Pairwise Linear Discriminant Analysis of Electromyographic signals

Oscar De Silva, George Mann and Ray Gosine
Faculty of Engineering and Applied Science
Memorial University of Newfoundland, Canada.
{oscar.desilva,gmann,rgosine}@mun.ca

Abstract—Electromyographic (EMG) signals are used as rich information sources for control of intelligent prosthetics. For efficient classification the machine learning algorithms used should allow the nonlinear nature of the multi class problem. For generalized application they should have the analytical ability to systematically tackle the problem in hand. To meet these requirements a pair-wise Linear Discriminant Analysis(LDA) is performed in a systematic manner on EMG signals captured from forehand muscles. A 6 class classification performance from 4 EMG channels are reported along with the ability of the algorithm to scale and visualize complex multidimensional cases.

Index Terms—EMG, Linear discriminant analysis, Upper limb prosthetics, Pattern Recognition

I. INTRODUCTION

Powered upper limb prosthetics are devices engineered for the class of amputee patients with partial or complete amputation, related to the arm. The devices should meet the requirements of day to day manipulation tasks amputee patients, thus considered a complex engineering problem both in terms of hardware design and control. Electromyography (EMG) pattern recognition based strategies have emerged as the dominant form of control for these devices during the last decade.

Pattern recognition based methods used for prosthetic control has enabled sequential execution of multiple operations by the patient. By exploiting the set of EMG patterns distinct to each discrete task, the method is able to continuously produce accurate control signals for the device. The classifiers have simplified from powerful Multi Layer Perceptron (MLP) neural nonlinear methods[1] to simpler Linear Discriminant (LD) methods[2] with comparable accuracy and speed for real time implementations. Products based on EMG pattern recognition has already made its way to active lower arm prosthetic market.

EMG based control of prosthetics become complex and highly constrained as the amputation is higher in the arm. As the level of amputation becomes higher the control input sources(muscles, nerves) reduces drastically. This issue is dominant in shoulder and forequarter amputees, requiring much cognitive demand from the patient to carry out simple manipulation tasks [3]. As a solution, complex strategies which employ EMG signals together with other information sources are sought to achieve multi functionality, simultaneous operation and reduced overall cognitive burden in control.

In developing such higher dimensional control methods it is necessary to employ algorithms which are scalable to different muscle groups and has the ability to provide analytical advantage over the conventional classifiers available. [4] reports the ability of partitioning schemes (Multiple Binary Classifications (MBC)) to meet the analytical advantage for EMG controller development. This study develops on the work by introducing modifications to the partitioning scheme with pair wised implementation and linear discriminant functions which is observed to perform accurately in the multidimensional nonlinear setting.

The study develops a feature projection based classification method for EMG signals to suit the analysis of algorithms in different muscle groups. The multi class classification is performed using a partitioning method which improves the ability to visualize the separability between tasks. A linear discriminant (LD) classifier is employed in the reduced partitioned space to preserve the simplicity present in the well established multi class linear discriminant classifiers. This method is found to posses features which would serve as a valuable analytical tool for complex EMG controller development.

II. BACKGROUND

Classifiers in the context of EMG signal processing is widely reported for multi functional control of prosthetics. It was first applied for transient EMG signal classification using invasive electrodes [1] and later developed for steady state EMG signals from surface electrodes [5]. The pattern recognition process can be separated in to three main steps which includes data segmentation, feature extraction/dimensionality reduction and classification. Table I summarizes the main methods reported in literature, which suggests that Linear Discriminant(LD) classifier on Time domain (TD) feature sets have become the popular choice for EMG signal classification. These have comparable accuracy to the much powerful yet resource demanding nonlinear neural classifiers with higher order wavelet packet feature sets.

Control methods available for upper extremity prosthetics (UEP) can be broadly identified as hybrid methods and surgical methods. The hybrid methods are developed by utilizing other body signals together with the available EMG signals. Establishing proprioception has been a concept studied for
the development of these devices [9] [10]. Hybrid methods reported for UEP control include shoulder position sensors [11] [12], foot mounted interfaces [13], cognitive vision [14] and Mechanomyography [16]. Recent development for UEP control is reported by surgical rerouting of the control nerves from brain to targeted muscle areas in the chest. This method (Targeted muscle Reinnervation (TMR)) allows the use of EMG classifiers for task discrimination using the reinnervated muscles instead of the actual physical muscle which is no longer available due to amputation [8]. These developments in TMR prosthetics and hybrid control schemes demand the EMG classifiers to provide analytical ability and robustness to handle nonlinearity in the higher dimensional setting.

Dimensionality reduction of EMG features, for classification is previously reported in both supervised [17] [4] and unsupervised settings [5]. Principal Componant Analysis (PCA) is reported to improve feature representation in wavelet and time domain feature sets [5]. Multi class Linear Discriminant Analysis (LDA) is reported for forearm EMG signal classification, but multi class implementation of the LDA has forced the use of complex neural methods to achieve classification accuracy [17]. Partitioned feature projection and classification methods are reported for transforming multi class problems to series of binary classification problems [4]. The partitioning was implemented in a “One against all” scheme with simple threshold selection for class boundaries. Improvement over LD classifier was not achieved but improved controllability and intuitive visualization suitable for analysis is reported. This research extends the partitioning scheme to a pair wised implementation and replaces the simple thresholding with statistical representations. The visualization capability in the lower dimensional domain and ability to handle nonlinearities of the approach serves as a valuable analytical tool for different EMG classifier development.

### III. Methodology

This study proposes pairwised partitioned feature projection approach for EMG classification. Fishers linear discriminant, also termed Linear Discriminant Analysis (LDA) is used for supervised feature projection in a partitioned feature space. It is important to note the difference between the Linear Discriminant classifier (LD) and the supervised feature projection method Linear Discriminant Analysis(LDA). The data acquisition, data segmentation and feature extractions steps of the pattern recognition process replicates the work of Englehart et al. [2]. The classifier developed for this study includes a pairwise LDA feature projection and a LD classifier for binary classification in the reduced feature space.

#### A. Reference Data set

Four surface electrodes were placed on wrist flexors, extensor, ulnar and radial deviators. These were used to capture resulting myoelectric signals during wrist flexion, wrist extension, ulnar deviation, radial deviation and grasp tasks. These tasks were performed sequentially and represented by indexes 0 to 5 in order. This one cycle of 5 tasks were repeated 20 times. The data sets were sampled at 1 KHz using an EMG acquisition system.

Although there was a labeled data set with an assigned class per 5 second window, the task was initiated from rest and terminated to rest during the 5 second period. By using threshold holds at the corresponding muscles EMG level the period which the arm was resting was identified. This captured labeled data set was used for training and validation of the developed classifier.

#### B. Data Segmentation

Data segmentation involves dividing data stream in to packets to be used for feature extraction purposes. The extracted features are then classified in to control signals corresponding to the data packet considered. In order to overcome the perceived delay in prosthetic operation it is necessary to produce control decision streams at a minimum speed of 3.3Hz (300 ms delay). Englehart et al. reports a continuous windowing scheme which uses a moving window to exploit the full process capability of the available signal processor. This produces denser decision streams which can improve classification error with maximum voting post processing operations. In this study the continuous windowing scheme is implemented with experimentally selected parameters.
C. Feature Extraction

Four time domain features were extracted from each EMG channel according to the data segmentation scheme selected. This includes Mean Absolute Value (MAV), Zero crossings (ZC), Sign changes (SC) and Signal length (SL). This constitutes a 16 dimensional feature vector for the pattern recognition task. The alternate options reported for EMG classifications include higher dimensional Wavelet Packet Transform (WPT) features and similar frequency domain features, but for overall simplicity of classification, Time Domain (TD) feature set was selected.

1) Mean Absolute Value: The estimate of the mean absolute value $MAV_i$ of window segment $i$ with a sample length of $L$ takes the form

$$MAV_i = \frac{1}{L} \sum_{k}^L \bar{x}_k$$

where $\bar{x}_k$ is the signal vector from EMG channels at the $k$th sampling time.

2) Zero Crossing: number of times the signal $\bar{x}_k$ crosses zero is used as a simple frequency measure. Increment the zero crossing $ZC_i$ if;

$$(x_k > 0 \text{ and } x_{k+1} < 0) \text{ or } (x_k < 0 \text{ and } x_{k+1} > 0)$$

and

$$|x_k - x_{k+1}| \geq \epsilon$$

where $\epsilon$ is a threshold hold to filter high frequency noise.

3) Slope Sign Changes: is another estimate of frequency by counting the number of derivative sign changes of signal $x_k$ in the window length of $L$. The $\epsilon$ threshold value is used to filter high frequency noise induced triggers. Increment the slope sign change $SC_i$ if;

$$(x_{k-1} > x_k \text{ and } x_k < x_{k+1}) \text{ or } (x_{k-1} < x_k \text{ and } x_k > x_{k+1})$$

and

$$|x_k - x_{k+1}| \geq \epsilon$$

4) Signal Length: A measure of signal complexity is calculated by the total signal length $SL_i$ in the window of length $L$

$$SL_i = \frac{1}{L} \sum_{k}^L |\Delta X_k|$$

D. Feature projection

Applying feature projection to the problem effectively reduces the dimensionality of the problem. In this study we use Linear Discriminant Analysis (LDA) for supervised feature projection. LDA reduces a $n$ dimensional feature space to $k-1$ dimensions where $k$ is the number of classes in the classification problem. For this particular study two partitioning methods were assessed for feature projection and classification. Partitioning method allows to reduce a multi class classification problem in to number of subproblems of binary classification. In a “one against all” approach the supervised projection is performed considering one class as the positive examples and all other classes as the negative examples. Therefore the original $n$ feature $k$ class problem transforms to $k$ subproblems of bi-classification in a 1 dimensional feature space.

The pairwise partitioning method considers only two classes at a time and performs the feature projection and classification. Therefore in pairwise LDA $nC_k$ subproblems of bi classification is performed. In partitioned feature projection, the supervised LDA assumes that the class given feature distributions $P(\bar{X_i}|C_j)$ are gaussian and of equal variance $\Sigma$. Since the variances of the two classes tend to change, a mean variance is used in place. Then the maximum separating hyperplane or in the two class case the maximum separating vector $\bar{w}$ is given by equation 3. The constants are neglected since we are only interested in the vector direction.

$$\bar{w} = (\Sigma_y = 0 + \Sigma_y = 1)^{-1}(\bar{\mu}_{y=1} - \bar{\mu}_{y=0})$$

E. Classification

The simplicity in the 1 dimensional 2 class classification problem allows application of the LD classifier. This classifier is a statistical analytical solution to the classification problem, which can be applied when sufficient knowledge is available of the underlying probability distributions $P(\bar{X}|C_i)$. Here the derivation is presented for a general multi class multi dimensional classification case which is easily simplified for binary classification problems.

The probability of class given feature distributions $P(\bar{X}|C_j)$ are assumed to be multivariate normal distributions with mean $\bar{\mu}_j$ and covariance matrix $\Sigma_j$. The probability of each class $P(C_j)$ is assumed to be multinomial distribution with equal probabilities for each class. Given a feature vector $\bar{X}$, the probability of belonging to class $P(C_j|\bar{X})$ is found by Bayes rule.

$$P(\bar{X}^i|C_j) = \frac{1}{(2\pi\Sigma)^{n/2}}\exp\left(-\frac{1}{2}(\bar{X} - \bar{\mu}_j)^T \Sigma^{-1}(\bar{X} - \mu_j)\right)$$

$$P(C_j|\bar{X}^i) = \frac{P(\bar{X}^i|C_j) \cdot P(C_j)}{P(\bar{X}^i)}$$

$$\hat{C}_j = \arg\max(P(C_j|\bar{X}^i))$$

The class with the maximum posterior probability $P(C_j|\bar{X})$ is selected as the correct class. The terms $P(C_j)$ and $P(\bar{X})$ are constants, thus neglected during logarithm maximization.

$$\hat{C}_j = \arg\max(ln[P(\bar{X}^i|C_j)])$$

$$= \arg\max(X^T \Sigma^{-1} \mu_j - \mu_j^T \Sigma^{-1} \mu_j)$$

For a two class classification problem the distribution $P(C_j|\bar{X})$ is a sigmoid function taking the form in equation 6. For verification of the performance of the simple LDA classifier, an algorithmic logistic classifier is also implemented with
maximum likelihood parameter estimation using the gradient decent method.

\[
P(C_j | \bar{X}^{(i)}) = \frac{1}{1 + e^{-\theta \bar{X}}} \\
\theta = \begin{bmatrix} \mu_2 - \mu_1 \\ 2\Sigma \\ \frac{\mu_1 - \mu_2}{\Sigma} \end{bmatrix}^T
\]

(6)

In the “one against all” partitioning scheme the probability distribution \(P(C_j | \bar{X}^{(i)})\) is directly estimated, but in the pairwise partitioning scheme only the feature given probability distribution between two classes, \(j\) and \(k\) are estimated \(P(C_j | \bar{X}^{(i)}, k)\). Thus the joint probability is found by multiplying pairwise probabilities \(P(C_j | \bar{X}^{(i)}, k)\) over all \(k\).

F. Validation

The last 10 trials(50%) of data were used as the validation data set. The classifier calculates posteriori probabilities of each feature set to select the class with maximum posteriori probability \(P(C_j | \bar{X}^{(i)})\). Since a dense decision stream is obtained using the continuous windowing scheme, the decision stream is combined with a maximum voting post processing to form 1 decision for the period of allowed time delay.

IV. RESULTS

The recorded dataset of 4 EMG channels were processed using a time domain feature extraction in a continuous windowing scheme, with 64ms window length and 16ms time delay for processing. Performance of 3 classifiers were studied, 1) LD classifier without feature projection and same with 2) simple partitioned LDA feature projection and 3) pairwise LDA feature projection. 

All three methods were implemented on the same training and validation data sets. Table 2 summarize the performance of the classifiers implemented i.e. LD- Linear Discriminant classifier, PLDA1- LDA classifier with "One against all” partitioning scheme, PLDA2- LDA classifier with pairwise classification scheme, PLDA3- LDA classifier with pairwise classification scheme (trained by gradient decent method) .The pairwise feature projection scheme out performs the other methods in classification accuracy. The LD classifier(a) was observed to achieve faster training ,thus a candidate choice for robust and simple implementations. The pairwise LDA feature projection (c) takes longer for training but achieves better accuracy and insight in to the problem under study. The longer time periods are acceptable during training because its usually carried out offline. The processing time for validation datasets were similar in all three methods. Therefor the methods can be implemented in a realtime embedded setting.

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Validation Accuracy</th>
<th>Training time (sec)</th>
<th>Validation time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LD</td>
<td>88.6%</td>
<td>3.0</td>
<td>3.3</td>
</tr>
<tr>
<td>PLDA 1</td>
<td>8.5%</td>
<td>78</td>
<td>5</td>
</tr>
<tr>
<td>PLDA 2</td>
<td>90.3%</td>
<td>67</td>
<td>4</td>
</tr>
<tr>
<td>PLDA 3</td>
<td>91.4%</td>
<td>500</td>
<td>4</td>
</tr>
</tbody>
</table>

TABLE II

PERFORMANCE OF THE IMPLEMENTED CLASSIFIERS FOR EMG TASK CLASSIFICATION.

The proposed method of combining the LD classifier with a pairwise LDA feature projection has allowed visualization of distributions between class pairs. This provides further information for feature selection and class selection in developing custom control methods for upper arm prosthetic cases. A rather interesting result was identified by analyzing the reduced dimensional space. All class given probabilities were resembling very closely to gaussian distributions except class 0 which was “rest”. This resembled a Gamma distribution.
This observation supports the assumptions of the LD classifier for most classes which assumes gaussian class given feature distributions.

V. CONCLUSION

A partitioned feature projection based EMG classifier is proposed in the study in order to handle nonlinearity in EMG signals and to provide analytical ability for EMG controller development. Both “one against all” and a pairwise partitioning scheme were implemented. A test dataset of forearm muscle signals were processed with the proposed methods and the Linear discriminant classifier. The pairwise feature projection based linear classifier performed better than the widely reported method. Additionally it provides visualization capability of probability distributions in the reduced dimensional space. This can be used both for analysis and as a visual tool for clinical technicians in the training process of prosthetics.

The pairwise LDA classifier is selected for further studies which involve hybrid controller development for upper extremity amputees. It is expected to apply the method for shoulder and chest muscle groups for shoulder movement discrimination. The analytical advantage and the accuracy of the method is achieved without relying on complex algorithmic classification schemes. Rather the simplicity of linear discriminant classification was preserved in the developed method.

ACKNOWLEDGMENT

This research is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) and Memorial University of Newfoundland, Canada. The data acquisition was supported by School of Human Kinetics and Recreation, Memorial University of Newfoundland.

REFERENCES