Feature Extraction and Classification for EMG Signals Using Linear Discriminant Analysis

SACHIN NEGI
Dept. of Electrical Engineering
G. B. Pant Engineering College
Pauri- Garhwal, Uttarakhand, INDIA
sachinnegi.coep@gmail.com

YATINDRA KUMAR
Dept. of Electrical Engineering
G. B. Pant Engineering College
Pauri- Garhwal, Uttarakhand, INDIA
kyatindra@gmail.com

V. M. MISHRA
Dept. of Electrical Engineering
G. B. Pant Engineering College
Pauri- Garhwal, Uttarakhand, INDIA
vmm66@rediffmail.com

***Abstract*—Analysis of EMG signal has been an interested topic in recent years for classifying surface myoelectric signal patterns. Myoelectric control is an unconventional method to control the upper limb prostheses, human-assisting robots and rehabilitation devices. The aim of present work is to assess the time-domain features of EMG signal for myoelectric control of upper extremity prostheses by utilizing scatter plot. Classification accuracy is calculated using linear discriminant classifier for different combination of feature vectors using principal component analysis (PCA) and uncorrelated linear discriminant analysis (ULDA) feature reduction techniques. Results show that willison amplitude and waveform length are the best features for separating the distinct upper-limb motions.**

Keywords—EMG; MES; PCA; ULDA; LDA.

I. INTRODUCTION

A record of bio-electric potentials associated with the activity of muscles at rest and during contraction constitutes the electromyogram. The simplest way to pick these bio-electric potentials from certain areas is the surface EMG (sEMG), where surface electrodes are placed over the muscle. Electromyography (EMG) is a technique for assessing and recording the electrical action produced by skeletal muscles. An EMG signal shows the activity level of a muscle, and may be used to analyse neuromuscular diseases such as myopathy and neuropathy [13]. Moreover, these myoelectric signals (MES) can be utilized for myoelectric control applications.

In the field of biomedical instrumentation, the use of MES is of utmost importance for controlling the assistive devices like upper limb prostheses, wheelchairs and rehabilitation devices. Many disabled people have difficulty in operating traditional user interfaces like joysticks and keyboard. So it has been a fascinating research area to develop an advanced human–machine interfaces to overcome the difficulty in accessing assistive devices. One solution to overcome this difficulty is the myoelectric control for such assistive devices. Myoelectric control refers to as the process of controlling an external device by utilizing myoelectric or sEMG signals from the human muscles. Surface myoelectric signals can be utilized as the effective input for controlling upper arm prostheses.

Hudgins et al. [1] described a novel method for controlling a multifunction prosthesis based on the classification of myoelectric patterns. They have implemented multifunction control strategy using an artificial neural network to classify myoelectric patterns. Zardoshti et al. [2] have evaluated eight EMG features for the control of myoelectric upper extremity prostheses. Englehart et al. [3] proposed an ensemble of time-frequency based representations in order to increase the accuracy of transient myoelectric signal pattern classification. Han et al. [4] proposed an EMG pattern classification method based on fuzzy pattern classification and fuzzy min-max neural networks (FMMNN) techniques to aid the disabled and the elderly handle rehabilitation robotic arm systems.

Zecca et al. [5] described the traditional methods for the control of artificial hand by using EMG signal and also described some pattern recognition techniques. Du et al. [6] presented feature extraction techniques for both temporal and spectral approaches. They have discussed multiple trapezoidal windows and multiple hamming windows for the classification of prehensile EMG signal. Reaz et al. [7] illustrated some methods and algorithms for EMG signal analysis and also highlighted hardware implementations using EMG for applications related to prosthetic hand control and grasp recognition. Chan et al. [8] developed a MATLAB library for myoelectric control. They used RMS and autoregressive coefficients as features and also compared two different feature reduction methods- PCA and ULDA.

Oskoei et al. [9] reviewed the progress in the pattern recognition based and non-pattern recognition based myoelectric control. Munteanu et al. [10] analyzed the time and frequency domains of EMG signals of healthy person and patients with muscular disorders. Phinyomark et al. [11, 12] used scatter plot to evaluate time domain and frequency domain features. They also investigated the classification accuracy for the features extracted from the first difference of EMG time series.

In present study authors have evaluated fourteen time domain features and extracted the best possible feature and their combinations to get high classification accuracy.

978-1-5090-3480-2/16/$31.00 ©2016 IEEE

II. MATERIALS

MES data were picked up by using Duotrode Ag-AgCl electrodes from eight locations on right hand (seven locations on the forearm and one location at bicep). For common ground reference an Ag-AgCl electrode was positioned on the wrist. Electrodes are used to convert ionic potential into electric potential. These signals were amplified with an amplifier having a gain of 60dB and bandwidth of 1 Hz to 1 kHz. By using an analog-to-digital converter these signals were sampled at 3 kHz. Before sending MES data for pattern classification these were downsampled from 3 kHz to 1 kHz.

MES data were collected for seven limb motions- hand close, hand open, pronation, supination, wrist flexion, wrist extension, and rest. For a person four sessions were performed on separate days. In each session six trials were completed. The subject have performed each limb motion four times within each trial and each motion was performed for three seconds. The order of each limb motions was random [8].

III. METHODOLOGY

A general block diagram of a myoelectric control system (MCS) and the mathematical definitions of various time-domain features are presented in this section.

A. Myoelectric Control System

In a MCS, different patterns of sEMG signals are apperceived and matched with the control commands. A general block diagram of MCS is shown in Fig.1. Primarily the sEMG signals are acquired from surface electrodes placed over the muscles. Usually, the acquisition process is performed together with pre-processing sEMG signals in order to reduce the effect of noises and improve spectral components of sEMG signals. Due to the small amplitude of sEMG signals, these are amplified with the gain of the amplifier set normally to 60dB. An analog-to-digital converter is used to sample the continuous sEMG signals. As a result, discrete sEMG signals are obtained from this process. Subsequently, using a band-pass filter having a high CMRR, filtering of sEMG signals are done to reduce motion artifact and other high-frequency random noises [8, 9].

Next block is the recognition of sEMG patterns. The pattern recognition process can be divided into feature extraction, feature reduction, and classification. In feature extraction the input signal is transformed into a set of representative signal features. Various time domain features are presented in this study. The process of feature extraction may result in high dimensionality feature vectors. Therefore feature reduction process is required to reduce the dimensionality, simplifying the task of the classifier [8]. The most popular feature reduction methods- PCA and ULDA are used. The classification algorithm or classifier is the next phase in sEMG pattern recognition. Reduced feature vector is sent to classifier and is classified into seven classes.

In present study the LDA classifier is used. The advantage of LDA classifier is that iterative training is not required and it avoids the under- or over-training [8]. Finally the commands are generated based on the decisions in the pattern recognition block which are used for myoelectric control of various applications.

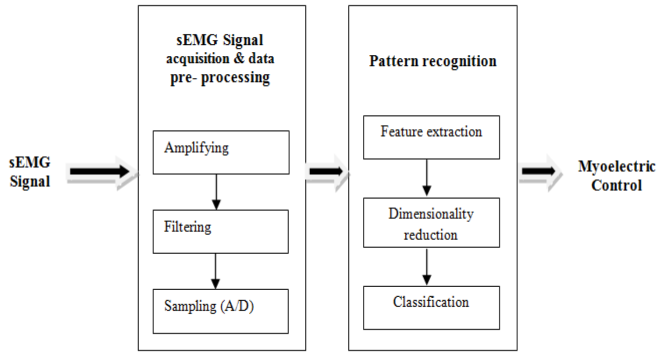


Fig. 1. Block diagram of myoelectric control system.

B. Time-Domain Feature Extraction

Time-domain features are the mostly used for myoelectric classification because they do not need a transformation and are based on signal amplitude. Mathematical definition of the time domain features [1, 2, 5, 6, 8, 11] are given here.

Root Mean Square (*RMS*): *RMS* of EMG signal is calculated as

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

where N is the length of the EMG signal and x_i represents the EMG signal in a segment i.

Integrated Absolute Value (*IAV*): *IAV* is calculated as

$$IAV = \sum_{i=1}^{N} |x_i| \tag{2}$$

Mean Absolute Value (*MAV*): *MAV* feature can be expressed as

$$MAV = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{3}$$

Modified Mean Absolute Value type 1 (*MAV1*): The extension of *MAV*, *MAV1* is calculated as

$$MAV1 = \frac{1}{N}\sum_{i=1}^{N} w_i |x_i| \tag{4}$$

;

$$\text{where} \quad w_i = \begin{cases} 1, & 0.25N \le i \le 0.25N \\ 0.5, & otherwise \end{cases}$$

Modified Mean Absolute Value type 2 (*MAV2*): Here the weighted window function w_i is a continuous function and is calculated as

$$MAV2 = \frac{1}{N}\sum_{i=1}^{N} w_i |x_i| \quad ; \tag{5}$$

$$\text{where} \quad w_i = \begin{cases} 1, & 0.25N \le i \le 0.75N \\ 4i/N, & i < 0.25N \\ 4(i-N)/N, & otherwise \end{cases}$$

Simple Square Integral (*SSI*): *SSI* is calculated as

$$SSI = \sum_{i=1}^{N} x_i^2 \tag{6}$$

Variance (*VAR*): *VAR* is calculated as

$$VAR = \frac{1}{N-1}\sum_{i=1}^{N} x_i^2 \tag{7}$$

The 3rd, 4th and 5th temporal moments: The 3rd, 4th and 5th order moments (*TM3*, *TM4* and TM5) can be expressed as

$$TM3 = \frac{1}{N}\sum_{i=1}^{N} x_i^3 \; ;$$

$$TM4 = \frac{1}{N}\sum_{i=1}^{N} x_i^4 \; ;$$

$$TM5 = \frac{1}{N}\sum_{i=1}^{N} x_i^5 \tag{8}$$

v-Order (*V*): The mathematical expression for *V* feature is defined as

$$V = \left(\frac{1}{N}\sum_{i=1}^{N} x_i^v\right)^{\frac{1}{v}} \tag{9}$$

Waveform Length (*WL*): *WL* feature can be expressed mathematically as

$$WL = \sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{10}$$

Average Amplitude Change (*AAC*): AAC can be calculated as

$$AAC = \frac{1}{N}\sum_{i=1}^{N-1} |x_{i+1} - x_i| \tag{11}$$

Difference Absolute Standard Deviation Value (*DASDV*): DASDV can be defined as

$$DASDV = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N-1} (x_{i+1} - x_i)^2} \tag{12}$$

Zero Crossing (*ZC*): *ZC* feature is used to count the number of times the EMG signal passes the zero amplitude axis. For two consecutive samples x_i and x_{i+1}, the *ZC* count is incremented, if

($x_i > 0$ and $x_{i+1} < 0$) Or ($x_i < 0$ and $x_{i+1} > 0$);

And ($|x_i - x_{i+1}| \ge$ threshold) (13)

Slope Sign Change (*SSC*): *SSC* feature shows the number of times the slope of EMG signal changes sign. For three consecutive samples x_{i-1}, x_i and x_{i+1}, the SSC count is incremented, if

($x_i > x_{i-1}$ and $x_i > x_{i+1}$) Or ($x_i < x_{i-1}$ and $x_i < x_{i+1}$)

And ($|x_i - x_{i+1}| \ge$ threshold or $|x_i - x_{i-1}| \ge$ threshold) (14)

Willison Amplitude (*WAMP*): *WAMP* feature counts the number of changes in the amplitude of EMG signal that exceed a pre-defined threshold and it can be expressed as

$$WAMP = \sum_{i=1}^{N} \left[f\left(|x_i - x_{i+1}|\right) \right]; \tag{15}$$

$$f(x) = \begin{cases} 1, & x \ge threshold \\ 0, & otherwise \end{cases}$$

Myopulse Percentage Rate (*MYOP*): *MYOP* feature can be expressed as

$$MYOP = \frac{1}{N}\sum_{i=1}^{N} [f(x_i)] \quad ; \tag{16}$$

$$f(x) = \begin{cases} 1, & x \ge threshold \\ 0, & otherwise \end{cases}$$

Threshold used in *ZC*, *SSC*, *WAMP* and *MYOP* features depends on the gain setting of instrument and on level of background noise.

Mean Absolute Value Slope (*MAVS*): *MAVS* feature i.e. the difference between MAV of the adjacent segments k and $k+1$, can be expressed as

$$MAVS = MAV_{k+1} - MAV_k \; ; \quad k = 1, \ldots K-1 \tag{17}$$

where K is the number of segments used in EMG signals.

Auto-Regressive (*AR*) Coefficients: AR prediction model is expressed as

$$x_i = \sum_{p=1}^{P} a_p x_{i-p} + e_i \tag{18}$$

where P is the order of the AR model, e_i is a white noise term and coefficients of the AR model (a_p) are used as feature vectors.

Multiple Time Windows Features: Multiple Hamming windows (*MHW*) and multiple trapezoidal windows (*MTW*) features can be defined as

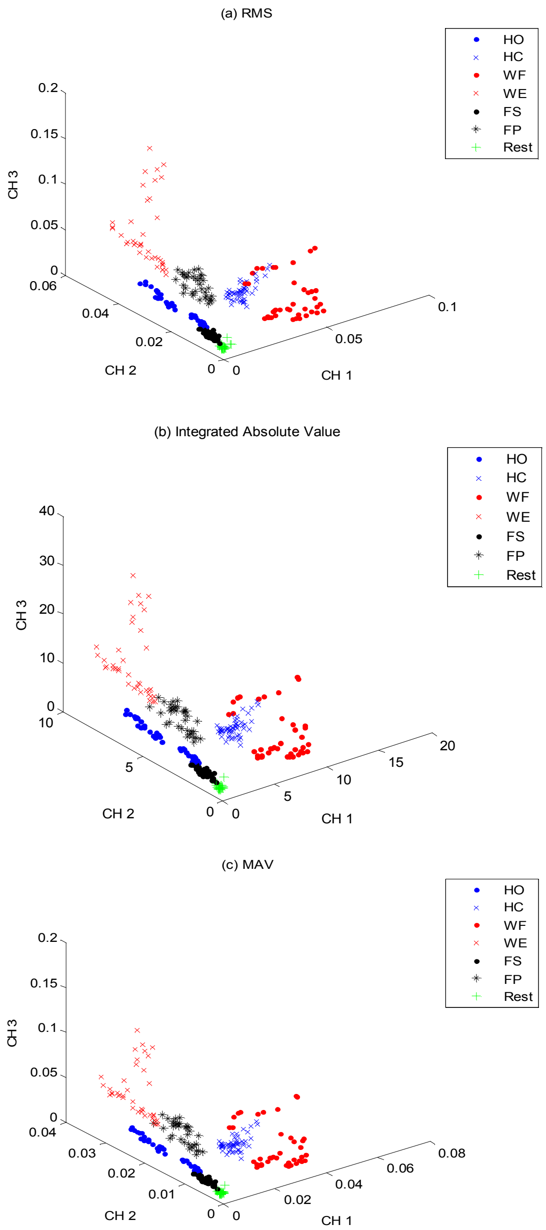
$$MHW_k = \sum_{i=0}^{N-1} \left(x_i w_{i-i_k}\right)^2 ; \quad k = 1, \ldots K \tag{19}$$

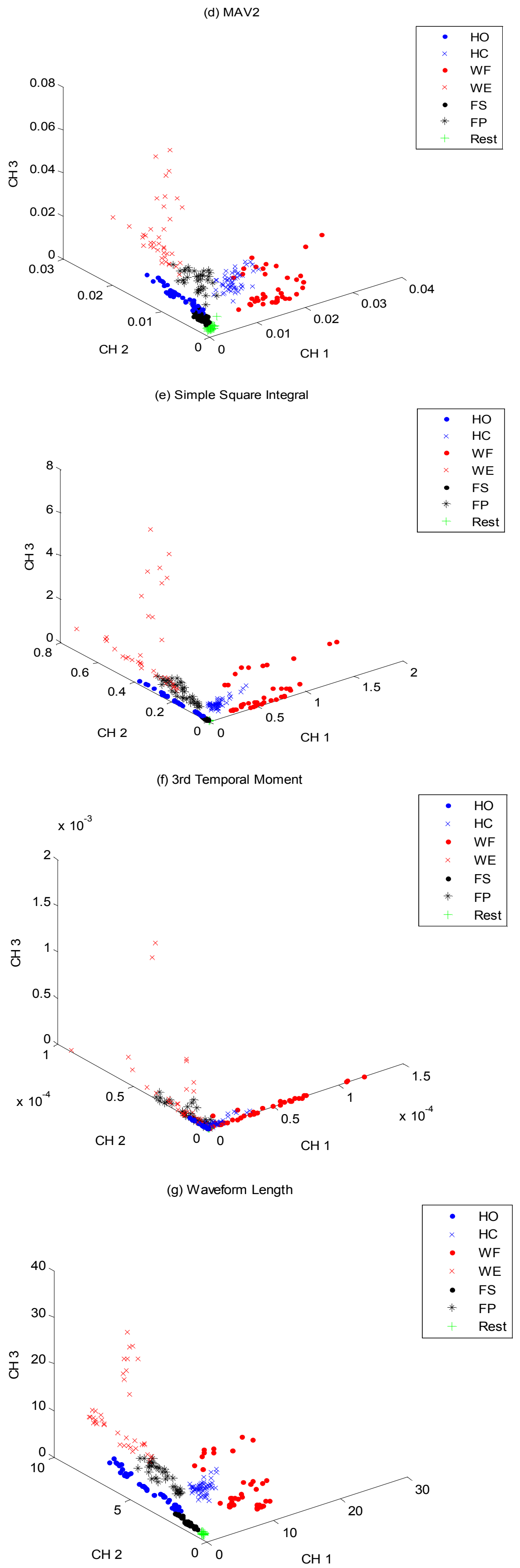
$$MTW_k = \sum_{i=0}^{N-1} \left(x_i w_{i-i_k}\right)^2 ; \quad k = 1, \ldots K \tag{20}$$

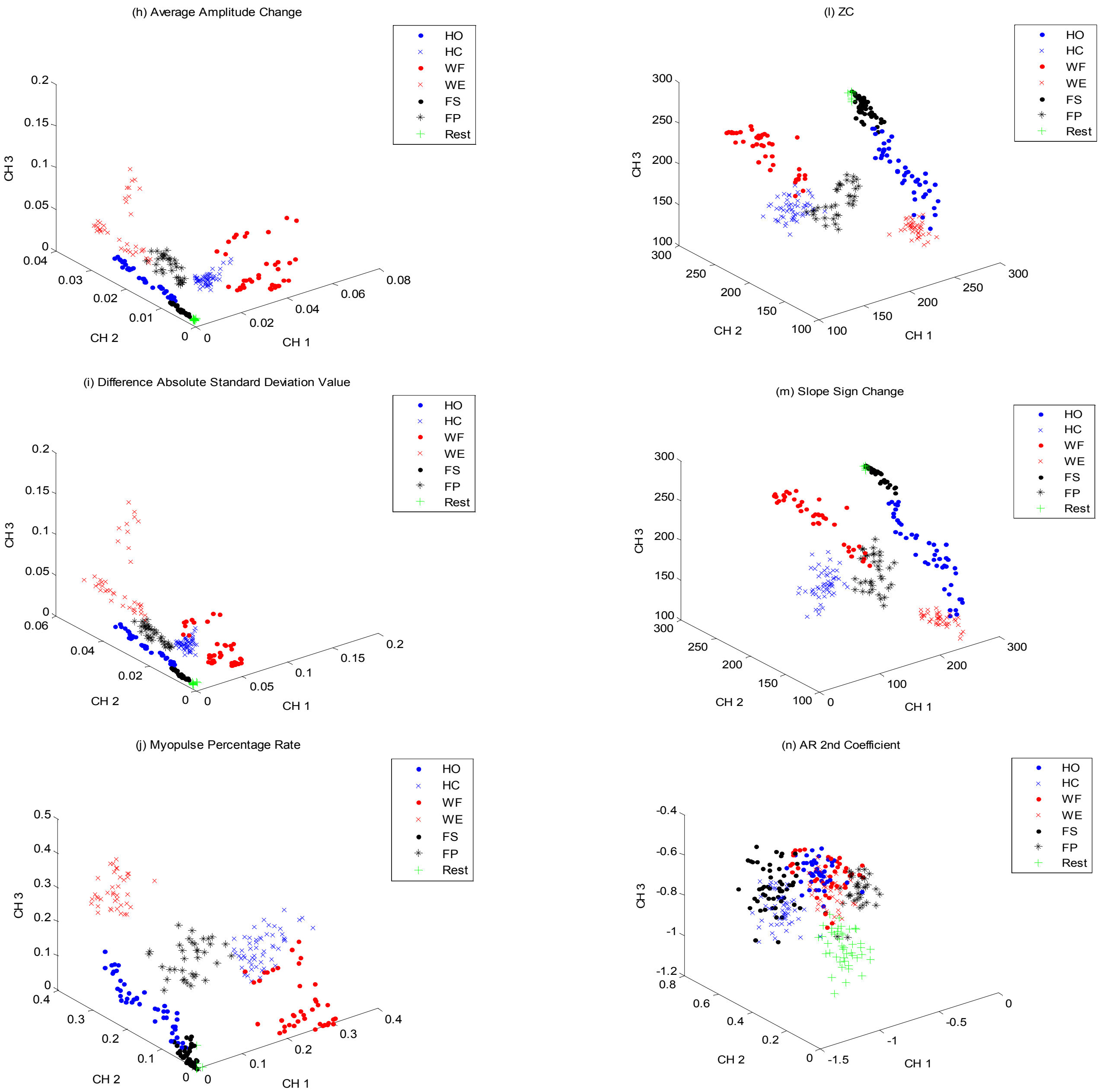
where w is respective windowing function.

IV. RESULTS AND DISCUSSION

The scatter plot between features extracted from three muscle channels for seven distinct movements (hand close, hand open, pronation, supination, wrist flexion, wrist extension, and rest) is used to evaluate the performance of classification. A scatter plot is a tool for analyzing relationships between two or three variables for a set of data. Three variables can be defined from three muscle channels (one feature per channel) or from single muscle channel (three features for a channel). In present study the three variables are defined for three muscle channels (CH-1, CH-2 and CH-3) data. The scatter plot of three channels and seven limb motions data are plotted for fourteen time- domain features as shown in Fig.2 (a-n).







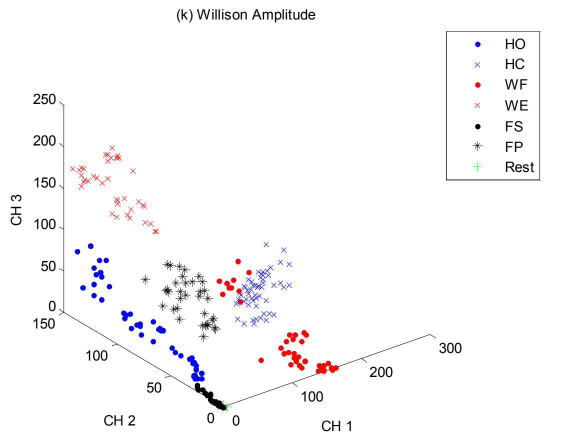


Fig. 2. Scatter plots for 7 upper-limb motion of a subject by applying individual feature: (a) RMS, (b) IAV, (c) MAV, (d) MAV2, (e) SSI, (f) TM3, (g) WL, (h) AAC, (i) DASDV, (j) MYOP, (k) WAMP, (l) ZC, (m) SSC, and (n) AR (2nd Coefficient).

Fig.2 of scatter plot for various time- domain features shows that distinct limb motions are significantly separable for willison amplitude and waveform length features. It concludes that the feature vectors combinations that include these features, may give the best classification accuracy.

The classification accuracy is obtained for some combination of feature vectors using PCA and ULDA feature reduction techniques [Table 1].

Table 1. CLASSIFICATION ACCURACY (%) OBTAINED FOR TIME DOMAIN FEATURE VECTORS.

Feature vectors	**(Mean ± standard deviation) of classification accuracy (%)**	
	PCA feature reduction	***ULDA feature reduction***
4-AR Coefficients/ RMS	92.84075 ± 1.55155	95.26652 ± 0.93092
4-AR Coefficients/ MAV	92.66076 ± 1.51334	95.26747 ± 0.94982
4-AR Coefficients/ WL	92.73626 ± 1.55765	95.43218 ± 0.98250
4-AR Coefficients/ DASDV	93.10343 ± 1.51430	95.56215 ± 0.77709
4-AR Coefficients/ WAMP (threshold 0.01)	92.12178 ± 2.00375	96.35613 ± 0.69265
RMS/MAV/SSI/TM3/DASDV	88.33430 ± 4.40186	89.81725 ± 3.40176
IAV/SSI/WL/SSC/ZC (threshold 0.01)	94.62583 ± 1.64647	93.46676 ± 1.58710
MAV1/WL/AAC/ZC/WAMP (threshold 0.01)	95.85695 ± 0.96225	95.39813 ± 1.44309
MAV2/WL/SSC/MTW/WAMP (threshold 0.01)	94.71765 ± 1.83791	94.40291 ± 1.43169
IAV/VAR/TM5/ZC/WAMP (threshold 0.01)	92.44743 ± 2.45073	93.37722 ± 2.67736

Linear discriminant classifier is used for classification purpose. In this paper, MES data of a subject is taken for 4 sessions and 6 trials are used in every session. For classification, the first trial data is used for training purpose and the data from all trials is used for testing purpose. The window size used was 256 ms. Training data had 50% overlap (128 ms) between windows whereas testing data had 87.5% overlap (32 ms) between windows.

The classification performances of most of the time domain features and some combinations of the features are estimated. Results show that a relatively simple pattern classification system using LDA can achieve good classification accuracy. The classification performance using two feature reduction methods: PCA and ULDA are also compared. Results clearly demonstrate that for feature vectors that include AR coefficients as features, ULDA gives better classification accuracy compare to PCA feature reduction.

V. CONCLUSION

A number of time domain features of EMG signals were presented. Scatter plots of fourteen time domain features were plotted to check their classification performances. The results showed that seven limb movements were significantly separable for WAMP and WL features. Also the feature combinations [4-AR Coefficients/ WAMP] and [MAV1/WL/ AAC/ZC/WAMP] showed the best classification accuracy. Effective feature reduction was demonstrated using ULDA for most of the feature vectors that include AR coefficients as features.

ACKNOWLEDGMENT

The authors would like to acknowledge Dr. Adrian D. C. Chan, Professor, Carleton University, CANADA, to provide EMG database.

REFERENCES

[1] Hudgins B, Parker P, Scott RN, "A new strategy for multifunction myoelectric control," IEEE Trans. Biomed. Eng., vol. 40, no.1, pp. 82-94, 1993.

[2] M. Zardoshti-Kermani, B. C. Wheeler, K.Badie, and R. M. Hashemi, "EMG feature evaluation for movement control of upper extremity prostheses," IEEE Trans. Rehabil. Eng., vol. 3, no. 4, pp. 324-333, Dec. 1995.

[3] K. Englehart, B. Hudgins, P. A. Parker, and M. Stevenson, "Classification of the myoelectric signal using time-frequency based representations," Med. Eng. Phys., vol. 21, no. 6-7, pp. 431-438, July 1999.

[4] Han J. S., Song W. K., Kim J. S., Bang, W. C., Lee, H., & Bien, Z., "New EMG pattern recognition based on soft computing techniques and its application to control a rehabilitation robotic arm," Proceedings of 6th International Conference on Soft Computing (pp. 890–897), 2000.

[5] Z. Zecca, S. Micera, M. C. Carozza, and P. Dario, "Control of multifunctional prosthetic hands by processing the electromyographic signal," Crit. Rev. Biomed. Eng., vol. 30, no. 4-6, pp. 459-485, 2002.

[6] Du, S., & Vuskovic M., "Temporal vs. spectral approach to feature extraction from prehensile EMG signals," Proceedings of IEEE International Conference on Information Reuse and Integration pp. 344-350, 2004.

[7] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," Biol. Proced. Online,vol.8, no.1, pp. 11-35, Mar. 2006.

[8] A. D. C. Chan and G. C. Green, "Myoelectric control development toolbox", in Proc. 30th Conf. Canadian Medical and Biological Engineering Society, Toronto, 2007, M0100.

[9] M. A. Oskoei and H. Hu, "Myoelectric control systems-A survey," Biomed. Signal Process Control, vol. 2, no. 4, pp. 275-294, Oct. 2007.

[10] M. Munteanu, C. Rusu, D. Moga, R. Moga, G. Tont, "EMG Signals Case Study: A Time and Frequency Domain Analysis," MEDITECH 2011, IFMBE Proceedings 36, pp. 202–205, 2011.

[11] A. Phinyomark, P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," Expert Syst. Appl., vol. 39, no. 8, pp.7420-7431, June 2012.

[12] A. Phinyomark, Franck Q., Sylvie C., Christine S., Franck T., Yann L., "Feature extraction of the first difference of EMG time series for EMG pattern recognition," computer methods and programs in biomedicine 117, pp.247-256, 2014.

[13] R. M. Rangayyan, "Biomedical Signal Analysis: A Case-Study Approach," Wiley India Pvt. Ltd., New Delhi (2014).