}} (f-MF)^2.PS_{cw}(f,t)df}{\int_{f_1}^{f_{\max}} PS_{cw}(f,t)df} \tag{9}$$

where $PS_{cw}(f,t)$ is the time-dependent power spectrum obtained from the Choi-Williams distribution, f_1 is 8 Hz and f_{max} is maximum frequency.

In DWT, the signals were decomposed into two different frequency sub-bands with different resolution. Scaling function (low pass filter) and wavelet function (high pass filter) were applied to decompose the signal into approximate and detail respectively. Myograms were decomposed up-to levels 5 using db5 and symlets5 (sym5) as mother wavelet. Five different features were extracted from all 3 myograms (M.González-Izal; *et al.*, 2009) (Table-1).

ARPN Journal of Engineering and Applied Sciences

Table-1. Five different indices were calculated using the DWT.

Feature name	Description	Wavelet name
WIRM1551	Wavelet index of ratio between moment-1 at scale 5 and moment 5 at scale 1	sym5
WIRM1M51	Wavelet index of ratio between moment -1 at maximum energy scale and moment 5 at scale 1	db5
WIRM1522	Wavelet index of ratio between moment -1 at scale 5 and moment 2 at scale 2	db5
WIRE51	Wavelet index of ratio of energies at scales 5 and 1	sym5
WIRW51	Wavelet index ratio between square waveform lengths at different scales	sym5

A total of 16 features were extracted from each myograms. In this work, feature level fusion is accomplished by a simple concatenation of the feature sets obtained from 1 second frame of each myograms (Ross and Govindarajan, 2005). The new feature set is generated by first augmenting features set of EMG, MMG and AMG. The total number of features in new feature set is 48.

Feature selection

Augmenting the features set of EMG, MMG and AMG, results in a new feature with high dimension. It is not necessarily the new feature set to produce a good result. Further, some of the features probably noise (Ross and Govindarajan, 2005). Therefore, feature reduction was used in this work. Feature reduction is important and necessary when a real world application has to deal with training data of high dimensionality. There are number of advantages using feature reduction: improve computational cost by reducing the feature dimension and improve classification results by removing irrelevant features (Mohamad *et al*., 2011; Sun *et al*., 2010). Therefore, the feature reduction technique was used in this work, to reduce the feature and to find the best features. Locally Linear Embedding (LLE) was used as feature reduction tool in this work.

Locally linear Embedding (LLE) is non-linear dimensionality reduction technique was introduced by Rowies and Saul. It based on simple geometric intuitions (Busa-Fekete and Kocsor, 2005; Roweis and Saul, 2000). LLE convert the dataset with non-linear manifold to approximated linearly using 2 stages. In stage, locally fitting hyper-planes around each sample x_i, based on its k-nearest neighbors, and calculating reconstruction weights (Busa-Fekete and Kocsor, 2005; De Ridder and Duin, 2002). In second stage, finding lower-dimensional coordinates y_i for each x_i, by minimizing a mapping function based on these weights (De Ridder and Duin, 2002).

Classification

The k-nearest neighbor (k-NN) was used to distinguish the non-fatigue and fatigue muscle condition (Venugopal *et al*., 2014). The k-nearest neighbor method has become a very efficient non-parametric approach classification purposes. Partly, because of its perfect mathematical theory, the k-NN method develops into several variations. In this work, the k-NN algorithm was implemented using cityblock distance metric to locate the nearest neighbor. The majority rule was used as the decision rule to derive a classification from the k-NN. The k value 1 was used to classify the muscle condition and it was decided based on empirical study (as in Figure-4).

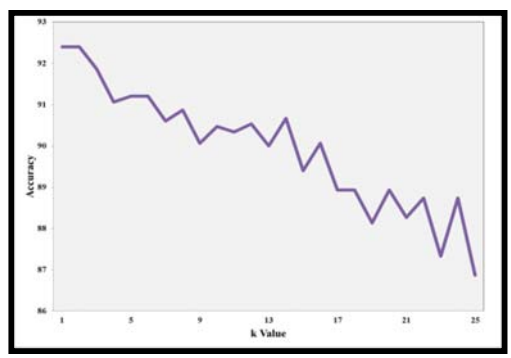

Figure-6. The accuracy of k-NN versus the increase of the k value.

RESULT AND DISCUSSIONS

A total of 48 features were obtained after fused the EMG, MMG and AMG features. The fused features were trained and tested by using k-NN with k values of 1 and distance metric as cityblock. The feature set was classified for 10 times with 10 fold cross validation and maximum accuracy; minimum accuracy, mean accuracy and standard derivation were tabulated in. Table-2.

ARPN Journal of Engineering and Applied Sciences

Table-2. Classification result before feature selection.

k-NN	Maximum	Minimum	Mean	S/derivation
k=1	93.00	91.00	92.07	0.60

Based on the Table-2, it is shown that the mean accuracy of the classification result was 92.07% with standard derivation of 0.60.

After that, feature selection was applied via the LLE method to reduce the high dimension of the feature spaces. In this phase, the effects of the individual feature weight operation by the LLE method. The weight was ranked in descending order with respect feature. The top 30 features were selected and divided into 6 different groups. The groups start with top 5 features and end with top 30 features with step size of 5 features. Each group was classified using k-NN and the classification results versus number of features were depicted in Figure-5.

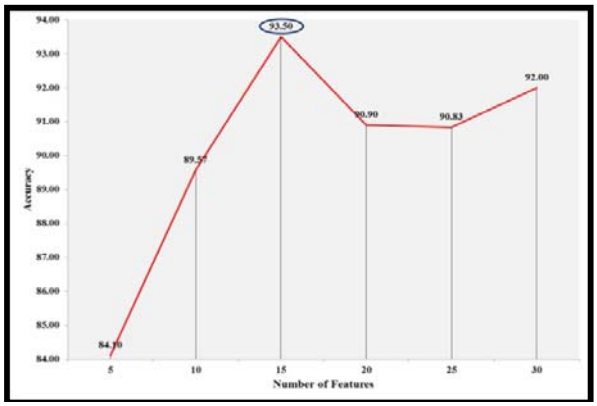

Figure-7. The accuracy of k-NN with the increase of the number of features.

Figure-7 shows accuracy with number of features selected via LLE. As shown in the figure above, the top 15 features shown the highest accuracy of 93.50% compared to other feature sets. Therefore, top 15 features with high accuracy was chosen.

CONCLUSIONS

This work presents the method to detect the muscle condition (non-fatigue and fatigue) by recording EMG, MMG and AMG signal from hamstring muscle. It is important to detect muscular fatigue before it is visible, not only to prevent future injuries but also to improve athletes' performance. Each myogram have its own specialty, as EMG represents the electrical activity and the MMG and AMG represent the mechanical activity. The myograms were recorded while preforming leg bend or curl with lifting a load. The 1st and last 10 seconds signals were segmented and labeled into fatigue and non-fatigue signal. The segmented signals were denoised using multisignal denoising using wavelet technique. The denoised signals were further segmented into 1 second segment. A total of 16 features were extracted from time domain, frequency domain and time-frequency domain for each myograms. The extracted features were fused into a new feature set. The fused feature set was classified using k-NN classifier and obtained maximum accuracy of 93.00%. LLE was used to reduce the number of features and the number of features reduces by 68.75% with an accuracy of 93.50%.

REFERENCES

Barry D. T., Geiringer S. R. and Ball R. D. 1985. Acoustic myography: a noninvasive monitor of motor unit fatigue. Muscle and nerve. 8(3): 189-194.

Barry D. T., Hill T. and Im D. 1992. Muscle fatigue measured with evoked muscle vibrations. Muscle and nerve. 15(3): 303-309.

Basmajian J. V. 1978. Muscles alive, their functions revealed by electromyography: Williams and Wilkins.

Beck T. W. 2010a. Surface mechanomyographic responses to muscle fatigue. ed. Kerala, India: Trivandrum, India: Transworld Research Network, 15.

Beck T. W. 2010b. Technical Aspects of Surface Mechanomyography. Applications of Mechanomyography for examining muscle function. Kerala: Transworld Research Network, 13.

Busa-Fekete R. and Kocsor A. 2005. Locally Linear Embedding and its Variants for Feature Extraction. Paper presented at the IEEE International Workshop on Soft Computing Applications (IEEESOFA 2005).

De Ridder D. and Duin R. P. 2002. Locally linear embedding for classification. Pattern Recognition Group, Dept. of Imaging Science & Technology, Delft University of Technology, Delft, The Netherlands, Tech. Rep. PH-2002-01, 1-12.

Dimitrov G. V., Arabadzhiev T. I., Mileva K. N., Bowtell J. L., Crichton N. and Dimitrova N. A. 2006. Muscle fatigue during dynamic contractions assessed by new spectral indices. Medicine and science in sports and exercise. 38(11): 1971.

Gradolewski D., Tojza P. M., Jaworski J., Ambroziak D., Redlarski G. and Krawczuk M. 2015. Arm EMG Wavelet-

Based Denoising System. Mechatronics-Ideas for Industrial Application. 289-296.

Grass. 2004. Acoustic Myography (AMG). Retrieved September 4 2013, 4.

Hamilton B. 2012. Hamstring muscle strain injuries: what can we learn from history? British journal of sports medicine. 46(13): 900-903.

Islam M. A., Sundaraj K., Ahmad R. B. and Ahamed N. U. 2013. Mechanomyogram for Muscle function assessment: A review. PloS one. 8(3): e58902.

Lee D. J., Stokes M. J., Taylor R. J. and Cooper R. G. 1992. Electro and acoustic myography for noninvasive assessment of lumbar paraspinal muscle function. European journal of applied physiology and occupational physiology. 64(3): 199-203.

Lim B., Khoo S. and Yusof A. 2014. Patient Satisfaction in the Treatment of Acute Hamstring Strain Injury. International Journal of Kinesiology and Sports Science. 2(2): 9-15.

M. González-Izal, I.Rodríguez-Carreño, F.Mallor-Giménez, A. Malanda and M.Izquierdo. 2009. New Wavelet Indices to Assess Muscle Fatigue during Dynamic Contractions. World Academy of Science Engineering and Technology. 55, 480-485.

Mohamad M. S., Omatu S., Deris S. and Yoshioka M. 2011. A modified binary particle swarm optimization for selecting the small subset of informative genes from gene expression data. Information Technology in Biomedicine, IEEE Transactions on. 15(6): 813-822.

Mohamed Irfan M., Sudharsan N., Santhanakrishnan S. and Geethanjali B. 2011. A comparative study of emg and mmg signals for practical applications. Paper presented at the 2011 International Conference on Signal, Image Processing Applications, with workshop of ICEEA.

Orizio C., Perini R., Diemont B. and Veicsteinas A. 1992. Muscle sound and electromyogram spectrum analysis during exhausting contractions in man. European journal of applied physiology and occupational physiology. 65(1): 1-7.

Rash G. S. 1999. Electromyography fundamentals. Retrieved September. 3 2013, 4.

Ross A. A. and Govindarajan R. 2005. Feature level fusion of hand and face biometrics. Defense and Security. 196-204.

Roweis S. T. and Saul L. K. 2000. Nonlinear dimensionality reduction by locally linear embedding. Science. 290(5500): 2323-2326.

Sarillee M., Hariharan M., Anas M. N., Omar M. I., Aishah M. N. and Oung Q. W. 2014. Non-invasive techniques to assess muscle fatigue using biosensors: A review. Paper presented at the Control and System Graduate Research Colloquium (ICSGRC), 2014 IEEE 5th.

Sullivan K. M., Silvey D. B., Button D. C. and Behm D. G. 2013. Roller‐massager application to the hamstrings increases sit‐and‐reach range of motion within five to ten seconds without performance impairments. International journal of sports physical therapy. 8(3): 228.

Sun Y., Todorovic S. and Goodison S. 2010. Local-learning-based feature selection for high-dimensional data analysis. Pattern Analysis and Machine Intelligence, IEEE Transactions on. 32(9): 1610-1626.

Venugopal G., Navaneethakrishna M. and Ramakrishnan S. 2014. Extraction and analysis of multiple time window features associated with muscle fatigue conditions using sEMG signals. Expert Systems with Applications. 41(6): 2652-2659.

Wang G., Ren X.-m., Li L. and Wang Z.-z. 2007. Multifractal analysis of surface EMG signals for assessing muscle fatigue during static contractions. Journal of Zhejiang University SCIENCE A. 8(6): 910-915.

Woolf A. D. and Pfleger B. 2003. Burden of major musculoskeletal conditions. Bulletin of the World Health Organization. 81(9): 646-656.