35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Two-Channel Surface Electromyography for Individual and Combined Finger Movements

Khairul Anam, *Student Member, IEEE* , Rami N Khushaba and Adel Al-Jumaily, *Member, IEEE*

Abstract— **This paper proposes the pattern recognition system for individual and combined finger movements by using two channel electromyography (EMG) signals. The proposed system employs Spectral Regression Discriminant Analysis (SRDA) for dimensionality reduction, Extreme Learning Machine (ELM) for classification and the majority vote for the classification smoothness. The advantage of the SRDA is its speed which is faster than original LDA so that it could deal with multiple features. In addition, the use of ELM which is fast and has similar classification performance to well-known SVM empowers the classification system. The experimental results show that the proposed system was able to recognize the individual and combined fingers movements with up to 98 % classification accuracy by using only just two EMG channels.**

I. INTRODUCTION

Myoelectric signals from surface electromyography (sEMG) signals have been used to control hand prosthetics for years. Many attempts have been done to achieve a dexterous control of hand prosthesis. Mostly, those efforts focus on hand movements [1-3]. In fact, to achieve a more dexterous prosthetic device, finger movements should be included in the control system [4, 5].

Tenore et al have successfully classified ten individual finger movements (flexion and extension treated as two movements) with more than 90% accuracy by using 32 sEMG channels[4]. Moreover, Al-Timemy et al used 16 channels for classifying individual finger movements and achieved 96% accuracy for 9 class finger movements[6]. The recent works focus on the use of as few channels as possible. Tsenov et al used two sEMG channels for 5 class finger movements i.e. the thumb, pointer, middle finger and hand closure. The accuracy was nearly 93 % [7]. The result was good but it only worked on limited fingers movements.

Khushaba et al [5] used two EMG channels to recognize 10 classes of fingers movements which consisted of five individual finger movements, four finger combinations and a hand closure. The recognition system extracting a number of time domain features and employing Linear Discriminant Analysis (LDA) for dimensionality reduction, Support Vector Machine (SVM) for classification and Bayesian vote for post-processing achieved 92 % accuracy. This result was the best accuracy produced so far for system with two EMG channels for individuated and combined fingers movements.

The main drawback of LDA which was used in Khushaba et al [8] is the costly time processing. As a consequence, it is impossible to improve the accuracy by extracting more features because it will not be feasible for real time application. In reality, the feature extraction is one of the most important factor in the classification process [9].

As in LDA, the multiclass classification using SVM is time consuming. One-against-all (OAL)-SVM [10] for instance, consists of *m* SVMs for *m* classes. As a result, *m* learning times were needed to do multiclass classification. Therefore, it will save much time if the learning process is done by only one machine learning instead of m learning machines.

To improve both the accuracy and processing time, this paper proposes a new recognition system for individual and combined finger movements. The proposed work used Spectral Regression Discriminant Analysis (SRDA)[8], the extension of LDA which is able to work on a large dataset. In addition, it has faster processing than LDA [8] and projects a large number of features to m-1 features where m is the number of classes. Because of its capability, more features can be extracted and projected to small number of features.

In the proposed system, the Extreme Learning Machine (ELM)[11] was used for classification instead of SVM. ELM is "generalized" single-hidden-layer feedforward networks (SLFNs) whose hidden layer does not need to be tuned. It needs fewer optimization constraint, has better generalization functioning and faster learning time than SVM[11]. This combination, SRDA and ELM along with the majority vote[12], provided a fast and good classification system for individuated and combined finger movements.

II. METHODS

A. Proposed Method

The proposed recognition system consisted of several stages. Firstly, the EMG signals were acquired by data acquisition device. The filtering and windowing were applied to the collected data before being extracted by using a time domain feature set. To reduce the dimension of the features, SDRA was employed. Then, the reduced data were classified using ELM and refined by using the majority vote.

B. Data Collection

The data in this work were acquired from eight subjects, two females and six males. All subjects, which were aged

Khairul Anam is with the University of Jember, Indonesia. Now he is PhD student at University of Technology Sydney, P.O. Box 123 Broadway, NSW 2007 Australia (Phone: (+61)295147962; fax: - ; e-mail: Khairul.Anam@student.uts.edu.au).

Rami N Khushaba and Adel Al-Jumaily is with School of Electrical, Mechanical and Mechatronic Systems, University of Technology, Sydney (e-mail: Rami.Khushaba@uts.edu.au and Adel.Al-Jumaily@uts.edu.au).

between 20 and 35 years, were normally limbed with no muscle disorder. To avoid the effect of position movement on EMG signals, subject's arm was supported and fixed at certain position[5].

Two Delsys DE 2.x series EMG sensors were used to collect EMG data. Next, the data were filtered and displayed by using Bagnoli Desktop EMG Systems from Delsys Inc. The signal sources were obtained from two electrodes with a conductive adhesive reference electrode placed in the wrist. The electrodes placement is shown in fig. 1. Channel 1 (left) captured the signals mainly from the Extensor carpi ulnaris and Extensor digiti minimi muscles whereas the channel 2 (right) from the Flexor digitorum superficials and palmaris longus muscles.

The EMG signals acquired were amplified to a total gain of 1000 and sampled by using a 12-bit analog-to-digital converter (National Instruments, BNC-2090) at 4000 Hz. The collected EMG signals were filtered by a bandpass filter between 20 and 500 Hz with a notch filter to remove the 50 Hz line interference. Finally, the EMG signals were down sampled to 1000 Hz.

Fig. 2 shows ten classes of the individual and combined finger movements consisting of the flexion of each of the individuated fingers, i.e., Thumb (T), Index (I), Middle (M), Ring (R), Little (L) and the pinching of combined Thumb–Index (T–I), Thumb–Middle (T–M), Thumb–Ring (T–R), Thumb–Little (T–L), and eventually the hand close (HC). In the experiment, the subjects performed a finger posture which was started from a relaxation state and then followed by holding certain posture for a period of 5 s. The subject repeated the same movement six times with 3 to 5 s resting period between trials. The collected data from six trials were divided into two groups, the training data and the testing data. The four trials were used as the training data and the remaining trials were as test data.

C. Features extraction

Various sets of time domain features were used to avoid a high computational complexity[13]. There are six time domain features involved, i.e. Slope Sign Changes (SSC), Number of Zero Crossings (ZC), Waveform Length (WL), Hjorth Time Domain Parameters (HTD), Sample Skewness (SS), Auto Regressive (AR) Model Parameters. The order of AR model was varied to investigate the accuracy and the time processing. These features were extracted by using myolectric toolbox [12]and Biosig toolbox [14].

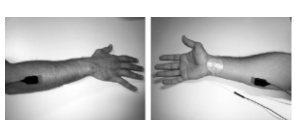

Figure 1. The placement of the electrodes on the right hand

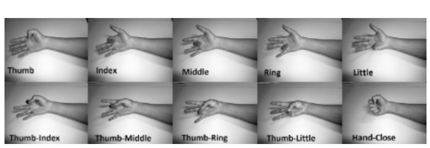

Figure 2. Different finger movements

The AR model parameters have been proven stable and robust to the electrode location shift and the change of signal level [13]. For this reason, the order of AR model was varied to achieve good accuracy without compromising the time processing.

Moreover, aforementioned time domain features were windowed by using disjoint window instead of sliding window to keep computational cost low. A 100 ms window length and a 100 ms window increment were used to form a system which is suitable for real time application.

D. SRDA for dimensionality reduction

SRDA is an improvement of LDA which is better than LDA in the computational aspect and the ability to cope with a large dataset[8]. Let eigen problem of LDA is

$$\bar{\mathbf{X}}\mathbf{W}\bar{\mathbf{X}}^{\mathrm{T}}\mathbf{a} = \lambda\bar{\mathbf{X}}\bar{\mathbf{X}}^{\mathrm{T}}\mathbf{a} \quad (1)$$

where $\bar{X}$ (1 x c) is a centered data matrix, W is an eigenvector matrix (m x m), λ = an eigenvalue, a = a transformation vector, c = the number of classes, and m = the number of total training data points. Modification of the equation (1) gives:

$$\mathbf{W}\bar{\mathbf{y}} = \lambda\bar{\mathbf{y}} \quad (2)$$

where $$\bar{\mathbf{X}}^{\mathrm{T}}\mathbf{a} = \bar{\mathbf{y}} \quad (3)$$

The solution of LDA problem by SRDA is to get y by solving eq (2) and then use the y obtained to find a. To solve a, the least square sense could be employed by using:

$$a = \arg\min_{a} \sum_{i=1}^{m} \left(a^T\bar{x}_i - \bar{y}_i\right)^2 \quad (4)$$

By regularizing least square problem of SRDA, we get:

$$a = \arg\min_{a} \sum_{i=1}^{m} \left(\left(\bar{\mathbf{X}}^{\mathrm{T}}\mathbf{a} - \bar{\mathbf{y}}\right)^{\mathrm{T}}\left(\bar{\mathbf{X}}^{\mathrm{T}}\mathbf{a} - \bar{\mathbf{y}}\right) + \alpha\mathbf{a}^{\mathrm{T}}\mathbf{a}\right) \quad (5)$$

Derivative of equation (5) gives:

$$\left(\bar{\mathbf{X}}\bar{\mathbf{X}}^{\mathrm{T}} + \alpha\mathbf{I}\right) = \bar{\mathbf{X}}\bar{\mathbf{y}}$$

$$\Rightarrow \mathbf{a} = \left(\bar{\mathbf{X}}\bar{\mathbf{X}}^{T} + \alpha\mathbf{I}\right)^{-1}\bar{\mathbf{X}}\bar{\mathbf{y}} \quad (6)$$

E. Extreme Learning Machine

ELM is a learning scheme for single layer feedforward networks (SLFNs). While the network parameters are tuned in classical SLFNs learning algorithms, most of these parameters are analytically determined in ELM. The hidden parameters can be independently determined from the training data, and the output parameters can be determined by pseudo-inverse method using the training data. As a result, the learning of ELM can be carried out extremely fast compared to the other learning algorithms[11].

The output function of ELM for generalized SLFNs (for one output node case) is:

$$f_L(x) = \sum_{i=1}^{L}\beta_i h_i(x) = \mathbf{h}(\mathbf{x})\boldsymbol{\beta} \quad (7)$$

where $\boldsymbol{\beta} = \left[\beta_1, ..., \beta_L\right]^T$ is the vector of the output weight between hidden layer of L nodes and the output node, $\mathbf{h}(\mathbf{x}) = \left[h_1(x), ..., h_L(x)\right]$ is the output vector of hidden layer.

The Objective of ELM is to minimize the error and the norm of weight:

$$\text{Minimize}: \left\|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\right\|^2 \text{ and } \left\|\boldsymbol{\beta}\right\| \tag{8}$$

where **T** is the target. For classification purpose, the output function of ELM in equation (7) could be modified to be:

$$f(x) = \mathbf{h(x)}\boldsymbol{\beta} = \mathbf{h(x)H}^T\left(\frac{1}{C} + \mathbf{HH}^T\right)^{-1}\mathrm{T} \tag{9}$$

where

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(x_1) \\ \vdots \\ \mathbf{h}(x_N) \end{bmatrix} = \begin{bmatrix} h_1(x_1) & \cdots & h_L(x_1) \\ \vdots & \vdots & \vdots \\ h_1(x_N) & \cdots & h_L(x_N) \end{bmatrix} \tag{10}$$

as well as C is a user-specified parameter and N is the number of the training data. In the equation (10), h(x) is a feature mapping (hidden layer output vector) which can be :

$$h(x) = \left[G(a_1, b_1, x), \ldots, G(a_L, b_L, x)\right] \tag{11}$$

where G is a non-linier piecewise continuous function such as sigmoid, hard limit, Gaussian, and multi quadratic function.

If the feature mapping h(x) is unknown to the user, a kernel function can be used to represent h(x). Then, the equation (9) would be:

$$\begin{aligned} f(x) &= \mathbf{h(x)H}^T\left(\frac{1}{C} + \mathbf{HH}^T\right)^{-1}\mathrm{T} \\ &= \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}\left(\frac{1}{C} + \Omega_{ELM}\right)^{-1}\mathrm{T} \end{aligned} \tag{12}$$

where

$\Omega_{ELM} = \mathbf{HH}^T : \Omega_{ELMi,j} = h(x_i).h(x_j) = K(x_i, x_j)$ and K is a kernel function such that :

$$K(\mathbf{u}, \mathbf{v}) = \exp\left(-\gamma\left\|\mathbf{u} - \mathbf{v}\right\|^2\right) \tag{13}$$

F. Post-processing

The majority vote was used to refine the classification results. It utilizes the results from the present state and *n* previous states and makes a new classification result based on the class which appears most frequent. This procedure produces the finger movement class that removes specious misclassification. Besides majority vote, the transition states in the classification results are removed too. This method gives the recognition system that works in steady state only regardless of the transition state.

III. Result and Discussion

To investigate the finger movement recognition performances by using SRDA, ELM and the majority vote, some experiments were done on a 2.8 Ghz intel core i7 based with 4 GB RAM. Three experiments were performed to determine the optimal parameters of ELM, to compare the performance of LDA and SRDA and to investigate the performance of the proposed system in recognizing the individual and combined finger movements. All experiments were done on the data from 6 experiments across eight subjects.

The first experiment was to determine the optimal ELM parameters. This work employed the Gaussian kernel based ELM with two importance parameters, C and γ as showed in equation 9 and 12. The (C, γ) were decided by selecting a possible interval of C or γ in a grid space. Then, all grid points of (C, γ) were applied to the ELM to find the one which gave the minimum accuracy error. Three-fold cross validation was performed to get the accuracy error rates across eight subjects.

Fig. 3 (upper) presents the accuracy error rates of the ELM by varying the C parameters while γ was constant at 2^{-5}. Based on the graph, the $C=2^0$ gives the minimum accuracy error by regarding the accuracy error around it therefore it was chosen as the optimal parameter of C. By using the obtained C value, the γ was varied and then determined based on the accuracy error rates as described in Fig. 3 (lower). From the graph, the $\gamma=2^{-5}$ was selected to be an optimal parameter of ELM along with $C=2^0$.

The second experiment was the comparison of LDA and SRDA performance in projecting the features for classification process. Regularized LDA (RLDA)[8] replacing LDA method is used to cope with the singularity problem. SRDA was applied by using equation (6) with $\alpha=10^{-5}$. The original features were extracted from time domain features as mentioned in section 2.C. To get more features, the order of AR model was varied from 2 to 100. The experimental results from eight subjects are presented in table 1.

Table 1 shows that SRDA is faster than RLDA and becomes faster with increasing the amount of the training data up to nearly three times when AR model of order 100. The increasing data was able to improve the accuracy but increase the time processing. Fortunately, the slow process of LDA could be overcome by using SRDA. Even though the amount of the features are increased, the time processing is still fast compared to LDA. The best accuracy obtained is 98.45 % which takes 14.3 ms only.

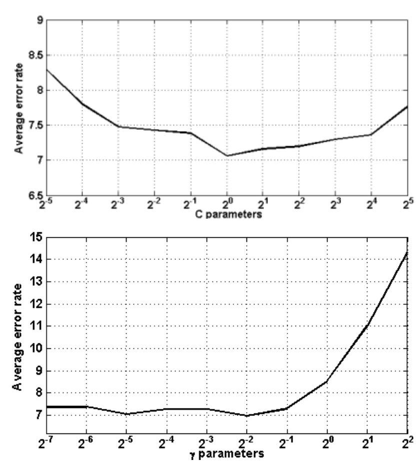

Figure 3. The accuracy error rates for $\gamma=2^{-5}$ (upper) and $C=2^0$ (lower)

TABLE I. LDA AND SRDA COMPARISON ACROSS EIGHT SUBJECTS

AR order	Average Classification Accuracy (%)		Average Time (ms)	
	RLDA	SRDA	RLDA	SRDA
2	89.89 ±8.56	91.23 ±7.56	1.40 ±0.27	1.30 ±0.37
4	94.70 ±4.53	93.68 ±6.27	2.30 ±0.41	1.60 ±0.37
8	94.50 ±5.56	94.66 ±5.47	3.00 ±0.09	1.80 ±0.42
16	95.73 ±5.08	96.45 ±4.69	4.50 ±0.46	2.60 ±1.10
32	97.20 ±3.66	97.57 ±2.78	8.10 ±0.62	3.60 ±0.69
64	97.48 ±2.94	97.71 ±3.44	19.80 ±1.40	7.30 ±0.98
100	98.48 ±2.17	98.45 ±2.64	41.20 ±3.30	14.30 ±3.30

The last experiment is the performance test of the combination of SRDA and optimized ELM along with the majority vote for pattern recognition of finger movements by using all parameters produced in the previous sections. They are 100 order of AR model, $\alpha=10^{-5}$ for SRDA, and ($\gamma=2^{-5}$ & $C=2^{0}$) for ELM. The experiment was done to eight subjects with 100 ms window length, 100 ms increment and 9 vote decisions[5]. The confusion matrix of the classification accuracy across eight subjects is presented in table 2.

TABLE II. THE CONFUSION MATRIX OF THE CLASSIFICATION RESULTS AVERAGED FOR EIGHT SUBJECTS (UNITS : %)

Classified task		Intended task									
		T	I	M	R	L	T-I	T-M	T-R	T-L	HC
	T	99.5	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	I	0.0	97.1	0.0	0.0	0.0	2.9	0.0	0.0	0.0	0.0
	M	0.0	0.0	100	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	R	0.0	0.0	0.0	99.3	0.0	0.0	0.0	0.7	0.0	0.0
	L	0.0	0.0	0.0	0.0	99.6	0.0	0.0	0.0	0.4	0.0
	T-I	1.3	0.7	0.0	0.0	0.0	98.0	0.0	0.0	0.0	0.0
	T-M	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0	0.0
	T-R	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100	0.0	0.0
	T-L	0.0	0.0	0.0	0.0	8.6	0.0	0.0	0.0	91.4	0.0
	HC	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	0.0	99.5

Table 2 shows that three classes, the M, T-M, and T-R are perfectly classified. On the other hand, the rest are slightly misclassified to another class. As an example the thumb-little (T-L) is misclassified to the little movement (L) by 8.6 % . In addition to the confusion matrix, the diagonal of the confusion matrix is presented in fig. 4. It can be seen from the figure that on average, the system was able to recognize the different fingers movements with classification accuracy 98.45 %. However, there were high variation in recognizing the index (I), thumb-index (T-I), and thumb-little (T-L) fingers movements across eight subjects. Especially for the T-L, it was the most difficult movement with wide range of classification accuracy from around 75% to 100 %.

IV. CONCLUSION

The proposed system which consists of the SRDA, the optimized ELM and the majority vote was able to recognize the individual and combined finger movements with the classification accuracy of 98.45 % by using only two EMG channels. The use of SRDA which is nearly three time faster than RLDA gives chance to add more features in order to increase the classification accuracy with reasonable processing time. However, there were some difficulties in recognizing some classes especially the thumb-little finger movement.

REFERENCES

[1]. Oskoei, M.A. and H. Huosheng, "Support Vector Machine-Based Classification Scheme for Myoelectric Control Applied to Upper Limb," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1956-1965, 2008.

[2]. Sang Wook, L., et al., "Subject-Specific Myoelectric Pattern Classification of Functional Hand Movements for Stroke Survivors," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 19, no. 5, pp. 558-566, 2011.

[3]. Micera, S., J. Carpaneto, and S. Raspopovic, "Control of Hand Prostheses Using Peripheral Information," *IEEE Reviews in Biomedical Engineering*, vol. 3, no., pp. 48-68, 2010.

[4]. Tenore, F.V.G., et al., "Decoding of individuated finger movements using surface electromyography," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 5, pp. 1427-1434, 2009.

[5]. Khushaba, R.N., et al., "Toward improved control of prosthetic fingers using surface electromyogram (EMG) signals," *Expert Systems with Applications*, no. 0, 2012.

[6]. Al-Timemy, A., et al., "Finger Movements Classification for the Dexterous Control of Upper Limb Prosthesis Using EMG Signals," *Advances in Autonomous Robotics*, no., pp. 434-435, 2012.

[7]. Tsenov, G., et al. *Neural Networks for Online Classification of Hand and Finger Movements Using Surface EMG signals*. in *8th Seminar on Neural Network Applications in Electrical Engineering (NEUREL)*. 2006.

[8]. Cai, D., X. He, and J. Han, "SRDA: An efficient algorithm for large-scale discriminant analysis," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 1, pp. 1-12, 2008.

[9]. Phinyomark, A., P. Phukpattaranont, and C. Limsakul, "Feature reduction and selection for EMG signal classification," *Expert Systems with Applications*, vol. 39, no. 8, pp. 7420-7431, 2012.

[10]. Hsu, C.W. and C.J. Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415-425, 2002.

[11]. Huang, G.B., et al., "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513-529, 2012.

[12]. Chan, A.D.C. and G.C. Green. *Myoelectric control development toolbox*. in *Proceedings of 30th Conference of the Canadian Medical & Biological Engineering Society*. 2007.[Online]. Availabel at http://www.sce.carleton.ca/faculty/chan/matlab.html.

[13]. Tkach, D., H. Huang, and T.A. Kuiken, "Study of stability of time-domain features for electromyographic pattern recognition," *Journal of NeuroEngineering and Rehabilitation*, vol. 7, no. 1, pp. 21, 2010.

[14]. Schlogl, A. and C. Brunner, "BioSig: a free and open source software library for BCI research," *Computer*, vol. 41, no. 10, pp. 44-50, 2008.[Online]. Available at : http://biosig.sourceforge.net/.

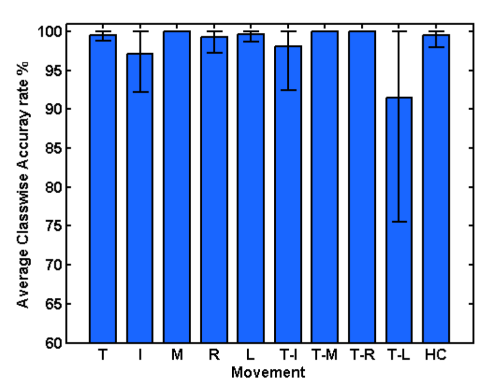

Figure 4. Average diagonal of the confusion matrix achieved accros eight subjects using the optimized ELM