Feature Selection and Ranking in EMG Analysis for Hand Movement Classification

Parvatam Ramya Chandrika
Department of Electrical and Electronics Engineering,
The National Institute of Engineering,
Mysuru, India – 570008
Email: parvatamramyachandrika@gmail.com

Dr Omkar S Powar
Department of Biomedical Engineering,
Manipal Institute of Technology,
Manipal Academy of Higher Education,
Manipal, Karnataka, India – 576104
Email: omkarpowar1991@gmail.com

Dr Krishnan Chemmangat
Department of Electrical and Electronics Engineering,
National Institute of Technology Karnataka,
Surathkal, Mangalore, India – 575025
Email: cmckrishnan@nitk.edu.in

Abstract — **Surface Electromyography has gained tremendous significance in the recent years due to its suitability and reliability in a wide range of applications like automatic prosthetic control, diagnosis of neuromuscular disorders, in robotics and many such fields. Considering such applications, identification of various muscular movements is necessary and hence, EMG pattern recognition is needed. This paper focusses on a generalised EMG pattern recognition of various hand movements. The data from Ninapro Database - 4 has been used for pattern recognition. The database has Surface Electromyogram (sEMG) data of 52 various hand movements. The data was subjected to pre-processing, feature extraction and classification. An average accuracy of 64.87% was obtained for a combination of seven features in the time (temporal) domain, using Linear Discriminant Analysis (LDA) as the classification model. The obtained classification accuracies are compared and discussed with respect to the state-of-the-art literature.**

Keywords—sEMG, Feature extraction, Classification

I. INTRODUCTION

Electromyogram signals are the biomedical signals that are generated due to muscular contraction, and the process of measuring these signals for analysis is called Electromyography. These signals contain information about the muscular state. They are measured using electrodes that are placed at the location of the desired muscular activity. Signal acquisition can be carried out in an invasive or non-invasive way. The most preferred technique is the non-invasive method, and the process is called Surface Electromyography. Surface Electromyogram (sEMG) signals have numerous applications in various domains, such as prosthetics, determination of muscle fatigue and neuromuscular disorders, and other clinical and industrial applications.

It is seen that hand movement recognition based on sEMG signal processing has attained significance in recent years because of its ability to achieve automatic control of hand prosthetic devices [10]. Hence, pattern recognition methods were proposed to achieve hand movement classification by *Englehart et al.* [12] and many others. It is observed that features in the time domain gave a good performance in hand gesture identification [11]. It is also to be noted that the number of hand movements chosen can affect the accuracy of classification [12]. Some of the pioneering works carried out on the Ninapro Database include the analysis of Ninapro database by *Manfredo Atzori et al.* [14]; the comparison of six EMG acquisition setups for hand movement classification by *Stefano Pizzolato et al.* [2]; Fast Signal Feature Extraction Using Parallel Time Windows by *Manfredo Atzori et al.* [8].

In general, sEMG pattern recognition method involves these steps – Segmentation of data, Extraction of features and Classification [1]. In this work, the analysis of sEMG signals was carried out through the following four steps: (i) Data acquisition, where the acquired sEMG signal data of various subjects is imported from the database; (ii) Data pre-processing, to segregate the data for batch processing and filter out the noise signals; (iii) Feature extraction, to extract relevant time domain features and (iv) Classification, to classify various hand movements based on the extracted features.

In this paper, the main aim is to carry out EMG pattern recognition by hand movement classification using sEMG signal processing for various subjects and various feature combinations. The obtained classification accuracies are then to be compared with the available literature. This paper is subdivided into various sections: Section I contains an Introduction; Section II describes the methodology utilized in this work; Section III includes the results and relevant discussion and Section IV comprises the conclusion of the work carried out.

II. METHODOLOGY

The EMG pattern recognition is carried out by sEMG signal processing using MATLAB software, and this involves four steps as mentioned above: (i) Data acquisition; (ii) Data pre-processing; (iii) Feature extraction and (iv) Classification. Fig. 1 represents the block diagram of the methodology employed.

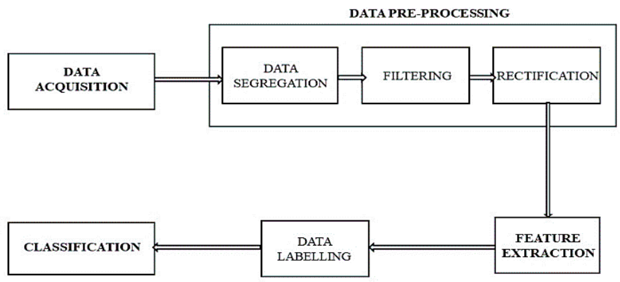

Fig. 1. Block diagram representing EMG Pattern Recognition method

A. Data Acquisition

The sEMG signal data used in this work has been obtained from Ninapro (Non–Invasive Adaptive Hand Prosthetics) Database 4 [2]. Ninapro is an open-source database publicly available for EMG data. In database 4, the EMG data corresponding to 52 hand movements was acquired from 10 intact subjects [2]. These 52 movements included basic finger movements, basic wrist movements, isometric configurations and grasping movements. These movements were grouped into three different sets of exercises. The first exercise set, Exercise A, includes 12 basic finger movements. Exercise B contains 17 isotonic, isometric and basic wrist movements. Exercise C includes 23 different grasping and functional configurations. Figure 2 indicates some of the hand movement configurations used in the Ninapro database. As a part of signal acquisition, the subjects were asked to carry out each movement for a duration of 5 seconds, followed by a rest period of 3 seconds. Each movement was repeated six times [2].

Fig. 2. Some of the hand movement configurations used in Ninapro database – 4

The electromyogram signals were obtained from the forearm using Cometa Wave Plus wireless sensors connected to Dormo SX-30 ECG electrodes [2]. 12 electrodes were used: eight electrodes were placed on the forearm corresponding to the joint between the radius and humerus; two electrodes were positioned at the activity locations of flexor digitorum and extensor digitorum muscles; and two electrodes were positioned at the activity spots of biceps and triceps [2]. Since 12 electrodes were used, the data acquired has 12 channels. The database has 30 MATLAB files corresponding to the three exercises for 10 subjects. Each file contains the EMG data, the stimulus data representing the repetition of movements, and other general information such as height, weight, gender etc.

B. Data Pre-processing

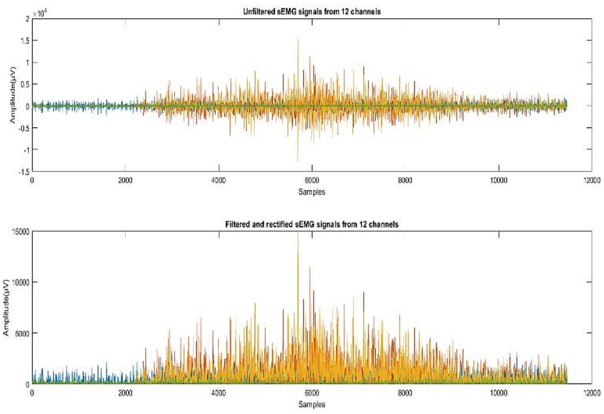

Fig. 3. Unfiltered and filtered sEMG signals from 12 channels

The raw EMG data present in the Ninapro database-4 is imported into MATLAB software for processing. The EMG data corresponding to the stimulus data (repetition of movements) is mapped and stored in order to be used in batch processing. The data is sampled at a frequency of 2kHz [2]. During signal acquisition, there can be interference due to power lines, motion artifacts etc. Hence, in order to eliminate the noise signals, filtering of the EMG data is carried out. A high pass, Butterworth filter of fourth order, having a 10Hz cut-off frequency is used, followed by a low pass, Butterworth filter of fourth order, having a cut-off frequency of 1000Hz [2]. As the sEMG signals have both positive and negative amplitudes, it becomes arduous to carry out feature extraction. Also, there can be a loss of information while using certain formulas during feature extraction. Hence, the filtered EMG data is full wave rectified in order to convert the negative amplitude values into positive values, thereby preventing the loss of information [3]. Figure 3 shows the unfiltered and filtered EMG signals.

C. Feature Extraction

In order to acquire useful information from the raw sEMG data, feature extraction is carried out. These features are extracted by the process of windowing, where the entire signal is divided into small segments of fixed size called windows. Windowing technique is employed as EMG is a stochastic signal, and hence less information is available in any instantaneous sample of the signal [4]. The standard window size for EMG Signal processing is 200ms [5]. Hence, in this work, a window size of 200ms has been used. Further, these windows can be of two types – disjoint and overlapping. In this work, an overlapping sliding window with a window increment of 20Hz is used to extract the required features [8]. The windows can be of different types such as Rectangular, Hamming, Hanning etc. Here, the type of window chosen is rectangular.

The features to be extracted can belong to any of these three categories – time domain, frequency domain and time-frequency domain [6]. This paper focuses on features from the time domain as time domain features are easy to extract and implement in real-time, compared to other domains. The time domain features extracted here are Root Mean Square (RMS), Mean Absolute Value (MAV), Zero Crossing rate (ZCR), Waveform Length, Average Power, Skewness and Kurtosis.

(i) Root Mean Square (RMS) [6]: The Root Mean Square value is the square root of the mean of instantaneous values of the signal.

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N}|x_n|^2} \tag{1}$$

(ii) Mean Absolute Value (MAV) [6]: The Mean Absolute Value represents the moving average of the signal that is full wave rectified.

$$MAV = \frac{1}{N}\sum_{n=1}^{N}|x_n| \tag{2}$$

(iii) Zero Crossing Rate (ZCR) [15]: Zero Crossing Rate represents the count of zero axis signal crossings, i.e., the number of times, the signal undergoes a sign change.

$$ZCR = \frac{1}{2N}\sum_{n=1}^{N}|sgn(x_n) - sgn(x_{n-1})| \tag{3}$$

where, *sgn* is the sign function such that,

$$sgn(x_n) = \begin{cases} 0, & x_n < 0 \\ 1, & x_n \geq 0 \end{cases} \tag{4}$$

(iv) Waveform length (WL) [6]: Waveform length indicates the cumulative length of the wave or signal over a segment.

$$WL = \sum_{n=1}^{N}|x_{n+1} - x_n| \tag{5}$$

(v) Average Power (AVGPOW) [16]: The average power of a signal is the average value of instantaneous power of a signal.

$$AVGPOW = \frac{1}{N}\sum_{n=1}^{N} {x_n}^2 \tag{6}$$

(vi) Skewness (SKEW) [6]: Skewness indicates how the signal is spread about the mean value.

$$SKEW = \frac{1}{N}\sum_{n=1}^{N}\left[\frac{x_n-\bar{x}}{\sigma}\right]^3 \tag{7}$$

(vii) Kurtosis (KURT) [6]: Kurtosis represents how the peaks of the signal are distributed in comparison to the normal Gaussian distribution.

$$KURT = \left(\frac{1}{N}\sum_{n=1}^{N}\left[\frac{x_n-\bar{x}}{\sigma}\right]^4\right) - 3 \tag{8}$$

D. Classification

Once the features were extracted, classification was carried out following the work of *Englehart et al.* [9, 17]. Labels were added to the EMG data and the extracted feature data to facilitate classification using machine learning algorithms. The extracted features were then used in various combinations for classifying the hand movements. According to *Cipriani C et al.*, it is seen that for temporal features, Linear Discriminant Analysis (LDA) has a better performance than Multi-Layer Perceptron (MLP) [7]. Hence, Linear Discriminant Analysis (LDA), which is a supervised machine learning algorithm has been used for classification. LDA projects the data onto a lower-dimensional space to maximize the separation between the classes. In this work carried out, LDA of type 'pseudolinear' has been employed. During classification, the signals corresponding to the 52 movements are classified into different classes. 60% of the data was utilized for training the classification model and 40% of the data was utilized for testing the classifier. The different feature combinations used are: (i) RMS; (ii) MAV; (iii) ZCR; (iv) WL; (v) RMS, MAV, ZCR; (vi) RMS, MAV, WL; (vii) RMS, ZCR, WL; (viii) RMS, ZCR, WL, AVGPOW; (ix) RMS, MAV, WL, AVGPOW and (x) RMS, MAV, ZCR, WL, AVGPOW, SKEW, KURT.

III. Results And Discussion

The EMG data was imported from the database. The data was then subjected to pre-processing, followed by temporal feature extraction and classification using LDA. For classification, ten different feature combinations have been used. The average accuracies of all the ten subjects for the ten feature combinations were obtained separately for Exercises A, B and C. Then, an average of all the three exercises for the ten feature combinations was considered. The following table represents the average accuracies of the three exercises obtained for the various feature combinations for Database-4.

TABLE I: AVERAGE CLASSIFICATION ACCURACIES OF DATABASE-4 FOR 10 FEATURE COMBINATIONS

Sl .no	Feature Combination	Average Accuracy (%)
1	RMS	48.14
2	MAV	48
3	ZCR	24.35
4	WL	47.65
5	RMS, MAV, ZCR	54.66
6	RMS, MAV, WL	57.11
7	RMS, ZCR, WL	56.69
8	RMS, ZCR, WL, AVGPOW	60.44
9	RMS,MAV,WL, AVGPOW	60.89

10	RMS, MAV, ZCR, WL, AVGPOW, SKEW, KURT	64.87

Table I records the average accuracies of all the ten feature groups. Figure 4 shows the graphical representation of average accuracies of all the ten feature groups. Figure 5 indicates the Confusion Matrix of Subject 1 for Exercise A and 9th feature group. From the obtained accuracies, it is observed that the tenth feature group yields a maximum average accuracy of 64.87%. Also, the sixth feature group involving RMS, MAV and WL yields an accuracy of 57.11%. Among individual features, RMS and MAV give an accuracy of 48.14% and 48.00% respectively. The average accuracy for Exercises B and C for the tenth feature group was obtained as 59.36%. It is seen that *Stefano Pizzolato et al.* obtained an accuracy of ≃ 66% when they used a feature combination of RMS, Time domain statistics (TD), Marginalised discrete wavelet transform (mDWT) and Histogram together, with Support Vector Machine (SVM) as the classifier for classifying 41 different hand movements (Exercises B and C along with Rest hand configuration) [2]. *Manfredo Atzori et al.* obtained an accuracy of ≃ 65.5% for the same feature combination using Random Forests as the classifier [8].

It is observed that the accuracy obtained for the tenth feature group is less than the accuracy obtained by *Stefano Pizzolato et al.* [2] and *Manfredo Atzori et al.* [8]. One of the reasons for this could be the choice of features used in this work. The type of window and window increment used could have affected the accuracy. It is also to be noted that the number of movements chosen can affect the accuracy of classification [12]. Hence, for a particular dataset, feature and classifier, it is seen that the accuracy of classification reduces as the number of movements (classes) increases [13]. However, the accuracy obtained can be improved by the use of certain methods. It was observed by *Omkar S et al.* that the classification accuracy is enhanced by the application of Minimum Entropy Deconvolution Adjusted (MEDA) during the pre-processing stage [6]. *Rami N et al.* suggested that the output misclassifications by the classifier can be reduced by the use of Bayesian fusion post-processing method [4]. *Krasoulis A. et al.* suggested that the use of Inertial Measurements (IM) along with sEMG can boost the classification accuracy [10].

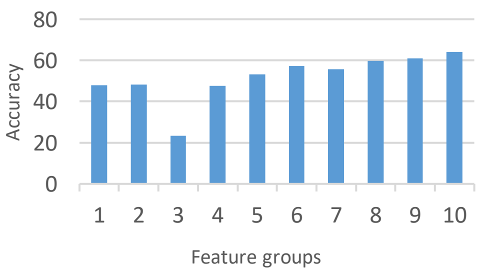

Fig 4. Plot of average accuracies for 10 feature groups

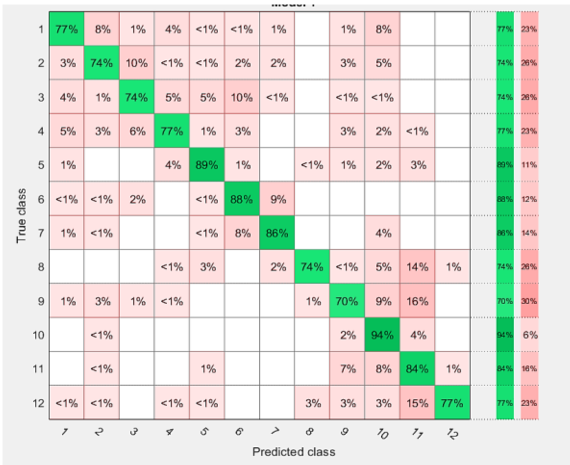

Fig 5. Confusion Matrix of Subject 1 for Exercise A and 9th feature group

IV. Conclusion

EMG pattern recognition was carried out where the data from the Ninapro database - 4 was subjected to segregation, filtering and rectification. After pre-processing, seven temporal features were extracted. The features were then used for classification into various hand movements using LDA as the classifier. The maximum average accuracy obtained was 64.87% for the tenth feature group. The obtained accuracies have been compared and discussed with respect to the literature.

References

[1] Li, C., Li, G., Jiang, G. et al. "Surface EMG data aggregation processing for intelligent prosthetic action recognition". Neural Comput & Applic 32, 16795–16806 (2020).

[2] Pizzolato, Stefano, et al. "Comparison of six electromyography acquisition setups on hand movement classification tasks." PloS one 12.10 (2017): e0186132.

[3] Reaz, M.B.I., Hussain, M.S. & Mohd-Yasin, F. "Techniques of EMG signal analysis: detection, processing", classification and applications. Biol. Proced. Online 8, 11–35 (2006).

[4] Rami N. Khushaba, Sarath Kodagoda, Maen Takruri, Gamini Dissanayake, "Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals, Expert Systems with Applications", Volume 39, Issue 12, 2012, Pages 10731-10738, ISSN 0957-4174

[5] R.N. Khushaba, A. Al-Timemy, A. Al-Ani, A. Al-Jumaily, "Myoelectric feature extraction using temporal-spatial descriptors for multifunction prosthetic hand control", IEEE 38th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC), IEEE, 2016, pp. 1696–1699.

[6] Powar, Omkar S., Krishnan Chemmangat, and Sheron Figarado. "A novel pre-processing procedure for enhanced feature extraction and characterization of electromyogram signals." Biomedical Signal Processing and Control 42 (2018): 277-286.

[7] Cipriani C, Controzzi M, Carrozza MC (2009) "Towards the development of the smart hand transradial prosthesis", Proceeding of IEEE 11th international conference on rehabilitation robotics, Kyoto, Kyoto, vol 35, pp 682–687

[8] Atzori Manfredo, Müller Henning, "PaWFE: Fast Signal Feature Extraction Using Parallel Time Windows", Frontiers in Neurorobotics, VOLUME 13, 2019, ISSN=16625218

[9] Englehart K, Hudgin B, Parker PA. "A wavelet-based continuous classification scheme for multifunction myoelectric control", IEEE Transactions on Biomedical Engineering. 2001;48(3):302–311. pmid:11327498

[10] Krasoulis, Agamemnon, Iris Kyranou, Mustapha Suphi Erden, Kianoush Nazarpour, and Sethu Vijayakumar. "Improved prosthetic hand control with concurrent use of myoelectric and inertial measurements." *Journal of neuroengineering and rehabilitation* 14, no. 1 (2017): 1-14.

[11] Resnik, Linda, Katherine Etter, Shana Lieberman Klinger, and Charles Kambe. "Using virtual reality environment to facilitate training with advanced upper-limb prosthesis." Journal of Rehabilitation Research & Development 48, no. 6 (2011).

[12] Englehart, Kevin, and Bernard Hudgins. "A robust, real-time control scheme for multifunction myoelectric control." IEEE transactions on biomedical engineering 50, no. 7 (2003): 848-854.

[13] Deng J, Berg AC, Li K, Fei-Fei L. "What does classifying more than 10,000 image categories tell us?", Computer Vision–ECCV 2010. Springer; 2010. p. 71–84.

[14] Atzori, Manfredo, Arjan Gijsberts, Claudio Castellini, Barbara Caputo, Anne-Gabrielle Mittaz Hager, Simone Elsig, Giorgio Giatsidis, Franco Bassetto, and Henning Müller. "Electromyography data for non-invasive naturally-controlled robotic hand prostheses." Scientific data 1, no. 1 (2014): 1-13.

[15] Sharma, Garima, Kartikeyan Umapathy, and Sridhar Krishnan. "Trends in audio signal feature extraction methods." Applied Acoustics 158 (2020): 107020.

[16] Gregd'Eon, Calculating the power of a signal, https://electronics.stackexchange.com/q/154218 , Version: 2015-02-14

[17] Phinyomark, Angkoon & Nuidod, Asan & Phukpattaranont, P. & Limsakul, Chusak. (2012). Feature Extraction and Reduction of Wavelet Transform Coefficients for EMG Pattern Classification. Elektronika ir Elektrotechnika. 122. 27-32. 10.5755/j01.eee.122.6.1816.