

Optimal feature set for finger movement classification based on sEMG

Ahmad A. Al-Tae, Adel Al-Jumaily, Senior IEEE Memeber

Abstract—One of the most important electrophysiological signal is the Electromyography (EMG) signal, which is widely used in medical and engineering studies. This signal contains a wealth of information about muscle functions. Therefore, the EMG signal is becoming increasingly important and has started to be used in many applications like finger movement rehabilitation. However, an advanced EMG signal analysis method is required for efficient usage of such applications. This signal analysis can include signal detection, decomposition, processing, and classification. There are many approaches in studying the EMG signals, however, one of the important factor of analyzing is to get the most efficient and effective features that can be extracted from the raw signal. This paper presents the best feature extraction set compared to previous studies. Where eighteen well-known features algorithm has been tested using the sequential forward searching (SFS) method to get excellent classification accuracy in a minimum processing time. Among these novel features only four combinations have been selected with perfect results, which are; Hjorth Time Domain parameters (HTD), Mean Absolute Value (MAV), Root Mean Square (RMS) and Wavelet Packet Transform (WPT). The superiority of this feature set has been proven experimentally, and the results show that the classification accuracy could reach up to 99% to recognize the individual and combined for ten classes of finger movements using only two EMG channels.

I. INTRODUCTION

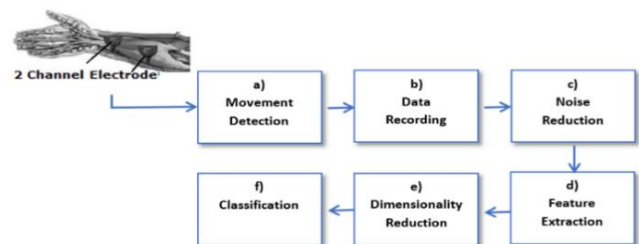
Loss of fingers function presents a major obstacle that required upper limb rehabilitation. The growth of spontaneous and accurate myoelectric systems would give an opportunity for people with upper limb disability, aged people and people with congenital defects to live and work in a way that was previously difficult.

Myoelectric control systems mainly use the Electromyography (EMG) signal, which is the electrical illustration of the neuromuscular activation correlated with an activity of skeletal muscle. Therefore, the EMG signal is related to the anatomical and physiological properties of muscles from which it originates, and this is the reason behind using it in clinical diagnosis. The EMG signal normally is a function of time and is represented in terms of its amplitude, frequency and phase.

The EMG (AKA myoelectric) signal is essentially detected in two ways. The first one uses the non-invasive surface electrode placed at the skin surface. The second way uses a fine needle, wire or an implanted electrode, which is more invasive as it is taken from the belly of the muscle. The invasive way gives a strong signal towards motor units near the electrode. However, surface detection gives more general information of the activities in the underlying motor units. Recently, researchers focused on the surface EMG (sMEG) as it is easy to record, simple in operation and less painful for the patients [1].

The raw EMG signal provides a valuable information; however, it is in a useless form. This signal information will

be useful only if it can be quantified. When compared with the other bio-signals, the EMG is a noisy signal as it gains the noise during traveling through different types of muscle tissue. Therefore, classifying and analyzing the EMG signals is not easy [2]. To maximize the quality of the EMG signal, it is required to increase the signal-to-noise ratio and minimize distortion of the EMG signal as much as possible. Different signal-processing stages are required to apply to the raw signal to get the accurate and useful information. Fig. 1 shows the stages implemented in the proposed pattern



recognition system.

Figure 1: Block diagram for EMG pattern recognition System

The pattern recognition system can be divided into three main stages; feature extraction, feature reduction and classification. The most important stage is the classification stage. To get better classification, researchers have used different types of EMG features as an input to the classifier. The selection of good features has been shown to have a higher effect on classification accuracy rather than the type of the selected classifier itself [3]. Selecting more features leads to more classification accuracy which is suitable in off-line processing. However, more feature extraction leads to more time consumption and it will be not relevant for real time application, therefore, the selected features should have less time consumption (it was agreed that delay less than 300ms is acceptable for myoelectrical control) [4].

In this paper, 18 feature combinations have been tested by using the sequential forward searching (SFS) method, and found the best combination set of features which consist of three-time domain features and a one time-frequency domain feature. This combination yields an accurate and fast classification system for individuated and combined finger movements. In this study, these features have been applied practically to recognize ten classes of fingers movement. These classes consisted of five individual finger movements; four finger combinations and a hand closure. To reduce the dimensionality, the extracted features applied to Spectral Regression Discriminant Analysis (SDRA), then using the Extreme Learning Machine (ELM) for classification and Bayesian vote for post-processing to smooth the data.

II. BACKGROUND FOR MOST WELL KNOW FEATURES

This section is providing information about common features used with EMG analysis and the types of these features. Features can be categorized under three main types; time domain (TD), frequency domain (FD) and time-frequency domain(TFD).

A. Time Domain Feature

Time domain (TD) features have been used universally in EMG pattern recognition system. Its advantages are; can be used in real-time application, required low computational complexity, easy implementation and having a good performance in low noise environment [5]. Herewith a mathematical definition for the thirteen most powerful TD features:

| Time Domain Feature Extraction | Expression |
|--|---|
| Integrated EMG (IEMG): It perform as the summation of the absolute values of the EMG signal amplitude. | $\sum_{n=1}^N x_n $ |
| Mean Absolute Value (MAV): It is like Average Rectified Value (ARV). It performs as the area under the EMG signal once it has been rectified. | $\frac{1}{N} \sum_{n=1}^N x_n $ |
| Modified Mean Absolute Value (MAV1): It is an extension of MAV using weighting window function w_n | $\frac{1}{N} \sum_{n=1}^N w_n x_n $ $w_n = \begin{cases} 1, & \text{if } 0.25N \leq n \leq 0.75N \\ 0.5, & \text{Otherwise} \end{cases}$ |
| Modified Mean Absolute Value2 (MAV2): is like MMAV1. However, it is using the continuous weighting window function w_n to improve the smooth window. | $\frac{1}{N} \sum_{n=1}^N w_n x_n $ $w_n = \begin{cases} 1, & \text{if } 0.25N \leq n \leq 0.75N \\ \frac{4n}{N}, & \text{if } 0.25N > n \\ \frac{4(n-N)}{N}, & \text{if } 0.75N < n \end{cases}$ |
| Mean Absolute Value Slope (MAVSLP): is a modified version of MAV. The differences between the MAVs of adjacent segments are determined. | $MAV_{i+1} - MAV_i$ |
| Simple Square Integral (SSI): This method use the energy of the EMG signal. | $\sum_{n=1}^N x_n ^2$ |
| Variance of EMG (VAR): This method use the power of the EMG signal as a feature | $\frac{1}{N-1} \sum_{n=1}^N x_n^2$ |
| Root Mean Square (RMS): This method uses the Amplitude Modulated Gaussian Random process whose RMS is related to the constant force and non-fatiguing contraction | $\sqrt{\frac{1}{N} \sum_{n=1}^N x_n^2}$ |
| Waveform length (WL): This method uses the cumulative length of the waveform over the time segment. | $\sum_{n=1}^{N-1} x_{n+1} - x_n $ |
| Zero crossing (ZC): This | $\sum_{n=1}^{N-1} sgn(x_n \times x_{n+1}) $ |

| | |
|--|---|
| method uses the number of times that the amplitude value of EMG signal crosses the zero y-axis | $(x_n - x_{n+1} \geq Theshod)$ $sgn = \begin{cases} 1, & \text{if } x > threshold \\ 0, & \text{otherwise} \end{cases}$ |
| Slope Sign Change (SSC): This method like ZC. However, it is another method to represent the information of EMG signal. | $\sum_{n=1}^{N-1} f[(x_n - x_{n-1}) \times (x_n - x_{n+1})]$ $f(x) = \begin{cases} 1, & \text{if } x \geq threshold \\ 0, & \text{otherwise} \end{cases}$ |
| Willison amplitude (WAMP): This method uses the number of times that the difference between EMG signal amplitude among two adjacent segments that exceeds a predefined threshold same as ZC and SSC | $\sum_{n=2}^{N-1} f(x_n - x_{n+1})$ $f(x) = \begin{cases} 1, & \text{if } x \geq threshold \\ 0, & \text{otherwise} \end{cases}$ |
| Hjroth time domain parameters (HTD): This time domain feature introduced by Hjorth (1970). The parameters for this method are Activity, Mobility, and Complexity. | <i>Activity:</i> $m_0 = var(x(t))$ <i>Mobility:</i> $m_1 = \sqrt{\frac{m_0(\frac{dx(t)}{dt})}{m_0(x(t))}}$ <i>Complexity:</i> $m_2 = \frac{m_1(\frac{dx(t)}{dt})}{m_1(x(t))}$ |

B. Frequency Domain Features

Spectral or Frequency domain features are frequently used in fatigue analysis. The features are mostly extract from power spectral density (PSD). Herewith a mathematical definition for the four most powerful FD features:

| Frequency Domain Feature Extraction | Expression |
|--|---|
| Autoregressive (AR): It describe each sample as a linear combination of previous samples plus a white noise error term. | $x_n = -\sum_{i=1}^p a_i x_{n-i} + W_n$ |
| Modified Median Frequency (MMDF): In this method, the spectrum is divided into two regions with equal amplitude | $\sum_{j=1}^{MMDF} A_j = \sum_{j=MMDF}^M A_j = \frac{1}{2} \sum_{j=1}^M A_j$ where A_j is the signal amplitude spectrum at frequency bin j . |
| Modified Mean Frequency (MMNF): It is the average frequency. | $\sum_{j=1}^M f_j A_j / \sum_{j=1}^M A_j$ |
| Traditional median frequency (MDF): It is calculated based on power spectrum | $\sum_{j=1}^M f_j P_j / \sum_{j=1}^M P_j$ |

C. Time Frequency Domain Features

Time-frequency (TFD) domain features are more accurate than the time domain and frequency domain features separately. The reason is that TFD can determine the energy of the EMG signals both in time and in frequency. However, the TFD transformation needs heavy computation; which make it not be reasonable for real-time application. Most well-known TFD features are; short-time Fourier transform (STFT), wavelet transform (WT) and wavelet packet transform (WPT).

D. Background for the proposed Feature Set

As mentioned the best combination of the feature set proposed in this paper, consists of three-time domain and one time-frequency domain features which are Hjorth time domain parameters (HTD), Root Mean Square (RMS), Mean absolute value (MAV) and Wavelet Packet Transform

(WPT). HTD proposed three time-domain features for the EEG signal: activity, mobility, and complexity. These features can be extended to the EMG signals as well and give good results (Mouzé-Amady & Horwat, 1996).

The wavelet transforms are generally classified into two types: Discrete Wavelet Transforms (DWT) and Continuous Wavelet Transform (CWT). DWT is used for non-situational application and it is suitable for time-scale approach. This method has been used in many applications like in analysing EMG signals [6-7]. However, DWT produces a high dimensional feature, which leads to increases in the learning parameters of a classifier [8]. To overcome this drawback, the DWT algorithm was developed to produce the Wavelet Packet Transform (WPT) algorithm. In this algorithm, the wavelet filter has been used to dissolve the input signal into wavelet coefficients, and uniform frequency bands. Therefore, this method has an advantage over DWT algorithm and can be used to analyse the EMG signal.

The WPT can be demonstrated as a tree. The original data set is the root of the tree. The next levels of the tree are the result of the previous step of the wavelet transform. Subsequent levels in the tree pass the EMG signal into multiresolution subsets of coefficients and pass through high-pass and low-pass filters. According to the wavelet function type the coefficients of the filter has been calculated [9].

III. FEATURE BASED FINGER MOVEMENT CLASSIFICATION

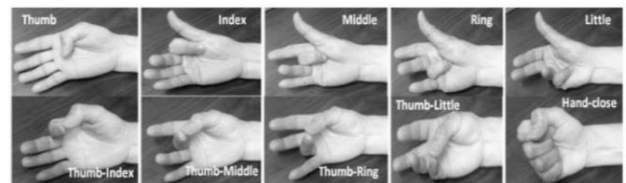
A feature maybe not relevant individually, but it becomes relevant when it joins other features [9]. Well-known feature sets have been combined and tested by famous researchers, such as the feature set of Hudgin (MAV, MAVS, SSC, ZC, WL) (Hudgins et al., 1993), Englehart's feature set (MAV, ZC, SSC, WL) (Kevin Englehart & Hudgins, 2003) and Hargrove's set (MAV, MAVS, SSC, ZC, WL, RMS, AR5) (L.J. Hargrove et al., 2007).

New EMG pattern recognition system has been developed by R. N Khushaba et al. (2012) for finger movements using support vector machine (SVM). Six time-domain features were used; slope sign changes (SSC), zero crossings (ZC), sample skewness (SS), waveform length (WL), Hjorth time domain parameters (HTD) and autoregressive (AR) model parameters. Then, the dimension of the extracted features was reduced using linear discriminant analysis (LDA). Experimental results show that the accuracy of the system reached 92%.

Recently K. Anam et al. (2017) extracted features from two EMG channels in addition to one extra channel by perform the summation of these two channels. The features were extracted from a time domain feature set which consists of (SSC), (WL), (ZCC), (SS), (AR) Model Parameters and some parameters from (HTD). By using this combination, the recognition system achieves accuracy of 97.07% [10].

IV. PROPOSED SYSTEM

In this study the data was collected from eight subjects; two females and six males, with ages between 25-35 and all the subjects are normal and don't have any muscle issues or disorder [10]. The proposed pattern recognition system, consists of multi stages as shown in Fig1. In the first stage data acquisition device has been used to collect the EMG signals. Where a FlexComp Infinity™ was used to collect the surface Electromyography signal through two MyoScan™ Sensors (T9503M). These electrodes placed directly on two groups of hand muscles; extensor digitorum (ED) and planaris longus (PL), as shown in Fig1. In the second stage, the collected signal was filtering and windowing before the extraction stage. Sampling frequency was set to 2000Hz and a window size of 256ms was used, as it should be less than 300ms for real time control [5]. After that, the feature extraction method has been applied, using the feature set. Then dimensionality of the features has been reduced using Spectral Regression Discriminant Analysis (SDRA). Extreme Learning Machine (ELM) has been used to classify the data and refined using the majority vote. This combination of dimensionality reduction, feature extraction and the majority vote, produce a fast and accurate classification system for individuated and combined finger movements. Fig. 2 shows the ten classes of the individual and combined finger movements consisting of five individual finger movements, i.e., Thumb (Th), Index (In), Middle (Mi), Ring (Ri), Little (Li) and four finger combinations Thumb - Index (Th - In), Thumb - Middle (Th - Mi), Thumb - Ring (Th - Ri), Thumb - Little (Th - Lia), and the hand close. Experimentally the subjects achieved the



finger combination movement starting with hand opening and then follow the configuration above, for a period of 5 seconds. This should be repeated six times. During these six trials, the collected data were divided into two groups; four trials were used as training and remaining trials were used as testing.

Figure 2: Different Finger Movement [10]

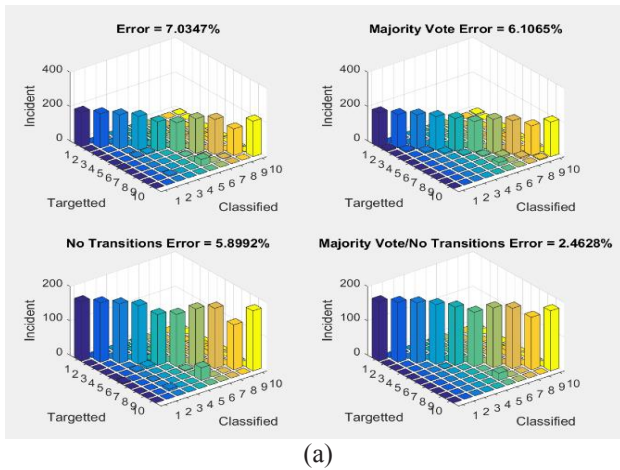
To select the best features combination that is most relevant to the problem, we used a sequential forward searching (SFS) method. SFS is the best method when the subsets are not big. The way that this method works is by creating two groups: group A and group B. Group A is empty from any feature at the beginning and group B contains all the features. Initially select the first best feature (with a minimum classification error) and move it from group B to A. Then start adding features from group B to A, sequentially and iteratively by pairing each feature from group B with all features in group A to give a minimum classification error with minimum testing time.

V. RESULTS

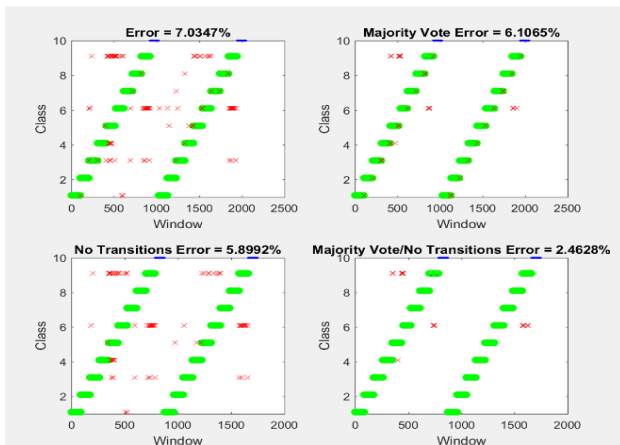
In this proposal, SFS method has been used to test eighteen most important features and applied to the proposed system (in Section IV). There are four feature combination which give the best results in terms of accuracy (more than 99%) and less processing time (around 46.9ms). These feature set are; MVA-3 features, WPT-7 features, RMS-2 features and HTD-9 features. Table1 shows the multi-feature set. Set1 (S1) contributes better classification than the feature set used in previous studies when applied to the same pattern recognition system. The information from the classification results as a function of time has been plotted in Fig. 3(a). Also, the confusion matrix for these multi-features set with the original and noisy EMG signal has been plotted as shown in Fig. 3(b).

TABLE1: FEATURES Set APPLIED EXPERIMENTALLY TO THE PROPOSED SYSTEM

| Group | # Feat. | Feature Set | % Accuracy | Time (ms) |
|-----------------------|---------|-----------------------------------|------------|-----------|
| S1 - The Proposed Set | 21 | MAV, RMS, WPT, HJORTH | 99.1% | 46.9 |
| S2 - K. Annam | 24 | SSC, ZC, WL, MAV, SS, HJORTH | 97.07% | 46.9 |
| S3 - K. Rami | 36 | SSC, ZC, WL, MAV, MAVS, RMS, 6AR | 97.2% | 93.8 |
| S4 - Hargrove | 42 | SSC, ZC, WL, MAV, SS, HJORTH, 6AR | 96.7% | 99.4 |



(a)



(b)

Figure 3: (a) The classification results as a function of time.
(b) The confusion matrix for the selected multi-feature set.

A statistical one-way ANOVA test has been used with $p = 0.05$ showed that the proposed feature set gives better performance than other sets.

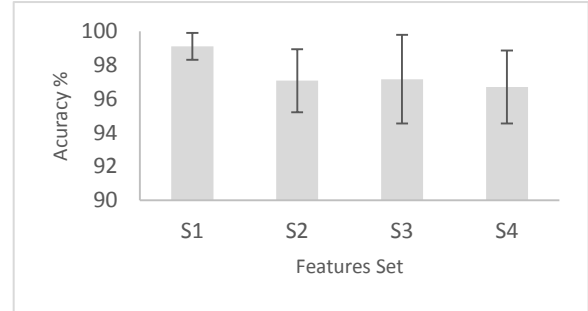


Figure 4: Classification accuracy for different feature set

VI. DISCUSSION & CONCLUSION

The objective of this paper is to present a study for using two sEMG channels to classify ten individual and combined finger movements. Eighteenth state-of-the-art features have been tested using SFS method to obtain the optimal and powerful feature set. The founded feature set in this paper has been achieved higher accuracies than reported in previous studies. The obtained feature set enhanced the classification accuracy in both offline and online system. The results in this paper can be used in a wide range of clinical and engineering applications. Future work is recommended to consider more subject data for further result enhancement.

REFERENCES

- [1] R. Merletti, D. Luca, C.J., "New Techniques in Surface Electromyography. In Computer Aided Electromyography and Expert Systems," *Sensors (Basel)* 2013 Sep; 13(9): 12431–12466.
- [2] A. Alkan, M. Günay," Identification of EMG signals using discriminant analysis and SVM classifier," *Expert System with Application* 2012, 39, 44–47.
- [3] G.M. Cavalcanti, T.M. Vieira, "Surface electromyography: Why, when and how to use it," *Rev. Andal. Med. Deport.* 2011, 4, 17–28.
- [4] G. Gokmen, "Wavelet Based Reference Current Calculation Method for Active Compensation Systems," *Electronics and Electrical Engineering*, 2011, No.2(108). P. 61–66.
- [5] A. Popov, A. Kanaykin, M. Zhukov, O. Panichev, O. Bodilovsky, "Adapted Mother Wavelets for Identification of Epileptiform Complexes in Electroencephalograms," *Electronics and Electrical Engineering*, 2010, No. 8(104), P. 89–92. 18.
- [6] V. M. Georgieva, "An Approach for Computed Tomography Images Enhancement," *Electronics and Electrical Engineering*, 2010, No. 2(98), P.71–74.
- [7] J. U. Chu, I. Moon, M. S. Mun, "A Real-Time EMG Pattern Recognition System Based on Linear–Nonlinear Feature Projection for a Multifunction Myoelectric Hand," *IEEE Transactions on Biomedical Engineering*. IEEE, 2006. –Vol. 53(11). – P. 2232–2239.
- [8] R. N. Khushaba, and A. Al-Jumaily, "Fuzzy Wavelet Packet based Feature Extraction Method for Multifunction Myoelectric Control," *International Journal of Biological and Medical Sciences*, Vol. 2, No. 3, pp. 186-194, 2007.
- [9] Khairul Anam, Rami N Khushaba and Adel Al-Jumaily, "Two-Channel Surface Electromyography for Individual and Combined Finger Movements," *35th Annual International Conference of the IEEE EMBS Osaka, Japan*, 3 - 7 July 2013.