Swarmed Discriminant Analysis for Multifunction Prosthesis Control

Rami N. Khushaba *, Ahmed Al-Ani, and Adel Al-Jumaily,

Abstract—One of the approaches enabling people with amputated limbs to establish some sort of interface with the real world includes the utilization of the myoelectric signal (MES) from the remaining muscles of those limbs. The MES can be used as a control input to a multifunction prosthetic device. In this control scheme, known as the myoelectric control, a pattern recognition approach is usually utilized to discriminate between the MES signals that belong to different classes of the forearm movements. Since the MES is recorded using multiple channels, the feature vector size can become very large. In order to reduce the computational cost and enhance the generalization capability of the classifier, a dimensionality reduction method is needed to identify an informative yet moderate size feature set. This paper proposes a new fuzzy version of the well known Fisher's Linear Discriminant Analysis (LDA) feature projection technique. Furthermore, based on the fact that certain muscles might contribute more to the discrimination process, a novel feature weighting scheme is also presented by employing Particle Swarm Optimization (PSO) for estimating the weight of each feature. The new method, called PSOFLDA, is tested on real MES datasets and compared with other techniques to prove its superiority.

Keywords—Discriminant Analysis, Pattern Recognition, Signal Processing.

I. INTRODUCTION

THE myoelectric signal (MES), also known as the Electromyogram (EMG), is a non-stationary signal that carries the distinct signature of the voluntary intent of the central nervous system. It is usually recorded in a noninvasive scheme utilizing a set of surface electrodes mounted on the human forearm. One of the most important applications of the MES is its use in controlling prosthetic devices functioning as artificial alternatives to missing limbs [1]. Advances in myoelectric signal studies revealed that the MES exhibits different temporal structure for different kinds of the arm movements. This in turn facilitates the use of pattern recognition in myoelectric control for prosthetic control. To this end, a wide set of pattern recognition methods were proposed in the literature to to accuratly classify the MES into one of a predefined set of movements [2].

In order to capture the complete muscles activity and maximize the amount of available information, a multi channel approach is usually utilized when measuring the MES signal. However, this will increase the number of extracted features (variables that describe these movements) and hence it will increase the learning parameters of the classifier and may degrade its performance. A straight forward solution to these problems is to project the data onto low-dimensional subspaces to extract the most significant features. Many feature projections techniques were used in myoelectric control with the aim to produce a statistically uncorrelated or independent feature set, a desirable goal in any pattern recognition system. Various approaches of dimensionality reduction were utilized in myoelectric control, these include: principal component analysis (PCA) [3], a combination of PCA and self organizing feature map (SOFM) [4] and linear discriminant analysis (LDA) [5].

Despite being a well known projection technique, the classical LDA suffers from a number of limitations [6]. The first is that it requires the scatter matrices to be nonsingular, while in real world problems they can be singular. The second limitation with LDA is that it treats all data points equivalently, whereas in real world problems each sample may belong to each of the different classes with a certain degree. Finally, classical LDA pays no attention to the decorrelation of the data, which is a desirable property in many applications. One possible approach to overcome the first and the third problems is to use the uncorrelated linear discriminant analysis (ULDA) that requires the reduced features to be statistically uncorrelated with one another [7].

As a variation to the ULDA approach which is based on Singular Value Decomposition (SVD) that is known to be expensive in terms of time and memory requirements for large datasets, this paper proposes a new mixture of fuzzy logic and discriminant analysis as a novel dimensionality reduction technique. The proposed method aims to reduce the dimensionality of the extracted feature set and cluster features, such that the classification accuracy is improved. Due to the fact that most of the biosignals generated by the human body tend to produce patterns that are fuzzy in nature (i.e., belongs to different classes with certain degrees), then the incorporation of the concept of fuzzy memberships is required to reduce the effect of overlapping and outliers points. Unlike the current available variations to Fisher's linear discriminant analysis (LDA), the new method, called PSOFLDA, accounts for the different contribution of different muscles into the discrimination process. Thus, it assumes that the extracted features vary in their importance. In order to implement this muscle importance concept, a novel feature weighting scheme that employs Particle Swarm Optimization (PSO) for the estimation of features' weights is introduced. Also in order to overcome the singularity problem, a regularization parameter is included within each particle (where each particle represents one member of the population).

This paper is structured as follows: Section 2 explains the

All the authors are with the Faculty of Engineering and Information Technology, University of Technology, Sydney (UTS). E-mail: (Rkhushab, Ahmed, Adel@eng.uts.edu.au)

proposed methodology. Section 3 presents the swarm-based weight optimization. The experimental results are given in section 4. Finally the conclusion is given in section 5.

II. Methodology

A variation of fisher's classical LDA is the fuzzy discriminant analysis (FLDA) [8] that emerged as a classification tool, proving to present better performance than LDA. The goal of FLDA is to determine the linear discriminant function that provides the maximum separation of fuzzy groups in a real space. Although the initial work on FLDA dates back to 1986 [8], there were only few attempts in the literature to propose variations to the original FLDA classifier. Inspired by the work of [8], an interesting approach in fuzzy linear discriminant analysis was proposed for use with chemical datasets [9]. In their approach a K-nearest neighbor (KNN) rule was utilized in order to estimate the required memberships. The approach was compared with classical LDA and proved to outperform classical LDA on chemical datasets. The authors mentioned explicitly that several runs should be made in order to decide the best value of K and the other membership parameter utilized in their approach. A different approach in estimating the FLDA memberships using the KNN rule was proposed in [10]. The main difference between the two approaches presented in [9] and [10] is that a preprocessing step employing PCA is utilized by the latter. This in turn forms a fuzzy variation to classical subspace LDA [11]. It is generally known from the literature that the limitation of this approach is that the application of PCA may leads to loosing useful information. In another attempt and inspired by fuzzy support vector machine principles, another technique for estimating the memberships in FLDA was presented first, then a kernel technique was introduced to perform the nonlinear mapping [12]. The presented comparison between the kernel based FLDA (KFDA) and FLDA indicated that better results were achieved by the former, but with the additional computational requirements of the kernel matrix.

In spite of the good performance achieved by FLDA in a number of applications, there are two main issues that both LDA and FLDA suffer from. Firstly, both techniques pay no attention to the decorrelation of the data. Hence they may not always give the optimal results especially when the feature set contains a large degree of redundancy. Secondly, both LDA and FLDA require the total scatter matrix to be nonsingular, while in real problems the scatter matrix can be singular. In order to overcome these problems with LDA and FLDA, a number of approaches were proposed, such as the subspace LDA (mentioned earlier) and regularized LDA [13]. In the regularized LDA a constant z is added to the diagonal elements of S_W, S_W = S_W + zI, for some $z > 0$, where I_N is an identity matrix. It is easy to see that the new S_W is nonsingular, but the main problem is finding the optimal value of z.

This paper proposes a new weighted fuzzy linear discriminant analysis technique (termed as PSOFLDA) for feature projection. A novel approach is presented within PSOFLDA to overcome the earlier mentioned problems by assigning a weighting factor w_j for each feature. In such approach, the features that are most relevant to the problem will be weighted with higher importance than other features. Thus rather than applying PCA to minimize redundancies, we employ a weighting process that aims to maximize the interaction between features. In order to identify these weights, particle swarm optimization is utilized, as explained in the next section.

A. Weighted Fuzzy Discriminant Analysis

Consider a classification problem with c classes, in which the data set of labelled training samples is given as:

$$S = \{(x_1,y_1),(x_2,y_2),...,(x_l,y_l)\} \subseteq (X,Y)^l \tag{1}$$

Where X is the input space and Y is the output space. $X \subseteq \Re^n, Y \subseteq \Re^n$, and l is the number of samples. Each training point x_i originally belongs to one of the c classes and is given a label $y_i \in \{1,2,3,\ldots,c\}$ for $i = \{1,2,3,\ldots,l\}$. The goal is to find an optimal hyper-plane using the training samples that can recognize the test points, i.e., the classifier will have a good generalization capability. In PSOFLDA each point, x_i, belongs to each of the c classes with a certain degree of membership. The fuzzy within class scatter matrix S_W, fuzzy between class scatter matrix S_B, and the fuzzy total class scatter matrix S_T are given as follows:

$$S_W = \sum_{i=1}^{c}\sum_{k=1}^{l} u_{ik}^m (x_k - v_i)(x_k - v_i)^T \otimes (ww^T) \tag{2}$$

$$S_B = \sum_{i=1}^{c}\sum_{k=1}^{l} u_{ik}^m (v_i - \overline{x})(v_i - \overline{x})^T \otimes (ww^T) \tag{3}$$

$$S_T = \sum_{i=1}^{c}\sum_{k=1}^{l} u_{ik}^m (x_k - \overline{x})(x_k - \overline{x})^T \otimes (ww^T) \tag{4}$$

where u_{ik} is the membership of pattern k in class i, m (given that $m > 1$) is the fuzzification parameter, x_{kj} is the value of the k'th sample across the j'th dimension, v_i is the mean of the patterns belonging to class i, and v_{ij} is its value across the j'th dimension. $\otimes$ refers to the *Kronecker product* operation, w is the weight vector associated with all features, i.e., $w = \{w_1, w_2, ..., w_f\}$, where f is the total number of features. $\overline{x}$ is the mean of the training samples which is given in Eq.(5) below.

$$\overline{x} = \frac{1}{l}\sum_{k=1}^{l} x_k \tag{5}$$

In this paper, the value of the membership u_{ik} is calculated using a possibilistic fuzzy clustering approach. The cost function of the possibilistic clustering approach is adopted from [14], as given in Eq.(6) below.

$$J(\theta,U) = \sum_{k=1}^{l}\sum_{i=1}^{c} u_{ik}^m (x_k - \theta_i)^2 + \sum_{i=1}^{c} \eta_i \sum_{k=1}^{l} (1-u_{ik})^m \tag{6}$$

where θ_i is the i'th cluster center, η_i are positive constants that are suitably chosen. The first term in Eq. (6) is the same objective function used in probabilistic clustering approach, while the second term is added to reduce the effect of outliers. In order to find the membership values from the above

International Science Index Vol:5, No:3, 2011 waset.org/Publication/9317

equation, then the values of the clusters centers are needed. A direct way would be to differentiate Eq. (6) with respect to θ_i, but this in turn would cancel the second term leaving only the first term. A general look at the first term of Eq. (6) reveals that it represents the classical within class scatter matrix S_W given in Eq. (2) if the weight is removed. Thus applying the values of the clusters means ensures that the objective function given by Eq. (6) would settle at a global optimum value. Then in order to compute the membership values, a differentiation of the resultant function with respect to u_{ik} needs to be done as follows.

$$\frac{\partial J(\theta,U)}{\partial u_{ik}} = m u_{ik}^{m-1}(x_k - v_i)^2 - m\eta_i(1-u_{ik})^{m-1} = 0 \quad (7)$$

This would in turn result in the following function

$$u_{ik} = \frac{1}{1+\left(\frac{(x_k-v_i)^2}{\eta_i}\right)^{\frac{1}{m-1}}} \quad (8)$$

The values of η_i, where $i = \{1,2,3,\ldots,c\}$ were chosen to be equal to the maximum distance between the samples belonging to that class and the class center.

After computing all the variables, PSOFLDA finds the vector G that would maximize the ratio of the between class scatter matrix to the within class scatter matrix by solving the following equation:

$$G = arg \max_{G} \; trace\left((G^T S_W G)^{-1} G^T S_B G\right) \quad (9)$$

The solution can be readily computed by applying an eigen-decomposition on $S_W^{-1}S_B$, provided that the within class scatter matrix S_W is nonsingular. In this paper, we are using a regularized version of S_W given by $S_W = S_W + zI$, for some $z > 0$ that is included in the particle representation of the weights, where I is an identity matrix. In this way the scatter matrix is guaranteed to be nonsingular. Since the rank of the between class scatter matrix is bounded from above by c - 1, there are at most c - 1 discriminant vectors by PSOFLDA.

III. Swarm based Weight Optimization

One possible solution for finding the best values of the weights is to employ evolutionary algorithms, or EAs. Powerful EA algorithms include genetic algorithm (GA) and Particle Swarm Optimization (PSO). PSO is an effective continuous function optimizer that encodes the parameters as floating-point numbers and manipulate them with arithmetic operators. By contrast, GAs are often better suited for combinatorial optimization because they encode the parameters as bit strings and modify them with logical operators. There are many variants to both approaches, but because PSO is primarily a numerical optimizer, thus PSO is considered in this paper.

A. Particle Swarm Optimization

Particle swarm optimization, is a population based stochastic optimization technique developed by Eberhart and Kennedy in 1995 [15]. It represents an example of a modern search heuristics belonging to the category of *Swarm Intelligence* methods. PSO mimics the behavior of a swarm of birds or a school of fish. The swarm behavior is modelled by particles in multidimensional space that have two characteristics: position (p) and velocity (s). These particles wander around the hyper space and remember the best position that they have discovered. A particle's position in the multi-dimensional problem space represents one solution for the problem. They exchange information about good positions to each other and adjust their own position and velocity with certain probabilities based on these good positions. The original formula developed by Kennedy and Eberhart was improved by Shi and Eberhart with the introduction of an inertia weight ϖ that decreases over time, (typically from 0.9 to 0.4), to narrow the search that would induce a shift from an exploratory to an exploitative mode. Though the maximum velocity of a particle (s_{max}) was no longer necessary for controlling the explosion of the particles, Shi and Eberhart continued to use it, often setting s_{max} = p_{max} that is the maximum velocity is equalled to the maximum value along the specific dimension, in order to keep the system within the relevant part of the search space. This was found to be a good idea that significantly improves the PSO performance and at the same time it costs very little computationally. During iterations each particle adjusts its own trajectory in the space in order to move towards its best position and the global best according to the following equations:

$$s_{ij}(t+1) = \varpi s_{ij}(t) + c_1 r_1(pbest_{ij} - p_{ij}) + c_2 r_2(gbest_{ij} - p_{ij}) \quad (10)$$

$$p_{ij}(t+1) = p_{ij}(t) + s_{ij}(t+1) \quad (11)$$

Where
i: is the particle index
j: is the current dimension under consideration
p_i: is the current position,
s_i: is the current velocity
ϖ: is the inertia weight
t: is the current time step

r_1 and r_2 are two random numbers uniformly distributed in the range (0,1), c_1 and c_2 are cognitive and social parameters respectively, . $pbest_i$ is the local best position, the one associated with the best fitness value the particle has achieved so far, and $gbest_i$ is the global best position, the one associated with the best fitness value found among all of the particles.

The personal best of each particle is updated according to the following equation:

$$pbest_i(t+1) = \left\{ \begin{array}{ll} pbest_i(t) & \text{if } (pbest_i(t)) \leq f(p_i) \\ p_i(t) & \text{if } (pbest_i(t)) > f(p_i) \end{array} \right\} \quad (12)$$

Finally, the global best of the swarm is updated using the following equation:

$$gbest(t+1) = arg \min_{pbest_i} f(pbest_i(t+1)) \quad (13)$$

Where $f(.)$ is a function that evaluates the fitness value for a given position. This model is referred to as the *gbest* (global best) model. In this paper, each particle will represent a vector whose elements are the weights assigned to each feature plus the regularization parameter z. The idea here is to generate a new weight vector by utilizing a set of particles that wander through the solution space searching for the best possible representation achieving the minimum error rates. In such a system, the fitness function was chosen to be the error rates achieved by a suitable classifier. The details of the classifiers chosen will be given in the experiments section.

B. Application of PSO in Fuzzy Discriminant Analysis

The aim of this section is to help the reader understand how to apply PSO in discriminant analysis. As mentioned before, the equation for both the between-class scatter matrix S_W and the within-class scatter matrix S_B require the multiplication by the weight vector w that consists of the weights values associated with all features, i.e., $w = \{w_1, w_2, ..., w_f\}$. This paper propose to utilize PSO to find the optimal values for w, as shown in Fig. 1.

In the first steps, the particles and their associated velocities are initialized with random numbers. Each of the particles will hold one possible representation for the weight vector based on which the scatter matrices are computed. After computing the scatter matrices, the datasets (training and validation) are projected. The optimality of each particle, i.e., the fitness, is evaluated using a classical wrapper approach in which a suitable classifier is chosen as will be mentioned later in the experiments section. Then using the fitness values for all particles, the local best ($pbest_i$) and global best ($gbest_i$) for the whole swarm are updated. The stopping criterion chosen in this paper was to stop the PSO optimization after reaching a certain number of iterations. In such a procedure, the weight vector that will correspondingly produce the minimum training and validation errors will be the optimal one. When the procedure finishes the testing dataset is employed to test the generalization capability of this technique on completely unseen data.

IV. Experiments and Results

In order to present a fair comparison with the available techniques, we include many of them in the experiments. The details of the experiments carried on are listed below:

- **Comparison with other methods**: The PSOFLDA will be compared against two groups of other techniques: the first has already been applied into myoelectric control like ULDA [6], and PCA [3]. The second group include technique that were not used within the myoelectric control problems, like Orthogonal Linear Discriminant Analysis (OLDA) [16], and Fuzzy Discriminant Analysis (FLDA) [9]. These were included because they represent new variations to Fihser's LDA.
- **Datasets employed**: Since this work aims to present a novel variation to the existing techniques, a comparison with the existing techniques is necessary on different datasets before employing it on an MES dataset. For this reason, two sets of experiments are conducted. In the first, the proposed PSOFLDA is tested on datasets acquired from the Machine Learning Repository (www.ics.uci.edu/~mlearn/mlrepository.html) with different number of samples and numbers of features. In the second experiment, an MES dataset collected from thirty subjects is employed.
- **Testing method employed**: The general testing scheme employed is a three way data split. The dataset utilized is divided into three sets: training, validation, and testing. An initial projection matrix is calculated based on the training set. Then a validation set is used in order to optimize the weights to produce the optimum projection matrix that can minimize the mean of the training and validation errors. Finally a completely unseen testing set is utilized to measure the generalization capability of the proposed system.
- **Parameters of PSO**: Specifically the following parameters values were used: maximum number of generations, 30; maximum velocity s_{max}: 20% of the range of the corresponding variable; maximum value along a specific dimension p_{max} = 1 and minimum p_{min} =0; w decreases linearly from 0.9 to 0.4; and acceleration constants c_1, c_2 are set to 2.0.
- **Classifier Type**: Three different classifiers will be utilized to validate the results. The first is a K-Nearest Neighbor classifier (k-NN). The second is Support Vector Machine classifier (SVM) for which the LIBSVM package available online at (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) is utilized. The third classifier is the Linear Discriminant Analysis classifier (LDA).

A. Experiments on UCI datasets

Each dataset taken from the UCI Repository is subdivided into three parts, with the percentage of the data forming each of the training, validation and testing given as 20%, 20% and 60% respectively. The classification error results shown in Table. I were acquired by utilizing a k-NN classifier with the number of neighbors (k) being 5, while Table. II gives the classification error results obtained using the SVM classifier. The results indicate that there are significant variations between the results obtained by FLDA, ULDA, and OLDA on the different datasets, in comparison with PCA. Meanwhile, the proposed PSOFLDA proves to outperform all of those techniques on all of the datasets utilized in a significant manner. This also proves the effectiveness of the PSOFLDA while training and validating on a very small number of samples and testing on a bigger number of samples, thus better generalization is achieved by PSOFLDA.

B. Experiments on MES datasets

The MES dataset utilized in this research was originally collected and used by Chan et al [7]. Eight channels of surface MES were collected from the right arm of thirty normally limbed subjects (twelve males and eighteen females). Each session consisted of six trials. Seven distinct limb motions were used, hand open (HO), hand close (HC), supination (S), pronation (P), wrist flexion (WF), wrist extension (WE), and rest state (R). Data from the first two trials were used

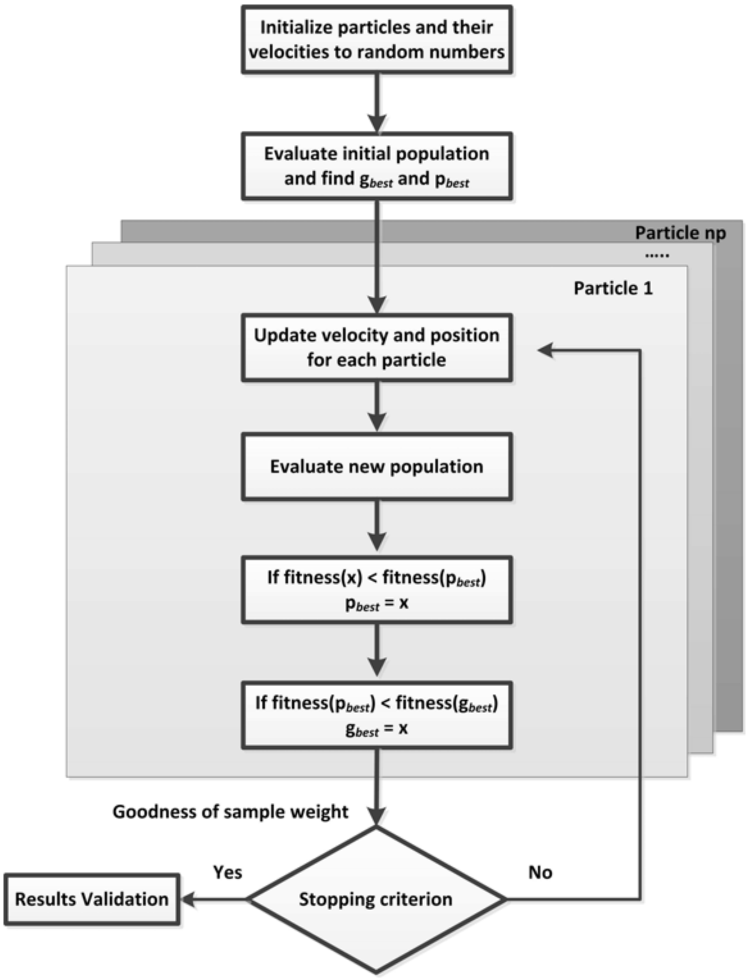


Fig. 1. Steps for finding the optimal weight vector value using PSO

as training set and data from the remaining four trials were divided equally into two trials for validation (trails 1 and 2) and two trails for testing (trails 3 and 4).

As a first part of the MES pattern recognition system, two sets of features were extracted from the raw data in order to test the performance of the proposed method with different feature extraction techniques. The first set of extracted features included a combination of the first four autoregressive (AR) coefficients and the root mean square value (Time-Domain (TD) feature), i.e., number of features = 40 (8 channels $\times$ 5 features/channel). This feature set was referred to as the **TDAR** feature set. The second feature set extracted included the mean of the square values of the wavelet coefficients using a Symmlet wavelet family with five levels of decomposition, i.e., number of features = 48 (8 channels $\times$ 6 features/channel). This feature set was referred to as **WT** feature set. The analysis window size was 256 msec. Data that were 256 msec before or after a change in limb motion were removed from the training set to avoid transitional data. As a dimensionality reduction part, all of the folowwing five methods: PSOFLDA, ULDA, OLDA, FLDA, and PCA were utilized to compare their performance. The final step of the MES recognition system involves a suitable classifier that can be chosen at the disposal of the designer. In the current experiments a Linear Discriminant Analysis (LDA) classifier was chosen. The advantage of this classifier is that it does not require iterative training, avoiding the potential for under- or over-training [7].

The classification accuracy results averaged across thirty subjects (with one standard deviation) using both the TDAR and the WT feature sets reduced in dimensionality with PSOFLDA, ULDA, OLDA, FLDA, and PCA are shown in Fig.2. The number of extracted features from all methods was set to c - 1, where c is the number of classes, as the discriminant analysis based techniques usually ends up with c - 1 features. The results shown for both the validation and testing sets were given first without post processing (referred to as Initial), then with a majority vote (MV) as a post processing step, followed by the transitional data between classes removed (NT), and finally with both majority vote and the removal of the transitional data (MV+NT). The results for both the validation and testing sets are given in the Table-III.

It is clear from the results that the PSOFLDA was able to outperform all other methods. This is due to the fact the PSOFLDA is assigning higher importance to good features compared to those that are less useful. At the same time, the

International Science Index Vol:5, No:3, 2011 waset.org/Publication/9317

TABLE I
RESULTS ON DATA OBTAINED FROM THE UCI REPOSITORY AVERAGED ACROSS 10 RUNS, USING KNN CLASSIFIER, WITH K=5

Dataset	Divisions	PCA	FLDA	ULDA	OLDA	NPE	LPP	PSOFLDA
	Train	26.0101	18.8889	11.7172	29.3435	18.5354	7.1774	6.2096
German	Validate	31.4357	26.4356	27.7228	30.0495	26.6832	13.6508	6.3492
	Test	30.1833	26.6833	27.0667	29.7333	26.7834	13.4133	11.6533
	Train	3.5870	0.7126	0.5737	0.4288	1.1353	0.4288	1.2863
Dermatology	Validate	9.5322	6.6078	9.3555	6.4796	8.4235	7.9037	1.5593
	Test	9.3479	5.5652	8.9861	6.5222	9.1671	6.5232	5.5322
	Train	16.4286	11.1905	10.4762	9.2857	7.8572	6.9048	3.5715
Glass	Validate	32.9546	28.1818	31.8182	22.0455	23.4091	20.6818	3.6364
	Test	29.3750	25.0782	26.7969	21.1719	20.8594	19.0625	8.6719
	Train	20.2362	12.3150	12.3622	12.3307	7.5433	12.4724	11.3543
Splice	Validate	33.0577	24.0874	24.7114	24.2746	30.5772	24.3994	14.7426
	Test	33.3647	24.7492	24.7127	24.6186	32.1055	24.6238	17.9990
	Train	18.7817	16.0914	20.9137	17.5127	16.7513	17.1066	19.4416
Vowel	Validate	46.8844	42.5628	52.6633	43.1156	44.1709	42.7136	28.0905
	Test	45.9259	40.1515	50.8249	42.7441	42.8283	41.2963	37.5421
	Train	4.3750	0.3125	0.0000	0.0000	1.5257	0.0000	0.0000
Wine	Validate	6.9369	5.6687	5.6687	5.9251	6.4379	6.6943	0.0000
	Test	5.7943	3.2710	2.9906	4.0186	5.8878	4.2990	2.8971

TABLE II
RESULTS ON DATA OBTAINED FROM THE UCI REPOSITORY AVERAGED ACROSS 10 RUNS, USING SVM CLASSIFIER (LIBSVM PACKAGE)

Dataset	Divisions	PCA	FLDA	ULDA	OLDA	NPE	LPP	PSOFLDA
	Train	17.0161	18.8889	29.3435	18.5354	13.9899	18.5354	18.0808
German	Validate	31.4285	26.4356	30.0495	26.6832	28.4159	26.6832	19.5049
	Test	31.1200	26.6833	29.7333	26.7834	27.9000	26.7834	26.1667
	Train	19.2647	2.2059	0.8823	8.2353	35.2941	8.2353	5.2941
Ionosphere	Validate	25.1388	18.7500	23.1944	21.9444	36.1111	21.9444	8.0555
	Test	24.7393	17.8672	20.8530	21.6587	35.7820	21.6587	14.7867
	Train	27.1338	15.7795	48.0157	15.7008	3.2283	15.7952	15.9055
Splice	Validate	30.2808	23.4009	48.1747	23.3697	21.0764	23.3229	16.7082
	Test	30.4545	23.0668	48.0093	23.2288	20.6844	23.2236	18.8140
	Train	3.0745	2.2187	4.1680	2.2662	2.2504	2.2662	1.2520
Thyrode	Validate	3.2965	2.6498	4.5110	2.7760	3.0757	2.7760	1.4195
	Test	3.3403	2.8292	4.5416	2.8977	3.0611	2.8977	2.1443
	Train	19.6791	20.9625	13.7967	13.7967	13.7433	21.8716	14.4919
Vowel	Validate	43.8756	41.4354	38.4210	38.4210	37.9904	42.0574	22.3923
	Test	40.8922	39.4275	35.3030	35.3030	35.0336	39.8148	29.7474
	Train	4.3363	2.3009	3.9823	3.7168	12.8319	3.7168	0.9735
WDBC	Validate	7.8261	6.0000	8.4348	7.7391	12.0000	7.7391	1.3044
	Test	7.6246	5.4252	8.5044	7.6833	12.9912	7.6833	3.8123

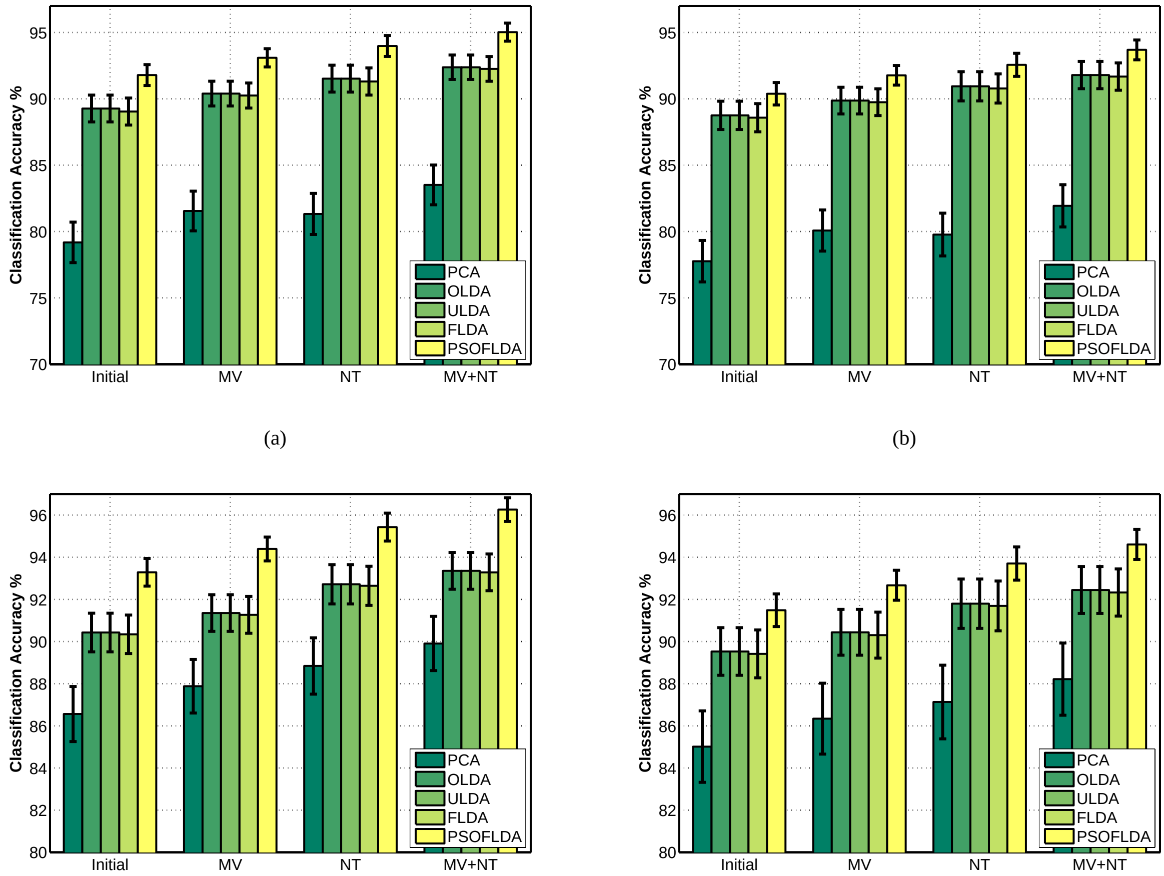
TABLE III
CLASSIFICATION RESULTS ACHIEVED BY DIFFERENT METHODS

Feature	Divisions	**PSOFLDA**	**FLDA**	**ULDA**	**OLDA**	**PCA**
TDAR	Validate	95.02	92.25	92.38	92.38	83.51
TDAR	Test	93.68	91.67	91.79	91.79	81.93
WT	Validate	96.26	93.28	93.35	93.35	89.90
WT	Test	94.60	92.32	92.44	92.44	88.21

use of the classification accuracy as a judgment criterion on the weight values moved the projection matrix closer toward the optimal projection matrix than all other techniques. Also the PSOFLDA assigns lower fuzzy membership values to the outlier points, thus reducing their effect. Another issue to be mentioned here is that with all of the feature projection techniques, the WT features achieved higher accuracies than that achieved using the simple TDAR features. But from computational cost point of view the performance of the system with the TDAR features is still highly accepted.

In order to provide a rigorous validation or comparison with existing techniques for dimensionality reduction, the confusion matrix for all the subjects was also computed for the different feature sets. A plot of the diagonal values of the confusion matrices (class wise classification accuracy) validation and testing sets are presented in Fig. 3 . All the results indicate that there were more significant enhancements when applying the PSOFLDA method than that of the other techniques.

Finally, we also provide the statistical significance test results for the proposed PSOFLDA against all other method. A two-way-analysis of variance (ANOVA) test was carried out with the significance level set to 0.05 and the test results



(a) (b)

(c) (d)

Fig. 2. Classification accuracies using different feature sets averaged across 30 subjects with different dimensionality reduction techniques (a) Using the TDAR validation set and (b) Using the TDAR testing set (c) Using the WT validation set and (d) Using the WT testing set

shown in Table.IV. These results proves the significance of the achieved classification results with PSOFLDA in comparison to the available methods on different feature sets.

V. Conclusion

In this paper, a novel feature projection technique based on a mixture of fuzzy logic and Fisher's LDA was developed. Unlike the typical variations to LDA, The new technqiue assigned higher importance to good features compared with others. The importance was based on a weighting scheme that was optimized with PSO technqiue. This in turn caused the PSOFLDA's projection matrix to be closer to the optimal. The proposed PSOFLDA technique was fairly compared with other techniques like FLDA, ULDA, OLDA, and PCA proving to present better results on real MES datasets. This in turn proves the ability of the proposed technique in enhancing the performance of the multifunction myoelectric hand control system.

Acknowledgment

The Authors would like to acknowledge the support of Dr. Adrian D. C. Chan from Carleton University for providing us with the MES datasets.

References

[1] P. A. Parker, K. B. Englehart, and B. S. Hudgins, "Control of powered upper limb prostheses", in Electromyography Physiology, Engineering, and Noninvasive Applications, R. Merletti and P. Parker Eds. John Wiley and Sons, 2004.

[2] Asghari Oskoei, M., & Hu, H. (2007). Myoelectric Control Systems–A Survey. *Biomedical Signal Processing and Control*, 2, 275–294

[3] Englehart, K., Hudgins, B., Parker, P. A., Stevenson, M. (1998). Time-Frequency Representation for Classification of the Ttransient Myoelectric Signal. In: The 20th EMBS Annual International Conference, pp. 2627–2630. China

[4] Chu, J. U., Moon, I., Mun, M. S. (2006). A Real-Time EMG Pattern Recognition System based on Linear-Nonlinear Feature Projection for a Multifunction Myoelectric Hand. *IEEE Transactions on Biomedical Engineering*, 53, 2232-2239

[5] Chu, J. U., Moon, I., Mun, M. S. (2006). A Supervised Feature Projection for Real-Time Multifunction Myoelectric Hand Control. In:The 28th IEEE EMBS Annual International Conference, pp. 2417–2420. New York City

[6] Ye, J., Janardan, R., Li, Q., Park, H. (2006). Feature Reduction via Generalized Uncorrelated Linear Discriminant Analysis. *IEEE Transactions on Knowledge and Data Engineering*, 18(10), 1312–1322

[7] Chan, A. D. C., & Green, G. C. (2007). Myoelectric Control Development Toolbox. In. Proceedings of The 30'th Conference of the Canadian Medical & Biological Engineering Society, Toronto, ON

[8] Watada, J., Tanaka, H., Asai, K. (1986). Fuzzy Discriminant Analysis in Fuzzy Groups. *Fuzzy Sets and Systems*, 19, 261–271.

[9] Chen, Z. P., Jiang, J. H., Y. Li, Liang, Y. Z. and Yu, R. Q. (1999). Fuzzy

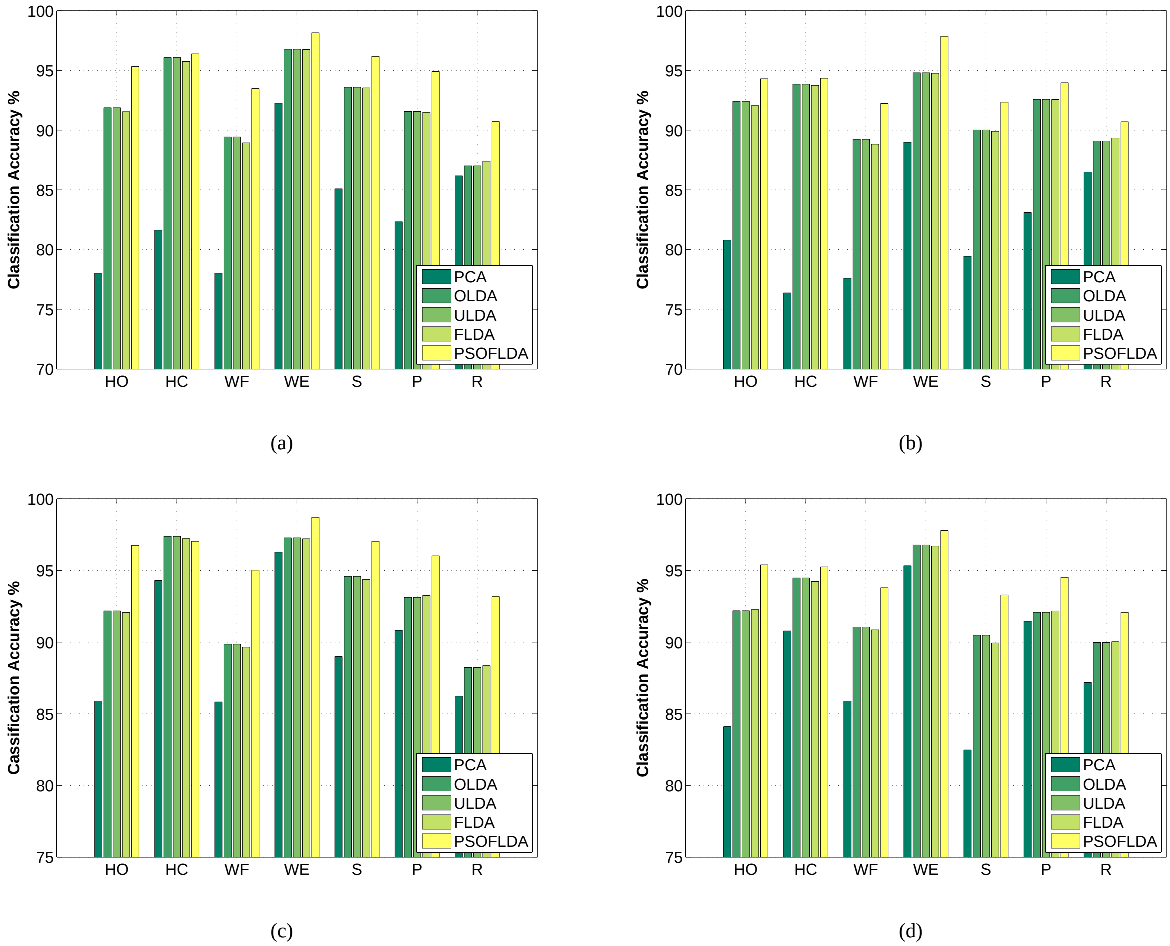


Fig. 3. Diagonal values of the confusion matrix averaged across 30 subjects using the proposed PSOFLDA in comparison with FLDA, ULDA, OLDA, and PCA (a) Using the TDAR validation set and (b) Using the TDAR testing set (a) Using the WT validation set and (b) Using the WT testing set

TABLE IV
TWO-WAY ANALYSIS OF VARIANCE TEST RESULTS

PSOFLDA vs. →	PCA	OLDA	ULDA	FLDA
TDAR	0.00E+00	3.17E-05	3.17E-05	1.44E-05
WT	1.75E-04	0.0013	0.0013	9.56E-04

Linear Discriminant Analysis for Chemical Data Sets. *Chemometrics and Intelligent Laboratory Systems*, 45(1-2), 295-302.

[10] Kwak, K. C. & Pedrycz, W. (2005). Face Recognition Using a Fuzzy Fisherface Classifier. *Pattern Recognition*, 38, 1717-1732.

[11] Swets, D. L. & Weng, J. (1996). Using Discriminant Eigenfeatures for Image Retrieval. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 18, 831–836.

[12] Wu, X.H. & Zhou, J. J. (2006). Fuzzy Discriminant Analysis with Kernel Methods. *Pattern Recognition*, 39, 2236–2239.

[13] Lu, J., Plataniotis, K. N. & Venetsanopoulos, A. N. (2003). Regularized Discriminant Analysis for the Small Sample Size Problem in Face Recognition. *Pattern Recognition Letters*, 24, 3079–3087.

[14] Oliveira, J. V. d., Pedrycz, W. (2007): A Comprehensive, Coherent, and in Depth Presentation of the State of the Art in Fuzzy Clustering. John Wiley & Sons Ltd.

[15] Kennedy, J., Eberhart, R. C., Shi, Y. (2001). Swarm Intelligence. The Morgan Kaufmann Series in Artificial Intelligence, Morgan Kaufmann Publishers, London

[16] Ye, J. (2005). Characterization of a Family of Algorithms for Generalized Discriminant Analysis on Undersampled Problems. *Journal of Machine Learning Research*, 6, 483–502