

# Comparative study of PCA in classification of multichannel EMG signals

P. Geethanjali<sup>1</sup>

Received: 30 April 2014 / Accepted: 30 March 2015 / Published online: 10 April 2015  
© Australasian College of Physical Scientists and Engineers in Medicine 2015

**Abstract** Electromyographic (EMG) signals are abundantly used in the field of rehabilitation engineering in controlling the prosthetic device and significantly essential to find fast and accurate EMG pattern recognition system, to avoid intrusive delay. The main objective of this paper is to study the influence of Principal component analysis (PCA), a transformation technique, in pattern recognition of six hand movements using four channel surface EMG signals from ten healthy subjects. For this reason, time domain (TD) statistical as well as auto regression (AR) coefficients are extracted from the four channel EMG signals. The extracted statistical features as well as AR coefficients are transformed using PCA to 25, 50 and 75 % of corresponding original feature vector space. The classification accuracy of PCA transformed and non-PCA transformed TD statistical features as well as AR coefficients are studied with simple logistic regression (SLR), decision tree (DT) with J48 algorithm, logistic model tree (LMT),  $k$  nearest neighbor ( $k$ NN) and neural network (NN) classifiers in the identification of six different movements. The Kruskal–Wallis (KW) statistical test shows that there is a significant reduction ( $P < 0.05$ ) in classification accuracy with PCA transformed features compared to non-PCA transformed features. SLR with non-PCA transformed time domain (TD) statistical features performs better in accuracy and computational power compared to other features considered in this study. In addition, the motion control of three drives for six movements of the hand is

implemented with SLR using TD statistical features in off-line with TMSLF2407 digital signal controller (DSC).

**Keywords** EMG · Feature extraction · Classification · Principal component analysis · Digital signal controller

## Introduction

Several researchers are attempting to develop assistive devices for the individuals with disabilities [1, 3]. Assistive devices are developed using bio-electric signals, speech signals, gestures, etc. Electromyography (EMG) based prosthetic devices have been used significantly to improve the day-to-day activity of disabled people resulted from accidents, peripheral vascular disease, diabetes, etc. EMG is a muscular technology, provides the information about the intention of the users from the recording of signals on the individual skin's surface. EMG signals are informative, however the idea of control of assistive devices using EMG is sophisticated as well as complex.

The common methodology underlying EMG based pattern recognition control are: (i) extraction of the most relevant features (ii) the classification/identification of extracted features for the intended movement. In addition to identification of intention, it is vital that this recognition does not use much computational power. EMG features can be extracted from transient signals, steady state signals and combining both the signals to discriminate the intentions of the user. Till date numbers of researchers classify the information using transient [4, 11], steady state [4] and also both transient and steady state EMG signals [9]. Hargrove et al. [9] found that the effectiveness of the prosthetic devices improves by considering both the transient and the steady state EMG signals. In this paper, transient and

✉ P. Geethanjali  
pgeethanjali78@hotmail.com

<sup>1</sup> School of Electrical Engineering, VIT University,  
Vellore 632 014, TN, India

steady state EMG signals are considered to extract time domain statistical features as well as auto regression (AR) coefficients.

The accuracy of pattern recognition is based on selection of the features and the technique applied for classification. The efficacy of classifier depends on the choices of features used to discriminate the functionality. Feature vector with high class separability will augment the classifier efficacy. In the recent years, many researchers are using features like mean absolute value (MAV) [4,11,14], number of zero crossings (ZC) [4,11], number of slope sign changes (SSC) [4,11], waveform length (WL) [4,11,14] and AR coefficients. In this study also, MAV, ZC, SSC, WL and fourth order AR coefficients are considered for classification of EMG.

It is desirable to extract increased number of features that helps in identification of control signals accurately for the actuation of prosthetic device. However, the dimension of feature space burgeon with increase of number of electrodes as well as a number of features. In order to eschew the computation time as well as to ameliorate the classification performance, it is indispensable to fore-shorten the irrelevant/redundant features using feature reduction technique. In feature reduction techniques, the high-dimensional feature space are projected into a low-dimensional feature space. These feature reduction techniques include principal component analysis (PCA), independent component analysis (ICA) [15], genetic algorithm (GA) [16,18] etc. PCA is a transformation technique that has been typically used due to its simplicity [12] in pre-processing [10], feature reduction [2] and design of robotic hands [8]. In recent years, different variation in PCA has been developed [3]. In this paper, PCA transformed using eigenvectors are used to study the performance of 25, 50 and 75 % transformed time domain statistical features and AR coefficients for classification of six movements.

In classification, a recurring theme is to utilize constructively the information available with the relevant features. In such cases, all the features in feature spans do not guarantee an optimum performance and it exacerbates the curse of dimensionality especially with the increased number of electrodes. Researchers used different classifiers such as fuzzy logic, neural network, state vector machine, etc. [1,2,7,17]. In the literature, it has been found that Simple Logistic Regression (SLR) technique, is more effective and efficient in the classification of the continuous EMG signals [6]. In addition to SLR, other classifiers such as logistic model tree (LMT), decision tree (DT), neural network (NN) and  $k$  nearest neighbor ( $k$ NN) are considered to study the influence of classification with PCA transformed TD statistical features and AR coefficients. The block diagram of EMG pattern recognition based schema is shown in Fig. 1. The block with dotted line may/may not

present in EMG pattern recognition schema, depending on the number of features.

## EMG signals acquisition

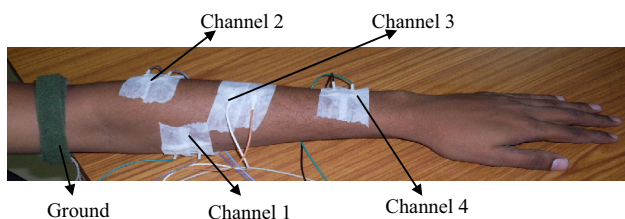
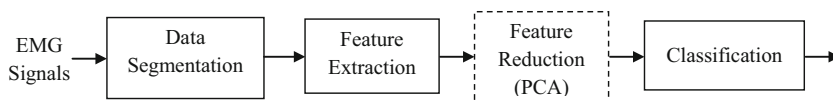
An indigenous, four channel EMG signal acquisition system with the necessary protection circuit has been used for acquisition of surface EMG signals. In this work, EMG signals are collected, using a pair of Ag/AgCl (Argentum/Argentum Chloride) disc electrodes of 10 mm diameter. Since hand motion results from contraction of the muscles in the forearm, the surface electrodes for four channels are placed on flexor digitorum superficialis, supinator, extensor digitorum communis and extensor indicis as shown in Fig. 2. A signal conditioning circuit is designed for each channel with high input impedance instrumentation amplifier. A high pass filter of 9 Hz cut-off frequency is connected after the instrumentation amplifier to prevent DC voltage offsets caused by skin impedance and reaction between the skin and electrode gel. Further, the output from the high pass filter passes through the variable gain amplifier whose gain can be varied from 6 to 100 and to band pass filter of 10–500 Hz to obtain signals in a dominant energy band. A digital Signal Processing and Control engineering (dSPACE) software with the DS1104 board is used to sample the EMG signals at a rate of 1000 samples per Sec and stored in personal computer having Intel core Duo Processor with 2.93 GHz and 1.96 GB of RAM.

In this study, ten healthy male subjects are considered for acquisition of continuous EMG signals for different limb movements. Signals acquired from the subjects without any skin preparation. The subjects were instructed to perform each movements, i.e. hand close (HC), hand open (HO), wrist flexion (WF), wrist extension (WE), ulnar deviation (UD) and radial deviation (RD) for a duration of 5 s. with the elbow resting on the table. All the subjects performed each trial for 30 s. duration with brief relaxation periods between each trial to avoid fatigue. Eight trials of data are used for classification of motions. Transient period lasted approximately for 50 ms after every change in movement. However, it varies with the speed with which the subject does the movement.

## Data segmentation

Features are computed from the conditioned EMG signal data stored in personal computer from a segment (window), where a single feature vector is produced from each segment. Window size and increment in window size is chosen, considering the real-time constraint of myoelectric

**Fig. 1** Block diagram of EMG based pattern recognition



**Fig. 2** Location of the electrodes on the limb for four different channels

hand control system which should be within 300 ms [11]. For real time myoelectric control, all processes including the generation of control must be completed within 300 ms. Therefore, analysis is performed on 256 ms window in all cases with a window increment of (overlap) 128 ms.

EMG data, so segmented contain both transient and steady state data to identify the intended motion, needs to be classified. To discriminate the different intention in data, it is necessary to extract the features for identifying the different classes of motion. A classifier to classify the intention it is necessary to extract the feature to discriminate the data. The effect of classification also depends on the proper extraction of the relevant features which should provide good separation between classes. The next section describes about the feature extraction considered in this work.

**Feature extraction**

EMG signals carry information about the intention of the subject for a specific motion. To identify the specific motion, four time domain statistical features such as mean absolute value (MAV), zero crossings (ZC), slope sign changes (SSC) and waveform length (WL) and fourth order AR coefficients are extracted for a window size of 256 with an overlap of 128. Feature span dimension of the TD statistical feature as well as AR coefficients is  $234 \times 16$  for one trial. 39 patterns are obtained for every 5 s movement from the four channel data of size  $5000 \times 4$ . Each pattern consists of 16 features.

From the EMG data, the MAV and WL are calculated as in (1) and (2). ZC occurs when both (3) and (5) are satisfied. SSC occurs when conditions (4)–(6) are satisfied.

$$MAV = \frac{1}{L} \sum_{i=1}^L |y_i| \tag{1}$$

$$WL = \sum_{i=1}^L |y_i - y_{i-1}| \tag{2}$$

$$(y_i > 0 \text{ and } y_{i+1} < 0) \text{ or } (y_i < 0 \text{ and } y_{i+1} > 0)$$

$$WL = \sum_{i=1}^L |y_i - y_{i-1}| \tag{3}$$

$$(y_i > y_{i-1} \text{ and } y_i > y_{i+1}) \text{ or } (y_i < y_{i-1} \text{ and } y_i < y_{i+1}) \tag{4}$$

$$|y_i - y_{i+1}| \geq \epsilon \tag{5}$$

$$|y_i - y_{i-1}| \geq \epsilon \tag{6}$$

where  $y_i$  is  $i$ th the EMG data,  $\epsilon$  is the threshold,  $L$  is the EMG data segment length,  $i$  is the sample in a segment.

Auto regression, model EMG signals as a linear autoregressive time series depends on previous outputs and defined as

$$y_k = \sum_{i=1}^p \lambda_i y_{k-i} + e_k \tag{7}$$

where  $y_k$  is the estimated data,  $\lambda_i$  are the AR-coefficients,  $e_k$  is the estimation error and  $p$  is the order of the model (Number of coefficients).

**Feature reduction**

To reduce overloading of the classifier, 16 features from each pattern are transformed to 4 (25 %), 8 (50 %) and 12 (75 %) features using PCA. PCA extracts the predominant data from the input feature sets. The first step of calculation in PCA is to compute covariance matrix of feature variables. Eigenvalues are calculated for the covariance matrix. The eigenvectors corresponding to eigenvalues are computed which account for most of the variance in features. Eigenvector with the largest eigenvalue is the principle component of the feature vector.

In this work, four features are extracted from each channel. Since the observation from each channel is four, the covariance matrix is  $4 \times 4$ . Eigenvalues and eigenvectors are calculated from the covariance matrix. Ordering the eigenvalue from largest to smallest gives, order of significance of principle components. Transformed feature set can be obtained by ignoring the components of lesser significance. To reduce feature set to 25 % of original feature vector, the first principal component alone is considered. To find the transformed/projected new feature vector, the chosen eigenvector is multiplied

with the original feature vector. In a similar way, feature vector size can be reduced to 50 % and 75 % of the original feature vector size with the consideration of two and three principal components respectively in the order of significance. A transformed/reduced feature set is obtained by multiplying the considered eigenvector to the original feature set. Figures 3 and 4 shows the scatter plot of 25 % transformed statistical features and 25 % transformed AR coefficients using PCA. The scatter plots clearly evince that, TD transformed using PCA has the better class capability than the PCA transformed AR features for six different movements of the hand. Features with better class separation yields good classification performance. The classification performance by transforming TD as well as AR and non-transformed features are discussed in the results and discussion section. The next section discusses various types of classification techniques considered in pattern recognition. The widely used linear discriminant analysis (LDA) classifier is not considered in pattern recognition, since LDA suffered from the problem of singularity with PCA transformed features, but not with non-PCA transformed features. In order to study the performance of the classifiers with and without PCA transformed features, LDA is not considered in pattern recognition.

**Classification**

In this section, the different classification techniques, namely simple logistic regression (SLR), decision tree (DT) with J48 algorithm, logistic model trees (LMT),

Neural network (NN) and *k* nearest neighbor (*k*NN) approaches are described for identification of six different movements from PCA transformed and non-PCA transformed TD statistical features as well as AR coefficients. The classification was divided into two stages: the training stage and the testing stage. Three trials out of 8 trials have been considered in training. During testing, the remaining five trials are used. The following subsections discuss the various classifiers used in classifying the EMG data.

**Simple logistic regression (SLR)**

Logistic regression performs a least-square fit of a parameter to a target feature vector to form a model for each class as shown in equation below.

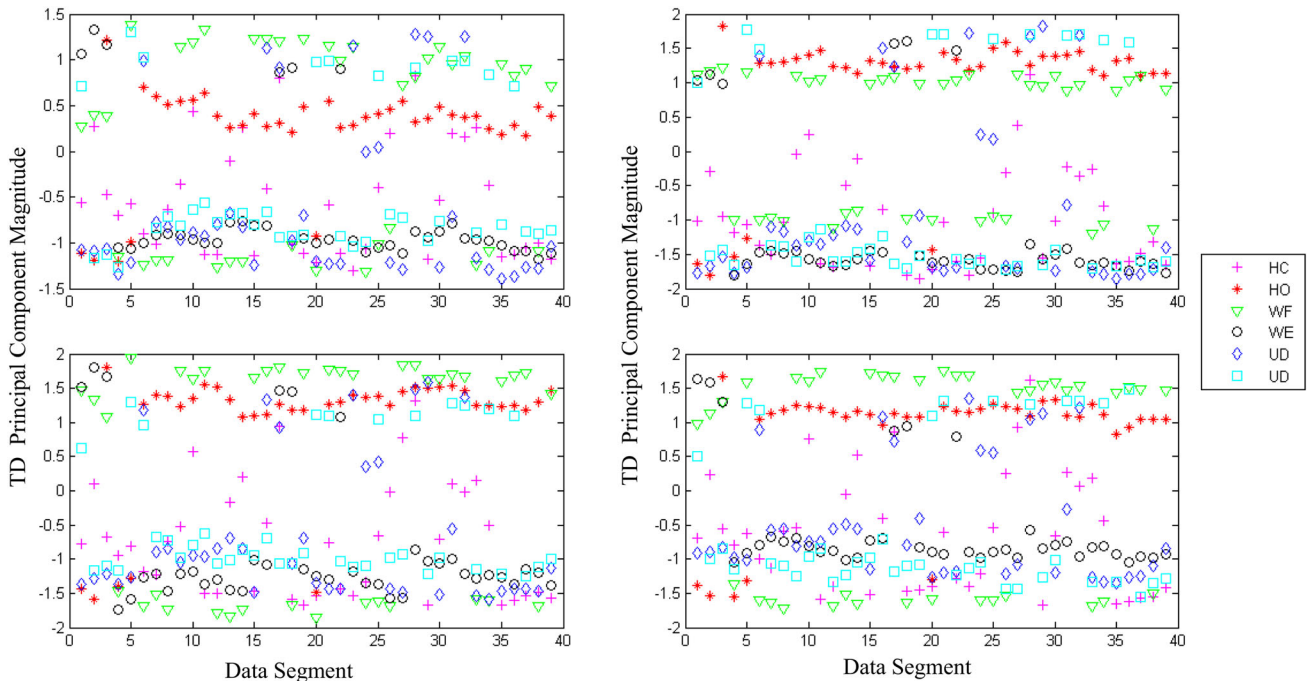
$$i_{class}^* = \arg \max_i P(C = i|X = x) \tag{8}$$

where *C* is a class variable of motion, *x* is the feature set which represents the class and *P(C = i|X = x)* is the posterior class probability for an instance *x*.

Logistic regression models the posterior probabilities for *I* classes using linear function in *x* ensuring that they sum to 1 and remain in [0, 1]. The linear regression model is specified in terms of *I-1* log-odds that separate each class from the “base class” *I*, such that,

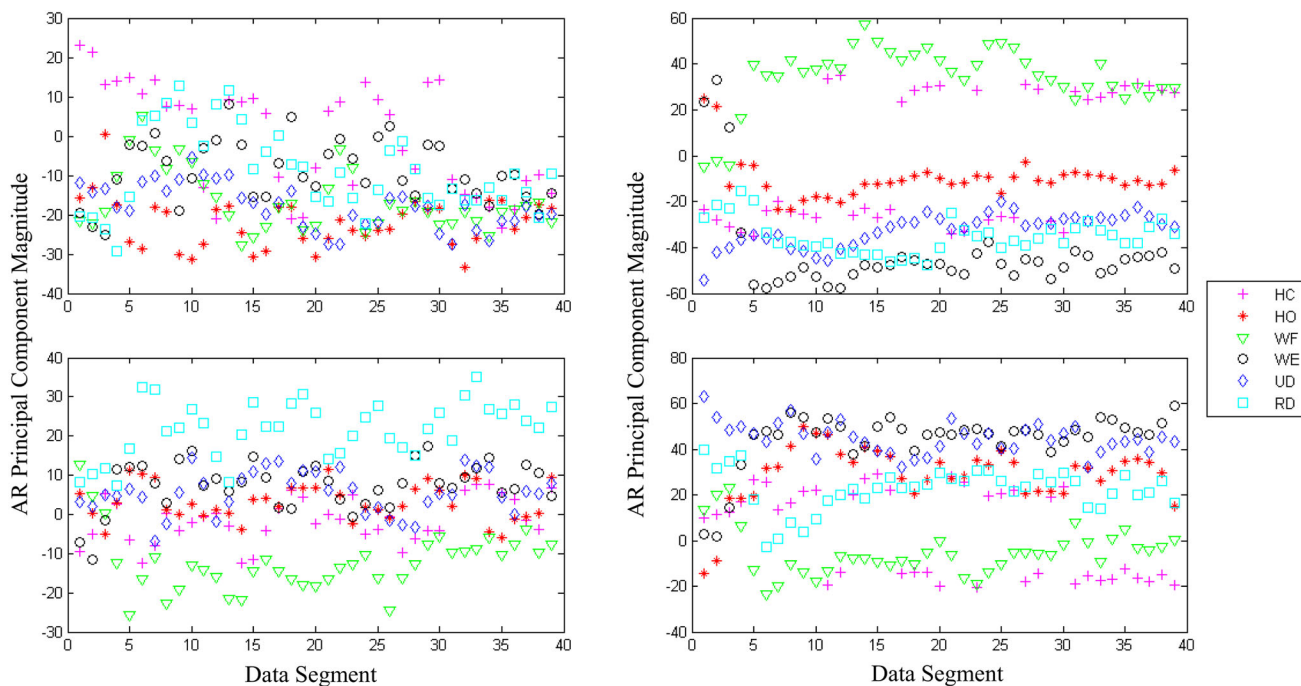
$$\log \frac{P(C = i|X = x)}{P(C = I|X = x)} = \alpha_i^T x, \quad i = 1, 2, \dots, I - 1 \tag{9}$$

where



**Fig. 3** Scatter plot of PCATransformed TD statistical feature vector to 4 features (25 %)





**Fig. 4** Scatter plot of PCATransformed AR coefficients to 4 features (25 %)

$$P(C = i|X = x) = \frac{e^{\alpha_i^T x}}{1 + \sum_{k=1}^{I-1} e^{\alpha_k^T x}} \quad \text{for } i = 1, 2, \dots, (I - 1) \tag{10}$$

$$P(C = I|X = x) = \frac{1}{1 + \sum_{k=1}^{I-1} e^{\alpha_k^T x}} \tag{11}$$

and,  $\alpha_i$  is the parameter vector in logistic regression model.

This logistic model produces linear boundaries between the regions in the feature space corresponding to different classes. The logistic regression model fitting means, estimating the parameter vector  $\alpha_i$ . The model is built for each class and the class of test feature vector is identified using the maximum value of class model  $\alpha^T x$  [13]. The standard procedure in statistics is to look for the maximum likelihood in logistic regression using numeric optimization algorithms that approach the maximum likelihood solution iteratively and reach it in the limit is in practice. Friedman et al. [5] proposed the LogitBoost algorithm for fitting additive logistic regression models by maximum likelihood.

**Decision tree (DT)**

Decision Tree (DT) is based on the values of feature vector from available training data to identify the class of the test feature vector. In this work, DT is built using J48 algorithm

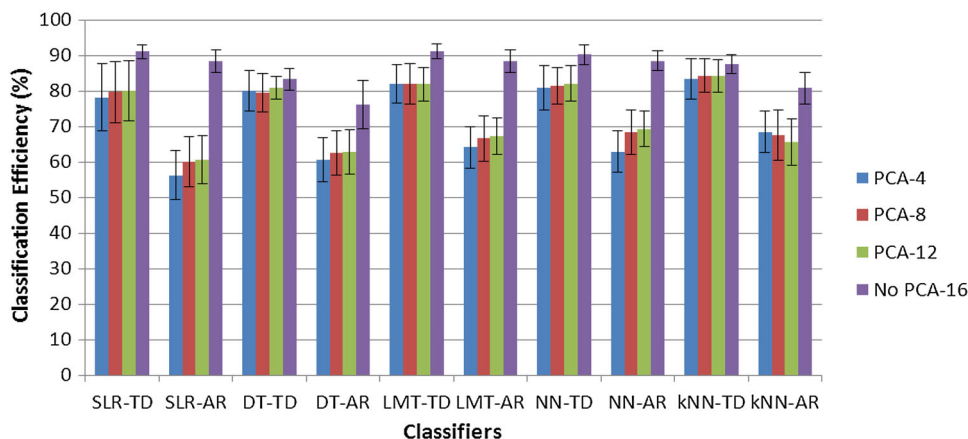
and nodes of trees are selected based on information gain from the concept of entropy [19,20].

In this pattern recognition system, simple if-then rule is applied for classification. One of the advantages of tree classification is that they can be constructed easily. Decision tree begins with a set of cases or training data described by a set of attributes or features which can have numeric or symbolic values. Each training data is associated with a label representing the name of a class it belongs to. Tree induction splits feature space repeatedly, and stops splitting when the feature subspace contains training data with mostly identical class labels. Splitting consists of selection of feature and decision on a threshold value for the feature. A path in a DT basically corresponds to Boolean expression of the form “Feature > threshold”, “Feature ≤ threshold”, so a tree can be seen as a set of rules to identify feature set. In the tree, the nodes specify the features, for which data are branched. The branches between the node refer to the possible value of which data is partitioned. Branch may connect either two internal nodes or a node and a leaf. Leaf node represents the name of the class of the input feature vector. For classification, a new feature set is sorted down to a leaf in the tree to predict class of the test data.

**Logistic model tree (LMT)**

In the logistic model tree, logistic regression and decision trees are special cases basically consists of a tree structure with logistic regression function at the leaves. LogitBoost

**Fig. 5** Classification Performance of classifier with Different Input



algorithm has been used to induce tree with linear, logistic regression models at the leaves. LMT growth would first involve building a standard classification tree and afterwards building a logistic regression model at every node. For building simple logistic regression at the child node, simple logistic regression at the parent node is used as a basis for fitting simple logistic regression at the child node using LogitBoost algorithm. LogitBoost algorithm iteratively changes the simple linear regression class functions to improve the fitting of the data by adding a simple regression function, to fit to the response. After splitting a node, LogitBoost iterations continue running by fitting to the response variable of the training data on the child node only [13].

### Neural network (NN)

The multilayer neural networks are built separately for TD statistical data and AR coefficients with PCA transformed and non-PCA transformed data. The Neural network is composed of an input layer, hidden layer and output layer. The number of neurons in the input layer is same as a number of features with an additional threshold unit. The neural network is trained with the back-propagation algorithm.

### $k$ Nearest neighbour ( $k$ NN)

$k$  nearest neighbour ( $k$ NN) is a very simple classification technique based on Euclidean distance measurements. Euclidean distance is measured between the test feature vectors with all stored prototype vectors. Each class is characterised by prototypes obtained from the three trials of feature vectors. In these conditions, the class of test feature vector is obtained by majority voting amongst  $k$  nearest neighbours. In this simple  $k$ NN, a number of neighbours  $k$  is set to 9.

The class probability of  $k$ NN classifier is calculated using the Eq. (12) given below.

$$P(C = I|X = x) = \frac{K_i}{K} \quad (12)$$

where  $K$  is the nearest neighbour, which is 9,  $K_i$  is the number of neighbours of class  $i$ , The test feature vector belongs to class  $i$ , the largest probability estimate from Eq. (12). The performance of the classifier was not much different for the other higher values of  $k$ .

## Results and discussion

The classification error was found to be more during the transition of posture, for example HC to HO, compared to steady state posture. This could be observed approximately for initial four feature vectors which are not clustered and found wide apart from the remaining steady state feature vectors from the Figs. 3 and 4. Further the classification performance was found to be varying with feature vectors.

The classification performance of 25 % (PCA-4), 50 % (PCA-8) and 75 % (PCA-12) PCA transformed statistical as well as AR coefficients along with non-PCA transformed features are studied with five trials of data. Figure 5 shows the classification performance of SLR, DT, LMT, NN and  $k$ NN classifiers with transformed and non-transformed features. Tables 1, 2, 3, 4, 5 shows the confusion matrix of classifiers with PCA transformed statistical and AR data for five trials of test data. The bold values in the Table indicate the efficiency in identifying the corresponding movement. Also, Tables indicate the measure of sensitivity and specificity for each movement in different classifiers.

From the Tables and Fig. 5, it is clear that, the classification accuracy of the considered classifiers with PCA transformed AR coefficients is less than 70 % irrespective of number of transformed features. The percentage classification accuracy of  $k$ NN, NN and LMT is found to be more than 80 % with PCA transformed statistical features.

**Table 1** Confusion Matrix of SLR with reduced features using PCA

Type of feature	Motion	SLR-4						SLR-8						SLR-12					
		HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD
Reduced TD feature	HC	<b>81.8</b>	7.7	1.1	0.3	6.7	2.5	<b>82.4</b>	5.8	3.7	0.2	6	1.9	<b>83.6</b>	5.5	2.7	0.2	5.9	2
	HO	10.5	<b>73.1</b>	8.5	1.3	4.1	2.5	8.7	<b>75</b>	9.4	0.7	3.9	2.3	8.6	<b>74.3</b>	9	0.7	4	3.5
	WF	10.3	5.3	<b>79.1</b>	1.1	2.6	1.6	6.2	8	<b>81.4</b>	1	2.3	1	6.6	8.8	<b>80.4</b>	1	2.1	1.1
	WE	0.6	1.3	0.3	<b>86.9</b>	3.4	7.6	0.4	1.5	0.5	<b>87.1</b>	4.1	6.3	0.4	1.4	0.7	<b>87.3</b>	4.1	6
	UD	4.6	6.7	2.4	2.2	<b>74.7</b>	9.4	3	7.3	3.8	2.6	<b>76.5</b>	6.8	2.6	7.1	3.9	3.5	<b>77</b>	5.9
Sensitivity	RD	2.7	2.3	0.1	9.4	6.1	<b>79.4</b>	1.7	2.4	0.6	11.2	6.6	<b>77.5</b>	1.7	1.7	0.9	11.1	6.7	<b>77.9</b>
		0.82	0.73	0.88	0.87	0.78	0.82	0.82	0.75	0.87	0.88	0.79	0.79	0.84	0.74	0.86	0.88	0.79	0.79
Specificity		0.94	0.95	0.97	0.97	0.95	0.95	0.96	0.95	0.96	0.97	0.95	0.96	0.96	0.95	0.97	0.97	0.95	0.96
Reduced AR feature	HC	<b>54.5</b>	12.5	6	7.2	10.8	9.2	<b>65.2</b>	11.3	4.4	4.6	7.1	7.4	<b>66.4</b>	9.1	4.3	4.1	7.9	8.2
	HO	20	<b>44.1</b>	12.3	4	8.4	11.2	18.1	<b>46.8</b>	10.1	3.2	11	10.9	16.6	<b>47.2</b>	10.2	3.8	11.7	10.5
	WF	5	10.3	<b>72.8</b>	1.4	7	3.4	4.4	9.5	<b>75.6</b>	1.5	6.4	2.6	4.6	9.8	<b>75.5</b>	1.6	5.8	2.7
	WE	2.2	3.3	2.3	<b>73.5</b>	4	14.8	1.8	3.1	2.2	<b>74.8</b>	4.9	13.1	2.3	3.5	2.1	<b>74.4</b>	5.8	11.9
	UD	13	9.9	6.7	11.6	<b>47.9</b>	10.9	8.6	7.2	4.9	10.5	<b>58.1</b>	10.7	9	7.3	5.6	10.6	<b>57.1</b>	10.3
Sensitivity	RD	5.6	12.5	3	22.7	11	<b>45.1</b>	5.8	8	2.9	23.5	13.7	<b>46.2</b>	5.6	7.2	2.7	22.1	13.7	<b>48.7</b>
		0.54	0.44	0.77	0.75	0.55	0.48	0.65	0.47	0.79	0.76	0.64	0.49	0.66	0.47	0.79	0.76	0.63	0.52
Specificity		0.91	0.90	0.94	0.91	0.92	0.9	0.92	0.92	0.95	0.91	0.91	0.91	0.92	0.93	0.95	0.92	0.91	0.91

**Table 2** Confusion matrix of DT with reduced features using PCA

Type of feature	Motion	DT-4						DT-8						DT-12					
		HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD
Reduced TD feature	HC	<b>85</b>	6.2	0.9	0.1	2.2	5.5	<b>82.2</b>	8.5	1.7	0.1	2.7	4.7	<b>84.7</b>	6.2	1.7	0.2	2.3	4.9
	HO	9.6	<b>76.9</b>	6	0	5.9	1.6	8.1	<b>78.1</b>	7	0	4.7	2.1	7	<b>77.5</b>	7.9	0.2	4.8	2.7
	WF	5	8.3	<b>82</b>	0.9	2.7	1.2	4.5	8.2	<b>82.6</b>	0.8	2.7	1.2	4.6	8.8	<b>82</b>	1	2.5	1.2
	WE	1	0.8	0.6	<b>83.3</b>	3.9	10.4	0.8	1	0.7	<b>84.9</b>	3.3	9.3	0.8	0.9	1	<b>86.6</b>	2.9	7.9
	UD	4	3.8	3.3	2.6	<b>76.1</b>	10.3	3	4.3	3.1	3.6	<b>76.9</b>	9	3.4	4.3	4	4.5	<b>79.1</b>	4.8
Sensitivity	RD	1	2	0.7	6.5	11.4	<b>78.4</b>	1.1	2.8	0.7	9.2	11.8	<b>74.4</b>	1.6	1.3	1	10.2	9.6	<b>76.4</b>
		0.85	0.77	0.86	0.84	0.79	0.79	0.82	0.78	0.87	0.86	0.79	0.75	0.85	0.77	0.86	0.87	0.82	0.78
Specificity		0.96	0.96	0.98	0.98	0.95	0.94	0.96	0.95	0.97	0.97	0.95	0.95	0.97	0.96	0.97	0.97	0.96	0.96
	HC	<b>65.1</b>	11.9	5	3.4	9.1	5.6	<b>69</b>	10.9	4.6	3.4	6.7	5.5	<b>68.4</b>	11.5	5.3	3.6	5.6	5.7
Reduced AR feature	HO	11.7	<b>54.6</b>	11.3	3	11.3	8	11.9	<b>54.9</b>	11.7	2.9	10.2	8.3	11.4	<b>54.8</b>	10.6	3	10.9	9.3
	WF	5.4	13.5	<b>68.2</b>	1.1	7.5	4.3	5.1	11.1	<b>72.5</b>	1.3	7	3.1	6.3	11.2	<b>70.7</b>	1.3	6.9	3.6
	WE	3	3.5	1.7	<b>71</b>	6.5	14.3	4.1	4.3	1.7	<b>69.4</b>	7.6	13	4.2	3.9	1.3	<b>71.6</b>	7.9	11.1
	UD	9.3	9.7	7.8	6.4	<b>58.2</b>	8.5	5.1	9.2	7.3	7.1	<b>61.1</b>	10.2	6.3	9.8	5.6	6.9	<b>62.8</b>	8.7
	RD	6.8	11.4	2.7	17.7	11.1	<b>50.4</b>	5.8	12	3.3	17.1	10.7	<b>51</b>	6.7	12	2.5	15.5	11.8	<b>51.5</b>
Sensitivity		0.65	0.55	0.72	0.73	0.64	0.54	0.69	0.55	0.76	0.72	0.64	0.54	0.68	0.55	0.75	0.75	0.67	0.55
	UD	0.93	0.9	0.94	0.94	0.91	0.92	0.94	0.9	0.94	0.94	0.92	0.92	0.93	0.9	0.95	0.94	0.91	0.92



**Table 3** Confusion Matrix of LMT with reduced features using PCA

Type of feature	Motion	LMT-4										LMT-8										LMT-12									
		LMT-4					LMT-8					LMT-8					LMT-12					LMT-12									
		HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD						
Reduced TD feature	HC	<b>87.6</b>	5.3	0.6	0	1	5.3	<b>88.1</b>	5.1	0.9	0.6	0.9	4.4	<b>86.7</b>	6	1.1	0.3	1	4.8												
	HO	7.6	<b>79.9</b>	6.2	0.4	3.4	2.4	7.8	<b>79.6</b>	5.9	1.3	3.4	2.1	7.2	<b>78.4</b>	6.7	0.6	3.3	3.8												
	WF	3.4	7.9	<b>84.7</b>	1.1	1.9	1	4.5	8.9	<b>82.7</b>	1.2	1.5	1.2	3.9	7.9	<b>83.2</b>	1.3	2.2	1.5												
	WE	0.7	0.7	0.9	<b>84.6</b>	4.4	8.7	0.7	1	0.7	<b>86.4</b>	4	7.2	1.2	1.7	7.3	<b>79.1</b>	3.4	7.3												
	UD	3.3	3.7	3.2	2.9	<b>78.3</b>	8.6	2.9	4.8	2.3	3.2	<b>79.4</b>	7.4	2.5	3.6	4	3.2	<b>80.1</b>	6.7												
Sensitivity	RD	0.4	1.9	0.3	9	10.1	<b>78.3</b>	0.7	1.6	0.3	9.6	10	<b>77.8</b>	0.9	0.9	0.5	9.7	<b>77.7</b>													
		0.88	0.8	0.88	0.85	0.81	0.79	0.88	0.8	0.87	0.87	0.82	0.78	0.87	0.78	0.87	0.8	0.82	0.78												
Specificity		0.97	0.96	0.98	0.97	0.96	0.95	0.97	0.96	0.98	0.97	0.97	0.96	0.97	0.96	0.96	0.97	0.96	0.95												
	HC	<b>67.6</b>	9.9	4.4	3.3	7.4	7.4	<b>70.8</b>	9.4	3.1	3.3	6.7	6.6	<b>71.8</b>	11.1	4.1	2.7	4.7	5.7												
AR feature	HO	11.3	<b>59.4</b>	8	2.3	9.2	9.8	11.3	<b>60.4</b>	9	2	9	8.3	10.8	<b>59.6</b>	8.4	2.4	10.3	8.5												
	WF	4.6	11	<b>70.4</b>	1.4	8.6	4.1	5.5	9.5	<b>74.6</b>	1.7	5.6	3.2	4.4	10.7	<b>74.3</b>	2	6.1	2.5												
	WE	1.9	3.1	1.1	<b>76.3</b>	5.3	12.3	1.8	3	2.1	<b>75.4</b>	6.8	10.9	2.1	2.5	1.4	<b>76.9</b>	7.1	10												
Sensitivity	UD	9.6	8.9	5.7	6.1	<b>59.1</b>	10.6	6.3	8.6	4.4	5.2	<b>63.9</b>	11.6	6.6	8	4.8	5.7	<b>65.1</b>	9.8												
	RD	4.1	9.9	2.6	18.7	11.2	<b>53.5</b>	4.4	8	3	17.6	10.2	<b>56.8</b>	4.3	9	1.6	16.3	10.9	<b>57.9</b>												
Specificity		0.68	0.59	0.74	0.78	0.65	0.56	0.71	0.6	0.79	0.77	0.68	0.59	0.72	0.6	0.78	0.79	0.7	0.61												
	HC	0.94	0.91	0.96	0.94	0.92	0.91	0.94	0.92	0.96	0.94	0.92	0.91	0.94	0.92	0.96	0.94	0.92	0.93												

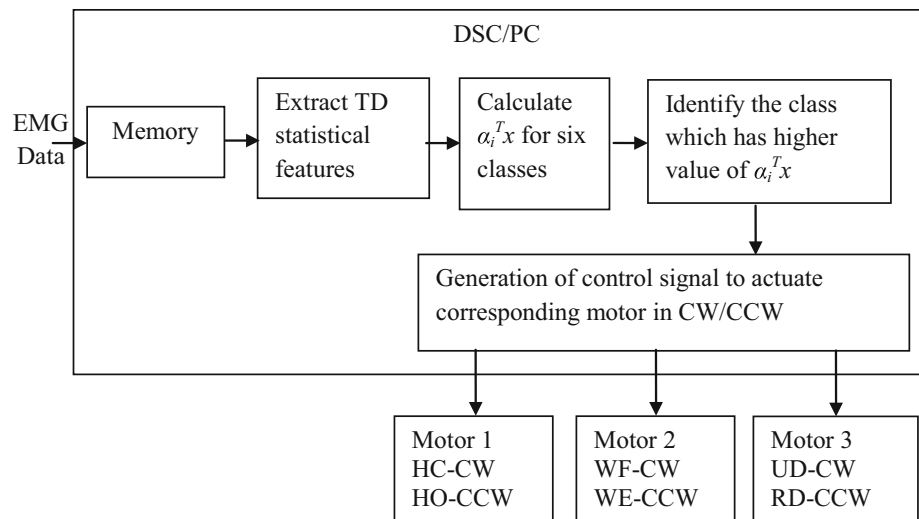
**Table 4** Confusion Matrix of NN with reduced features using PCA

Type of feature	Motion	NN -4										NN -8										NN -12																																							
		HC					HO					WF					WE					UD					RD					HC					HO					WF					WE					UD					RD				
		HC	HO	WF	WE	UD	RD	UD	WE	WF	HO	HC	HO	WF	WE	UD	RD	UD	WE	WF	HO	HC	HO	WF	WE	UD	RD	UD	WE	WF	HO	HC	HO	WF	WE	UD	RD	UD	WE	WF	HO	HC	HO	WF	WE	UD	RD	UD	WE	WF	HO										
Reduced TD feature	HC	85.5	6.5	0.9	0	1	6	86.6	5.5	1.7	0	0.8	5.5	89.4	5.2	1.2	0.1	1.1	3	86.6	5.5	1.7	0	0.8	5.5	89.4	5.2	1.2	0.1	1.1	3	86.6	5.5	1.7	0	0.8	5.5	89.4	5.2	1.2	0.1	1.1	3																		
	HO	8.4	78.1	6.6	0.3	4.3	2.3	6.8	79.1	6.6	0.3	4.6	2.5	7.6	78.1	5.6	0.7	4.9	3	6.8	79.1	6.6	0.3	4.6	2.5	7.6	78.1	5.6	0.7	4.9	3	6.8	79.1	6.6	0.3	4.6	2.5	7.6	78.1	5.6	0.7	4.9	3																		
	WF	3.8	11.6	79.5	1.3	2.2	1.5	4.7	7.5	82.6	0.8	3.2	1.2	3.6	9	80.6	2.5	3	1.3	4.7	7.5	82.6	0.8	3.2	1.2	3.6	9	80.6	2.5	3	1.3	4.7	7.5	82.6	0.8	3.2	1.2	3.6	9	80.6	2.5	3	1.3																		
	WE	0.7	0.8	0.5	87.5	3.7	6.9	0.5	1.1	0.6	86.6	3.7	7.4	0.5	1.2	0.7	87	29	7.7	0.5	1.1	0.6	86.6	3.7	7.4	0.5	1.2	0.7	87	29	7.7	0.5	1.2	0.7	87	29	7.7																								
	UD	4.8	4.3	1.7	2	77.9	9.4	2.7	3.7	2.9	2.5	80.5	7.6	2.8	4.8	2.1	2.7	82.4	5.2	2.7	3.7	2.9	2.5	80.5	7.6	2.8	4.8	2.1	2.7	82.4	5.2	2.7	3.7	2.9	2.5	80.5	7.6	2.8	4.8	2.1	2.7	82.4	5.2																		
	RD	0.5	2.1	0	8.7	10.6	78	0.7	2.4	0.4	9.6	11	76	0.5	1.1	0.1	10	10.6	77.5	0.7	2.4	0.4	9.6	11	76	0.5	1.1	0.1	10	10.6	77.5	0.7	2.4	0.4	9.6	11	76	0.5	1.1	0.1	10	10.6	77.5																		
Sensitivity		0.86	0.78	0.83	0.88	0.82	0.78	0.87	0.79	0.87	0.87	0.77	0.89	0.78	0.84	0.87	0.84	0.78	0.87	0.79	0.87	0.87	0.83	0.77	0.89	0.78	0.84	0.87	0.84	0.87	0.84	0.78	0.87	0.79	0.87	0.87	0.83	0.77	0.89	0.78	0.84	0.87	0.84	0.78																	
		0.96	0.95	0.98	0.98	0.96	0.95	0.97	0.96	0.98	0.97	0.95	0.97	0.96	0.98	0.97	0.97	0.95	0.96	0.97	0.96	0.98	0.97	0.95	0.95	0.97	0.96	0.98	0.97	0.97	0.95	0.96	0.98	0.97	0.95	0.96	0.97	0.96	0.98	0.97	0.95	0.95	0.97	0.96	0.98	0.97	0.95	0.96													
Reduced AR feature	HC	65.7	8.8	3.8	5.1	8.5	8.2	67.2	10.6	3.4	2.8	10	6	70.3	10.2	3.3	2.6	5.3	8.3	67.2	10.6	3.4	2.8	10	6	70.3	10.2	3.3	2.6	5.3	8.3	67.2	10.6	3.4	2.8	10	6	70.3	10.2	3.3	2.6	5.3	8.3																		
	HO	13.7	57.3	8.4	3.5	8	9	11.5	62.4	8.3	1.3	8.6	7.9	8.3	64.5	6.6	2.7	9.4	8.6	11.5	62.4	8.3	1.3	8.6	7.9	8.3	64.5	6.6	2.7	9.4	8.6	11.5	62.4	8.3	1.3	8.6	7.9	8.3	64.5	6.6	2.7	9.4	8.6																		
	WF	4.2	13.2	70.3	1.3	7.2	3.8	3.4	9.6	77.1	1.4	6.2	2.4	3.1	10.1	77.9	1	5.1	2.7	3.4	9.6	77.1	1.4	6.2	2.4	3.1	10.1	77.9	1	5.1	2.7	3.4	9.6	77.1	1.4	6.2	2.4	3.1	10.1	77.9	1	5.1	2.7																		
	WE	1.7	2.5	0.7	79	4.6	11.6	1.7	2.5	2.2	2.2	77.7	5.6	1.2	2	1.2	80.2	6	9.5	2.2	2.5	2.2	2.2	77.7	5.6	1.2	2	1.2	80.2	6	9.5	2.2	2.5	2.2	2.2	77.7	5.6	1.2	2	1.2	80.2	6	9.5																		
	UD	13.3	6.8	6	8.5	54.8	10.7	6	7.5	4.8	6.3	66.7	8.7	5.7	8.5	5.9	5.3	66.1	8.5	6	7.5	4.8	6.3	66.7	8.7	5.7	8.5	5.9	5.3	66.1	8.5	6	7.5	4.8	6.3	66.7	8.7	5.7	8.5	5.9	5.3	66.1	8.5																		
	RD	5.6	11.4	3	20.3	9.5	50.2	3.5	6.2	3	20	11.5	55.8	3.2	9.6	2	16.5	10	58.7	3.5	6.2	3	20	11.5	55.8	3.2	9.6	2	16.5	10	58.7	3.5	6.2	3	20	11.5	55.8	3.2	9.6	2	16.5	10	58.7																		
Sensitivity		0.66	0.57	0.73	0.80	0.63	0.53	0.67	0.62	0.8	0.79	0.71	0.58	0.7	0.64	0.8	0.81	0.7	0.6	0.67	0.62	0.8	0.79	0.71	0.58	0.7	0.64	0.8	0.81	0.7	0.6	0.67	0.62	0.8	0.79	0.71	0.58	0.7	0.64	0.8	0.81	0.7	0.6																		
		0.92	0.91	0.96	0.92	0.92	0.91	0.95	0.93	0.96	0.94	0.92	0.93	0.96	0.92	0.96	0.94	0.93	0.92	0.95	0.93	0.96	0.94	0.92	0.93	0.96	0.92	0.96	0.94	0.93	0.92	0.95	0.93	0.96	0.94	0.92	0.93	0.96	0.92	0.96	0.94	0.93	0.92																		
Specificity		0.92	0.91	0.96	0.92	0.92	0.91	0.95	0.93	0.96	0.94	0.92	0.93	0.96	0.92	0.96	0.94	0.93	0.92	0.95	0.93	0.96	0.94	0.92	0.93	0.96	0.92	0.96	0.94	0.93	0.92	0.95	0.93	0.96	0.94	0.92	0.93	0.96	0.92	0.96	0.94	0.93	0.92																		
		0.66	0.57	0.73	0.80	0.63	0.53	0.67	0.62	0.8	0.79	0.71	0.58	0.7	0.64	0.8	0.81	0.7	0.6	0.67	0.62	0.8	0.79	0.71	0.58	0.7	0.64	0.8	0.81	0.7	0.6	0.67	0.62	0.8	0.79	0.71	0.58	0.7	0.64	0.8	0.81	0.7	0.6																		

**Table 5** Confusion Matrix of *k*NN with reduced features using PCA

Type of Feature	Motion	<i>k</i> NN -4						<i>k</i> NN -8						<i>k</i> NN -12					
		HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD	HC	HO	WF	WE	UD	RD
Reduced TD feature	HC	<b>90</b>	51	0.6	0	0.7	3.6	<b>91</b>	4.7	0.6	0	0.7	3	<b>91</b>	4.6	0.6	0	0.7	3
	HO	7.3	<b>80.9</b>	6.2	0.2	3.9	1.6	6.7	<b>81.7</b>	6.3	0.2	3.3	1.8	6.8	<b>81.6</b>	6.3	0.2	3.3	1.8
	WF	3.5	7.2	<b>85</b>	0.8	2.3	1.2	3.1	8	<b>85</b>	0.8	2	0.9	3.1	8.1	<b>85</b>	0.8	2	1
	WE	0.4	0.9	0.6	<b>85.7</b>	4.2	8.2	0.5	1	0.5	<b>86.3</b>	3.9	7.8	0.5	1	0.5	<b>86.2</b>	4	7.9
Sensitivity	UD	2.4	3.7	1	1.7	<b>81.6</b>	9.7	1.6	4	0.9	1.5	<b>83.4</b>	8.7	1.7	3.9	0.9	1.3	<b>83.5</b>	8.7
	RD	0.5	2.1	0.1	7	11.6	<b>78.7</b>	0.6	2.1	0.1	7.2	10.1	<b>79.9</b>	0.6	2.1	0.1	6.9	10.2	<b>80.1</b>
		0.9	0.81	0.88	0.86	0.83	0.79	0.91	0.82	0.88	0.87	0.85	0.8	0.91	0.82	0.88	0.87	0.85	0.8
		0.97	0.96	0.98	0.98	0.95	0.95	0.97	0.96	0.98	0.98	0.96	0.96	0.97	0.96	0.98	0.98	0.96	0.96
Reduced AR feature	HC	<b>69.4</b>	9.9	4.1	2.3	7.4	6.9	<b>69.8</b>	10.4	2.4	4	6.3	7.2	<b>66.6</b>	10.7	2.7	3	8.2	8.7
	HO	9.1	<b>60.9</b>	8.8	1.8	10	9.4	8.1	<b>62.1</b>	7.6	2.3	10.8	9.1	8	<b>59.8</b>	7.5	2.4	11.8	10.4
	WF	3.9	9.8	<b>75.1</b>	0.9	6.2	4.2	3.2	9.7	<b>76.8</b>	1.2	6.1	3.1	3.6	9.6	<b>75.2</b>	1.1	6.9	3.6
	WE	1.1	2	1.2	<b>77.5</b>	4.9	13.2	1.4	1.7	1.4	<b>75</b>	6	14.5	1.6	1.8	1.2	<b>72.2</b>	9.2	14
Sensitivity	UD	5.4	6.6	5.7	4.5	<b>65.3</b>	12.5	4.4	7.9	5.2	6.6	<b>63.9</b>	12	5.4	8.3	5	6.7	<b>62</b>	12.7
	RD	3.3	8.3	2.3	16	8.3	<b>61.8</b>	3	6	1	18	11	<b>61.1</b>	3.3	7.7	1.9	15.2	12.1	<b>59.8</b>
		0.69	0.61	0.78	0.78	0.69	0.64	0.7	0.62	0.79	0.76	0.67	0.63	0.67	0.6	0.78	0.73	0.65	0.62
		0.95	0.93	0.96	0.95	0.93	0.91	0.96	0.93	0.96	0.94	0.92	0.91	0.96	0.92	0.96	0.94	0.9	0.9

**Fig. 6** Block diagram of interfacing motor with TMSLF2407 DSP



Surprisingly, DT classifier with 25 and 75 % of PCA transformed TD features is more than 80 % and performed better than SLR classifiers with transformed features. Further, DT classifier with 50 % of PCA transformed classifier performance is almost close to 80 %. There is no significant difference in the performance of DT classifiers with PCA transformed TD statistical features, compared to NN and LMT. Also, it is found, no significant difference in the performance of DT with transformed and non-transformed statistical features. But, there is a significant difference in performance between transformed and non-transformed AR features with DT. However, the performance of DT classifiers with PCA transformed features is superior than SLR classifier performance with PCA transformed features. The performance of LMT classifier with transformed features is superior than DT, may be due to the combination of DT and SLR. Only the performance of SLR classifiers with 25 % of PCA transformed AR features is inferior than the other classifiers with less than 60 % of classification efficiency. It is clearly evinced that, SLR classifier performs very poor with PCA transformed features than any other classifier. From the Fig. 5 and from Tables 1, 2, 3, 4, 5 it is also evident that the classification results ameliorates with the increase in the transformed AR feature data except *kNN* classifier. Furthermore, there is a slight improvement/decrease and no significant difference in the classification efficiency while increasing PCA transformed data from 4 (25 %) to 12 (75 %) in statistical features with the considered the classifiers.

There is a significant difference in performance of considered classifiers compared to PCA and non-PCA transformed features. SLR and LMT classifier with non-transformed TD statistical feature is superior than other classifier. Though the performance of SLR, LMT and NN with non-transformed AR features is higher among the other classifiers, but less than non-transformed statistical

features. The performance of *kNN*, classifier is always less than 90 %, even though *kNN*, outperforms other classifier with PCA transformed statistical features. From the analysis, it can be divulged that, the performance of classifiers degrades with PCA transformed time domain features compared to non-transformed time domain features. In addition the performance all classifiers with non-transformed statistical features is superior than non-transformed AR features. Further, it has been observed that the performance of classifiers with AR coefficients less than that of statistical features irrespective of the transformation. Though SLR and LMT performs equally better with statistical features, SLR with the statistical features outperforms other classifier in computation power as well. LMT is a combination of DT and SLR and may require more computing power than SLR.

## DC motor actuation

The different motions of hand are implemented in off-line, to understand the application pattern recognition in actuation of drives. The block diagram of hardware implementation for the actuation of DC motor is shown in Fig. 6. This DC motor actuation is also implemented in the personal computer (PC) using MATLAB.

In this research during off-line stage, the stored data of subjects for different motions of the hand are used to test the developed pattern recognition model of the controller using TMSLF2407. A DSC programme is developed to extract the features namely mean absolute value (MAV), number of zero crossings (ZC), number of slope sign changes (SSC) and waveform length (WL) from the EMG data stored in the memory. The extracted features are classified using simple logistic regression (SLR) classifier. The output of the classifier is used to generate the driving

signals for the actuation of the specific drive in the specified direction, through the configuration of general purpose input and output port (GPIO). The pattern recognition algorithm is carried out in the processor. The input to the controller is the EMG data and the output of the processor is the digital signals that are fed to the bidirectional driver KA3082. Control signal from two GPIO pins of port2 of DSC is fed to pin 5 and pin 6 of the bidirectional driver. GPIO port P2.3 and port P2.4 are used to obtain control signals for hand close (HC) and hand open (HO) motions. GPIO port P2.5 and GPIO port P2.6 are used to obtain control signals for wrist flexion (WF) and wrist extension (WE). GPIO port P2.7 and GPIO port P2.8 are used to obtain control signals for ulnar deviation (UD) and radial deviation (RD). All the six motions are thus realised to perform the intended limb motions by driving three DC motors in both forward and reverse direction in off-line. The time period of execution in offline was found to be less than 150 ms, which is an acceptable delay in real-time implementation of prosthetic hand [4].

## Conclusion

The influence of PCA based feature reduction on pattern recognition of different classifier is tested with 25, 50 and 75 % PCA transformation of TD statistical features as well as AR coefficients. The performance is compared with non-PCA transformed features obtained from the continuous entire length of EMG signals for ten healthy subjects as well. From the results, it has been observed that principal component with the highest Eigenvalue contributes in classification and other components does improve/decrease classification accuracy slightly but not significantly. Performance analysis also shows that *k*NN classifier is better at classifying the EMG signal with PCA transformed statistical data compared to other classifiers in accuracy, sensitivity and specificity. It is also found that there is not much improvement in classification accuracy with non-transformed and transformed TD statistical feature in DT as well as *k*NN. But there is much significant difference in DT as well as *k*NN with with non-transformed and transformed AR coefficients. However, SLR performs better with the statistical features compared to AR and PCA transformed features both in accuracy and computational power. Further, transformed features using PCA decreases accuracy, sensitivity and specificity compared to non-transformed features in the identification of movements from EMG using SLR, NN and LMT. The efficient SLR classifier is implemented in off-line to study motion control of three drives using TMSLF2407 processor to understand the pattern recognition in control of drives.

## References

- Ajiboye AB, Ff Weir RF (2005) A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthetic control. *IEEE Trans Neural Syst Rehabil Eng* 13:280–291
- Chu J, Moon I, Mun M (2006) A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand. *IEEE Trans Biomed Eng* 53:2232–2239
- Das K, Osechinskiy S, Nenadic Z (2007) A classwise PCA-based recognition of neural data for brain computer interfaces. *Proc IEEE EMBS, Lyon*, pp 6519–6522
- Englehart K, Hudgins B, Parker PA (2003) A robust real-time control scheme for multifunction myoelectric control. *IEEE Trans Biomed Eng* 50:848–854
- Friedman J, Hastie R, Tibshirani (2000) Additive logistic regression: a statistical view of boosting. *J Ann Stat* 38:337–374
- Geethanjali, P and Ray KK (2013) Statistical pattern recognition technique for improved real-time myoelectric signal classification. *Biomed Eng—Appl, Basis Commun.* 25(2):1–9
- Geethanjali P, Ray KK, Vivekananda Shanmuganathan P (2009) Actuation of prosthetic drive using EMG signal. *Proceedings of IEEE International Conference TENCON, Singapore*, pp 1–6
- Gioioso G, Salvietti G, Malvezzi M, Prattichizzo D (2013) Mapping synergies from human to robotic hands with dissimilar kinematics: an approach in the object domain. *IEEE Trans Robotics* 29:825–837
- Hargrove L, Losier L, Lock B, Englehart K, Hudgins B (2007). A real-time pattern recognition based myoelectric control usability study implemented in a virtual environment. *Proceedings of 29th Annual International Conference of the IEEE EMBS*, pp 4842–4845
- Hargrove LJ, Li G, Englehart KB, Hudgins BS (2009) Principal components Analysis preprocessing for improved classification accuracies in pattern-recognition-based myoelectric control. *IEEE Trans Biomed Eng* 56:1407–1414
- Hudgins B, Parker PA, Scott RN (1993) A new strategy for multifunction myoelectric control. *IEEE Trans Biomed Eng* 40:82–94
- Kruger U, Zhang J, Xie L (2007). Developments and applications of non-linear principal component analysis a review, in *Principal Manifolds for Data Visualization and Dimension Reduction*, 58:1–44
- Landwehr N, Hall M, Frank E (2005) Logistic model trees. *J Mach Learn* 59:161–205
- Bin Ahmad Nadzri AA, Ahmad SA, Marhaban MH, Jaafar H (2014) Characterization of surface electromyography using time domain features for determining hand motion and stages of contraction. *Australas Phys Eng Sci Med* 37(1):133–137
- Naik GR, Arjunan S, Kumar DK (2011) Applications of ICA and fractal dimension in sEMG signal processing for subtle movement analysis: a review. *Australas Phys Eng Sci Med* 34(2):179–193
- Naik GR, Kumar DK, Palanizwami M (2010) Classification of low level surface electromyogram using independent component analysis. *IET Signal Proc* 5:479–487
- Naik GR, Kumar DK (2010) Twin SVM for gesture classification using the surface Electromyogram. *IEEE Trans Inform Technol Biomed* 14:301–308
- Phinyomark A, Phukpattaranont P, Limsakul C (2012) Feature reduction and selection for EMG signal classification. *Expert Syst Appl* 39:4420–7431
- Quinlan JR (1996) Improved use of continuous attributes in C4.5. *J Artif Res* 4:77–90
- Witten LH, Frank E (2005) *Data Mining: Practical Machine Learning Tolls and Techniques*. Morgan Kaufman Press, San Francisco