ADVANCED SIMPLE AND ROBUST TECHNIQUE FOR DETECTING LOCALIZED EARLY MUSCLE FATIGUE

RUBANA H. CHOWDHURY1,*, MAMUN B. I. REAZ2, MUHAMMAD L. ALI1, SULVIA M.1

1Department of Electrical and Electronic Engineering, Southern University Bangladesh, 739/A Mehedibag Road, Chittagong
2Department of Electrical, Electronic and Systems Engineering, Universiti Kebangsaan Malaysia, Bangi, Selangor 43600, Malaysia
*Corresponding Author: rubana1986@hotmail.com

Abstract
The purpose of the study was to establish a simple and robust technique that can be used to predict fatigued and non-fatigued muscle by surface electromyography (sEMG). Muscle fatigue is often caused by unhealthy work practices and recovery may take longer when the fatigue tolerable limit is exceeded. Instantly, it is difficult for the person to realize that there is a complicity in the muscle (symptoms such as a sore throat, muscle or joint pain, and headache) before it is visible. In this study, it is shown that by using the proposed fatigue model, a near identification of the presence of fatigue is possible. Biceps curl exercises with a dumbbell (7 kg) were used to develop fatigue in the biceps brachii muscle of the upper arm. In this work, four different feature sets were considered from sEMG signals and they were considered as an input to the dimensionality reduction method PCA and then the classifier SVM. sEMG signals were recorded from a total of 34 subjects (from 21 to 32 years). A classification success rate of 94.12% was achieved by using only two time-frequency features (average instantaneous frequency after Empirical mode decomposition and total energy by using Wavelet packet transforms) as input to the proposed model. These results also show that fewer features can be used as an alternative to more features for differentiating between fatigued and healthy muscle using the proposed method.

Keywords: Empirical mode decomposition, Muscle fatigue, Surface electromyography (sEMG), Wavelet transform.
1. Introduction

Excessive muscle usage over a period can result in a decline in the ability to generating muscle force and a temporary reduction in muscle strength, power, or endurance; this condition is called fatigue. From the perspective of physiology and sports science, muscle fatigue is ‘any exercise-induced reduction in the maximal capacity to generate force or power output’ [1]. Electromyography (EMG) is the signal that records electrical activity generated by the muscle. Local muscle fatigue usually occurs when continuous activities (eccentric or concentric) are applied to a localized group of muscle(s). It is either static or dynamic, based on the frequency of loading [2].

Basing on the placement of electrodes, EMG is divided into two types: invasive and non-invasive. Both types of EMG techniques are equally valid for the assessment of muscle fatigue. Whenever a fatigue stage arises, different categories of biological changes occur in the subjects, for example, in the muscle fibre conduction velocity, number of motor units, and so on. Measurement of the biological properties of the muscle is considered most reliable for monitoring the fatigue state. These fatigue states are determined by invasive electrode but this is an uncomfortable technique for subjects because these types of electrodes are inserted directly into the muscle. Surface EMG (sEMG) signals are recorded at the surface of the skin, above the inspected muscle during contraction. sEMG is found to be more comfortable in muscle fatigue realization because its principle advantage is non-invasiveness [3]. Moreover, another advantage of sEMG is that during the performance of the exercise, real-time fatigue monitoring of a particular muscle is possible [2]. This paper aims to present a simple and robust technique for the classification of local muscle fatigue by using sEMG information.

Much useful information can be gathered from sEMG signals by feature extraction methods [4, 5]. This paper accesses the time, frequency, a combination of time and frequency, and time-frequency domains as the features of the sEMG signal. There are 11 time-domain (TD) features of the sEMG signal, namely the mean absolute value (MAV), slope sign changes (SSCs), integrated EMG (IEMG), simple square integral (SSI), variance of EMG (VAR), root mean square (RMS), waveform length (WL), log detector (LOG), zero crossing (ZC), v-order 2 (V2) and sample entropy (SampEn) [6, 7]. Meanwhile, there are five frequency domains (FD), five major features of the sEMG signal; namely, median frequency (MDF), mean frequency (MNF), average instantaneous value (AIF), total power (TTP) and average power of EMG (mean power) [8, 9]. Moreover, empirical mode decomposition (EMD) is a recently excelled time-frequency representation technique, which has proven to be a very dependable method [10]. In this study, within the time-frequency domain, instantaneous frequency based on the EMD approach and total energy by the use of wavelet packet transforms are adopted. The dimensionality of these feature vectors is then condensed by using principal component analysis (PCA).

After dimensionality reduction, Support Vector Machine (SVM) is implemented in order to classify fatigued and non-fatigued muscle. This paper presents an early fatigue recognition model with high classification accuracy, which can be used and applied in many potential sEMG applications, including sports science and engineering applications.
2. Methodology

2.1. Measurement procedure

Measurements were performed at normal room temperature. Subjects were relaxed, seated on a chair, and rested for at least 5 minutes before EMG signals were collected. Here, disposable monopolar neutral sEMG electrodes (45×42 mm) were used in the right biceps brachii muscle to gather data. The abrasive paste was used to clean the surface before the electrode placement. In this experiment, a g.tec device (g.BSamp) was used during data collection. Figure 1 shows the experimental Simulink model before the whole experimental setup. A pair of electrodes were placed at the biceps brachii muscle and the inter-electrode gap was 2 cm [11]. The reference electrode was placed at the bone above the scapula and on the clavicle. Figure 2 shows the electrode placement on the upper arm muscle.

Three-four subjects in three age groups (mean ± S.D: 22.5 ± 0.47 years, 26.5 ± 0.9 years, and 30.5 ± 1.30 years) participated in the experiment. The average BMIs of the respective groups were 22.07 kg/m², 24.94 kg/m², and 25.8 kg/m². All participants were male, non-smokers, and had no muscle injury recorded previously; they were duly informed about the experiment and procedure and signed a disclaimer form before participating. Figure 3 shows the exercise actions chosen for our study. Biceps brachii muscles are very important in weightlifting motor tasks [12]. Recording of signals was done under two protocols.

Fig. 1. Experimental Simulink model.
Firstly, data were collected from subjects while they moved their lower arm rapidly (every 1 second) from a folded position to an elongated position. The recording period lasted 6 seconds.

Secondly, subjects performed a dumbbell curls exercise. Subjects tried to lift a 7-kg weight with the intention of developing muscle fatigue. Subjects brought the dumbbell up by bending their arm at the elbow in a curling motion. Subjects flexed their arm to raise the dumbbell towards their shoulder. They were asked to try to hold the load in each of the positions for 1 second and to perform the repetitions continuously. Subjects tried to perform the maximum number of repetitions until they felt a loss of the energy required to perform another repetition. It was found that the maximum number of repetitions performed by participants varied from 30 to 40. The number of participants who performed more than 35 repetitions was very
few. Signals were collected every after five repetitions. This recording period also lasted 6 seconds.

The sampling frequency during data acquisition was 1200 Hz. Signals were filtered with a 20 to 500 Hz bandpass filter with the aim of decreasing the digitization error and removing unwanted noise [13]. A conventional offline notch filter eliminated occasional interference of 50 Hz.

All of the processing was done using Matlab 2012a. All data were analysed in the following three basic stages:

- Extraction of feature vectors.
- Dimensionality reduction of feature vector using the principal component approach.
- Classification with SVM.

### 2.1.1. Feature selection

Regardless of the effort made to reduce noise at the source by apposite skin preparation and the use of well-made active electrodes and signal recording instrumentation, some noises will always accompany the desired signal.

The study of feature extraction properties in the TD and the FD has become an important subject matter in EMG signal classification [13]. Eighteen features were considered in this evaluation study, as listed in Table 1. Most of them are defined in the TD. A five features are computed in the FD which are listed in Table 2 and two are computed in the time-frequency domain.

### Table 1. Mathematical equations of commonly expressed sEMG TD feature-extraction methods.

<table>
<thead>
<tr>
<th>Feature extraction</th>
<th>Mathematical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEMG</td>
<td>( IEMG = \sum_{n=1}^{L}</td>
</tr>
<tr>
<td>MAV</td>
<td>( MAV = 1/L \sum_{n=1}^{L}</td>
</tr>
<tr>
<td>SSI</td>
<td>( SSI = \sum_{n=1}^{L}</td>
</tr>
<tr>
<td>VAR</td>
<td>( VAR = 1/(L-1) \sum_{n=2}^{L}</td>
</tr>
<tr>
<td>RMS</td>
<td>( RMS = \sqrt{1/L \sum_{n=1}^{L}</td>
</tr>
<tr>
<td>WL</td>
<td>( WL = \sum_{n=1}^{L}</td>
</tr>
<tr>
<td>LOG</td>
<td>( LOG = e^{1/n \sum_{n=1}^{L} \log</td>
</tr>
</tbody>
</table>
SSC  
$$SSC = \frac{1}{L} \sum_{n=2}^{L} f[(S_n - S_{n-1}) \ast (S_n - S_{n+1})]$$

$$f(s) = \begin{cases} 
1, & \text{if } S \geq \text{threshold} \\
0, & \text{otherwise}
\end{cases}$$

$$ZC = \sum_{n=1}^{L-1} \{\text{sgn}(S_n \ast S_{n+1}) \cap (S_n - S_{n+1})\} \geq \text{threshold}$$

$$\text{sgn} = \begin{cases} 
1, & \text{if } S \geq \text{threshold} \\
0, & \text{otherwise}
\end{cases}$$

$$V2 = \left( \frac{1}{L} \sum_{n=1}^{L} S_n^2 \right)^{1/2}$$

SampEn

$$SampEn(s, d, t) = -\ln \frac{X_m(r)}{Y_m(r)}$$

For $$r = 0, 1 \ldots d-1$$ with $$Y(0) = t$$, the length of the input series.

Here $$d$$ = dimension and $$t$$ = tolerance, where $$Y_m(r)$$ is the probability that two sequences coincide for $$d$$ points and $$X_m(r)$$ is the probability that they coincide for $$d + 1$$ points.

### Table 2. Mathematical equations of commonly expressed sEMG FD feature-extraction methods.

<table>
<thead>
<tr>
<th>Feature extraction</th>
<th>Mathematical representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDF</td>
<td>$$\frac{1}{2} \sum_{j=1}^{m} P_j$$</td>
</tr>
<tr>
<td>MNF</td>
<td>$$\frac{\sum_{j=1}^{m} f_j P_j}{\sum_{j=1}^{m} P_j}$$</td>
</tr>
<tr>
<td></td>
<td>where $$f_j$$ is the frequency of the sEMG power spectrum at frequency bin $$j$$</td>
</tr>
<tr>
<td>Average Instantaneous Value (AIF)</td>
<td>$$S(t) = a(t) \exp[i\phi(t)]$$</td>
</tr>
<tr>
<td></td>
<td>where $$a(t)$$ is the amplitude and $$\phi(t)$$ is the instantaneous phase of the analytic signal. The derivation of $$\phi(t)$$ is the instantaneous frequency. The requisite average instantaneous value is the average value of all the instantaneous frequencies</td>
</tr>
<tr>
<td>Total power (TTP)</td>
<td>$$\sum_{j=1}^{m} P_j$$</td>
</tr>
<tr>
<td></td>
<td>where $$P_j$$ is the sEMG power spectrum at frequency bin $$j$$</td>
</tr>
<tr>
<td>Average power of EMG (mean power)</td>
<td>$$\frac{\text{norm} \left( \sum_{n=1}^{L} S_n^2 \right)}{L}$$</td>
</tr>
</tbody>
</table>

### Time-frequency representation methods

In this study, AIF is obtained by using the EMD technique introduced by Phinyomark et al. [10]. The EMD method decomposes a raw signal into a number of components of intrinsic mode functions (IMFs) [11]. The acquired raw EMG signal is divided into different IMF components. The number of extracted IMF components was found to be 19-25 and varied from subject to subject. The IMFs of one subject are shown in Fig. 4.
Fig. 4. IMF components of EMG signal from right biceps brachii muscle before exercise.
This decomposition is based on the local time scale of the data. Marcolin et al. [13] proposed that the fundamental frequency of information estimated from the EMG is very useful. First, the fundamental frequency of each IMF component is calculated using fast Fourier transform and those components, which are similar to the fundamental frequency of the raw signal, are selected. By computing all the experimental signals (34 subjects), two components (IMF2 and IMF3) were selected. Thereafter, the signals were reconstructed by using Eq. (1). Now, the average instantaneous frequency was chosen from those signals to generate the feature vector after EMD analysis.

\[
S(t) = \sum_{i=1}^{k} C_i + m_n
\]  

(1)

where \(m_n\) is the mean of the final residual component and \(C_i\) is the \(i\)th IMF component.

### Total energy by using wavelet packet transforms

The aptitude of DWT is to decompose the original signal into multi-resolution components according to a fundamental function and to extract the significant features from the signal [14, 15]. The mathematical representation of the total energy \(E_n\) of wavelet packet transform is expressed by Eq. (2):

\[
E_n = \sum_k |C_{i,n}(k)|^2
\]  

(2)

where each node of the wavelet packet tree is indexed with the pair of integers \((i, n)\): \(i\) represents the level, where \(i = 4\), and \(n\) represents the number of nodes per level with \(n = 0, 1, \ldots, 2^i - 1\). A vector of WP coefficients, \(C_{i,n}\), corresponds to each node \((i, n)\). ‘db4’ was used as a wavelet function.

#### 2.1.2. Principle component analysis

According to Phinyomark et al. [16], PCA was applied to the time domain feature set, the frequency domain feature set, compound features (a combination of time and frequency domain features) and time-frequency domain features.

#### 2.1.3. Classification with support vector machine

Classification is one of the most significant stages, where automatic recognition of fatigued and non-fatigue muscle is done. In this research by Joliffe [17], the non-linear robust classifier SVM was selected over other techniques due to its ability to simplify complex data.

The effectiveness of SVM depends on the type of kernel function, the choice of the regularization parameter \(C\), and the kernel bandwidth parameter. Different kernels can be selected to create the SVM. The most commonly used kernel functions are the polynomial, linear and Gaussian radial basis kernel functions (RBFs) [15]. The Gaussian radial basis function (GRBF) is one of the most popular kernels; it generally outperforms polynomial kernels when dealing with overlapped class distributions and it requires fewer parameters to be set than other kernel functions [18]. In this research, the Gaussian radial basis kernel function was chosen in SVM using 10-fold cross-validation. 140 data were randomly partitioned...
into 10 equally sized sub-datasets, where 9 sub-datasets are employed to train the SVM classifier. The remaining dataset is used as test samples for the muscular fatigue classification.

The mathematical representation of the kernel function is given below by Eq. (3).

\[ k(x, x_n) = \exp\left(-\frac{1}{2} \left\| x_n - x \right\|^2 / \sigma^2 \right) \]  

(3)

2. Experimental results and Discussion

First, it was determined whether the experimental protocol could cause muscle fatigue or not before applying the proposed model to differentiate the sEMG signals between subjects with fatigued and non-fatigued muscle. This was done by statistical analysis using FD and TD features.

Statistical analysis

In this study, all subjects performed the maximum number of repetitions until it was challenging for them to perform another repetition; that is until they did not have enough strength to perform another repetition (rep.). Subjects decided their own capability by performing the maximum possible number of repetitions. In this study, data were collected every after five repetitions. Figures 5 and 6 represent three frequency-and time-domain feature values as an example (after five repetitions) versus the number of repetitions, respectively, for 34 subjects.

In the same age group (assuming almost similar strength), the frequency-or time-domain features share a general tendency after a certain number of repetitions. However, it is observed that after 30 repetitions the feature value for all subjects does not change significantly. The statistical analysis was done to obtain confirmation of whether or not the experimental protocol could cause muscle fatigue. In this study, considering above this level as fatigue, a statistical investigation of the difference between fatigued and non-fatigued muscle features was done by ANOVA test (SPSS 21). Furthermore, the results of the ANOVA statistical analysis test on all features for 34 subjects demonstrated that there was a significant difference between sEMG features before and after exercise (p < 0.05).

One person's muscle strength may differ from another person's, but the characteristics should have some similarity. Past researchers have reported several features of muscle fatigue as an enhancement in amplitude and a conversion from the high-frequency range to low frequency [19]. It can be observed from Fig. 5 that the AIF, MDF, and MNF values decreased significantly under the fatigue condition compared to the values of these features under the non-fatigue condition.

Moreover, the statistical result obtained from Fig. 6 showed that the time domain features MAV, V-vector, and LOG shifted towards higher values under the fatigue condition compared to the non-fatigue condition. The statistical test performed on this feature resulted in a P-value of 0.001. Figure 7 presents a scatter plot of seven features of fatigued and non-fatigued muscle. From the figures, it can be observed that the difference between the feature values of fatigued and non-fatigued muscle is significant.
Fig. 5. Frequency domain features values vs. number of repetitions (reps).

Fig. 6. Time-domain features values vs. number of repetitions (reps).
Applying PCA to feature sets

Four types of feature sets (TD features, FD features, a combination of FD and TD features, and time-frequency features) were formed by using sEMG as described in Section 2.1.1. The feature vectors of all participating subjects (34 subjects with fatigued and non-fatigued muscle) were set as column vectors. The eigenvalues and eigenvectors of the correlation matrix were then solved. Then the principal components (PCs) were solved using a linear model [15].

The first and second PCs performed better in differentiating between fatigued and non-fatigued muscle. For this reason, these two components were used for the classification of subjects.
Classification

The first two PCs explain more than 90% of the total variance for the TD, FD, TD+FD, and time-frequency feature sets, which were used as input into the SVM classifier. Table 3 represents the confusion matrixes for the TD, FD, TD+FD, and time-frequency domain features. Table 4 tabulates the success rate of correct classifications for four different types of feature sets for the SVM.

In this study, before applying the proposed model to these real data, it was ensured by statistical analysis that the muscles were fatigued. This proposed model gave the best decision by combining three steps. Firstly, three sets of features were extracted from the sEMG signal and used individually as input to the PCA method. The SVM classifier gave four types of classification results for the individual output of PCA. The different feature sets were applied to the SVM classifier to compare the classification results of all methods with the same feature reduction method (PCA). These analyses lead to a prescription of the best combination of feature extraction with the same dimensionality reduction for the sEMG signal classification problem. In the SVM classifier, the data set is arbitrarily divided into two equal subsets, with half of the data being selected for training and the other half for testing. Here, Table 1 reports the classification rate and error rates of the SVM classifier. From the table, it is clear that the classification rate is high (94.12%) for time-frequency domain features with the SVM classifier.

The results indicate that only two time-frequency domain features with the SVM as a classifier provide a suitable EMG pattern identifier in the application of muscle fatigue recognition. Based on the rate of correctness, the proposed fatigue detection system (time-frequency feature → PCA→ SVM) proved to be superior. Surface EMG is a non-invasive source of information about neuromuscular conditions. Sarille et al. [20] studied three types of myograms; the EMG, the mechanomyogram (MMG) and the acoustic myogram (AMG), for detecting muscle fatigue. Among these three types of myograms, they showed that sEMG assessed muscle fatigue most accurately. Moreover, they extracted features from the three myograms and classified them using k-nearest neighbour (k-NN), with a classification accuracy of 92.07% [21]. In the last few decades, researchers have carried out studies, which have mainly focused on computing muscle fatigue in dynamic contraction conditions. The detection of early muscle fatigue helps to improve the performance of sportsmen and prevents the risk of injury in sports science. Marri et al. classified sEMG signals as non-fatigue and fatigue in dynamic contraction using k-NN and logistic regression (LR) classifiers. sEMG signals were recorded from biceps brachii muscles of normal subjects while performing a curl exercise. They showed that the maximum accuracy increased to 88% when using k-NN, while it was 82% when LR was used [22]. In the study reported in this paper, an accuracy of 94.12% is achieved using SVM. The proposed technique successfully recognizes fatigue with fewer channels (four channels). Thus, it is now not necessary to use more channels to detect fatigue. Moreover, it is not necessary to extract more features from the sEMG signal. As the accuracy of this methodology was found to reliably higher than that of other methods, it can be routinely used in case of a prognosis of muscle fatigue.

The model acquired from this research may be beneficial and routinely used in sports science and diagnosis of neuromuscular pathology. The symptoms of the fatigue created by this type of over-exercise and irregular exercise arise gradually.
Instantly it is difficult for the person to realize that there is a complicacy. In this study, prevention is not an issue. It is shown that by using a simple fatigue protocol model, successful early detection of fatigue is possible.

### Table 3. Confusion matrixes for TD, FD, TD+FD, and time-frequency domain feature sets.

<table>
<thead>
<tr>
<th>Feature type</th>
<th>Predicted class</th>
<th>True class</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-fatigued</td>
<td>Fatigued</td>
</tr>
<tr>
<td>TD set</td>
<td>31</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>FD set</td>
<td>26</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>Compound feature set</td>
<td>31</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>32</td>
</tr>
<tr>
<td>Time-frequency domain set</td>
<td>31</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>33</td>
</tr>
</tbody>
</table>

### Table 4. Summary of the rate of accuracy and error of different sEMG features with the SVM classification system for identifying fatigue.

<table>
<thead>
<tr>
<th>Feature type</th>
<th>Dimensionality reduction method</th>
<th>Classification accuracy rate (%)</th>
<th>Error rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FD set</td>
<td></td>
<td>86.76</td>
<td>13.24</td>
</tr>
<tr>
<td>TD set</td>
<td>Principle</td>
<td>92.65</td>
<td>7.35</td>
</tr>
<tr>
<td>Compound feature set</td>
<td>Component</td>
<td>92.65</td>
<td>7.35</td>
</tr>
<tr>
<td>Time-frequency domain set</td>
<td>Analysis (PCA)</td>
<td>94.12</td>
<td>5.88</td>
</tr>
</tbody>
</table>

*Results taken as averages of all subjects

3. Conclusions

sEMG signals represent valuable non-invasive diagnostic information about the localized muscle fatigue condition. In this study, a time-domain, frequency-domain, amalgamation of time and frequency domains, and time-frequency domain representation method has been used as a feature extraction method. Then the extracted features of sEMG signals have been used as input to SVM, which showed comparatively exceptional results when only two time-frequency domain features were used for classification. In this study, we have proposed a simple and robust technique (time-frequency feature → PCA → SVM) with a reduced number of electrodes. It is our conclusion that the presented method can successfully detect muscle fatigue under dynamic contraction and the presence of muscle fatigue at an early stage.

References


