ANALYTICAL COMPARISON OF VARIOUS TYPES OF CLASSIFIERS FOR SURFACE EMG SIGNAL IN ORDER TO IMPROVE CLASSIFICATION ACCURACY

Ram Murat Singh, S. Chatterji
National Institute of Technical teachers Training and Research, Chandigarh, India

Amod Kumar
Chief Scientist, (Biomedical Instrumentation Department, C.S.I.O., Chandigarh, India

ABSTRACT
Surface EMG is an important signal originating from human body while doing different movements. This can be utilized for various applications like movement classification, diagnosing neuromuscular disorders, prosthetic control and many more. Surface EMG signal analysis is complex in nature because of its random nature. Several researchers are trying to provide solutions for tackling this problem in the form of improving acquisition circuit for surface EMG signal, increasing the density of sensors during acquisition process, extracting novel features which could give more information and so on. One of the crucial stages while analyzing surface EMG signal is selection of feature sets and classification algorithm. In present work the authors tried different time domain feature sets and their combinations to improve classification accuracy. It was observed that a combination of feature sets improves classification accuracy but response time is increased. The present study explains the optimized solution for the aforesaid problem.

Key words: Acquisition of sEMG, Time domain Features, Classification Algorithms, Pattern Recognition and sEMG control.


1. INTRODUCTION
Electromyography (EMG) signals are the physical phenomena of electrical activity emanated by muscle acupressure points. Motor Unit Action Potential (MAUP) is a key element of any EMG signal. Electrical signals are generated by the muscle voluntary or involuntary contraction effect during the desired movement performed by the subject in the range of few µV due to the
resulted sum of MAUP provided by pickup area electrode being used. Generally 20-50 motor units are conjoined together to form muscle fibers throughout the muscle area occupied by the surface electrode used for EMG signal for a particular muscle. sEMG signal originated the hidden concept of Human Robot Interaction (HRI) for rehabilitation process.

Surface electromyographic (sEMG) signals are the resultant of the motor unit action potential which is collected over the skin in the form of electrical activity during contraction. The sEMG signals are acquired corresponding to some predefined movement performed by the subject with the help of multi-channel sEMG acquisition device to discriminate various motion [1]. The concept of sEMG based assistive robotic device has been applied for developing the myoelectric prostheses which utilized the result of classified sEMG signal for generating the desired control signal. The complete process consists of signal acquisition, feature extraction, and classification followed by the control of the robotic device [2]. In the modern era of patterns recognition, EMG signal based robotic device attracted the lots of research attention which utilized the various classifier namely Support Vector Machines (SVM), Self-Organizing Map (SOM), Linear Discriminant Analysis (LDA) and Self-Organizing Map (SOM) for developing the different EMG applications [3].

Better performance can be achieved in pattern recognition by considering the main key point i.e feature extraction and feature selection which decide the reliability of the developed system. The sEMG signals feature can be categories in three type: first on is time domain (TD), second is frequency domain (FD) and Final is time-frequency domain (TFD) which is also known as time scale feature [4-6].

2. DATA ACQUISITION OF SEMG SIGNAL

The MYOTRACE 400 device is used for data acquisition with some signal processing facilities.

After acquiring the raw EMG signals from the subjects, Signal processing is done. Feature extraction and selection are then done in order to form the feature vector. For achieving the better and reliable output of the classifier, features set is divided into training as well as testing sets. In this study, three-fold cross-validation scheme is carried out for accurate computation. The data is acquired from 5 healthy subjects of the age group from 20 to 24 years. The hand movements are performed by the subjects under the predefined protocol. The EMG signal is sampled at 1000 Hz. Fifteen trails have been performed by each subject for each activity (flexion and extension) in a time period of 20 seconds. The surface electrodes are used for acquiring the data. The location of surface electrodes is fixed at the acupressure point of the right hand upper arm. The experimental set up is shown in Fig.1.
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

In the present work, non-invasive type electrodes have been used. Generally used electrodes are pre-gelled in the form of Ag/Agcl electrode which has been recommended in the various literature. Ag/Agcl electrodes are easy for handling as well as fulfill all desired medical hygienic aspects as shown in Fig. 2. The diameter of electrode (conductive area) is 1 cm or even smaller than that, and the overall dimension is 50mm x 35mm x 1mm.

![Figure 2 Non Invasive Pre Gelled Electrodes](image)

The SEMG signal was acquired for two movements: elbow flexion and elbow extension.

The muscles selected during elbow flexion are Flex Carp U., Flex Carp. R., Brachiorad and Biceps Br. as shown in Fig. 3.

![Figure 3 Muscle Selection in Myotrace 400](image)

The muscles selected during elbow extension are Lat. Triceps and Med. Triceps.

All processing steps such as feature extraction, selection and classification have been carried out using MATLAB.

3. CLASSIFICATION OF SEMG SIGNAL INTO MOVEMENTS

Generally three main steps are followed for classifying the SEMG data for hand movements: i) SEMG signal acquisition; ii) Feature Extraction; iii) Classification. The Time domain features are extracted as mentioned below in Table1 [7].
Table 1 Time Domain Features.

<table>
<thead>
<tr>
<th>Features</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Root Mean Square</td>
<td>RMS</td>
</tr>
<tr>
<td>Waveform length</td>
<td>WL</td>
</tr>
<tr>
<td>Slope sign change</td>
<td>SSC</td>
</tr>
<tr>
<td>Skewness</td>
<td>SKEW</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>KURT</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>SD</td>
</tr>
<tr>
<td>Variance</td>
<td>VAR</td>
</tr>
<tr>
<td>Auto-regressive coefficients</td>
<td>AR</td>
</tr>
</tbody>
</table>

4. RESULTS AND DISCUSSION

In the presented research work four classifiers namely Decision tree, Ensemble classifier, K nearest neighbor (KNN) and support vector machine (SVM) has been used for classification of SEMG signal into two classes (Elbow flexion and extension) [8]. They are explained below:

- Decision Tree: Decision tree is commonly used in operation research and is based on conditional probabilities approach. In this study three variants of decision tree classifier has been used namely: simple tree, Medium tree and complex tree. We achieved Maximum classification accuracy of 92.3 percent was achieved. Results are given in table 2.

Table 2 Decision Tree Classification Results

<table>
<thead>
<tr>
<th>S. No</th>
<th>Type of Decision Tree</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Complex Tree</td>
<td>92.3</td>
</tr>
<tr>
<td>2</td>
<td>Medium Tree</td>
<td>92.3</td>
</tr>
<tr>
<td>3</td>
<td>Simple Tree</td>
<td>91.1</td>
</tr>
</tbody>
</table>

Parallel coordinate plots for decision tree classifier are shown below in Fig. 4(a), Fig. 4(b) and Fig. 4(c). A parallel coordinate plot maps each row in the data table as a line, or profile. Each attribute of a row is represented by a point on the line. This makes parallel coordinate plots similar in appearance to line charts, but the way data is translated into a plot is substantially different. The plots shown in Fig. 4(a), Fig. 4(b) and Fig. 4(c) may be interpreted as two color plot representing the two output classes (orange and blue) for each individual feature.

Fig 4(a) Parallel Coordinate Plot for Complex tree  Fig 4(b) Parallel Coordinate Plot for Medium Tree
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

**Fig. 4(c) Parallel Coordinate Plot for Simple Tree**

Confusion matrices for various decision tree are shown in below given Fig. 5(a), Fig. 5(b) and Fig. 5(c) respectively. The confusion matrix represents the performance of classifier by demonstrating the true classified sample and false classified sample. The number of the correctly classified sample is shown by the diagonal element whereas misclassified sample by off-diagonal elements.

**Fig. 5(a) Confusion Matrix for Complex Tree**

**Fig. 5(b) Confusion Matrix for Medium Tree**

**Fig. 5(c) Confusion Matrix for Simple Tree**

Scatter plots for decision tree are shown in below given Fig. 6(a), Fig. 6(b) and Fig. 6(c) respectively. Scatter plots are shown between the two features which is also known as scatter graphs. It represents the scatter of two different features in a continuous manner on an X-Y axis showing the similarity between them. Basically, this types of plots demonstrate the how two different variables or features are associated with each other.
Receiver operating characteristic (ROC) curve for decision tree classifiers are shown in below given Fig. 7(a), Fig. 7(b) and Fig. 7(c) respectively. ROC curve is the graphical way of representation between the false positive rate versus the true positive rate of a diagnostic test of classification. The perfect ROC curve shows the perpendicular graph between the true positive rate and false positive rate. Higher the perfectness more is the accuracy. It is also used to calculate the various other parameters of the classifier such as sensitivity and specificity. The closer the curve follows the left-hand border and then the top border of the ROC space, more accurate is the test.
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

- Ensemble Classifier: this type of classification method utilize a set of the different weak classifier to discriminate between various classes. After the classification performed by the weak classifier, output of these classifiers are used to vote for the final decision of classification. Initially, the Bayesian averaging method was very famous but nowadays different others methods like bagging as well as boosting can also be considered.

In the present study the authors used 5 types of ensemble classifiers namely: Boosted trees, Bagged trees, Subspace discriminant trees, Subspace KNN Trees and RUS boosted trees. Highest classification accuracy of 95.7% was achieved in bagged trees type of ensemble classifier as shown in Table 3.

<table>
<thead>
<tr>
<th>S.NO</th>
<th>Type</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Boosted Trees</td>
<td>66.6</td>
</tr>
<tr>
<td>2</td>
<td>Bagged Trees</td>
<td>95.7</td>
</tr>
<tr>
<td>3</td>
<td>Subspace Discriminant</td>
<td>92.6</td>
</tr>
<tr>
<td>4</td>
<td>Subspace KNN</td>
<td>81.1</td>
</tr>
<tr>
<td>5</td>
<td>RUS Boosted Trees</td>
<td>93.1</td>
</tr>
</tbody>
</table>

Parallel coordinate plot for ensemble classifier are shown below in Fig. 8(a), Fig. 8(b), Fig. 8(c), Fig. 8(d) and Fig. 8(e) respectively.
Confusion matrices for ensemble classifier are shown below in Fig. 9(a), Fig. 9(b), Fig. 9(c), Fig. 9(d) and Fig. 9(e) respectively.

**Fig. 9(a)** Confusion Matrix for Bagged Trees

**Fig. 9(b)** Confusion Matrix for Boosted Trees

**Fig. 9(c)**: Confusion Matrix for RUS Boosted Trees

**Fig. 9(d)** Confusion Matrix for Subspace Discriminant Trees
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

**Fig. 9(e) Confusion Matrix for Subspace KNN Trees**

Various scatter plots for ensemble classifier are shown below in Fig. 10(a), Fig. 10(b), Fig. 10(c), Fig. 10(d) and Fig. 10(e) respectively.

**Fig 10(a) Scatter Plot for Bagged Trees**

**Fig 10(b) Scatter Plot for Boosted Trees**

**Fig 10(c) Scatter Plot for RUS Boosted Trees**

**Fig 10(d) Scatter Plot for Subspace Discriminant Trees**

**Fig 10(e) Scatter Plot for Subspace KNN Trees**
ROC plots for above results are shown below in Fig. 11(a), Fig. 11(b), Fig. 11(c), Fig. 11(d) and Fig. 11(e) respectively.

- **Fig. 11(a)** ROC Plot for Bagged Trees
- **Fig. 11(b)** ROC Plot for Boosted Trees
- **Fig. 11(c)** ROC Plot for RUS Boosted Trees
- **Fig. 11(d)** ROC Plot for Subspace Discriminant Trees
- **Fig. 11(e)** ROC Plot for Subspace KNN Trees

- **K- Nearest neighbor (KNN):** the KNN classifier is a very popular method of biomedical signal classification which utilizes the density estimation technique to find the suboptimal output. the algorithm is based on the strength of the nearest neighbor rule as explained below. for the unknown feature vector $x$ and a distance measure, then: Out of the $N$ training vectors, identify the $k$ nearest neighbors, regardless of the class label. $k$ is chosen to be odd for a two class problem, and in general not to be a multiple of the number of classes $M$. 

http://www.iaeme.com/IJARET/index.asp 223  editor@iaeme.com
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

- Out of these k samples, identify the number of vectors, \( K_i \), that belong to class \( \omega_i \), \( i = 1, 2, \ldots, M \). Obviously, \( \sum_i K_i = k \).
- Assign \( x \) to the class \( \omega_i \) with the maximum number \( K_i \) of samples [9].

In the present study the authors used six variants of KNN namely Fine KNN, Medium KNN, Coarse KNN, Cosine KNN, Cubic KNN and weighted KNN. Highest classification accuracy of 95.4 % was achieved in case of Fine KNN as shown in Table 4.

<table>
<thead>
<tr>
<th>S.NO</th>
<th>Type</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Fine KNN</td>
<td>95.4</td>
</tr>
<tr>
<td>2</td>
<td>Medium KNN</td>
<td>91.4</td>
</tr>
<tr>
<td>3</td>
<td>Coarse KNN</td>
<td>88.3</td>
</tr>
<tr>
<td>4</td>
<td>Cosine KNN</td>
<td>93.4</td>
</tr>
<tr>
<td>5</td>
<td>Cubic KNN</td>
<td>91.4</td>
</tr>
<tr>
<td>6</td>
<td>Weighted KNN</td>
<td>94</td>
</tr>
</tbody>
</table>

Various parallel coordinate plots for KNN classifier are shown in Fig. 12(a), Fig. 12(b), Fig. 12(c), Fig. 12(d), Fig. 12(e) and Fig. 12(f) respectively.
Different confusion matrices for KNN classifier are shown in Fig. 13(a), Fig. 13(b), Fig. 13(c), Fig. 13(d), Fig. 13(e) and Fig. 13(f).

Fig. 12(e) Parallel Coordinate Plot for Medium KNN  
Fig. 12(f) Parallel Coordinate Plot for Weighted KNN
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

**Fig. 13(e)** Confusion Matrix for Medium KNN  
**Fig. 13(f)** Confusion Matrix for Weighted KNN  
Different scatter plots for KNN classifier are shown in Fig.14 (a), Fig. 14(b), Fig. 14(c), Fig. 14(d), Fig. 14(e) and Fig. 14(f) respectively.

**Fig. 14(a)** Scatter Plot for Coarse KNN  
**Fig. 14(b)** Scatter Plot for Cosine KNN  
**Fig. 14(c)** Scatter Plot for Cubic KNN  
**Fig. 14(d)** Scatter Plot for Fine KNN
Various ROC plots for KNN classifier are shown in Fig. 15(a), Fig. 15(b), Fig. 15(c), Fig. 15(d), Fig. 15(e) and Fig. 15(f) respectively.

**Fig. 15(a)** ROC Plot for Coarse KNN

**Fig. 15(b)** ROC Plot for Cosine KNN

**Fig. 15(c)** ROC Plot for Cubic KNN

**Fig. 15(d)** ROC Plot for Fine KNN
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

- Support Vector Machine (SVM) classifier: SVM algorithm attempts to identify the separating hyper plane, which is found to be the best in separating each class of signals from all the other classes [10]. In the present work 6 variants of SVM classifiers has been used namely: linear SVM, quadratic SVM, cubic SVM, fine Gaussian SVM, medium Gaussian SVM and coarse Gaussian SVM. The maximum classification accuracy of 95.1 percent has been achieved in the case of cubic SVM, as shown in Table 5.

Table 5. SVM Classification Results

<table>
<thead>
<tr>
<th>S.NO</th>
<th>Type</th>
<th>Classification Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Linear SVM</td>
<td>93.7</td>
</tr>
<tr>
<td>2</td>
<td>Quadratic SVM</td>
<td>94</td>
</tr>
<tr>
<td>3</td>
<td>Cubic SVM</td>
<td>95.1</td>
</tr>
<tr>
<td>4</td>
<td>Fine Gaussian SVM</td>
<td>94.6</td>
</tr>
<tr>
<td>5</td>
<td>Medium Gaussian SVM</td>
<td>92.6</td>
</tr>
<tr>
<td>6</td>
<td>Coarse Gaussian SVM</td>
<td>91.7</td>
</tr>
</tbody>
</table>

Various parallel coordinate plot for SVM classifier are shown in Fig. 16(a), Fig. 16(b), Fig. 16(c), Fig. 16(d), Fig. 16(e) and Fig. 16(f) respectively.
Various confusion matrices for SVM classifier are shown in Fig. 17(a), Fig. 17(b), Fig. 17(c), Fig. 17(d), Fig. 17(e) and Fig. 17(f) respectively.
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

**Fig. 17(c)** Confusion Matrix for Fine Gaussian SVM  
**Fig. 17(d)** Confusion Matrix for Linear SVM  
**Fig. 17(e)** Confusion Matrix for Medium Gaussian SVM  
**Fig. 17(f)** Confusion Matrix for Quadratic SVM

Various scatter plots for SVM classifier are shown in Fig. 18(a), Fig. 18(b), Fig. 18(c), Fig. 18(d), Fig. 18(e) and Fig. 18(f) respectively.

**Fig. 18(a)** Scatter Plot for Coarse Gaussian SVM  
**Fig. 18(b)** Scatter Plot for Cubic SVM
Various ROC plots for SVM classifier are shown in Fig. 19(a), Fig. 19(b), Fig. 19(c), Fig. 19(d), Fig. 19(e) and Fig. 19(f).

Fig. 19(a) ROC Plot for Coarse Gaussian SVM  
Fig. 19(b) ROC Plot for Cubic SVM
Analytical Comparison of Various Types of Classifiers For Surface EMG Signal In Order To Improve Classification Accuracy

CONCLUSION AND FUTURE WORK

In the present work an analytical comparative study has been discussed for four classifier techniques namely: Decision Tree, Ensemble classifier, KNN classifier and SVM classifier. Various Time domain features as listed in Table1 have been extracted and feature vectors have been formed before applying the classification algorithms. The major outcome of the study is mentioned in the form of salient points provided below:

- Ensemble classifier technique provides highest classification accuracy of 95.7 in bagged trees type followed by 95.1 in case of SVM classifier in cubic SVM technique.
- Ensemble classifier technique although provides the highest classification accuracy, consumes much more training time as well as response time.
- SVM classifier is best suited when compared classification accuracy, training time and also response time.
- The experimental result shows that increasing the feature vector increases the classification accuracy but the training and the response time are also increased.
- The experimental result also shows that increasing the length of feature vector also creates chances of error.

Application of time-domain, frequency domain and time-frequency domain features may be combined to frame the future scope of the present study in order to achieve even higher classification accuracy.

Also, a novel approach may be tried to generate the control signal by reducing the size of feature vector of EMG signals for robot having the more number of degree of freedom can be achieved depending on the requirements.

Fig. 19(c) ROC Plot for Fine Gaussian SVM

Fig. 19(d) ROC Plot for Linear SVM

Fig. 19(e) ROC Plot for Medium Gaussian SVM

Fig. 19(f) ROC Plot for Quadratic SVM
There is a great need of a reliable, safe and precise control scheme for the Myoelectric control of an Exoskeleton robots which establishes the majority of extensive research work for implementing the model in Real-time applications.

ACKNOWLEDGEMENT
The authors wish to thank Department of Science and technology, Government of India for providing financial assistanceship.

REFERENCES