MEASUREMENT SCIENCE REVIEW, Volume 12, No. 3, 2012

Application of Linear Discriminant Analysis in Dimensionality Reduction for Hand Motion Classification

A. Phinyomark[1,*], H. Hu[2], P. Phukpattaranont[1], C. Limsakul[1]

[1]Department of Electrical Engineering, Faculty of Engineering, Prince of Songkla University,
15 Kanjanavanich Road, Kho Hong, Hat Yai, Songkhla 90112 Thailand

[2]School of Computer Science & Electronic Engineering, University of Essex, Colchester CO4 3SQ U.K.

[*]angkoon.p@hotmail.com, hhu@essex.ac.uk, pornchai.p@psu.ac.th, chusak.l@psu.ac.th

The classification of upper-limb movements based on surface electromyography (EMG) signals is an important issue in the control of assistive devices and rehabilitation systems. Increasing the number of EMG channels and features in order to increase the number of control commands can yield a high dimensional feature vector. To cope with the accuracy and computation problems associated with high dimensionality, it is commonplace to apply a processing step that transforms the data to a space of significantly lower dimensions with only a limited loss of useful information. Linear discriminant analysis (LDA) has been successfully applied as an EMG feature projection method. Recently, a number of extended LDA-based algorithms have been proposed, which are more competitive in terms of both classification accuracy and computational costs/times with classical LDA. This paper presents the findings of a comparative study of classical LDA and five extended LDA methods. From a quantitative comparison based on seven multi-feature sets, three extended LDA-based algorithms, consisting of uncorrelated LDA, orthogonal LDA and orthogonal fuzzy neighborhood discriminant analysis, produce better class separability when compared with a baseline system (without feature projection), principle component analysis (PCA), and classical LDA. Based on a 7-dimension time domain and time-scale feature vectors, these methods achieved respectively 95.2% and 93.2% classification accuracy by using a linear discriminant classifier.

Keywords: Electromyography signal, EMG, uncorrelated LDA, orthogonal LDA, orthogonal fuzzy neighborhood discriminant analysis, kernel discriminant analysis, QR decomposition, feature extraction, feature projection

1. Introduction

Electromyography (EMG) signals are among the most useful electrophysiological signals widely used in medical and engineering applications [1]. In the context of engineering, EMG is used to generate control commands for many rehabilitation and human computer interface (HCI) applications. Control systems based on the classification of EMG signals are usually known as Myoelectric Control Systems (MCSs) [2]. Powered upper-limb prostheses [3] and electric powered wheelchairs [4] are two of the main potential applications of MCSs. Most commercial EMG based prostheses recognize the user's movements by comparing magnitude features of EMG signals with a pre-determined threshold [5]. However, such systems can only generate a small number of control commands, such as an open and close control scheme with a single propulsion speed.

Since the early 1990s, a number of multifunction EMG prostheses have been developed [6-10]. Most research groups try to increase the efficiency of these multifunction EMG prostheses by increasing the number of movements recognized, which can directly increase the number of control commands. However, this leads to a need for increased information to be extracted from the EMG signals. There are two major ways used to increase the information derived from EMG recognition systems: obtaining information from different muscle positions and utilizing the information present in features of the signal [3]. However, whilst increasing the number of EMG channels and EMG features yields a high dimensional feature vector, it also yields the curse of dimensionality problem [6]. As a result, an effective dimensionality reduction technique [8-17] is required to yield an efficient result in all related costs, i.e., measurement, storage and computation, and in classification performance.

Feature projection is a popular way to reduce the dimensions of the EMG feature vector [2]. Several studies have demonstrated that feature projection performs better than other dimensionality reduction techniques used in feature selection including the Euclidean distance [8]. Feature projection creates an appropriate subset of new features from an original feature set such that the learning criterion is optimized. This method not only reduces the dimensions of the feature vector but also increases the power of the classifier [8-17]. Englehart et al. [8] extracted a feature vector through a Discrete Wavelet Transform (DWT) and a Wavelet Packet Transform (WPT), and used Principal Component Analysis (PCA), a linear unsupervised method, as the dimensionality reduction method. Several recent studies have employed PCA as the feature reduction method in a number of EMG applications [9-12]. Chu et al. [13] employed a linear-nonlinear unsupervised method combining PCA and a Self-Organizing Feature Map (SOFM). However, the ability of the classifier is reduced by the reduction of dimensionality if the PCA-reduced dimension is less than twenty orders [14].

The linear supervised method, Linear Discriminant Analysis (LDA), is competitive in its performance with Nonlinear Discriminant Analysis (NLDA) in terms of their class separability [15]. In addition, LDA has a better classification performance compared with PCA and SOFM. However, it is much more efficient than NLDA and SOFM from the point of view of processing time. Recently, several extensions of LDA-based algorithms have been proposed, i.e., Uncorrelated Linear Discriminant Analysis (ULDA) [16], Orthogonal Fuzzy Neighborhood Discriminant Analysis (OFNDA) [17], Generalized Discriminant Analysis (GDA) [18], and a combination of LDA, Fuzzy Logic and the Differential Evolution optimization technique (DEFLDA) [19]. Further, several simple time domain and frequency domain extraction methods [20-21] such as energy, variance, mean absolute value, zero crossing rate, mean power frequency, median power frequency, and auto-regressive coefficients can be deployed as the feature reduction method [22-23].

In this study, simple pattern recognition systems based on one linear discriminant (LD) classifier and seven EMG multi-feature sets, representing both a time domain and time-scale approach, were utilized to evaluate the performance of the feature projection methods proposed, all of which are described in Section 2. The EMG data used in the experiments were recorded from twenty subjects performing eight upper-limb movements through four EMG channels. Based on the experiments described in Section 3, an optimal LDA method with seven multi-feature sets is recommended and the performance of this optimal LDA method compared with a baseline system (BS) using the full feature set without projection and PCA is presented and discussed in Section 4. In addition, EMG multi-feature sets are also evaluated and the set most suitable for implementation with the feature projection technique is described. Finally, concluding remarks and recommendations for future research are given in Section 5.

2. Discriminant analysis based projection methods

LDA is a well-known method for feature extraction and dimension reduction [24]. It has been widely used in many applications such as brain tissue analysis [25], face recognition [26], speech recognition [27] and text classification [28]. LDA takes as its input a set of high-dimensional features grouped into classes by finding an optimal transformation (projection) that maps the raw features into a lower-dimensional space while preserving the class structure. It minimizes the within-class distance and simultaneously maximizes the between-class distance, thus achieving maximum discrimination. This transformation is readily computed by applying the eigen-decomposition on the scatter matrices of a training data set.

Recently, a number of novel LDA-based algorithms have been proposed which are competitive in terms of both classification accuracy and computational costs in time and space compared with classical LDA.

The development of extended LDA-based algorithms is based on both linear and nonlinear structures. Four feature projection methods based on linear structure are evaluated in this study including ULDA [29], orthogonal LDA (OLDA) [30], OFNDA [17] and linear discriminant analysis via QR-decomposition (LDA/QR) [31].

Both ULDA and OLDA are designed to solve the problem of undersampling [32], where the feature dimension is much larger than the sample size. It should be noted, however, that this problem is unlikely to arise with MCS. A key property of ULDA is that the features in the reduced space are uncorrelated to each other; while in OLDA the discriminant vectors are orthogonal to each other. Chan et al. [16], reported on an experiment using EMG data findings that ULDA outperforms PCA. In contrast, Ye [30], using a variety of real-world data sets showed that OLDA is much better than ULDA in terms of its classification accuracy. However, these findings do not guarantee that OLDA will perform better than ULDA in experiments on EMG data sets; therefore, both ULDA and OLDA were evaluated in this study.

Khushaba et al. [17] experimented with the use of the OFNDA method of EMG classification, combining useful properties from many different projection techniques. OFNDA first applies PCA to remove redundancies, then the contribution of the different data points is split into different classes based on fuzzy memberships. Using a time-domain multi-feature set with a two-channel EMG classification system, OFNDA was found to perform better than ULDA and OLDA in this work [17].

However, in the present study, all methods in the literature comprising classical LDA, ULDA, OLDA and OFNDA were evaluated again with a different EMG data set and a different EMG classification system, based on seven multi-feature sets both in the time domain and the time-scale domain.

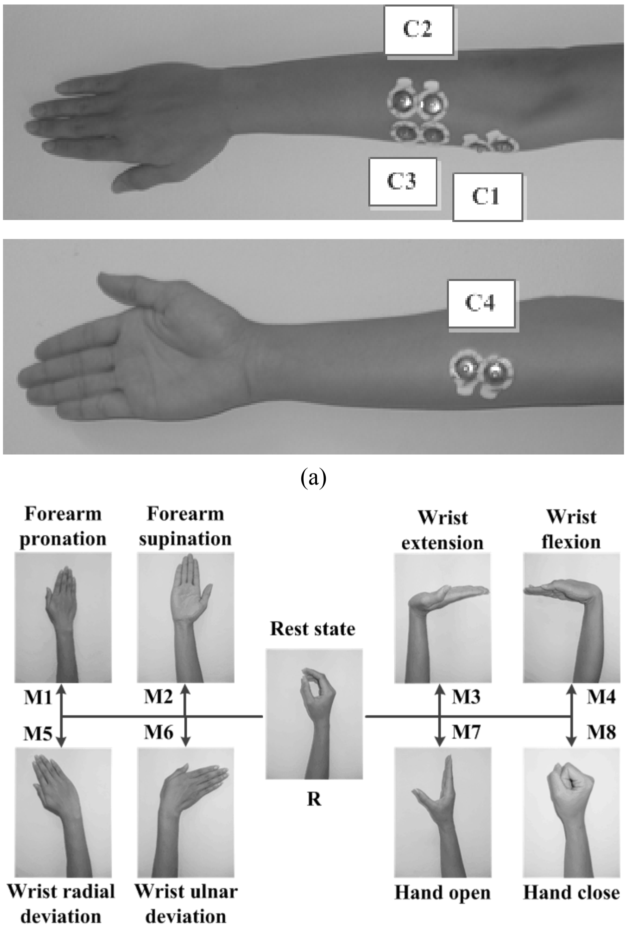
Ye and Li [31] proposed an efficient variant of LDA, named LDA/QR, based on utilizing QR-decomposition on a small size matrix. Thus, the time complexity of LDA/QR is smaller than that of LDA. However, its classification accuracy has been found to be competitive with LDA in several experiments conducted using real-world data.

Although LDA works well for linear problems, it may be less effective when severe non-linearity is involved. To deal with this limitation, the use of non-linear extensions through kernel functions has been suggested. GDA is a general kernel-based non-linear extension of LDA [33]. It was applied as an EMG feature projection technique by Lui et al. [18]. GDA, however, entails high computational cost, and was not, therefore, included in this comparative study. As an alternative, Approximate Kernel Discriminant Analysis via QR-decomposition (AKDA/QR) was included instead of GDA and is the only kernel-based non-linear extension of LDA applied in this comparative study. Its time complexity is much less than that of GDA, but it is similar to the extended LDA method based on linear structure, as can be seen in Table 1. It utilizes QR-decomposition to minimize time complexity by using Gaussian kernels and in an experiment on face image data, Tao et al. [34] showed that the classification accuracy of AKDA/QR is competitive with GDA.

Table 1. A comparison of the time complexity of five dimension reduction methods: n is the number of samples, d is the dimension, and c is the number of classes.

	PCA	LDA	LDA/QR	GDA	AKDA/QR
Time	$O(n^2d)$	$O(n^2d)$	$O(ndc)$	$O(n^2d+n^3)$	$O(ndc)$

Based on the time complexity mentioned in Table 1, it should be noted that many of these LDA methods based on the linear structure have the same computational cost including ULDA and OLDA. Moreover, PCA and LDA also have the same computational cost.



(a)

(b)

Fig.1. (a) Four EMG channels on the right forearm (b) Eight upper-limb movements and a rest state [35].

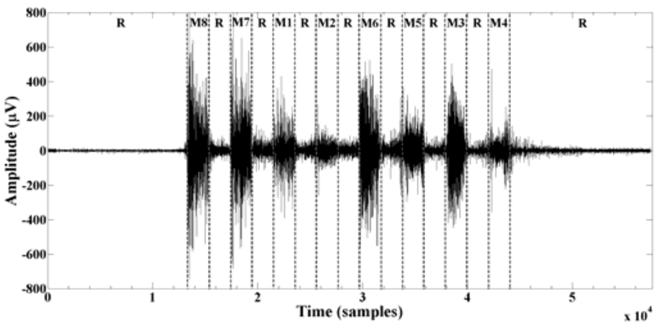


Fig.2. The example EMG data of 56-second in length with the movement order [35].

3. Performance evaluation

Classical LDA was compared with two groups of extended techniques: the first group included methods that have already been applied in MCS: ULDA, OLDA, and OFNDA. The second group included methods that have not previously been used within MCS: LDA/QR and AKDA/QR. Moreover, PCA and the BS were also included in the comparison.

A. EMG Data Acquisition

The EMG data which were used to evaluate the proposed feature projection methods were recorded from eight upper-limb movements through four EMG channels. The Mobi6-6b wireless EMG measurement system (TMS International B.V.) was used with a built-in amplifier of 19.5x and band-pass filter of 20-500 Hz. EMG data were sampled at 1024 Hz with a high resolution of 24 bits. Data were acquired from 20 normally limbed individuals (10 males and 10 females) with ages ranging from 20 to 23 years. Four EMG channels were recorded from electrodes placed on the extensor carpi radialis longus muscle (C1), extensor carpi ulnaris muscle (C2), extensor digitorum communis muscle (C3) and flexor carpi radialis muscle (C4) of the right forearm, as depicted in Fig.1(a). For each channel, two 24 mm circular Ag/AgCl electrodes (H124SG, Kendal ARBO) were placed 20 mm apart. In addition, an Ag/AgCl electrode (Red Dot 2223, 3M Health Care) was placed on the wrist to provide a common ground reference. This was also a disposable pre-gelled self-adhesive surface electrode but its diameter was 43.1 mm.

Each subject generated eight different movement classes: forearm pronation (M1), forearm supination (M2), wrist extension (M3), wrist flexion (M4), wrist radial deviation (M5), wrist ulnar deviation (M6), hand open (M7) and hand close (M8), as shown in Fig.1(b). EMG data were recorded from each subject on four different days. Each day, three sessions were conducted in each of which five data sets were collected from each subject. Within each data set, the subjects performed each movement for a duration of 2 seconds and each movement was separated by a 2-second period in the rest state (R) to avoid there being a transitional stage during movement changes. A 13-second rest period was also allowed at the start and at the end of each data set to allow the subject to prepare and to avoid recording incomplete data by cutting off the recording before the final action was completed. Thus, each data set covered a period of 56 seconds, as can be observed in Fig.2. On each of the four days on which data were collected, a total of fifteen data sets per subject were collected based on which the EMG fluctuations were studied.

B. Classification and Feature Extraction Methods

For each data set, the EMG data were divided into discrete 256-sample (250 ms) records with an increment of 128-sample (125 ms, 50% overlap), and classification was performed on each discrete window. This mode of sampling and segmentation makes the response time achieved in this study acceptable for real-time MCS, in which a delay shorter than 300 ms is acceptable [13-15]. Based on 15 data sets, a feature vector of 1800 class decisions was obtained per channel for each day and each subject. It should be noted that 1800 windows are obtained from 8 movements × 15 datasets × 15 windows/dataset. Fifteen windows for each dataset are computed from the first 256-sample window plus fourteen 128-sample overlapping windows (256·1 + 128·14 = 2048 samples).

The EMG pattern classification algorithms were tested using a 10-fold cross-validation procedure for each day and each subject. Nine out of every ten data groups were used as learning data and the remaining data group was used as the testing data Thus, training was conducted on 1620 patterns and testing on 180 patterns per feature and channel and in this way, the average classification accuracy was calculated for each subject and each day. It should be noted that the same proportion of classes in both training and testing sets was enforced, and in each time the dimensionality reduction transformation matrices were re-computed.

A Linear discriminant (LD) classifier was used to yield the classification accuracy on which the performance of the various feature projection methods was evaluated. The LD classifier was chosen because it is a computationally efficient real-time operation and has been shown to be one of the best classifiers for MCS under stationary conditions [14]. Its classification performance is similar to more complex classification algorithms and it is also a simple statistical approach without any parameter adjustment [36]. Moreover, it has been shown to be the best classifier for MCS based on considerations of reusability and robustness [37]. The LD classifier was trained using learning sets of features extracted by one of seven multi-feature sets (MS) previously reported on in the literature. These were:

MS1- Histogram of EMG (HEMG) and Auto Regressive (AR) coefficients: HEMG and AR are respectively implemented with nine components and with the 4th-order [18]. By using these two time domain methods, a 13-dimension feature vector is obtained from each channel.

MS2 - Hudgins' time domain approach: A popular time domain feature set [6] consisting of five features: Mean Absolute Value (MAV), Mean Absolute Value Slope (MAVS), Waveform Length (WL), Zero Crossing (ZC) and Slope Sign Change (SSC). A 5-dimension feature vector is obtained from each channel. MS2 is usually used as a baseline feature set when researchers evaluate other EMG recognition components, e.g., Geethanjali and Ray [38].

MS3 - Root Mean Square (RMS) and AR coefficients: The first four AR coefficients and RMS were used as a feature vector (5-dimension feature vector per channel) [16, 19].

MS4 - Six time domain methods: Integrated EMG (IEMG), Variance of EMG (VAR), Willison Amplitude (WAMP), WL, ZC, and SSC. A 6-dimension feature vector was obtained from each channel, and evaluated by using the EMG signals recorded from the hand and finger movements [39].

MS5 - Eight time domain methods: RMS, MAV, IEMG, WL, ZC, SSC, sample skewness (SKW) and AR. It should be noted that AR is implemented with the 6th-order [40] for this set. These methods form a feature vector that consists of 13 dimensions per channel (7 dimensions from the first 7 features and 6 dimensions from AR).

MS6 - Reduced coefficients of WPT: The WPT method was computed using 4-decomposition levels with the Symlet wavelet of the 5th-order [8, 14]. Full-length wavelet coefficients (384-dimension vector per channel) were used as an input vector, and the reduced features were obtained by applying the feature projection method.

MS7 - Relative WPT energy (RWPE): This method is computed based on a mother wavelet and a decomposition level similar to that in MS6 [41, 42]. It is calculated as shown in Table 2. This set yielded a 16-dimension feature vector per channel.

Table 2. Mathematical definitions of feature extraction methods which were used for EMG pattern recognition. Let x_n represent the EMG data in a segment. a_i is the auto-regressive coefficient. w_n is the white noise error. N denotes the length of the signal. I is the number of segments covering the EMG signal. d_j^p is the WPT coefficient of the subspace W_j^p . j is the decomposition level. p is the index of the subspace occurring at the j^{th} level.

Feature Extraction	Definition	Feature set
AR (a_i)	$\hat{x}_n = -\sum_{i=1}^{p} a_i x_{n-i} + w_n$	1, 3, 5
RMS	$\text{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} x_n^2}$	3, 5
MAV	$\text{MAV} = \frac{1}{N}\sum_{n=1}^{N} \lvert x_n \rvert$	2, 5
WL	$\text{WL} = \sum_{n=1}^{N-1} \lvert x_{n+1} - x_n \rvert$	2, 4, 5
ZC	$\text{ZC} = \sum_{n=1}^{N-1}\left[\operatorname{sgn}(x_n \times x_{n+1}) \cap \lvert x_n - x_{n+1}\rvert \ge threshold\right]$ $\operatorname{sgn}(x) = \begin{cases} 1, & if\ x < 0 \\ 0, & otherwise \end{cases}$	2, 4, 5
SSC	$\text{SSC} = \sum_{n=2}^{N-1}\left[f\left[(x_n - x_{n-1}) \times (x_n - x_{n+1})\right]\right]$ $f(x) = \begin{cases} 1, & if\ x \ge threshold \\ 0, & otherwise \end{cases}$	2, 4, 5
IEMG	$\text{IEMG} = \sum_{n=1}^{N} \lvert x_n \rvert$	4, 5
SKW	$\text{SKW} = \frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^3 \Big/ \left(\frac{1}{N}\sum_{n=1}^{N}(x_n - \bar{x})^2\right)^{3/2}$	5
MAVS	$\text{MAVS}_i = \text{MAV}_{i+1} - \text{MAV}_i;\quad i = 1,\ldots,I-1$	2
VAR	$\text{VAR} = \frac{1}{N-1}\sum_{n=1}^{N} x_n^2$	4
WAMP	$\text{WAMP} = \sum_{n=1}^{N-1} f\left(\lvert x_n - x_{n+1}\rvert\right)$ $f(x) = \begin{cases} 1, & if\ x \ge threshold \\ 0, & \text{otherwise} \end{cases}$	4
RWPE ($P_{j,p}$)	$E_{j,p} = \sum_k \lvert d_j^p(k) \rvert^2$ $P_{j,p} = E_{j,p} \Big/ \sum_p E_{j,p}$	7

Table 3. The number of dimensions (Dim) of the seven multi-feature sets (MS1-MS7) based on four EMG channels.

	MS1	MS2	MS3	MS4	MS5	MS6	MS7
Dim	52	20	20	24	52	1536	64

The first five multi-feature sets are calculated based on the time domain, whereas the last two multi-feature sets are calculated based on a time-scale approach. The dimensions of the seven multi-feature sets based on four EMG channels are summarized in Table 3.

In order to make a fair comparison between the feature projection methods, the dimensions of the projected feature vector were set to $c - 1$, where c denotes the number of classes, because this is the default value and limitation of a supervised method. The number of dimensions after the dimension reduction step is equal to or less than seven for the eight classes considered in this paper. In order to establish the effect of dimension reduction on the ability of the classifier, after finding an optimal feature projection and multi-feature set, the number of dimensions was varied between 1 and 7 and the outcome evaluated for each value.

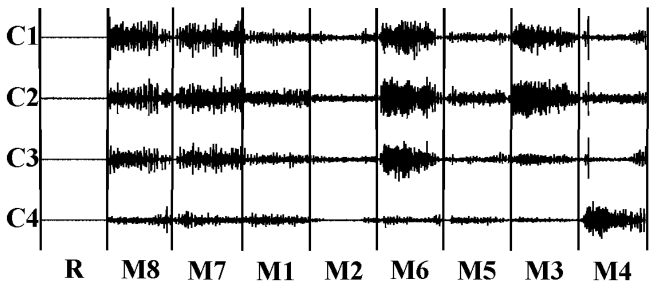


Fig.3. Four-channel surface EMG from 8 upper-limb movements and in the rest state in the time domain. Sample from subject 1.

4. RESULTS AND DISCUSSION

Fig.3 shows data from one data set of the first subject performing eight upper-limb movements recorded through four EMG channels. It can be observed that, particularly in the rest state, the raw EMG signals were less likely to be contaminated by extraneous noise, i.e., movement artifacts and power-line interference [43]. The signal-to-noise ratios (SNRs) were all higher than 20 dB. SNRs are defined as the power ratio between the EMG signal (M1-M8) and the background noise (R). The EMG amplitudes were found to be significantly different according to the movement being performed and were also significantly different based on the EMG channels [44]. Due to the low-level of noise and high SNRs of the interested movements, the rest (R) was not included as a class in the study.

In this experiment, the classification performance of six LDA-based algorithms: classical LDA, ULDA, OLDA, OFNDA, LDA/QR, and AKDA/QR together with that of the BS and PCA were calculated based on seven EMG multi-feature sets. The results are summarized in Table 4, from which, the performance of the different feature projection techniques can be clearly categorized into four groups.

1). The first group includes ULDA, OLDA and OFNDA, which offer the best performance in terms of classification accuracy across different multi-feature sets. The difference in classification accuracy between these techniques is less than 0.5% in all cases. These three techniques consistently outperformed the BS.
2). The second group consists solely of classical LDA, whose classification performance was slightly better than the BS for six of the multi-feature sets although not for the seventh, MS5, and its accuracy was equal to or only slightly lower than that of ULDA, OLDA and OFNDA for five of the multi-feature sets but not for MS5 and MS6.
3). The third group consists of both LDA/QR and ALDA/QR. Both achieved similar results while showing lower performance than that of other LDA techniques.
4). The last group is composed of only PCA which offered the worst performance across all the multi-feature sets.

Moreover, the third and the fourth groups both had lower classification accuracy than the BS.

Based on the seven multi-feature sets and the three feature projection methods in the first group, MS5 produced the highest classification accuracy and much lower computational time than MS6 [45]. MS5 was therefore clearly the optimal multi-feature set, and is recommended to be used in future MCS studies. Based on the results of multi-feature sets MS5 and MS6, the main finding from this experiment is that OFNDA is competitive with ULDA and OLDA in terms of class separability and is much more efficient than both ULDA and OLDA in terms of time complexity during the offline training phase [17].

However, the execution time during the online testing phase for all the techniques is almost equal, as the input only needs to be multiplied by the transformation matrix. The findings of the present study are slightly different to those of Khushaba et al. [17]. In that study OFNDA was found to consistently show 1-2% better recognition accuracy than ULDA and OLDA based on MS5. However, in this study, OFNDA did not always perform better than ULDA and OLDA which showed similarly competent classification results (the difference being less than 0.5%) for different EMG data and EMG multi-feature sets. In fact, the performance of ULDA and OLDA in this study was almost indistinguishable. However, further studies could usefully focus on the analysis of and comparisons between OFNDA and ULDA/OLDA.

The classification performance of ULDA, OLDA and OFNDA is slightly better than that of classical LDA for MS1 and MS7, and is much higher than that of classical LDA for MS5 and MS6 because these techniques are designed to avoid the undersampling problem which LDA entails. Where there is a high degree of redundancy among the features (e.g., MS5 and MS6), then the within-class scatter matrix will be singular, and thus, classical LDA will fail because it does not consider de-correlation of the data.

Moreover, the classification performance of ULDA, OLDA and OFNDA is much higher than that of PCA which projects the original feature set into a new representation with the same number of features as the original feature set. Thus, the use of a smaller number of features as implemented in this study, i.e., $c - 1$ features does not necessitate providing good classification accuracy, as the information in the transformed domain may be dispersed along some of the remaining dimensions. This, however, results in the loss of some information required for classification. This finding is similar to those of previous studies [14, 16]. As found by Englehart et al. [14], the ability of the classifier is reduced if the PCA-reduced

Table 4. Average classification accuracy (and standard deviation) of seven multi-feature sets (MS) with a baseline system, a PCA and six LDA methods from 20 subjects and 4 days (Unit: %). Note that the highest classification accuracy for each MS is highlighted in bold.

MS	BS	PCA	LDA	ULDA	OLDA	OFNDA	LDA/QR	AKDA/QR
1	80.97 (7.31)	48.40 (8.46)	81.86 (6.63)	84.60 (6.30)	**84.61 (6.29)**	84.55 (6.27)	63.89 (7.62)	61.60 (7.11)
2	91.29 (5.70)	86.03 (7.29)	**91.67 (5.52)**	91.64 (5.57)	91.63 (5.59)	91.21 (5.68)	88.50 (6.87)	87.83 (6.81)
3	91.80 (5.24)	74.09 (8.90)	**92.19 (5.03)**	**92.19 (5.06)**	**92.19 (5.04)**	**92.19 (4.90)**	82.19 (7.02)	79.70 (7.25)
4	92.40 (5.52)	88.42 (6.89)	92.76 (5.30)	92.73 (5.29)	**92.77 (5.24)**	92.67 (5.23)	89.60 (6.60)	88.68 (6.53)
5	92.98 (5.15)	82.11 (8.50)	70.49 (14.46)	95.19 (4.11)	95.18 (4.11)	**95.38 (3.89)**	88.74 (6.28)	87.18 (6.53)
6	14.48 (1.57)	15.24 (1.93)	47.87 (8.62)	93.21 (1.95)	93.23 (1.96)	**93.62 (1.94)**	48.53 (4.02)	50.74 (5.16)
7	72.05 (6.61)	61.37 (6.54)	72.92 (6.17)	76.31 (5.73)	**76.32 (5.75)**	76.24 (5.76)	70.09 (5.93)	68.23 (6.00)

Table 5. Average classification accuracy (and standard deviation) of two multi-feature sets (MS5 and MS6) with a baseline system, a PCA, a classical LDA and OLDA methods from 20 subjects and 4 days for different dimensions of feature vectors (Unit: %)

Method	Dim	MS5	Dim	MS6
BS	52	92.98 (5.15)	1536	14.48 (1.57)
PCA	7	82.11 (8.50)	7	15.24 (1.93)
	6	80.11 (8.83)	6	14.86 (1.60)
	5	77.93 (8.78)	5	14.19 (1.47)
	4	74.97 (9.19)	4	13.60 (1.35)
	3	69.19 (9.45)	3	13.29 (1.36)
	2	55.31 (9.63)	2	13.07 (1.14)
	1	36.75 (9.20)	1	12.91 (1.06)
LDA	7	70.49 (14.46)	7	47.87 (8.62)
	6	67.99 (14.62)	6	46.96 (8.44)
	5	65.23 (13.47)	5	45.97 (8.52)
	4	62.01 (13.94)	4	44.27 (8.53)
	3	56.83 (13.49)	3	41.60 (8.70)
	2	49.70 (11.24)	2	37.89 (9.08)
	1	35.90 (6.90)	1	30.64 (9.64)
OLDA	7	95.18 (4.11)	7	93.23 (1.96)
	6	94.11 (4.34)	6	88.69 (3.00)
	5	93.25 (4.67)	5	83.22 (4.09)
	4	91.80 (5.49)	4	75.28 (4.87)
	3	89.14 (6.83)	3	65.88 (5.86)
	2	83.33 (9.19)	2	53.32 (5.95)
	1	59.07 (7.13)	1	37.04 (6.43)

dimension is less than twenty orders. On the other hand, ULDA, OLDA, and OFNDA do not present this problem while trying to quantify the suitability of the reduced feature space. From the results shown in Table 5, it can be seen that the classification accuracy of OLDA is still higher than that of the BS at the five-dimension feature vector which is therefore a sufficient level to provide accurate classification.

One interesting result is in the case of MS6 where the BS (using raw wavelet coefficients) is very low at 14.48%. On the other hand, after applying ULDA, OLDA or OFNDA, the classification accuracy increased to approximately 93%. This is because the raw wavelet coefficients contain both meaningful and unwanted information. In future work, noise or unwanted information should be removed before performing the classification task [22, 23, 46]. Hence, it is necessary to apply the feature projection technique for the time-scale feature vectors DWT and WPT. This investigation clearly confirmed the rapid development of the feature projection technique for MCS based on the time-scale approach [8-19, 47] and time-scale features should not be used without implementing the feature projection technique.

Another interesting result is in the case of MS5, where the classification accuracy of the OFNDA-reduced feature set is slightly higher than that of the original feature set (BS), and is consistent with other time domain feature sets. For online applications, the feature projection technique can in fact reduce the time needed to classify unseen patterns, and this becomes more apparent in problems with a large number of EMG channels and EMG time-domain features.

5. CONCLUSIONS AND FUTURE RESEARCH

Based on the classification results with a linear discriminant classifier, and a 4-channel, 8-movement EMG system, ULDA, OLDA and OFNDA are suitable for use as dimensionality reduction techniques in MCS. They not only reduce the computational complexity but also increase the classification accuracy. These techniques produce better classification performance for both the time domain and time-scale feature methods. Among the seven multi-feature sets considered in previous studies, the optimal feature vector which provides the highest classification accuracy consists of eight time domain methods: RMS, MAV, IEMG, WL, SKW, ZC, SSC, and 6th-order AR. In addition, the dimensions of the feature vector can be reduced to five without loss of classification accuracy.

ACKNOWLEDGMENT

This work was supported in part by the Thailand Research Fund through the Royal Golden Jubilee Ph.D. Program (Grant No. PHD/0110/2550), and by NECTEC-PSU Center of Excellence for Rehabilitation Engineering and Faculty of Engineering, Prince of Songkla University.

REFERENCES

[1] Merletti, R., Hermens, H. (2004). Detection and conditioning of the surface EMG signal. In Merletti R., Parker, P. (eds.) *Electromyography: Physiology, Engineering, and Noninvasive Applications.* New Jersey: John Wiley & Sons, 107-132.

[2] Oskoei, M.A., Hu, H. (2007). Myoelectric control systems—A survey. *Biomed. Signal Process. Control*, 2 (4), 275-294.

[3] Parker, P.A., Englehart, K.B., Hudgins, B.S. (2004). Control of powered upper limb prostheses. In Merletti, R., Parker, P. (eds.) *Electromyography: Physiology, Engineering, and Noninvasive Applications.* New Jersey: John Wiley & Sons, 453-476.

[4] Phinyomark, A., Phukpattaranont, P., Limsakul, C. (2011). A review of control methods for electric power wheelchairs based on electromyography (EMG) signals with special emphasis on pattern recognition. *IETE Technical Review*, 28 (4), 316-326.

[5] Jacobson, S.C., Knutti, D.F., Johnson, R.T., Shears, H.H. (1982). Development of the Utah artificial arm. *IEEE Trans. Biomed. Eng.*, 29 (4), 249-269.

[6] Hudgins, B., Parker, P., Scott, R.N. (1993). A new strategy for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.*, 40 (1), 82-94.

[7] Light, C.M., Chappell, P.H. (2000). Development of a lightweight and adaptable multiple-axis hand prosthesis. *Med. Eng. Phys.*, 22 (10), 679-684.

[8] Englehart, K., Hudgin, B., Parker, P.A., Stevenson, M. (1999). Classification of the myoelectric signal using time-frequency based representations. *Med. Eng. Phys.*, 21 (6-7), 431-438.

[9] Khezri, M., Jahed, M. (2009). An exploratory study to design a novel hand movement identification system. *Comput. Biol. Med.*, 39 (5), 433-442.

[10] Artemiadis, P.K., Kyriakopoulos, K.J. (2010). EMG-based control of a robot arm using low-dimensional embeddings. *IEEE Trans. Robotics*, 26 (2), 393-398.

[11] Rivera-Alvidrez, Z., Kalmar, R.S., Ryu, S.I., Shenoy, K.V. (2010). Low-dimensional neural features predict muscle EMG signals. In *Annual Internal Conference of the IEEE Engineering in Medicine and Biology Society*, 31 Aug. – 4 Sept. 2010. IEEE, 6027-6033.

[12] Khushaba, R.N., Al-Jumaily, A., Al-Ani, A. (2007). Novel feature extraction method based on fuzzy entropy and wavelet packet transform for myoelectric control. In *ISCIT '07 : International Symposium on Communications and Information Technologies*, 17-19 October 2007. IEEE, 352-357.

[13] 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 Trans. Biomed. Eng.*, 53 (11), 2232-2239.

[14] Englehart, K., Hudgin, B., Parker, P.A. (2001). A wavelet-based continuous classification scheme for multifunction myoelectric control. *IEEE Trans. Biomed. Eng.*, 48 (3), 302-311.

[15] Chu, J.–U., Moon, I., Lee, Y.–J., Kim, S.–K., Mun, M.–S. (2007). A supervised feature-projection-based real-time EMG pattern recognition for multifunction myoelectric hand control. *IEEE-ASME Trans. Mechatron.*, 12 (3), 282-290.

[16] Chan, A.D.C., Green, G.C. (2007). Myoelectric control development toolbox. In *30th Conference of the Canadian Medical & Biological Engineering Society*, 16-19 June 2007. CMBES, M0100.

[17] Khushaba, R.N., Al-Ani, A., Al-Jumaily, A. (2010). Orthogonal fuzzy neighborhood discriminant analysis for multifunction myoelectric hand control. *IEEE Trans. Biomed. Eng.*, 57 (6), 1410-1419.

[18] Liu, Y.H., Huang, H.P., Weng, C.H. (2007). Recognition of electromyographic signals using cascaded kernel learning machine. *IEEE-ASME Trans. Mechatron.*, 12 (3), 253-264.

[19] Khushaba, R.N., Al-Jumaily, A., Al-Ani, A. (2009). Evolutionary fuzzy discriminant analysis feature projection technique in myoelectric control. *Pattern Recognition Lett.*, 30 (7), 699-707.

[20] Khezri, M., Jahed, M. (2011). A neuro–fuzzy inference system for sEMG-based identification of hand motion commands. *IEEE Trans. Ind. Electron.*, 58 (5), 1952-1960.

[21] Rafiee, J., Rafiee, M.A., Yavari, F., Schoen, M.P. (2011). Feature extraction of forearm EMG signals for prosthetics. *Expert Syst. Appl.*, 38 (4), 4058-4067.

[22] Phinyomark, A., Limsakul, C., Phukpattaranont, P. (2011). Application of wavelet analysis in EMG feature extraction for pattern classification. *Meas. Sci. Rev.*, 11 (2), 45-52.

[23] Phinyomark, A., Nuidod, A., Phukpattaranont, P. Limsakul, C. (2012). Feature extraction and reduction of wavelet transform coefficients for EMG pattern classification. *Electronics and Electrical Engineering*, 122 (6), in press.

[24] Fukunaka, K. (1990). *Introduction to Statistical Pattern Recognition*. Academic Press.

[25] Schafer, K.C., Balog, J., Szaniszlo, T., Szalay, D., Mezey, G., Denes, J., Bognar, L., Oertel, M., Takats, Z. (2011). Real time analysis of brain tissue by direct combination of ultrasonic surgical aspiration and sonic spray mass spectrometry. *Analytical Chemistry*, 83 (20) 7729-7735.

[26] Belhumeour, P.N., Hespanha, J.P., Kriegman, D.J. (1997). Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Trans. Patt. Anal. Mach. Int.*, 19 (7), 711-720.

[27] Sakai, M., Kitaoka, N., Takeda, K. (2010). Acoustic feature transformation based on discriminant analysis preserving local structure for speech recognition. *IEICE Trans. Inform. Syst.*, E93D (5), 1244-1252.

[28] Howland, P., Jeon, M., Park, H. (2003). Structure preserving dimension reduction for clustered text data based on the generalized singular value decomposition. *SIAM J. Matrix Anal. Applicat.*, 25 (1), 165-179.

[29] Jin, Z., Yang, J.Y., Tang, Z.M., Hu, Z.S. (2001). A theorem on the uncorrelated optimal discriminant vectors. *Patt. Recog.*, 34 (10), 2041-2047.

[30] Ye, J. (2005). Characterization of a family of algorithms for generalized discriminant analysis on undersampled problems. *Journal of Machine Learning Research*, 6, 483-502.

[31] Ye, J., Li, Q. (2004). LDA/QR: An efficient and effective dimension reduction algorithm and its theoretical foundation. *Patt. Recog.*, 37 (4), 851-854.

[32] Krzanowski, W.J., Jonathan, P., McCarthy, W.V., Thomas, M.R. (1995). Discriminant analysis with singular covariance matrices: Methods and applications to spectroscopic data. *J. Roy. Statist. Soc. Ser. C-Appl. Stat*., 44 (1), 101-115.

[33] Baudat, G., Anouar, F. (2000). Generalized discriminant analysis using a kernel approach. *Neural Comput.*, 12 (10), 2385-2404.

[34] Tao, X., Ye, J., Li, Q., Janardan, R., Cherkassky, V. (2004). Efficient Kernel discriminant analysis via QR decomposition. In *NIPS 2004 : 18th Annual Conference on Neural Information Processing Systems*, 13-18 December 2004. NIPS, 1529-1536.

[35] Phinyomark, A., Phukpattaranont, P., Limsakul, C., Phothisonothai, M. (2011). Electromyography (EMG) signal classification based on detrended fluctuation analysis. *Fluctuation and Noise Lett.*, 10 (3), 281-301.

[36] Oskoei, M.A., Hu, H. (2008). Support vector machine-based classification scheme for myoelectric control applied to upper limb. *IEEE Trans. Biomed. Eng.*, 55 (8), 1956-1965.

[37] Kaufmann, P., Englehart, K., Platzner, M. (2010). Fluctuating EMG signals: Investigating long-term effects of pattern matching algorithms. In *Annual Internal Conference of the IEEE Engineering in Medicine and Biology Society*, 31 Aug. – 4 Sept. 2010. IEEE, 6357-6360.

[38] Geethanjali, P., Ray, K.K. (2011). Identification of motion from multi-channel EMG signals for control of prosthesis hand. *Australas. Phys. Eng. Sci. Med.*, 34 (3), 419-427.

[39] Du, Y.-C., Lin, C.-H., Shyu, L.-Y., Chen, T. (2010). Portable hand motion classifier for multi-channel surface electromyography recognition using grey relational analysis. *Expert Syst. Appl.*, 37 (6), 4283-4291.

[40] Khushaba, R.N., Al-Ani, A., Al-Jumaily, A. (2010). Orthogonal fuzzy neighborhood discriminant analysis for multifunction myoelectric hand control. *IEEE Trans. Biomed. Eng.*, 57 (6), 1410-1419.

[41] Yan, Z., Wang, Z., Xie, H. (2008). Joint application of rough set-based feature reduction and fuzzy LS-SVM classifier in motion classification. *Med. Biol. Eng. Comput.*, 46 (6), 519-527.

[42] Yan, Z., Wang, Z., Xie, H. (2008). The application of mutual information-based feature selection and fuzzy LS-SVM-based classifier in motion classification. *Comput. Method. Program. Biomed.*, 90 (3), 275-284.

[43] Clancy, E.A., Morin, E.L., Merletti, R. (2002). Sampling, noise-reduction and amplitude estimation issues in surface electromyography. *J. Electromyograph. Kinesiol.*, 12 (1), 1-16.

[44] Naik, G.R., Kumar, D.K., Arjunan, S.P. (2010). Pattern classification of Myoelectrical signal during different Maximum Voluntary Contractions: A study using BSS techniques. *Meas. Sci. Rev.*, 10 (1), 1-6.

[45] Boostani, R., Moradi, M.H. (2003). Evaluation of the forearm EMG signal features for the control of a prosthetic hand. *Physiol. Meas.*, 24 (2), 309-319.

[46] Phinyomark, A., Phukpattaranont, P., Limsakul, C. (2011). Wavelet-based denoising algorithm for robust EMG pattern recognition. *Fluctuation and Noise Lett.*, 10 (2), 157-167.

[47] Hussain, M.S., Mamun, Md. (2012). Effectiveness of the wavelet transform on the surface EMG to understand the muscle fatigue during walk. *Meas. Sci. Rev.*, 12 (1), 28-33.

Received December 2, 2011.
Accepted May 31, 2012.