Journal of Engineering and Applied Sciences 13 (Special Issue 8): 6341-6345, 2018
ISSN: 1816-949X

Support Vector Machine based Classification Improvement for EMG Signals using Principal Component Analysis

Vivek Ahlawat, Ritula Thakur and Yogendra Narayan
National Institute of Technical Teachers Training and Research, Chandigarh, India

Abstract: This study purposed and evaluates a method based on Support Vector Machine (SVM) classification of surface Electromyogram (sEMG) signals. The result shows that the best selection of time domain features gives better performance with quadratic SVM classifier. The sEMG signals are acquired from 15 healthy volunteers by placing the electrode on biceps and triceps muscles on the right arm. After the signal acquisition, pre-processing (denoising, rectification, filtering and amplitude normalization) is performed and suitable features were extracted for the classification purpose. In this research, sEMG data is acquired by two channels. PCA is used for dimension reduction of the feature vector. Higher classification accuracy is achieved by using the quadratic SVM classifier. The overall mean classification accuracy with selected time domain feature, i.e, Kurtosis, Skewness, Slop Sign Change (SSC), Mean Absolute Value (MAV), Autoregressive coefficient of the first order (AR1) is 99.04% with quadratic SVM classifier. This much accuracy can be used for designing assistive robotic device.

Key words: Quadratic SVM, sEMG, features extraction, design, mean absolute value, domain feature

INTRODUCTION

The information about the muscles activities can be analyzed by measuring the surface electromyogram signals. This technique is widely used in the application of clinical and industrial research (Reaz *et al*., 2006). It is also used to control the prosthetic devices for the amputees and partially paralyzed persons (Hudgins *et al*., 1993). The sEMG signals are generated by the muscles during the contraction (Lei *et al*., 2001). The amplitude of sEMG signal is too low in the range of millivolts and micro volts but it can be directly captured by the bio-signal acquisition devices for further processing and applications. These steps have an important role in classification and control applications. There is wide application range of sEMG signals including engineering, medical science, ergonomics and rehabilitation (Phinyomark *et al*., 2012). Human machine interface and robot control are the currently active research areas in the present scenario of engineering and technology. In medical engineering, electromyography is used as a tool to identify the neuromuscular disorder or disease (Phinyomark *et al*., 2009).

The movement identification of sEMG signals is a challenging task due to its non-stationary nature. This challenging task can be solved by the process of effective feature extraction for the sEMG signals. The features of sEMG signal can be extracted in various domains such as time domain, frequency domain and time-scale domain. An approach for having high accuracy with the combined features from different domains has been proposed (Tsai *et al*., 2015).

In the previous research, various time domain features have been studied such as Root Means Square (RMS), kurtosis, zero crossing, Wilson amplitude, etc. Yet, the classification accuracy achieved by researchers by the time domain signals was not satisfactory. The classification accuracy is an important parameter to distinguish the various muscles activities. Feature selection has a great influence to form the feature vector which is used as an input to the classifier. An appropriate selection of features provides the best result in classification. A control signal generated by the classifier would be better with the high accuracy of classification of the signal as much as possible.

A wide range of features has been investigated by the researchers. In this study, features are considered individually and in the group. All features are taken from time-domain, i.e., Mean Absolute Value (MAV), Root Means Square (RMS), Waveform Length (WL), Variance (VAR), Slope Sign Change (SSC), Kurtosis (KURT), Skewness (SKEW), Willson Amplitude (WAMP), Autoregressive coefficients of order 1-4 (AR1-4), Temporal Moments of order 1-7 (TM1-7) and Myopulse (MYOP) rate at 20 and 25 μV threshold value, etc.

First of all different time domain features are calculated and then more dominant are selected on the

Corresponding Author: Vivek Ahlawat, National Institute of Technical Teachers Training and Research, Chandigarh, India

basis of their accuracy. In this research, the higher classification accuracy is achieved by combining some time domain features.

MATERIAL AND METHODS

Data acquisition: The data were acquired from 15 healthy right hands dominated subjects of the age group from 20-24 years in National Institute of Technical Teachers Training and Research (NITTTR) Biomedical Instrumentation Lab. The hand movements were performed by the subjects under the predefined protocol. The device has two channel recording facility in which one channel was used for triceps and rest one was for biceps brachii muscles. The sEMG signal was sampled at 1000 Hz. Fifteen trials have been performed by each subject for each activity (flexion and extension) with a delay of 2 sec. Figure 1 represents the acquired raw sEMG signals. The location of electrodes is fixed at the acupressure point of the right-hand upper arm based on the myotrace manual as shown in Fig. 2. The MYOTRACE 400 device was used for data acquisition with some signal processing facilities.

The data acquisition process chart is shown in Fig. 3. Finally obtained feature set is normalized and dimensionality reduction also is done by PCA.

Signal processing: The amplitude of acquired sEMG signal is too low in the range of 5-20 μV, therefore, after amplification, rectification, smoothing, normalization of amplitude and filtering of the signal are done. A FIR bandpass filter with 20 Hz lower and 500 Hz higher cut off frequency was used for filtering. In rectification process, all sEMG signals are full wave rectified.

Feature extraction and reduction: Feature selection is a significant and essential stage for myoelectric control design. A lot of literature investigates and compares the various features in time domains, frequency domain and time-scale domain (Englehart *et al*., 2000, 2001; Oskoei and Hu, 2007). The features should be capable of presenting the properties or characteristics of sEMG signals for the different movements.

Feature selection is also a significant process in case of movement classification (Oskoei and Hu, 2006; Phinyomark *et al*., 2012a, b). In the case of more number of the features, there are some redundant features. So, the feature reduction is required for reducing the redundant feature. In this study, twenty-time domain features are calculated which makes a large feature vector set. The

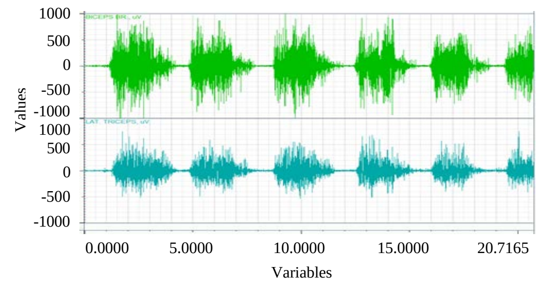

Fig. 1: Original sEMG signal acquired with the help of MyoTrace device

Fig. 2: EMG data acquisition system with the subject

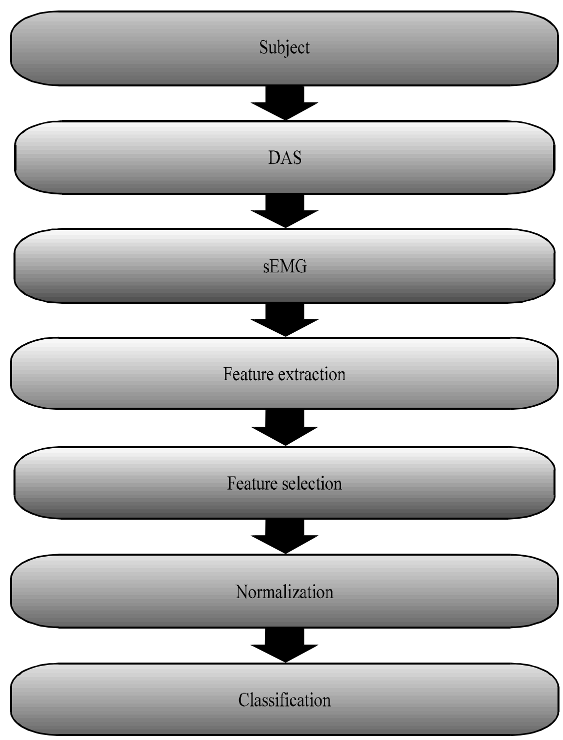

Fig. 3: Various steps of sEMG classification process

large feature vector will increase the classifier complexity and operation time. Principal Component Analysis (PCA) is used for feature reduction. Then the combination all of this feature is applied to the classifier. The formulation of the selected feature is given:

$$SSC = \sum_{i=2}^{N-1}\left[f\left[(X_i - X_{i-1}) * (X_i - X_{i-1})\right]\right] \quad (1)$$

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

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|X_i| \quad (2)$$

$$KUR = \frac{\frac{1}{n}\sum_{i=1}^{N}(X_i - \bar{X})^4}{\left(\frac{1}{n}\sum_{i=1}^{N}(X_i - \bar{X})^2\right)^2} \quad (3)$$

$$SKEW = \frac{\frac{1}{n}\sum_{i=1}^{N}(X_i - \bar{X})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{N}(X_i - \bar{X})^2}\right)^3} \quad (4)$$

Support vector machines classifier: Support vector machine is a classification method used in machine learning to solve the pattern recognition problems. In this study, the SVM is used for movement classification of sEMG data acquired from different subjects into their respective movements.

In this research, we used the SVM classifier for pattern classification of hand movements. Accuracy was testing on the linear, quadratic, cubic and fine Gaussian SVM classifiers (Table 1 and 2) with the PCA and various fold cross validations and holdout validations. Various classifiers have also been used in the previous literature (Phinyomark *et al*., 2012).

Two classes have been taken account, therefore, it is a binary classification problem. The SVM constructs a hyperplane of maximum margin that separates the sample points into the two classes as shown in Fig 4. The hyperplane of the classifier creates a boundary for the new data sets. Let w be normal to the hyperplane. The SVM labels the sample points x as +1 or -1 based on whether w, x+b is >1 or <-1. Where (w, b) is the chosen to maximize the margin of the decision boundary.

It is based on the statistical learning approach and has been widely used in pattern recognition. In a classification problem, a hyperplane distinct the classes of pattern based on the input space sample vector. The vectors near to the hyperplane are known as support vectors.

Table 1: Accuracy with PCA at different fold cross validation

	SVM classifier accuracy (%)			
Fold cross validation	Linear	Quadratic	Cubic	Fine Gaussian
2	90.90	98.50	98.20	94.50
3	93.90	99.40	98.20	97.90
5	94.20	99.10	99.40	98.20
7	94.50	98.80	98.80	97.90
9	94.20	99.40	98.80	98.20
Mean accuracy	93.54	99.04	98.68	97.34

Table 2: Accuracy with PCA at different holdout validation

	SVM classifier accuracy (%)			
Holdout validation (%)	Linear	Quadratic	Cubic	Fine Gaussian
15	95.9	99.0	99.0	99.0
20	86.4	97.0	97.0	97.0
25	95.1	99.0	98.8	98.0
30	92.9	99.0	97.0	98.0
35	96.5	99.1	99.1	96.5
Mean accuracy	93.36	98.62	98.18	97.7

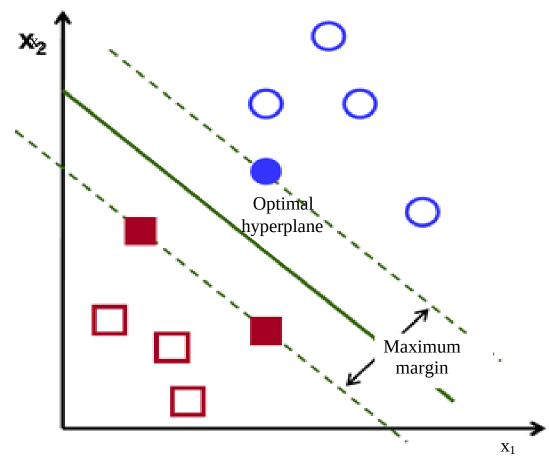

Fig. 4: Support vector machine classifier

RESULTS AND DISCUSSION

The samples of the selected features are shown in Fig. 5 by the scatter plot which indicates the good class separability, that will yield higher classification accuracy.

Evaluation of accuracy by SVM classifier: The accuracy of different SVM classifiers with time domain feature is given in Table 1 and 2. All classifier accuracy is evaluated with different fold cross validation and holdout validation as shown in Fig. 6 and 7.

It is observed from Fig 5 and 6 that, the quadratic SVM classifier has more stability and accuracy in comparison to others classifiers. The overall classification mean accuracy is 99.04 % by Quadratic SVM (Fig. 8).

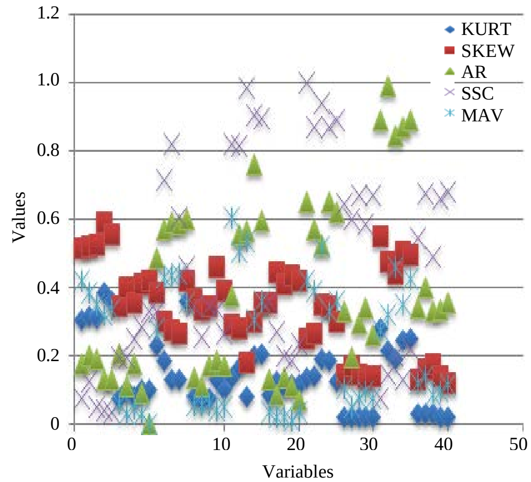

Fig. 5: Scatter plot of selected features

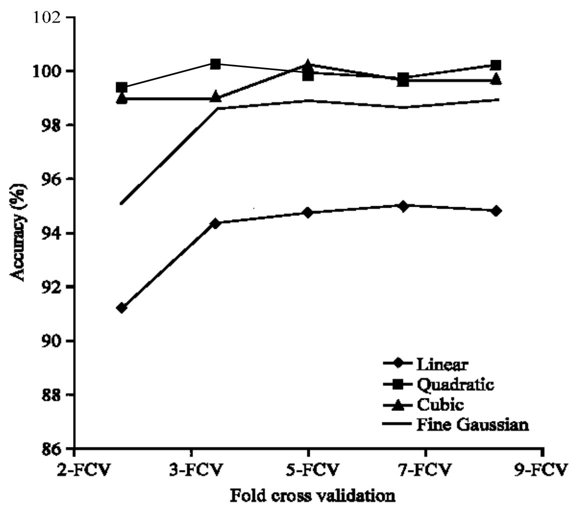

Fig. 6: Accuracy graph of SVM classifiers with fold cross validation

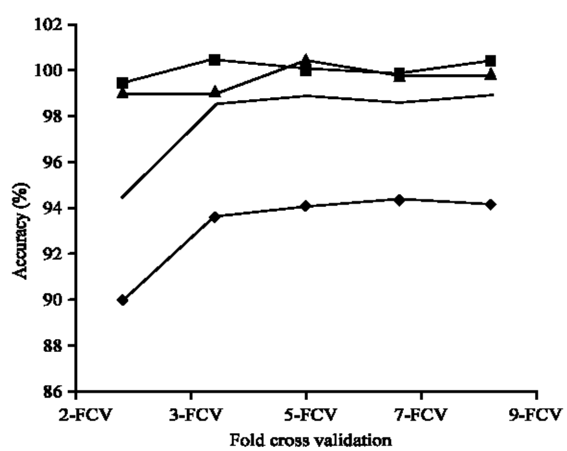

Fig. 7: Accuracy graph of SVM classifiers with holdout validation

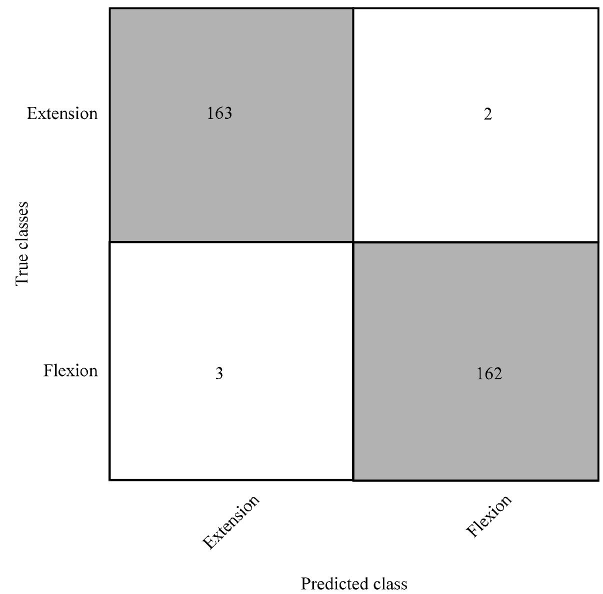

Fig. 8: Confusion matrix of quadratic SVM of extension and flexion movements by quadratic SVM

Evaluation by confusion matrix: The confusion matrix shows the number of samples in true classified and false classified for each class of movement. The overall mean accuracy for the flexion and extension movements of true class and predicted classes is 99.04%. The True Positive Rates (TPR) and True Negative Rates (TNR) can be obtained mathematically by the following formulation:

$$\text{TPR} = \text{TP}/(\text{TP+FN}) \tag{5}$$

And:

$$\text{TNR} = \text{FN}/(\text{TP+FN}) \tag{6}$$

Where:

TP = True Position
FN = False Negative

Figure 8 demonstrates the confusion matrix which indicates the truly classified and false classified sample of two movements. The 163 samples are truly classified in extension and 2 samples are wrongly classified. In the case of flexion 162 samples are truly classified in flexion and 3 samples are wrong classified in extension.

Evaluation through ROC: The ROC curve for flexion and extension are shown in Fig. 9. The curve depicts the relation between the TPR and FPR. The total area under the curve is 0.990421. It is given by the FPR of the current classifier and positive class and TPR of the current classifier and positive class. The overall accuracy can be formulated as.

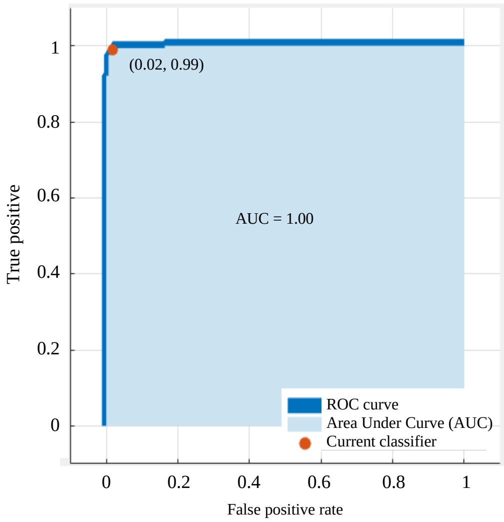

Fig. 9: ROC of extension and flexion movements by quadratic SVM classifier

CONCLUSION

This study investigates the classification of sEMG signals based on support vector machine. This methodology considers only the time-domain features. The result indicated the effectiveness of selected time domain features to improve the computational capacity of the classifier. PCA was used for feature dimension reduction. The experiment was conducted on the healthy subjects. The observations and conclusions are not definitive solutions but they can be a notable contribution to advance knowledge in this field. The main advantage of this approach is that it reduces the classification error, thereby improving accuracy. This way of classification can be more dominant with effective feature set consisted of KURT, SSC, SKEW, AR1, MAV features which yielded the best result by Quadratic SVM classifier.

RECOMMENDATIONS

In future, the classification accuracy can be improved by selecting the more effective set of the feature from time, frequency and time-scale domain. Some advanced techniques can be used for improving classification such as EMD and maximum overlap DWT based feature extraction and further improvement of classification algorithm by using hybrid optimization techniques.

REFERENCES

Englehart, K., B. Hudgins and P. Parker, 2001. Intelligent System and Technologies in Rehabilitation Engineering. CRC Press, Boca Raton, Florida,.

Englehart, K., B. Hugdins and P. Parker, 2000. Multifunction Control of Prostheses using the Myoelectric Signal. In: Intelligent Systems and Technologies in Rehabilitation Engineering, Teodorescu, H.N.L. and L.C. Jain (Eds.). CRC Press, New York, USA., pp: 153-208.

Hudgins, B., P. Parker and R.N. Scott, 1993. A new strategy for multifunction myoelectric control. IEEE. Trans. Biomed. Eng., 40: 82-94.

Lei, M., Z. Wang and Z. Feng, 2001. Detecting nonlinearity of action surface EMG signal. Phys. Lett. A., 290: 297-303.

Narayan, Y., L. Mathew and S. Chatterji, 2017. sEMG signal classification using discrete wavelet transform and decision tree classifier. Intl. J. Control Theory Appl., 10: 511-517.

Oskoei, M.A. and H. Hu, 2006. GA based feature subset selection of myoelectric classification. Proceeding of International Conference on Robot Biomimetics, December 17-20, 2006, IEEE, Kunming, China, ISBN:1-4244-0570-X, pp: 1465-1470.

Oskoei, M.A. and H. Hu, 2007. Myoelectric control systems: A survey. Biomed. Signal Process. Control, 23: 275-294.

Phinyomark, A., A. Nuidod, P. Phukpattaranont and C. Limsakul, 2012. Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification. Electron. Electr. Eng., 122: 27-32.

Phinyomark, A., C. Limsakul and P. Phukpattaranont, 2009. An optimal wavelet function based on wavelet denoising for multifunction myoelectric control. Proceedings of the 6th International Conference on Electrical Engineering Electronics, Computer, Telecommunications and Information Technology Vol. 2, May 6-9, 2009, IEEE, Pattaya, Chonburi, Thailand, ISBN:978-1-4244-3387-2, pp: 1098-1101.

Phinyomark, A., P. Phukpattaranont and C. Limsakul, 2012. Feature reduction and selection for EMG signal classification. Expert Syst. Appl., 39: 7420-7431.

Reaz, M.B.I., M.S. Hussain and F. Mohd-Yasin, 2006. Techniques of EMG signal analysis: detection, processing, classification and applications. Biol. Proced. Online, 8: 11-35.

Tsai, A.C., J.J. Luh and T.T. Lin, 2015. A novel STFT-ranking feature of multi-channel EMG for motion pattern recognition. Expert Syst. Appl., 42: 3327-3341.