On the use of Temporal and Spectral Central Moments of forearm surface EMG for Finger Gesture Classification

Sneha Sharma
Electronics and Communication Engineering Department
Amity University
Noida, Uttar Pradesh-201313, India
sjangid@amity.edu

Rinki Gupta
Electronics and Communication Engineering Department
Amity University
Noida, Uttar Pradesh-201313, India
rgupta3@amity.edu

***Abstract*—Analyzing the surface electromyogram (sEMG) signal is becoming increasingly popular in fields other than medical diagnostics, such as assistive technology and human machine interfaces. This work focusses on analysing data from three sEMG sensors placed on the forearm in an armband configuration for the purpose of identification of finger gestures in a sign-language recognition system. The higher order central moments defining the shape of the power spectral density (PSD) are found to be particularly useful for the considered application. A comparative study of temporal and spectral central moments derived from the probability density function (PDF) and PSD of sEMG signals, respectively, is carried out to study their utility in the aforementioned application. Practical experiments reveal that spectral moments along with the most prominently used set of features out-perform the temporal moments in the considered classification. An average classification accuracy of 82.1% is achieved with temporal moments, which is improved to 90.1% with spectral moments.**

Keywords*—sEMG, sign-language recognition, spectral moments, temporal moments, probability density function, power spectral density

I. Introduction

Surface Electromyogram (sEMG) measures the electrical potentials generated by muscle cells in a non-intrusive manner by placing sensors on the skin surface over the muscle being monitored. Surface electromyogram has enormous field of applications, some of which are sign language recognition, human supporting prosthetic arms, human machine interaction, medical field, gamming and smart interfaces [1]. The motivation of this work is towards developing a sign language translator using surface electromyogram signals from forearm muscles. Apart from sEMG, various assistive sensing technologies has been used for gesture pattern recognition and they can be broadly classified as vision-based, glove-based, and inertial measuring unit (IMU) based systems.

Vision-based systems detect the motion based on camera recordings. For instance, 3D depth data from hand gestures, produced from Microsoft's Kinect sensor have been used in [1]. However, the performance of vision-based recognition systems depends on factors such as lighting condition, background texture and color. Moreover, the system is challenging to wear. Whereas, glove-based systems consist of a glove that is equipped with flex-sensors to measure the bending of fingers, and inertial sensors, namely accelerometers and gyroscopes to measure the motion and rotation of hand and/or fingers. The combination of glove-based flex sensors and sEMG sensors for complete functional hand assessment has been proposed in [1]. However, wearing a sensor-equipped glove can hinder with natural finger motions and could lead to discomfort for a sign-language user.

To enhance the wearability of a sign-language recognition system, it seems appropriate to place the sensors on the forearm of the signer. Identification of finger motion from sensors placed on the forearm have been reported in literature [2-5]. Detection of different finger tapping motions have been analysed using vibration patterns recoded from accelerometers placed on the forearm [2]. The effect of hand orientation on sEMG signals for analyzing single and multiple finger flexion motions is presented in [3]. A post-processing based on Bayesian fusion to improve the accuracy of finger motions from signals recorded from two sEMG sensors placed on the forearm is proposed in [4], while neural networks have been used in [5] to classify six numeric finger gestures with an average classification accuracy of 69.17%. In both [4, 5], a set of prominently used time-domain features have been employed for classification.

In feature-based classifications, the primary task is to identify or design features that best suit the application. In this work, the challenge lies in detecting finger gesture from forearm arm muscles. The difficulty in predicting finger movements from forearm sEMG signals originate from the fact that sEMG signals are recorded from skin surface and hence, not only are they a superposition of various muscle potentials but, they are also influenced by different nonlinear components, for example, fat and tissue, before they are summed with different possibilities. Finger movements are mostly fast, with little scope

978-1-5386-6918-1/18/$31.00 ©2018 IEEE
DOI 10.1109/ICMETE.2018.00059

of movement, in this way the amplitude of finger-movement related surface myoelectric possibilities is minute and of low signal-to-noise ratio. The utility of first and second order temporal and spectral moments of sEMG signals have also been discussed for classification of hand gestures [6] and medical diagnostics [7]. The high-order central moments defined using time-domain sEMG signals have been used in literature, for instance in detection of Parkinson's disease [7]. In this paper, novel definition of third and fourth order moments derived from the power spectral density (PSD) of the sEMG signals are proposed for identification of the considered finger gestures. Their utility is compared to the central moments derived from sEMG signals in time domain.

The remaining paper is organized as follows. The experimental setup for recording the sEMG signals and the basic signal processing required for extraction of commonly used features in the literature are described in Section II. Section III contains a detail description of features proposed to be useful for the classification of finger gestures from forearm sEMG signals along with the approach used for testing the proposed hypothesis. The results supporting the proposed hypothesis are given in Section IV. Section V concludes the paper.

II. Materials and Method

A. Data acquisition and experimental setup

Five healthy, able-bodied subjects, all females, 4 right-handed and 1 left-handed, between the age of 22-31 years participated voluntarily in this study. Nine finger gesture activities as shown in *Fig.1* were selected that include one to nine digits in the Indian sign language. The signs given in the Indian Sign Language (ISL) dictionary published by Ramakrishna Mission Vivekananda University have been followed [8]. The sEMG signals were collected while the subjects were performing the finger gesture activities.

ISL1 ISL2 ISL3 ISL4 ISL5 ISL6 ISL7 ISL8 ISL9

Fig. 1. Finger gesture activities

Delsys Trigno wireless system *Fig.2a* with following specification has been used for data acquisition- transmission bandwidth 20m, sampling period 900 µsec, signal bandwidth 20-450 Hz and 16-bit resolution depth. Three sensors were placed in a forearm armband configuration on the muscles flexor digitorum, extensor carpi radialis and brachioradialis, as shown in *Fig.2b*. After following the basic skin preparation procedure, the Delsys sensors are placed on the skin surface using adhesive interfaces to keep the motion artifacts to a minimum. Each activity was designed for the duration of 2 min and 45 second in which 5 second was given as rest period and 3 second was set as performing time with 20 repetitions, additional 5 second is considered at the beginning of each activity. While performing the gestures other than hand rest, the

(a)Delsys Trigno wireless sensor

(b) Sensor lacement on forearm muscles

Fig. 2. Experimental setup for data aquisition

subjects were asked to maintain a moderate level of force for each repetition.

B. Pre-processing and Feature Extraction

Firstly, the raw sEMG signals are processed for baseline removal. The sEMG signals are known to be zero-mean signals [9]. Hence, a moving average filter with 125 ms window length is selected to determine the baseline, which is removed from raw sEMG signals. Next, the region of activity is determined according to the change in the signal envelope above a certain threshold. Now, signals from each sEMG channel are windowed into 4 segments and a set of most commonly used features [10] is extracted from the signals. This set of features will now be referred to as the base-set and the features in the base-set are listed in Table 1.

Base-set features consist of time-domain features including mean absolute value (MAV), standard deviation (STD), waveform length (WL), zero crossing (ZCR), slope sign change (SSC). In Table 1, x_i characterizes the sEMG signal in a sample of i and N denotes the total length of sEMG signals. Mean value is given by $\bar{x} = (1/N)\sum_{i=1}^{N} x_i$ and threshold (th) is evaluated

TABLE I. Base-set features extracted from each channel of sEMG

Base-set Features	Definition
Mean absolute value	$MAV = \frac{1}{N}\sum_{i=1}^{N}\lvert x_i\rvert$
Standard deviation	$STD = \left(\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2\right)^{\frac{1}{2}}$
Waveform length	$WL = \sum_{i=1}^{N-1}\lvert x_{i+1} - x_i\rvert$
Zero crossing	$ZCR = \sum_{i=1}^{N-1}[sgn(x_i \times x_{i+1}) \cap \lvert x_i - x_{i+1}\rvert] \geq th$
Slope sign change	$SSC = \sum_{i=2}^{N-1}[f[(x_i - x_{i-1}) \times (x_i - x_{i+1})]]$

as twice the standard deviation of sEMG signal recorded during the rest durations.

The sEMG signals are known to be non-stationary and non-Gaussian because of which high-order moments have often been proposed to aid in classification [10, 11]. In the following section, features describing the shape of the sEMG signal distribution in time and frequency are analyzed for their utility in finger gesture classification.

III. Finger Gesture Classification from Forearm sEMG

The mathematical definitions of central moments are given as follows.

A. Temporal and Spectral Moments

The n^{th} order moment of a random variable z with probability density function (PDF) $f(z)$ is defined as [12]

$$m_n = \int_{-\infty}^{\infty} z^n f(z)\, dz. \quad (1)$$

Hence, the for n=0, m_0=1 since the area under the PDF is unity. For n=1, $m_1=\eta$ is the mean of the random variable. Similarly, higher order moments may be defined. The central moments are defined as [12]

$$\mu_n = \sum_{k=0}^{n} \binom{n}{k} m_k (-\eta)^{n-k}, \quad (2)$$

where m_k and η are as defined above and

$$\binom{n}{k} = \frac{n(n-1)\ldots(n-k+1)}{k(k-1)\ldots 1}. \quad (3)$$

The 2nd order central moment or *variance* is the amount of dispersal about the mean, defined as

$$\mu_2 = m_2 - \eta^2. \quad (4)$$

The 3rd order central moment is defined as

$$\mu_3 = m_3 - 3\,\eta\, m_2 + 2\,\eta^3, \quad (5)$$

and the 4th order central moment is given as

$$\mu_4 = m_4 - 4\,\eta\, m_3 + 6\,\eta^2 m_2 - 3\,\eta^4 \quad (6)$$

The normalized 3rd order central moment is termed as *skewness,*

$$\text{skew} = \frac{\mu_3}{\mu_2^{3/2}}. \quad (8)$$

Skewness measures the symmetry of the shape of a distribution. If the skewness is positive then the PDF is dominant on left of the mean with a long tail on right. The normalized 4th order central moment is known as kurtosis, defined as

$$\text{Kurtosis} = \frac{\mu_4}{\mu_2^2}. \quad (9)$$

Kurtosis is a well-defined shape parameter it measures peakedness or flatness of the signal distribution, the PDF of a distribution with large kurtosis has more fatter tail, for the standard Gaussian distribution the kurtosis value remains equal to 3.

In case of temporal moments, the PDF $f(z)$ is considered to be the PDF of the sEMG signal in time domain and z are the amplitude levels. The PDF of sEMG signals is simply defined by normalizing the histogram of the sEMG signal quantized in some finite number of levels. Here, 1024 levels are considered between signal amplitudes of $\pm 1.5\times10^{-3}$. In literature it is reported that the PDF of the sEMG signal under low muscle contraction are more peaked, while when the contraction force level rises the sEMG feature tends to be more bell-shaped Gaussian distributed [13]. This property could be useful in determining the figure gestures, since signs involving extension of multiple fingers have higher muscle activity in comparison with individual finger movements.

To define the spectral moments, the normalized PSD of the sEMG signal is taken as $f(z)$ and the z is replaced with the frequency of the i^{th} bin f_i. The PSD, which is evaluated as the square of the magnitude Fourier spectrum of the sEMG signal, is also normalized so that central moments may be defined on it. The first and second order moments of PSD have been reported to be particularly useful in detection of muscle fatigue [10]. The third spectral moment has also been reported [10], however, the higher order central moments of PSD have not been discussed in the context of activity classification from sEMG signals. Since skew and kurtosis convey information about the shape of the PDF, replacing $f(z)$ with the normalized PSD of the sEMG signals should describe the shape of the PSD of the sEMG signals. In the following, the temporal moments defined using the PDF of the sEMG signal as $f(z)$ are compared with the spectral moments wherein the PSD of the sEMG signal as $f(z)$.

B. Finger gesture classification

Fig. 3 shows the block diagram of the approach followed for evaluation of the temporal and proposed spectral features for finger gesture classification. After acquiring the raw sEMG data from 3 channel sensor, pre-processing steps were followed which includes baseline removal and activity detection as explained in Section II*B*. Then, windowing and extraction of features from the base-set, as defined in Table I, are carried out. The features from the base-set are used along with the spectral moments or the temporal moments to compare their performance in finger gesture classification. In this study, the temporal moments used are- temporal variance (TV), temporal skew (TS) and temporal kurtosis (TK), while the spectral moments include spectral variance (SV), spectral skew (SS) and spectral kurtosis (SK).

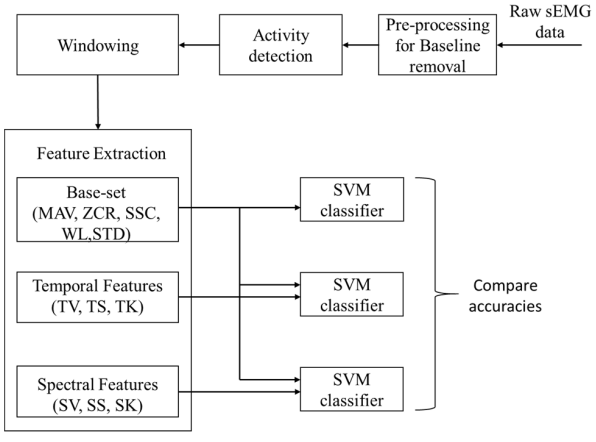
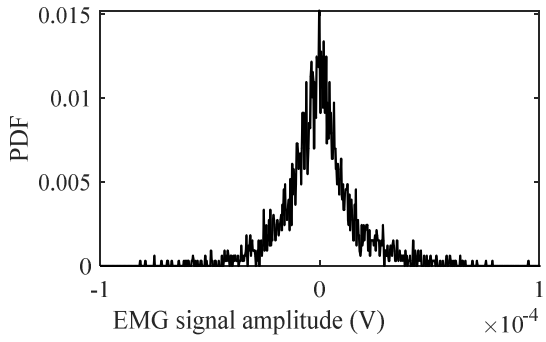


Fig. 3. Finger gesture classification approach

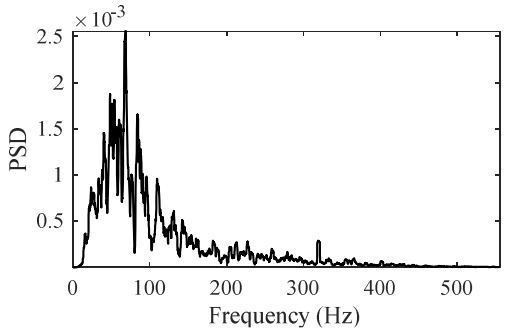
According to [14], SVM proves to be the most flexible classifier specialized in producing classification mapping, even for complex separating surfaces. Here, support vector machine (SVM) with radial basis function as kernel is used for classification. For multi-class classification the one-versus-all coding approach is used with binary learners. For each gesture, there are 100 observations available (20 observations each from 5 subjects). Five-fold cross validation has been employed to avoid over-fitting and the classification accuracies when spectral and temporal features are used along with the features from the base set are determined. The results are presented in the following section.

IV. Results

The PDF of the finger gesture corresponding to ISL 9 is plotted in *Fig.4a*. It can be seen that the PDF is centered around a zero-mean value with a low skew value, obtained as 0.2425, and kurtosis greater than 3 (determined as 5.3039) indicating a leptokurtic PDF. The PSD for the same gesture is shown in *Fig.4b*. For the PSD, the value of skew was obtained as 1.8287 which corresponds with the shape of the PSD. The kurtosis for PSD is obtained as 6.4026, which is again greater that 3 and the PSD is also non-Gaussian and leptokurtic.



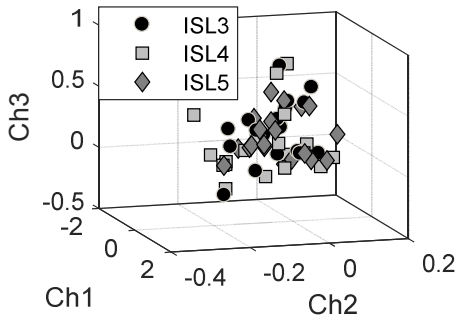
(a) Probability density function



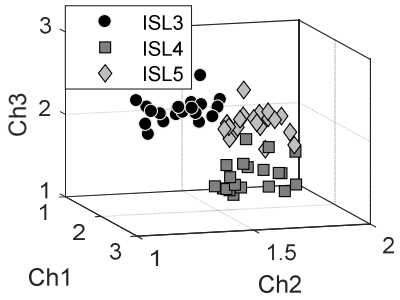
(b) Power spectral density

Fig. 4. PDF and PSD of sEMG signal for ISL 9

The scatter plots of the temporal and spectral skew determined for the sEMG signals recorded from the three channels (Ch 1,2,3) for the activities ISL3, ISL4, ISL5 are shown in *Fig. 5*. It is evident that the skew evaluated from PSDs (*Fig. 5b*) are more discriminative for the three finger gestures as compared to the skew evaluated from the PDFs shown in *Fig.*

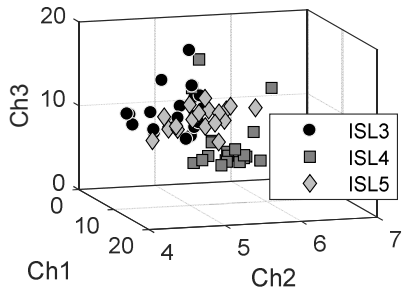


(a) Temporal Skew

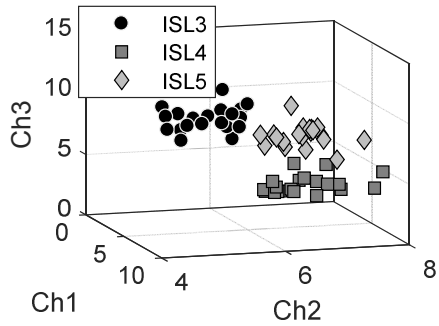


(b) Spectral Skew

Fig. 5. Temporal and spectral skew for ISL 3, ISL 4 and ISL 5



(a) Temporal Kurtosis



(b) Spectral Kurtosis

Fig. 6. Temporal and spectral kurtosis for ISL 3, ISL 4 and ISL 5

5a, which overlap for the three activities. Similarly, the temporal kurtosis shown in *Fig. 6a* for the three activities overlap and they cannot be expected to contribute towards classification of the activities. Whereas, the kurtosis evaluated from the PSDs plotted

in *Fig. 6b* show a clear separation between the classes and are useful for classification. Similar observation has been made for the temporal and spectral variance.

ISL 1	ISL 2	ISL 3	ISL 4	ISL 5	ISL 6	ISL 7	ISL 8	ISL 9	
87	0	3	1	1	2	3	3	0	**ISL1**
1	78	6	8	0	3	0	4	0	**ISL2**
1	4	83	1	7	0	1	1	2	**ISL3**
0	4	1	85	7	1	0	1	1	**ISL4**
0	0	4	9	80	1	0	1	5	**ISL5**
0	0	0	0	0	99	1	0	0	**ISL6**
2	3	2	1	0	2	76	12	2	**ISL7**
4	3	2	1	0	1	8	77	4	**ISL8**
0	1	2	1	2	0	1	4	89	**ISL9**

Fig. 7. Confusion matrix between output class and target class for base set features. Average classification accuracy over all activities is 83.8%.

ISL 1	ISL 2	ISL 3	ISL 4	ISL 5	ISL 6	ISL 7	ISL 8	ISL 9	
76	3	1	0	2	9	8	1	0	**ISL1**
1	82	8	2	2	2	0	2	1	**ISL2**
2	3	84	3	6	0	1	0	1	**ISL3**
1	4	1	86	6	1	1	0	0	**ISL4**
2	0	6	12	68	3	1	1	7	**ISL5**
3	1	0	0	1	93	2	0	0	**ISL6**
8	3	0	0	0	2	78	9	0	**ISL7**
0	2	0	0	1	0	3	86	8	**ISL8**
0	0	4	2	2	1	2	3	86	**ISL9**

Fig. 8. Confusion matrix between output class and target class for temporal moments along with base set features. Average classification accuracy over all activities is 82.1%.

ISL 1	ISL 2	ISL 3	ISL 4	ISL 5	ISL 6	ISL 7	ISL 8	ISL 9	
90	1	2	0	1	2	2	2	0	**ISL1**
1	90	2	1	0	0	1	5	0	**ISL2**
0	1	91	1	3	2	0	0	2	**ISL3**
0	2	1	92	3	2	0	0	0	**ISL4**
0	1	3	8	87	0	0	0	1	**ISL5**
0	0	0	0	0	100	0	0	0	**ISL6**
2	2	0	1	0	1	89	4	1	**ISL7**
2	6	0	0	1	0	7	80	4	**ISL8**
0	0	1	1	3	0	1	2	92	**ISL9**

Fig. 9. Confusion matrix between output class and target class for spectral moments along with base set features. Average classification accuracy over all activities is 90.1%.

In *Fig. 7-9* the average confusion matrix for the nine finger gestures i.e. ISL1, ISL2, ISL3, ISL4, ISL5, ISL6, ISL7, ISL8, ISL9 from all five subjects are stated. The average classification accuracies for all the activities for the base-set, base-set augmented with temporal moments and base-set augmented with spectral moments is 83.8%, 82.1% and 90.1%, respectively. On combining the base set features with the temporal moments, the average classification accuracy in fact reduces by 1.7%, whereas the spectral moments enhance the average classification accuracy by 6.3%. Individually, for all the nine signs, the spectral moments contribute more towards classification as compared to the temporal moments, with an exception of ISL 8.

Now, consider the signs ISL 3-5 and ISL 9 from *Fig.7-9*. As shown in *Fig.1*, the difference between ISL 5, and ISL 4 and ISL 9 is just that of the thumb or the little finger, respectively. Whereas the confusion of ISL 5 with ISL 3, 4 and 9 are 4, 9 and 5 times; on augmenting the base-set with temporal moments, these confusions increase to 6, 12 and 7 times. On the other hand, on augmenting the base-set with spectral moments, these confusions decrease to 3, 8 and 1 times, hence validating the utility of the proposed high-order spectral central moments in classification of finger gestures. In future, the database will be increased to include more subjects and more signs involving different finger gestures.

V. Conclusion

A three channel sEMG signal pattern recognition system was used in this paper to classify nine finger gestures movements corresponding with digits 0-9 in the Indian sign language. The challenge is to design such features that would aid the classification of finger gestures using sEMG signals from the forearm of the signer. The utility of a set of temporal and spectral central moment is tested. The proposed spectral central moments when used along with a set of commonly used time-domain features give an average classification accuracy of 90.1%, which is significantly higher as compared to that obtained when temporal central moments are used, which yield an average classification accuracy of 82.1%. Hence, the utility of central moments defined on the PSD of the sEMG signals is established.

Acknowledgment

The authors would like to recognize the funding support provided by the Science & Engineering Research Board, a statutory body of the Department of Science & Technology (DST), Government of India, SERB file number ECR/2016/000637.

References

[1] Xue, Y., Ju, Z., Xiang, K., Chen, J. , Liu, H., "Multiple sensors based hand motion recognition using adaptive directed acyclic graph". Applied Sciences, vol. 7, no. 4, pg. 358, 2017.

[2] Yu, Wenwei, et al. "Finger motion classification by forearm skin surface vibration signals." *The open medical informatics journal*, vol. 4, pg. 31, 2010.

[3] You, Kyung-Jin, Ki-Won Rhee, Hyun-Chool Shin. "Finger motion decoding using EMG signals corresponding various arm postures." *Experimental neurobiology*, vol. 19, no. 1, pp.54-61, 2010.

[4] R. N. Khushaba, S. Kodagoda, M.Takruri, G. Dissanayake, "Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10731-10738, 2012.

[5] Q. Li, B. Li, "Online Finger Gesture Recognition Using Surface Electromyography Signals," *Journal of Signal and Information Processing*, vol. 4, no. 02, pg. 101, 2013.

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

[7] Kugler, Patrick, Christian Jaremenko, Johannes Schlachetzki, Juergen Winkler, Jochen Klucken, and Bjoern Eskofier. "Automatic recognition of Parkinson's disease using surface electromyography during standardized gait tests," In *35th Annual International Conference of the IEEE on Engineering in Medicine and Biology* Society, pp. 5781-5784, 2013.

[8] Ramakrishna mission Vivekananda University, Coimbatore Campus, Indian Sign Language Dictionary, http://indiansignlanguage.org.

[9] C. J. De Luca, L. D. Gilmore, M. Kuznetsov, S. H. Roy, "Filtering the surface EMG signal: Movement artifact and baseline noise contamination," *Journal of Biomechanics*, vol. 43, no. 8 , pp. 1573-1579, 2010.

[10] A. Phinyomark, P. Pornchai, C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Systems with Applications*, Vol. 39, no. 8, pp. 7420-7431, 2012.

[11] R. Gupta, A. Kulshreshtha, "Analysis of dual-channel surface electromyogram using second-order and higher-order spectral features," *IEEE Int. Conf. Communication Control and Intelligent Systems*, pp. 49-53, 2016.

[12] A. Papoulis, S. U. Pillai, Probability, random variables, and stochastic processes, Tata McGraw-Hill Education, 2002.

[13] K. Nazarpour, A. H. Al-Timemy, G. Bugmann, A. Jackson, "A note on the probability distribution function of the surface electromyogram signal," *Brain research bulletin*, vol. 90, pp. 88-91, 2013.

[14] Abraham, L., Bromberg, F. and Forradellas, R., "Ensemble of shape functions and support vector machines for the estimation of discrete arm muscle activation from external biceps 3D point clouds". *Computers in biology and medicine*, *95*, pp.129-139, 2018.