Journal of Zhejiang University SCIENCE A
ISSN 1673-565X (Print); ISSN 1862-1775 (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

Joint application of feature extraction based on EMD-AR strategy and multi-class classifier based on LS-SVM in EMG motion classification*

YAN Zhi-guo†, WANG Zhi-zhong†‡, REN Xiao-mei
(*Departmen of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200030, China*)
†E-mail: hengdaoxiao@sjtu.org; zzwang@sjtu.edu.cn
Received Nov. 8, 2006; revision accepted Mar. 23, 2007

Abstract: This paper presents an effective and efficient combination of feature extraction and multi-class classifier for motion classification by analyzing the surface electromyografic (sEMG) signals. In contrast to the existing methods, considering the non-stationary and nonlinear characteristics of EMG signals, to get the more separable feature set, we introduce the empirical mode decomposition (EMD) to decompose the original EMG signals into several intrinsic mode functions (IMFs) and then compute the coefficients of autoregressive models of each IMF to form the feature set. Based on the least squares support vector machines (LS-SVMs), the multi-class classifier is designed and constructed to classify various motions. The results of contrastive experiments showed that the accuracy of motion recognition is improved with the described classification scheme. Furthermore, compared with other classifiers using different features, the excellent performance indicated the potential of the SVM techniques embedding the EMD-AR kernel in motion classification.

Key words: Electromyografic signal, Empirical mode decomposition (EMD), Auto-regression model, Wavelet packet transform, Least squares support vector machines (LS-SVM), Neural network
doi:10.1631/jzus.2007.A1246 **Document code:** A **CLC number:** TN911.7; R318.04

INTRODUCTION

Recently, bio-signals have been paid much attention due to their potential in providing convenient control channels between the disabled and rehabilitation engineering (Han *et al.*, 2000; Soares *et al.*, 2003; Sebelius *et al.*, 2005; Reddy and Gupta, 2007; Su *et al.*, 2007). By analyzing the EMG signals collected from remnant muscles, as control inputs for the artificial limb, the intelligent artificial limbs can implement the corresponding functions. Recognizing bio-signals correctly is the preliminary stage of controlling these prosthesis devices.

As far as motion classification is concerned, like all other classification problems, there exist two fundamental issues. One is how to extract the most representative feature set. In many cases, an effective set of features used in one classification task may not work as well when used in other classifications. This problem is subject-dependent. The other is how to design and construct the powerful capability classifier. Relatively, the feature extraction methods are responsible mainly for the classification accuracy rather than the classifiers.

In previous work, various approaches of extracting features of bio-signals were employed for discerning predefined human motions. In (Hudgins *et al.*, 1993; Englehart *et al.*, 1999; Englehart and Hudgins, 2003), Hudgins *et al.* adopted the classical time-domain indexes, including the Mean Absolute Value (MAV), the Mean Absolute Value Slope (MAVS), the Zero Crossings (ZC), the Slope Sign Changes (SSC)

‡ Corresponding author
* Project (No. 2005CB724303) supported by the National Basic Research Program (973) of China

and the Waveform Length (WL), etc. to form the feature set. In the above-mentioned work, researchers extracted the feature set through non-parametric methods. In the field of physiological signal processing, there are many parameter/modeling methods to exploit the feature set. Among these methods, the autoregressive model is a kind of convenient and widely used algorithm. References (Kang *et al.*, 1995; Soares *et al.*, 2003; Hu and Nenov, 2004) discussed some techniques based on the AR model and proposed the optimal selection of the order of AR model.

Those algorithms based on time domain for extracting feature set are very convenient, because they do not require any prior knowledge. Furthermore, the time expense for computing is very low and suitable for real-time recognition and classification. Nevertheless, the absence of the frequency resolution obscured the subtle presentation of the complicated EMG signals. Similarly, those feature extraction algorithm based fully on the frequency domain, such as the Fourier Transform (FT), cannot accurately represent the intrinsic nature of the biological signals due to the absence of time resolution. In recent years, some joint time-frequency algorithms have been developed to extract feature set, such as the Short-Time Fourier Transform (STFT), Wavelet Transform (WT), Wavelet Packet Transform (WPT) (Englehart *et al.*, 1999; 2001; Zecca *et al.*, 2002), etc. In comparison with the feature exaction methods fully based on time-domain or frequency-domain, these methods have the localization time and frequency resolution and can extract the detailed information of the biological signals.

Classifiers can generally be divided into linear classifiers and nonlinear classifiers. It is not necessarily the case that the nonlinear classifiers are more powerful and suitable than linear classifiers for specific classification problems. It depends on the practical application and the feature extraction method. In the field of bio-signal processing, the researches have successfully employed some classifiers for EMG signal recognition and classification, such as the Hidden Markov Model, Artificial Neural Network, Fuzzy Logic and Neuro-Fuzzy, etc. (Kwon *et al.*, 1998; Christodoulou and Pattichis, 1999; Karlik, 1999; Micera *et al.*, 1999; Ajiboye and Weir, 2005).

It is worth pointing out that in (Englehart *et al.*, 1999; 2001; Subasi *et al.*, 2005), the authors demonstrated the performance of different combinations of feature extraction algorithms and classifier structures. In this paper, what we are concerned with is the motion classification of the upper arms by studying the corresponding surface EMG signals, and we exploited such a combination consisting of EMD-AR-based feature extraction and LS-SVM classifier in EMG motion classification. Experiments verified its efficiency and superior effect over other combinations.

The rest of this paper is organized as follows. Section 2 addresses the acquisition of EMG signals, the preprocessing to EMG signals, the extraction of feature vectors, the construction of the multi-class LS-SVMs and the relevant background knowledge. In Section 3, we compare in detail the combinations of other classifiers and feature vectors with the EMD-AR based LS-SVMs and demonstrate the performance of different combinations. Finally, the conclusions are drawn in Section 4.

METHODS

Data acquisition and preprocessing

In this work, we focus on the motion recognition and classification by analyzing the EMG signals. The four functions to be classified and controlled are: (1) fist clench (FC), (2) fist stretch (FS), (3) wrist flexion (WF), and (4) wrist extension (WE). The poses of these four motions are shown in Fig.1.

The sEMG signals were acquired by two pairs of bipolar electrodes placed on the flexors and extensors in the forearm. The experiments were implemented in the EMG room at Huashan Hospital in Shanghai, China.

To measure EMG signals, clean and sensitive electrodes and shielded cables were used to reduce the

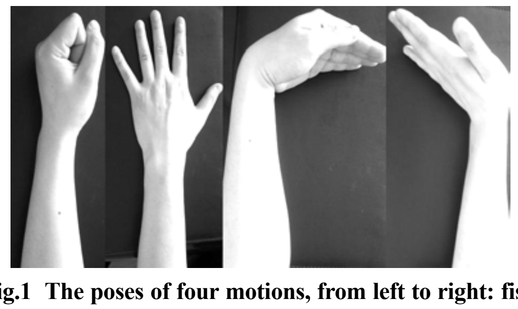

Fig.1 The poses of four motions, from left to right: fist clench (FC), fist stretch (FS), wrist flexion (WF), wrist extension (WE)

interference of noise in the environment. The detected differential EMG signals were fed to a band-passed filter with cutoffs at 10 Hz and 1 kHz, and a preamplifier with total gain of 2000. The amplified EMG signals were converted by A/D converter with 2 kHz sampling frequency and 12-bit resolution. As the useful information mainly distributes in the 20~500 Hz range, the sampling frequency can satisfy the Shannon Sampling theorem.

The hardware configuration was comprised of the Pentium-Mobile 1.5 GHz CPU, 768 M DDR 400 Memory, etc.

The subjects were 30 healthy college-student volunteers averaging 25.8 years old (ranging from 23 to 27 years old, 25 males and 5 females). They were all informed of the intention of the experience before the experiment implementation. Ethical approval for this research had been granted by the Ethical Committee of the Huashan Hospital. In the process of data acquisition, each subject generated four different classes of the above-mentioned motion and was asked to be consistent in reproducing the predefined motions 80 times. The EMG signals collected during isometric extension and flexion were all at 20 to 40 percent of the Maximal Voluntary Contraction, far from the extremity. To avoid the influence of fatigue on the Mean Frequency (Ravier *et al.*, 2005), after executing every 5 times the required movements, the subjects were allowed to have a rest for 2 min. The four patterns of EMG signals which lasted for at least 600 ms were recorded elaborately from every subject's forearm muscle groups. The number of the subjects was 30, so 120 datasets were obtained, 30 datasets for each motion, and 80 patterns in each dataset.

As the acquired signal may contain areas of inactivity, it is important to find out exactly when the EMG activity started. This was achieved by searching the beginning of the EMG activity and using a sliding-window method (in this case a 100 ms rectangular window) to extract the correct portion. Our strategy to find the limits of the EMG activity is based on a threshold for the variance of the signal. The underlying research about this strategy can be found in (Soares *et al.*, 2003).

Feature extraction

The EMG signal is an electric manifestation of neuromuscular activation associated with a contracting muscle and has the following characteristics (Nishikawa *et al.*, 1999).

First, EMG signals are non-stationary in the sense that their frequency spectra are time-varying. The EMG signal is summed from lots of motor unit action potentials (MUAPs), and the recruitment and firing of the MUAPs are stochastic. Therefore the intensity distribution of the surface EMG signal is nonlinear and time-varying. Second, the number of muscle fibers which make a single motor unit is person-dependent. Also the thickness of skins and other muscle tissues may modify EMG signal waveforms. Thus, many diverse factors exist in forming EMG signals, which makes it difficult to analyze. Third, EMG signals have weak amplitude of 0.1~1 mV, which means they can be easily contaminated by noise.

Considering the non-stationary and nonlinear essence of EMG signals, in the sequel, we executed the EMD to analyze the EMG signals. EMD was introduced by (Huang *et al.*, 1998b; Huang *et al.*, 1999) for nonlinear and non-stationary signal analysis. In bio-signal processing field, there are some successful applications (Huang *et al.*, 1998a; Neto *et al.*, 2002).

1. Principium of EMD

The general idea of this method is the sifting process to decompose any given signal into its intrinsic oscillations. With the EMD approach, the basis functions themselves are nonlinear functions which can be derived directly from the data, or in other words, an adaptive basis called Intrinsic Mode Function (IMF) can be found.

Assume $c(t)$ is an IMF, then it must satisfy two conditions as follows:

First, the number N_{extrema} of extrema and the number N_{zero} of zero point of $c(t)$ are equal or differ by one at most, i.e.

$$|N_{\text{extrema}}-N_{\text{zero}}|\leq 1. \tag{1}$$

Second, at any time, the mean value of the upper envelope $e_{\text{up}}(t)$ defined by the local maxima and the low envelope $e_{\text{low}}(t)$ defined by the local minima is zero, i.e.,

$$[e_{\text{up}}(t)+e_{\text{low}}(t)]/2\equiv 0. \tag{2}$$

In the case a time series is $s(t)$, EMD has the following major steps:

Initialization: $r_0(t)=s(t)$, $i=1$;

Step 1: Set $h_0(t)=r_{i-1}(t)$, $j=1$;

Step 2: Extract the local minima and maxima of time series $h_{j-1}(t)$. Interpolate the local maxima by a cubic spline to form upper envelope $e_{up}(t)$ of $h_{j-1}(t)$. And construct the lower envelope $e_{low}(t)$ of $h_{j-1}(t)$ by fitting all the local minima with cubic spline. The upper and lower envelopes should cover all the data between them and satisfy

$$e_{low}(t)\le h_{j-1}(t)\le e_{up}(t). \quad (3)$$

Step 3: Calculate the mean value of the envelopes:

$$m_{j-1}(t)=[e_{low}(t)+e_{up}(t)]/2. \quad (4)$$

The difference between $h_{j-1}(t)$ and its mean is

$$h_j(t)=h_{j-1}-m_{j-1}. \quad (5)$$

Step 4: If $h_j(t)$ meets the criteria of an IMF, designate this $h_j(t)$ as $imf_i(t)$.

If $h_j(t)$ is not an IMF, then increase j, return to Step 2 and repeat the procedure. If the amplitude of $h_j(t)$ is smaller than 10^{-8} times of the amplitude of r_{i-1}, the sifting process will be artificially stopped.

Step 5: Define the residue as

$$r_i(t)=r_{i-1}(t)-imf_i(t). \quad (6)$$

If $r_i(t)$ meets the stop criteria, the whole sifting procedure should stop. If not, increase i and return to Step 1. The authors set the final stop criterion to be that $r_i(t)$ has a predetermined number of extrema.

The essence of EMD is to identify the intrinsic oscillatory modes by their characteristic time scales in the data. Unlike the methods based on Fourier transform, EMD can be used to analyze data which are neither linear nor stationary. In comparison with the classical time-frequency analysis algorithm, WPT, which decomposes various signals using the specified and fixed base functions, i.e., the used base functions are object-independent, the EMD decomposition can obtain the essential and intrinsic features adaptively according to the signal amplitude and frequency information, so this algorithm is more flexible and adaptive.

On the other hand, EMD still has some inconveniences and drawbacks. While executing the EMD, the critical part is the envelope fitting. However, so far no strict theoretical evidence of fitting the envelope of EMD has been found but experience is still employed to fit the envelope. With a number of experiments, Huang *et al*.(1998b) recognized that using cubic spline fitting could obtain better result while others are non-ideal. However, there still exist two problems in cubic spline fitting: (1) The heavy computation load of sifting makes performing EMD very time-consuming; (2) The cubic spline fitting easily results in overshoot or undershoot, i.e., the fitted envelope is usually unable to satisfy Eq.(3), even leads to generation of a big error, thus the original essential structure of IMF is easy to be destroyed in EMD decomposition process.

To overcome these drawbacks, some researchers have put forward the improved EMD decomposition algorithm (Zhong *et al.*, 2004; Qin and Zhong, 2006).

2. Extracting the features from IMFs

Compared with the classifier, the classification performance is more profoundly affected by the choice of feature set (Englehart *et al.*, 1999). As the above sentence stated, EMD has a strong ability to extract the intrinsic characteristics from the nonlinear and non-stationary raw signals. To obtain the discriminative features, we used the combination of EMD and AR model as the feature extractor, which can be called as "the EMD-AR kernel to the SVM classifier".

What we are concerned with is getting some discriminative AR coefficients fed to the classifier instead of depicting the detailed frequency spectrum, so after having undermined the non-stationary characteristic of the EMG signal by the EMD transform, maybe the relatively low-order AR model is eligible for EMG classification. The four-order AR model has been considered adequate for modeling EMG signals in many previous applications, e.g. (Soares *et al.*, 2003; Hu and Nenov, 2004; Kim *et al.*, 2005). In this work, we adopt the four-order AR model to analyze EMG signals.

When AR model is mentioned, the high interdependence between the AR coefficients is a drawback. But this phenomenon exists mainly in the relatively high-order AR model. In our work, the four-order AR model is documented advisable for EMG classification, so this drawback can be ignored.

After having obtained the IMFs of each motion signal, we constructed the four-order AR model to extract AR coefficients from the first six IMFs as the feature vector. Fig.2 shows the waveforms of the first six IMFs.

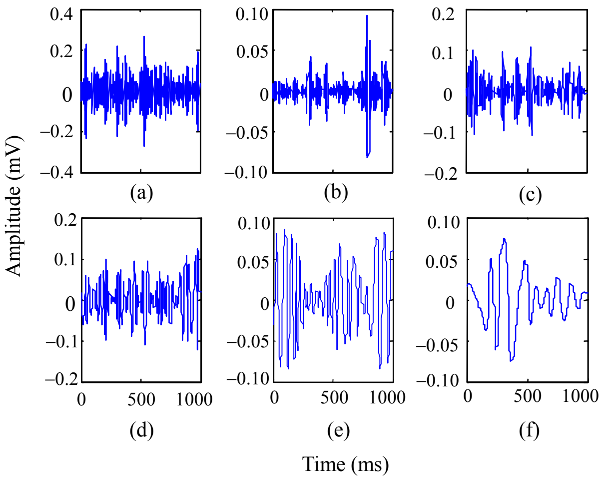

Fig.2 The first six IMFs of the fist clench. (a), (b), ..., (f) refer to the 1st, the 2nd, ..., and the 6th IMF, respectively

Classification implementation

Support vector machines (SVMs) (Vapnik, 1998) are powerful new tools for data classification and function estimation. SVM maps input data into a high dimensional feature space where it may become linearly separable. Recently SVM has been applied successfully to wide fields such as bio-information and pattern recognition. One reason that SVM often performs better than earlier methods is that SVM is designed to minimize structural risk whereas previous techniques are usually based on minimization of empirical risk, i.e., the minimization of the number of misclassified points on the training set. So SVM is usually less vulnerable to those over-fitting problems.

The training problem in standard SVM is reducible to solving a convex quadratic programming (QP) problem. The main drawback of standard SVM is its high computational complexity, therefore recently a new technique, the Least Squares SVM (LS-SVM) (Suykens and Vandewalle, 1999a; Suykens *et al.*, 2002), was introduced. This is algorithmically more effective, because the solution can be obtained by solving a linear equation set instead of a computation-intensive QP problem. In our experiments, by extending the binary LS-SVM classifier, the SVM technique was successfully applied in multi-class motion classification.

1. Principium of LS-SVM classifier

Given a training set $\{(x_i, y_i)\}$ with input data $x_i\in\mathbb{R}^n$ and corresponding binary class label $y_i\in\{-1, +1\}$, the SVM classifier satisfies

$$y_i[\boldsymbol{w}^{\mathrm{T}}\varphi(x_i)+b]\geq 1,\ \ i=1,\ldots,N, \tag{7}$$

where $\boldsymbol{w}$ and b are the weight vector and bias of the decision hyper-plane, respectively. The nonlinear function $\varphi(\cdot)$: $\mathbb{R}^n\rightarrow\mathbb{R}^m$ maps the input space to a high dimensional feature space.

The optimization problem is given as follows:

$$\min J(\boldsymbol{w},\xi)=\frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}+C\sum_{i=1}^{N}\xi_i, \tag{8}$$

$$\text{s.t.}\ \begin{cases} y_i[\boldsymbol{w}^{\mathrm{T}}\varphi(x_i)+b]\geq 1-\xi_i, \\ \xi_i\geq 0, \end{cases}\ \ i=1,\ldots,N. \tag{9}$$

The variables ξ_i are slack variables allowing for misclassification in the set of inequality. The positive real constant C should be considered as a tuning parameter in the algorithm.

The standard SVM classifier formulation was modified in (Suykens and Vandewalle, 1999a) into the following LS-SVM form:

$$\min_{\boldsymbol{w},b,e} J(\boldsymbol{w},e)=\frac{1}{2}\boldsymbol{w}^{\mathrm{T}}\boldsymbol{w}+\frac{\gamma}{2}\sum_{i=1}^{N}e_i^2, \tag{10}$$

$$\text{s.t.}\ \ y_i[\boldsymbol{w}^{\mathrm{T}}\varphi(x_i)+b]=1-e_i,\ \ i=1,\ldots,N, \tag{11}$$

where e is the classification error.

To keep the number of misclassified points as small as possible, and at the same time, make the margin of the hyper-plane as big as possible, the positive constant γ is introduced as a trade-off between the two competing terms.

The solution is obtained after constructing the Langrangian function:

$$L(\boldsymbol{w},b,e;\alpha)=J(\boldsymbol{w},b,e)-\sum_{i=1}^{N}\alpha_i\left\{y_i[\boldsymbol{w}^{\mathrm{T}}\varphi(x_i)+b]-1+e_i\right\}, \tag{12}$$

where $\alpha_i\in\mathbb{R}$ are the Lagrange multipliers that can be positive or negative in the LS-SVM form.

According to the above conditions, by solving the corresponding Karush-Kuhn-Tucker (KKT) system (Theodoridis and Koutroumbas, 2003), one can get the following equation:

$$\begin{bmatrix} 0 & \boldsymbol{Y}^{\mathrm{T}} \\ \boldsymbol{Y} & \boldsymbol{\Omega}+\gamma^{-1}\boldsymbol{I} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \tilde{\boldsymbol{1}} \end{bmatrix}, \tag{13}$$

where

$$\boldsymbol{\alpha}^{\mathrm{T}}=[\alpha_1,\ldots,\alpha_N],\ \boldsymbol{Y}^{\mathrm{T}}=[y_1,\ldots,y_N],\ \tilde{\boldsymbol{1}}^{\mathrm{T}}=[1,\ldots,1], \tag{14}$$

and

$$\Omega_{ij}=y_i y_j[\varphi(x_i)]^{\mathrm{T}}\varphi(x_j)=y_i y_j K(x_i,x_j). \tag{15}$$

For the kernel function K, one typically has the following choices: linear kernel, polynomial kernel, RBF kernel, MLP kernel, etc. In the following experiments, we execute the EMD to the original EMG signals and extract AR coefficients of each IMF. Here, we can denote the process of the EMD and AR modeling as EMD-AR kernel.

LS-SVM is a binary classifier, but we can extend it to solve multi-class classification problems. In a multi-class classification, by combining some basic binary classifiers, such as the one-to-one binary SVMs, the one-to-rest binary SVMs and the hierarchies of binary SVMs (Suykens and Vandewalle, 1999b; Schwenker, 2000), it is typically solved.

Although the flat binary SVMs of each level have high accuracy, the total accuracy of hierarchical SVM may be degenerated seriously because of the multiplication between each SVM on the same level. To some input data, the multi-class classifier based on one-to-one SVM and majority voting decision strategy may assign the same input data to more than one class label. Given the convenience of the one-to-rest LS-SVM, we constructed the multi-class classifier based on one-to-rest binary LS-SVM. To the k-class classification task, the number of constructed one-to-rest binary LS-SVM classifiers is k. To the ith one-to-rest binary classifier, it is trained with the ith class samples as positive ones and the rest as negative ones. To each input sample from the test set, we assign it the class label with maximum output. This decision mechanism can be called Max-Win strategy. In the following experiments, the advantage of using this kind of binary SVM to construct the multi-class will be discussed in detail.

Although the gain in efficiency is rather significant, for really large problems the computational burden of LS-SVM is still too high. Moreover, an attractive feature of SVM, its sparseness, is lost. As to online and fast adaptive signal processing and large-scale problems, the computational burden must be further reduced. The improved version of the LS-SVM was proposed by Valyon and Horváth (2003).

2. Classification scheme and contrast experiments

As Fig.3 shows, we first executed the EMD and AR modeling of the preprocessed EMG signals, then fed the extracted feature set to the multi-class LS-SVM classifier. Finally, we adopted the max-operator to decide the out class label.

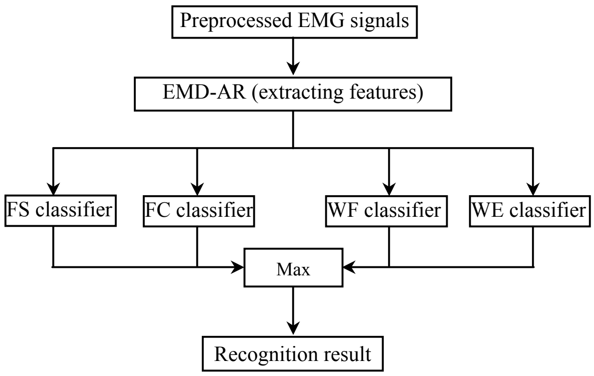

Fig.3 The classification strategy for multi-class classification

The FS classifier can be constructed by using the one-to-rest binary LS-SVM, which labelled the training data of fist stretch as +1 and the other training data as −1 in this case. The FC, WF and WE classifiers are constructed under the similar principle (see Table 1).

Table 1 The expected out label of the four binary LS-SVM classifiers

LS-SVM classifier	Fist clench	Fist stretch	Wrist flexion	Wrist extension
FC	+1	−1	−1	−1
FS	−1	+1	−1	−1
WF	−1	−1	+1	−1
WE	−1	−1	−1	+1

The EMG classification was divided into two stages: training and execution/test. At the training stage, a group of 20 patterns was chosen randomly for each of the four classes of movement (i.e., to each single one-to-rest LS-SVM binary classifier, total of

80 patterns for training) as the support vectors and the coefficient and bias scalar of the optimal hyper-plane were calculated. The rest of the datasets were used as the test dataset at the test stage.

As a contrast, we also used a neural network based on the Multi-Layer Perceptron (MLP) architecture as the classifier with the same feature vectors. The neural network classifier to be used was a three-layer MLP with 30 hidden nodes and 4 output nodes.

For further comparative study, the adaptive neuro-fuzzy inference system (ANFIS), which can be found in the Matlab Fuzzy Logic Toolbox, was exploited to design the fuzzy classifier. With the ANFIS aiming at the binary classification, we adopted the same strategy to extend it for multi-class classification as used in extending the binary LS-SVM classifier to multi-class LS-SVM classifier.

To exhibit the performance of EMD decomposition used in motion classification, we gave up the EMD decomposition and extracted the AR coefficients directly from the original EMG signals and then gave a contrast of the class separability between the two feature sets.

Given the existence of strong nonlinearity and non-stationarity, we divided the EMG signal into several equal segments to preserve the pattern structure. We could assume that each segment signal is stationary and the AR coefficients of each segment were taken as the features. In fact, the shorter the length of each segment is, the more reasonable is the assumption.

Additionally, we also figured out the quantification analysis on the separability of the two feature sets formed by EMD-AR and WPT respectively.

RESULTS AND DISCUSSION

Datasets from 30 subjects were used in the experiments. To verify the generalization ability of the LS-SVMs with small test dataset, the training dataset vs test dataset are 20 vs 20, 20 vs 40 and 20 vs 60, respectively.

In Table 2, the averaged classification performance of the LS-SVM classifiers between the 30 subjects is listed and it is clearly shown that the multi-class LS-SVM classifier based on the one-to-rest binary LS-SVM has good generalization with small training dataset.

Table 2 Performance of the multi-class LS-SVM classifier with different ratios of training vs test (averaged among the 30 subjects)

Training vs test	Correct rate			
	Fist clench	Fist stretch	Wrist flexion	Wrist extension
20 vs 20	0.993	0.983	0.987	0.993
20 vs 40	0.987	0.987	0.987	0.993
20 vs 60	0.983	0.983	0.985	0.987

Fig.4 shows the output of four LS-SVM classifiers against the motion type (Class) and the number of hand movements (Pattern) under observation for further understanding. As described above, four types of hand movements were studied and each subject was asked to perform 80 repetitions of each movement. While test vs. training is 20 vs 20, the pattern axis presents those repetitions as follows: (1) 1~20: fist clench; (2) 21~40: fist stretch; (3) 41~60: wrist flexion; (4) 61~80: wrist extension.

Fig.4 demonstrates the performance of the four binary LS-SVMs classifier while training dataset vs test dataset is 20 vs 20 and the subject is #10.

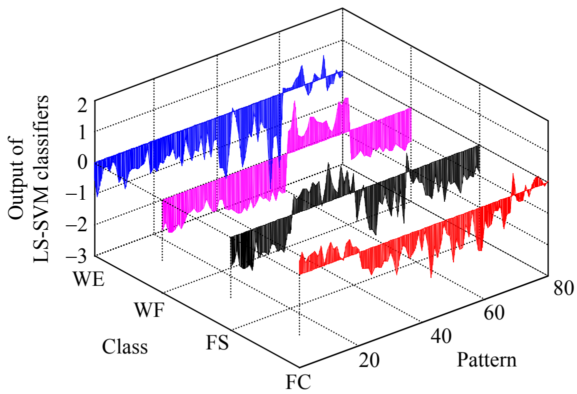

Fig.4 Output of the four one-to-rest LS-SVM classifiers. Trials 1~20: FC; 21~40: FS; 41~60: WF; 61~80: WE. (Subject #10)

From Fig.4, we can find the advantage of the multi-class classifier based on one-to-the-rest LS-SVM over that based on the hierarchical LS-SVM. Each LS-SVM classifier has some wrong output, but to the input sample from different test sets, such as 1~20 (FC), 21~40 (FS), 41~60 (WF), 61~80 (WE), we only adopted the maximum output of the four LS-SVM classifiers as the output of the multi-class classifier. To some extent, if there exists one LS-SVM with rarely poor performance, this strategy will prevent the great degeneration of the classification result, while in the case of hierarchical LS-SVM it degenerates seriously by multiplication.

In Eq.(10), when γ was adjusted from 1.0 to 4.0, the performance of the LS-SVM classifier only fluctuated slightly, which shows that the features extracted by the EMD-AR have good intrinsic separability. The following statement will verify it again.

Compared to the binary SVM classifier, the NN classifier is very convenient for multi-class classification. On the other hand, the disadvantages of the NN, such as over-fitting, under-fitting and vulnerability to the local minimum, etc., have impeded its application in some cases. Moreover, its computation load is larger than that of LS-SVM. While the test set vs training set is 20 vs 20, the performance of NN classifier is depicted in Fig.5. The subject is #10.

As the experiments proved, with the same feature set, the generalization of the NN classifier is not good compared with the LS-SVM classifier, as shown in Table 3.

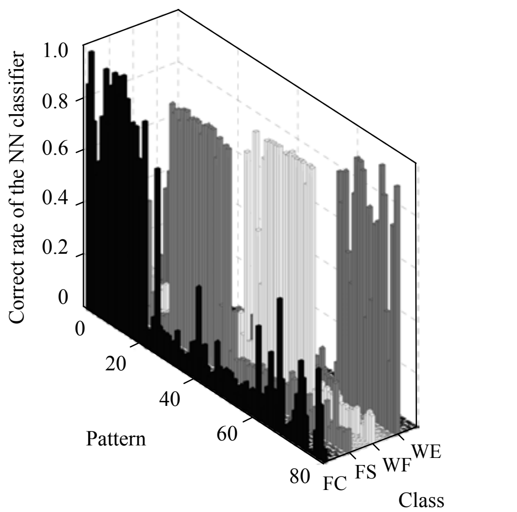

Fig.5 Performance of the NN classifier while training vs test is 20 vs 20 (Subject #10)

Table 3 The NN classifier's generalization ability in the case of getting global minimum in training phase (averaged among the 30 subjects)

Training vs test	Correct rate			
	Fist clench	Fist stretch	Wrist flexion	Wrist extension
20 vs 20	0.913	0.947	0.956	0.927
20 vs 40	0.867	0.923	0.843	0.873
20 vs 60	0.843	0.887	0.813	0.847

When training dataset vs test dataset is 20 vs 20, the consumed time of implementing the four-class EMG classification are 45 s and 98 s in Matlab environment for the four-class LS-SVM classifier and NN classifier, respectively. Relatively, the SVM classifiers are much simpler to implement and much faster to train. This result shows the superiority of LS-SVM classifier over the NN classifier.

Similarly, with the same feature set, by using the ANFIS-based multi-class fuzzy classifier, we executed the classification experiment. Through many tentative experiments, we found that because of the large input number (here it is 24) of the fuzzy classifier, the inference of the fuzzy system and rule extraction are tremendously time-consuming, mostly about two hours, even for offline classification, which is unacceptable. On the other hand, its classification performance is inferior to the LS-SVM classifier. For the practical use of the ANFIS in EMG classification, the feature reduction must precede the classification. Some algorithms of feature reduction can be referred to (Grzymala-Busse, 2003; Wang G. *et al.*, 2006; Wang X.Y. *et al.*, 2006).

As mentioned above, we tested three different combinations of classifier and feature vectors: the multi-class LS-SVM classifier with the feature set extracted by EMD-AR, the NN classifier and the ANFIS-based fuzzy classifier with the same feature set.

Furthermore, the experiment shows that, adopting the AR coefficients derived directly from the EMG signals as features, the NN classifier has worse performance than the above-mentioned combinations. This case shows that the performance of NN classifier is not so good as the SVM classifier, and that the feature vectors extracted directly by AR model are more difficult to separate than those extracted by AR model after performing the EMD.

Here the cluster separation index (CSI) (Kang *et al.*, 1995) is used as the criterion to measure the separability. Lower values of CSI imply a higher degree of cluster separability.

In the experiment, while we adopted the EMD-AR to form the feature set, the CSI was 1.074, with this result being comparable and competitive with the feature set obtained by the WPT algorithm in (Englehart *et al.*, 2001). This comparison shows that the combination of EMD and AR can also be used to extract subtle time-frequency joint information as well as the WPT. But if we extracted the features directly by the four-order AR model, the CSI was terribly 2.953. The experiments and the above result of changing γ of LS-SVM classifier prove consistently good separability of the features extracted by EMD-AR.

CONCLUSION AND FUTURE WORK

In this work, we develop an efficient combination of classifier and features, which prove to be applicable for recognition of the EMG signals by the contrast experiments. As the experiments showed, the combination represented as the LS-SVMs based on the EMD-AR can achieve better performance than other combinations over the four motion patterns. The methods proposed in this paper can also be extended to process other type of medical signals.

The method described here can help amputees fitted with prosthetic device train the artificial limbs. The experiment proves that the features extracted by EMD-AR have good subject-independence, and intrinsic good separability in contrast with the conventionally known time-domain features. And as a classifier, the LS-SVMs demonstrate better generalization ability and more rapid execution speed over the other two classifiers.

In this study, to solve the *k*-class classification problem, we design *k* one-to-rest SVM classifiers and adopt the Max-Win strategy to decide the output label. If the class number is rather large, the Max-Win strategy is roughly rigid. In further study, we will adopt the more flexible decision mechanism, for example, the information fusion algorithm or fuzzy theory.

Compared to the WPT tiling the various signals by the fixed base function, EMD is more adaptive and problem-relevant and therefore can get more subtle and specific information. Combined with the Hilbert Transform, EMD can describe elaborately the significant instantaneous time-frequency information. This algorithm is more popular and very promising.

It is clear that while adopting the multi-channel configuration, we can obtain more abundant information originated from sEMG. Englehart *et al.*(2001) verified that the four-channel configuration can offer improved recognition accuracy over that of the two-channel configuration. But in the multi-channel sampling process, there exists the coupling phenomenon between different channels. In further work, we will take the four-channel configuration and adopt the Independent Component Analysis (ICA) algorithm (Hyvärinen, 1999) to improve the EMG classification accuracy.

The whole classification program was written in Matlab and the training and testing phases of the mentioned classifiers were all executed in Matlab. For the online recognition, we can rewrite this algorithm in C or compile language. Now we are utilizing the specific DSP chip to develop the real-time bio-signal recognition.

References

Ajiboye, A.B., Weir, R.F., 2005. A heuristic fuzzy logic approach to EMG pattern recognition for multifunctional prosthesis control. *IEEE Trans. on Neural Syst. Rehabil. Eng.*, **13**(3):280-291. [doi:10.1109/TNSRE.2005.847357]

Christodoulou, C.I., Pattichis, C.S., 1999. Unsupervised pattern recognition for the classification of EMG signals. *IEEE Trans. on Biomed. Eng.*, **46**:169-178. [doi:10.1109/10.740879]

Englehart, K., Hudgins, 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. [doi:10.1016/S1350-4533(99)00066-1]

Englehart, K., Hudgin, B., Parker, P.A., 2001. A wavelet-based continuous classification scheme for multifunction myoelectric control. *IEEE Trans. on Biomed. Eng.*, **48**(3):302-311. [doi:10.1109/10.914793]

Englehart, K., Hudgins, B., 2003. A robust, real-time control scheme for multifunction myoelectric control. *IEEE Trans. on Biomed. Eng.*, **50**(7):848-854. [doi:10.1109/TBME.2003.813539]

Grzymala-Busse, J.W., 2003. A comparison of three strategies to rule induction from data with numerical attributes. *Electr. Notes in Theor. Computer Sci.*, **82**(4):1-9.

Han, J.S., Bien, Z., Bang, W.C., 2000. New EMG Pattern Recognition Based on Soft Computing. Techniques and its Application to Control of a Rehabilitation Robotic Arm. Proc. 6th IIZUKA 2000. Lizuka, Japan.

Hu, X., Nenov, V., 2004. Multivariate AR modeling of electromyography for the classification of upper arm movements. *Clin. Neurophysiol.*, **115**(6):1276-1287. [doi:10.1016/j.clinph.2003.12.030]

Huang, W., Shen, Z., Huang, N.E., Fung, Y.C., 1998a. Engineering analysis of biological variables: an example of blood pressure over 1 day. *PNAS,* **95**(9):4816-4821. [doi:10.1073/pnas.95.9.4816]

Huang, N.E., Shen, Z., Long, S.R., Wu, M.J.C., Shih, H.H., Zheng, Q.N., Yen, N.C., Tung, C.C., Liu, H.H., 1998b. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Royal Soc. London A*, **454**:903-995.

Huang, N.E., Shen, Z., Long, S.R., 1999. A new view of nonlinear water waves: the Hilbert spectrum. *Ann. Rev. Fluid Mech.*, **31**(1):417-457. [doi:10.1146/annurev.fluid.31.1.417]

Hudgins, B., Parker, P., Scott, R.N., 1993. A new strategy for multifunction myoelectric control. *IEEE Trans. on Biomed. Eng.*, **40**(1):82-94. [doi:10.1109/10.204774]

Hyvärinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Trans. on Neural Networks*, **10**(3):626-634. [doi:10.1109/72.761722]

Kang, W.J., Shiu, J.R., Cheng, C.K., Lai, J.S., Tsao, H.W., Kuo, T.S., 1995. The application of cepstral coefficients and maximum likelihood method in EMG pattern recognition. *IEEE Trans. on Biomed. Eng.*, **42**(8):777-785. [doi:10.1109/10.398638]

Karlik, B., 1999. Differentiating type of muscle movement via AR modeling and neural network classification. *Turk. J. Electr. Eng. & Computer Sci.*, **7**(1-3):45-52.

Kim, J.Y., Jung, M.C., Haight, J.M., 2005. The sensitivity of autoregressive model coefficient in quantification of trunk muscle fatigue during a sustained isometric contraction. *Int. J. Ind. Ergon.*, **35**(4):321-330. [doi:10.1016/j.ergon.2004.08.011]

Kwon, J., Lee, S., Shin, C., Jang, Y., Hong, S., 1998. Signal Hybrid HMM-GA-MLP Classifier for Continuous EMG Classification Purpose. Proc. 20th Annual Int. Conf. of the IEEE, **3**:1404-1407.

Micera, S., Sabatini, A.M., Dario, P., 1999. A hybrid approach to EMG pattern analysis for classification of arm movements using statistical and fuzzy techniques. *Med. Eng. Phys.*, **21**(5):303-311. [doi:10.1016/S1350-4533(99)00055-7]

Neto, E.P.S., Custand, M.A., Cejka, J.C., Abry, P., Frutoso, J., Gharib, C., Flandrin, P., 2002. Assessment of Cardiovascular Autonomic Control by the Empirical Mode Decomposition. 4th International Workshop on Biosignal Interpreatation, **43**:123-126.

Nishikawa, D., Yu, W., Yokoi, H., 1999. EMG Prosthetic Hand Controller Discriminating Ten Motions Using Real-Time Learning Method. Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems.

Qin, S.R., Zhong, Y.M., 2006. A new envelope algorithm of Hilbert-Huang transform. *Mech. Syst. Signal Processing*, **20**(8):1941-1952. [doi:10.1016/j.ymssp.2005.07.002]

Ravier, P., Buttelli, O., Jennane, R., Couratier, P., 2005. An EMG fractal indicator having different sensitivities to changes in force and muscle fatigue during voluntary static muscle contractions. *J. Electromyogr. Kinesiol.*, **15**(2):210-221. [doi:10.1016/j.jelekin.2004.08.008]

Reddy, N.P., Gupta, V., 2007. Toward direct biocontrol using surface EMG signals: control of finger and wrist joint models. *Med. Eng. Phys.*, **29**(3):398-403. [doi:10.1016/j.medengphy.2005.10.016]

Schwenker, F., 2000. Hierarchical Support Vector Machines for Multi-class Pattern Recognition. 4th Int. Conf. on Knowledge-Based Intelligent Engineering Systems & Allied Technologies. Brighton, UK, p.561-565.

Sebelius, F., Eriksson, L., Holmberg, H., Levinsson, A., 2005. Classification of motor commands using a modified self-organising feature map. *Med. Eng. Phys.*, **27**(5):403-413. [doi:10.1016/j.medengphy.2004.09.008]

Soares, A., Andrade, A., Lamounier, E., Carrijo, R., 2003. The development of a virtual myoelectric prosthesis controlled by an EMG pattern recognition system based on neural networks. *J. Intell. Inf. Syst.*, **21**(2):127-141. [doi:10.1023/A:1024758415877]

Su, Y., Fisher, M.H., Wolczowski, A., Bell, G.D., 2007. Towards an EMG-controlled prosthetic hand using a 3-D electromagnetic positioning system. *IEEE Trans. on Instrum. Meas.*, **56**(1):178-186. [doi:10.1109/TIM.2006.887669]

Subasi, A., Alkan, A., Koklukaya, E., Kiymik, M.K., 2005. Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing. *Neural Networks*, **18**(7):985-997. [doi:10.1016/j.neunet.2005.01.006]

Suykens, J.A.K., Vandewalle, J., 1999a. Least squares support vector machine classifiers. *Neural Processing Lett.*, **9**(3):293-300. [doi:10.1023/A:1018628609742]

Suykens, J.A.K., Vandewalle, J., 1999b. Multiclass Least Squares Support Vector Machines. Proc. Int. Joint Conf. on Neural Networks, p.900-903. [doi:10.1109/IJCNN.1999.831072]

Suykens, J.A.K., Gestel, T.V., Brabanter, J.D., Moor, B.D., Moor, B.D., Vandewalle, J., 2002. Least Support Vector Machines. World Scientific, Singapore. Http://www.worldscibooks.com/compsci/5089.html

Theodoridis, S., Koutroumbas., K., 2003. Pattern Recognition (2nd Ed.). Elsevier Science, p.77-82.

Valyon, J., Horváth, G., 2003. A weighted generalized LS-SVM. *Period. Polytechn. Electr. Eng.*, **47**(3-4):229-254.

Vapnik, V., 1998. The Support Vector Method for Function Estimation. Int. Workshop on Advanced Black-box Techniques for Nonlinear Modeling: Theory and Applications with Time-Series Prediction Competition, p.55-85.

Wang, G., Wang, Z.Z., Chen, W.T., Zhuang, J., 2006. Classification of surface EMG signals using optimal wavelet packet method based on Davies-Bouldin criterion. *Med. Biol. Eng. Comput.*, **44**(10):865-872. [doi:10.1007/s11517-006-0100-y]

Wang, X.Y., Yang, J., Jensen, R., Liu, X.J., 2006. Rough set feature selection and rule induction for prediction of malignancy degree in brain glioma. *Computer Methods and Programs in Biomed.*, **83**(2):147-156. [doi:10.1016/j.cmpb.2006.06.007]

Zecca, M., Micera, S., Carrozza, M.C., 2002. Control of multifunctional prosthetic hands by processing the electromyographic signal. *Crit. Rev. Biomed. Eng.*, **30**:459-485. [doi:10.1615/CritRevBiomedEng.v30.i456.80]

Zhong, Y.M., Qin, S.R., Tang, B.P., 2004. Research on theoretic evidence and realization of directly-mean EMD method. *Chin. J. Mech. Eng.*, **17**(3):399-404.