Identification Scheme of Surface Electromyography of Upper Limb Movement

Kexin Xing
Zhejiang University of Technology/College of Information Engineering, Hangzhou, China
Email: xkx@zjut.edu.cn

Qi Xu
Huazhong University of Science and Technology/Department of Control Science and Engineering, Wuhan, China
Email: xuqi@mail.hust.edu.cn

Yegui Lin
Zhejiang University of Technology/College of Information Engineering, Hangzhou, China
Email: lygxiaoxuanchang@126.com

Abstract—In this study four pairs of Ag/AgCl surface electrodes respectively that pasted on the muscle belly of the biceps, triceps in the upper arm, palmaris longus and the brachioradialis muscle in the forearm are used to collect the electromyography (EMG) signals of six patterns of upper limb movements. The feature vectors of the EMG are extracted by wavelet decomposition method, and then these features are classified by the method of the support vector machine. Experimental results show that the wavelet decomposition is an effective feature extraction method. For small samples, nonlinear and high dimensional classification problems, the support vector machine method has better classification effect and generalization performance than the traditional BP neural network. To speed up the computation of the SVM and adapt the increase of the samples, the incremental SVM is also introduced. By comparing with the traditional SVM method, the advantage of the incremental SVM is demonstrated.

Index Terms—EMG signal; wavelet decomposition; feature vector; support vector machine; incremental SVM

I. INTRODUCTION

Surface electromyography (EMG) is a technique for evaluating and recording the electrical activity produced by skeletal muscles. In fact the EMG pattern recognition system has been widely applied in many man-machine interface applications such as medical research and engineering practice [1]. For example, there is evidence that intensive therapy has beneficial effect on the rehabilitation progress of stroke patients, and such kind of therapies demand a detection of the intention of the patient’s activity [2, 3]. By extracting the intention, a human could control prosthetic legs with neuromuscular–mechanical fusion-based interface [4, 5].

However, the surface EMG signals are often mixed with signals generated by many muscles due to crosstalk and contain much noise [6], which causes difficulty identifying the intention of motivation from EMG signals. Thus the feature extraction is a key issue in pattern recognition or classification. At present, the signal feature extraction methods including time-domain [6], frequency domain [7, 8] and time-frequency domain feature extraction methods [9]. The time-domain method has the advantage of simplicity, but it assumes the EMG data as a stationary signal. When the surface EMG signal is recorded through dynamic movements, variation of features may be introduced in large scale. Frequency-domain features such as mean frequency and median frequency are more reliable and accurate than time-domain features. However, the frequency-domain methods are designed for analyzing stationary signals while EMG signals are actually non-stationary.

At present, the popular time-frequency feature extraction methods for EMG mainly include short-time Fourier transform (STFT), Vigner-Ville time frequency analysis and continuous wavelet transform (CWT) etc. [9-12]. Due to the non-stationary property of EMG signals, the CWT can better represent and analyze the signals than other methods as it describes the information in various time windows and frequency bands. The support vector machines (SVM) [13, 14] is a general learning method which has been developed on the basis of the statistical learning theory. According to the structural risk minimization (SRM) principle, the generalization ability of the classifier is maximized under the premise of classification error minimization. Compared with traditional BP neural network classifiers, the SVM method considers both the training error and the generalization ability. Thus, the SVM technology has many unique advantages in solving small sample, nonlinear and high dimensional pattern recognition problems[26, 27].

Although the SVM method has been successfully used in many fields [15-18], its slow computational speed still limits its application. To reducing the computational cost, Osuna have proposed an effective method to approximate the separation hypersurface with a subset of the support vectors by support vector regression machine (SVRM)[19]. However, this method still need a large number of support vectors to approximate the highly
convoluted separation hypersurface in the high-dimensional feature space. Later, some researchers improved this method [20, 21]. But these proposed methods are still the offline training SVM algorithm. In some application the samples always continue to be updated and it is difficult to obtain all the samples for the training of SVM at the start stage. An intuitive method is an incremental approach that utilizes the current SVM solution to simplify the solution of the next quadratic program in the search. Many incremental techniques to the SVM training have been developed to facilitate the online SVM training and speed up computations [22-25]. These incremental SVM methods have been used in face recognition, fault analysis and etc. In this work, the incremental SVM method also will be applied to the EMG recognition and compared with the traditional SVM method.

Based on the above understanding, this study has integrated wavelet feature extraction technology and SVM method to realize accuracy EMG pattern recognition. By the proposed method, 6 patterns of motion have been identified by using 4 channels of the EMG. By the analysis and comparison of the experimental data, the effectiveness of the method is illustrated. The rest of this paper is organized as follows: Section 2 describes the data acquisition experiment and system. In section 3, we present the wavelet algorithms for feature extraction. In Section 4, the principle and realization of SVM algorithm is described in detail. The experimental results are demonstrated and analyzed in section 5. Finally, the conclusion of current work and the possible future improvement are described.

2. DATA ACQUISITION

In this work the BioRadio 110 producted by Cleveland medical equipment company is selected as the EMG measurement instrument, shown in figure 1, which is mainly consist of three parts, the emitter, the receiver and the test module. Before the measurement, it is necessary to connect the test module and the transmitter, and then configure some parameters such as the sampling channel, the resolution of the input channels, the input signal range and a transmission frequency etc. During the actual measurement, the emitter is connected to the electrodes by a cable, and the EMGs are emitted to the receiver in a band of 902-928MHz. After the receiver detects and corrects the signal obtained, the EMG data are sent into the PC by serial port.

In our experiment, six kinds of upper limb motions are recognized by using four channels of EMG signals. As shown in Fig.2, four pairs of Ag/AgCl surface electrodes are respectively placed on the muscle belly of the biceps, triceps in the upper arm, palmaris longus and the brachioradialis muscle in the forearm. Four pairs electrodes are used to measure the EMG signal while the other one is used to provide a reference voltage. The 6 patterns are respectively elbow flexion, elbow extension, wrist internal rotation, wrist external rotation, hand close and hand open (mode1~mode6), shown as the Fig.3. The EMG signals are collected from 4 subjects (namely S1~S4). For each acquisition, two seconds preparation time is needed before starting, the action and the hold time are total of two seconds. When the acquisition is ended, the upper limb of the subject restores to the initial state and takes the rest for two seconds.
The continuous wavelet transform (CWT) is a strong time-frequency analysis method that developed in recent years and shows great potential in the field of biomedical signal processing. The CWT of the function \( f(t) \in L^2(R) \) is defined as:

\[
WF(a, b) = \langle f, \psi_{a,b}(t) \rangle = \int_{-\infty}^{\infty} \frac{1}{a^{1/2}} \psi\left(\frac{t-\tau}{a}\right) f(t) \, dt
\]

where \( a \) is the scale parameter which represents the size of dilation of the wavelet \( \psi \), and \( \tau \) is the shifting parameter. \( \psi_{a,b}(t) \) is the mother wavelet function, which is defined as follows.

\[
\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-\tau}{a}\right) \quad a, \tau \in R; a > 0
\]

The mother wavelet function is a continuous function in both the time domain and the frequency domain. In general, it is preferable to choose a mother wavelet that is continuously differentiable with compactly supported scaling function and high vanishing moments. The wavelet function is usually regarded as a band-pass filter and meets the permissible conditions:

\[
C_{\psi} = \int_{-\infty}^{\infty} \left| \hat{\psi}(\omega) \right|^2 \, d\omega < \infty
\]

### TABLE 1. THE FEATURE VECTORS OBTAINED BY THE WAVELET DECOMPOSITION

<table>
<thead>
<tr>
<th>Action</th>
<th>elbow flexion</th>
<th>elbow extension</th>
<th>wrist internal rotation</th>
<th>wrist external rotation</th>
<th>hand close</th>
<th>hand open</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palmaris longus</td>
<td>213.443</td>
<td>166.521</td>
<td>189.411</td>
<td>57.8143</td>
<td>490.201</td>
<td>69.0023</td>
</tr>
<tr>
<td>404.209</td>
<td>2</td>
<td>516.033</td>
<td>349.679</td>
<td>87.8271</td>
<td>400.726</td>
<td>164.257</td>
</tr>
<tr>
<td>647.152</td>
<td>2</td>
<td>356.826</td>
<td>178.233</td>
<td>93.7471</td>
<td>489.180</td>
<td>183.907</td>
</tr>
<tr>
<td>543.869</td>
<td>1</td>
<td>104.711</td>
<td>229.849</td>
<td>90.6168</td>
<td>248.975</td>
<td>423.534</td>
</tr>
<tr>
<td>729.711</td>
<td>2</td>
<td>240.194</td>
<td>477.304</td>
<td>156.358</td>
<td>522.379</td>
<td>862.453</td>
</tr>
<tr>
<td>Brachioradialis</td>
<td>985.720</td>
<td>315.182</td>
<td>485.354</td>
<td>155.069</td>
<td>489.984</td>
<td>819.598</td>
</tr>
<tr>
<td>Adialis muscle</td>
<td>282.576</td>
<td>136.951</td>
<td>69.612</td>
<td>311.138</td>
<td>48.8235</td>
<td>55.2744</td>
</tr>
<tr>
<td>715.167</td>
<td>8</td>
<td>204.827</td>
<td>158.639</td>
<td>178.29</td>
<td>141.597</td>
<td>154.765</td>
</tr>
<tr>
<td>682.154</td>
<td>3</td>
<td>321.988</td>
<td>194.366</td>
<td>329.436</td>
<td>355.228</td>
<td>324.587</td>
</tr>
<tr>
<td>Biceps</td>
<td>79.0194</td>
<td>419.791</td>
<td>93.4299</td>
<td>71.7001</td>
<td>18.1377</td>
<td>18.7661</td>
</tr>
<tr>
<td>Triceps</td>
<td>155.428</td>
<td>600.959</td>
<td>236.995</td>
<td>193.427</td>
<td>41.9086</td>
<td>51.3137</td>
</tr>
<tr>
<td>192.590</td>
<td>3</td>
<td>390.076</td>
<td>234.568</td>
<td>292.175</td>
<td>99.3003</td>
<td>130.740</td>
</tr>
</tbody>
</table>

Figure 4. The amplified and filtered EMG signal

The amplified and filtered EMG signals corresponding to 6 patterns are shown in Fig.4, where each 5 second of the wave corresponds to one pattern of upper limb motion.
In this paper the smy5 is selected as the mother wavelet function and the triplet of wavelet transform is used. The absolutely largest values of the wavelet detail coefficients are used to construct feature vector. In this way, 3 feature vectors can be obtained from the EMG signal of each muscle. So, for the four muscles in each sample we can get 12 feature vectors. For example, the feature vectors obtained from a set of actions by the wavelet decomposition are shown in Table 1.

B. Support Vector Machine

The original SVM algorithm was invented by Vladimir N. Vapnik and the current standard incarnation (soft margin) was proposed by Vapnik and Corinna Cortes in 1995 [13]. The SVM algorithm is mainly used for two-group classification problems in pattern recognition.

Define the linear separable sample set as \((X_i, y_i) | i = 1, 2, \ldots, N, X \in \mathbb{R}^n, y \in \{+1, -1\}\), which can be divided into the positive sample set \(X^+\) and a negative sample set \(X^-\) according to the different category \(y\). Shown as the Fig.5, the basic idea for the classification problem is to select an optimal decision surface which ensures to separate the two types of sample sets with a maximum interval.

![Figure 5. The basic principle of the support vector machine](image)

The two subsets could be divided by the hyperplane if a vector \(W^*\) and a constant \(b^*\) can be found and meet the following constraint:

\[
y_i (\langle x_i, W^* \rangle + b^*) \geq 1 - \xi_i , i = 1, \ldots, N
\]

where \(y_i = \text{sgn}(\langle x_i, W^* \rangle + b^*)\), \(\text{sgn}\) is the sign function, \(\xi\) is the slack variable and the vector \(W^*\) has the minimum norm

\[
\min \rho(W) = \min \left[ \frac{1}{2} \|W^*\|^2 + C \sum_{i=1}^{N} \xi_i \right]
\]

After the vector \(W^*\) and the constant \(b^*\) are determined, the optimal separating hyperplane is defined by the formula (5). Therefore, in order to get the optimal separating hyperplane, the quadratic programming problem must be solved in case of satisfying the linear constrained formula (4). According to the Lagrange method and introducing the kernel function, this problem can be transformed to the following optimization problem:

\[
\min H(a) = \min \left[ \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} a_i a_j y_i y_j K(X_i, X_j) \right] - \sum_{i=1}^{N} a_i
\]

\[\text{s.t. } C \geq a_i \geq 0, i = 1, \ldots, N, \sum_{i=1}^{N} y_i a_i = 0\]

The optimal solution \(a^*\) represents for the vector \(W^*\) and \(b^*\) can be obtained by looking up the Table 2.

![Table 2. The relationship of nowNo, subNo and SVMNo](image)
c) Calculate the vector \( W = \sum_{i=1}^{N} y_i a_i X_i \) according to the optimal Lagrange multiplier \( a^* \), and then select a support vector \( X \) from the training set. Substitute them into the equation (7), the left \( f(x) \) is equal to the category value (-1 or 1), so the deviation \( b^* \) can be obtained and stored into the SVM(SVMNo(m, n)]; \( b^* \).

d) Then go to step 2 until the calculation between all two types is completed.

(2) The classification steps of the basic SVM

a) Input the test sample \( X_{in} \).

b) Construct the binary tree between categories. Start to construct three pairs of nodes \{1, 2\}; \{3, 4\}; \{5, 6\} from the bottom of the tree, and define the former of each pair of nodes as the positive samples, the latter as a negative sample.

c) According to the result from above training step, substitute the test sample \( X_{in} \), the Lagrange multipliers \( a^* \), the deviation value \( b^* \) and the kernel function into the discriminant function (7), and then the category of the sample can be judged from each pair of nodes. At last a new nodes collection can be constructed as the higher layer of the tree, shown as the Fig. 6.

d) If there are an odd number of categories in the new collection during constructing the node pairs, the redundant node will be reserved until the next iteration. According to the new node configuration, repeated the steps 3 until the top layer.

e) The final result of the discrimination is given in the top layer.

![Figure 6. The principle of SVM classification using the binary tree structure](image)

**C. KKT Condition and Incremental Study of the SVM**

The equation (7) must be satisfied with the KKT condition as follows:

\[
\begin{align*}
[f(x)_i] & \geq 1 \quad \text{if } a_i = 0 \\
[f(x)_i] & = 1 \quad \text{if } 0 < a_i < C \\
[f(x)_i] & \leq 1 \quad \text{if } a_i = C
\end{align*}
\]  

(12)

The above KKT condition shows that the support vectors contain all the recognition information. When adding samples which include the new classification information, the support vectors will be changed by the learning process. So the following theorem can be derived.

**Theorem 1:** Support \( f(x) \) is the discriminant function and \((x_i,y_i)\) is the new adding sample. The samples which satisfy the KKT condition have no effect on the support vectors, while the other samples will change the support vectors. These samples that violate the KKT condition can be classified three categories.

1. The sample locates in the interval of the classification and can be recognized by the original SVM classifier, satisfying \( 0 \leq y_i f(x_i) < 1 \);
2. The sample locates in the interval of the classification and is incorrectly recognized by the original SVM classifier, satisfying \( -1 \leq y_i f(x_i) < 0 \);
3. The sample exceeds the margin of the classification and is incorrectly recognized by the original SVM classifier, satisfying \( y_i f(x_i) < -1 \);

Thus the incremental algorithm can be designed to only learn above three kinds of the samples for the new support vectors. The incremental algorithm step is shown as follows:

1. Input the six types of training sample \( X_{in}, i = 1, 2, \cdots, 25, j = 1, 2, \cdots, 6 \), where \( j \) represents the type, \( i \) is the index of the sample in the group.
2. Randomly select two samples of \( X_{in}, X_{in} \). Calculate the nuclear space distance \( d \) of them according to the following space distance equation.

\[
d(x_{in}, x_{in}) = \sqrt{|K(x_{in}, x_{in}) - K(x_{in}, x_{in}) + K(x_{in}, x_{in})|} \]  

(13)

Where \( K(.) \) is the corresponding kernel function.
3. Set a threshold of the nuclear space distance \( d \) as \( D \). Determine boundary vector sample sets \( B \) according \( B = \{x_{k}, \|f(x_{k}) - D, k = 1, \ldots, N + M\} \).
4. Using the basic SVM algorithm, Obtain the support vector sets \( SV0 \) and calculate the determinant function \( f_0 \);
5. Calculate \( y_i f(x_i) \) by Using the \( SV0 \) on the incremental training sample set \( B1 \) and find set \( B1' \) in where the new samples violate the KKT conditions.
6. Use of the sample set \( B \cup B1' \) training, get a new support vector sets \( SV1 \) and the new determinant function \( f_1 \).
7. Then go to step (5) until the incremental training sample set is null.
8. Go to step 2 until the calculation between all two types is completed.

**IV. EXPERIMENTAL RESULTS**

First, the feature vectors that previously obtained are normalized into the interval (-1, 1), then the data are divided into 6 groups according to the different actions. 50 samples in each group are divided to two parts, namely the front 25 samples are used as training samples, the others for testing samples.

In the first experiment, for each subject, 50 samples for
each pattern were recorded to train and test the system, so the total number of data to test the recognition system is 50*6*4=1200. According to the basic SVM training and classification algorithm, the samples are trained and test in case of different kernel function. The parameters of different kernel functions and classification results are shown in Table 3.

<table>
<thead>
<tr>
<th>kernel function</th>
<th>parameter</th>
<th>classification accuracy%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>none</td>
<td>96</td>
</tr>
<tr>
<td>Polynomial</td>
<td>d=1</td>
<td>96.0</td>
</tr>
<tr>
<td>Polynomial</td>
<td>d=3</td>
<td>98</td>
</tr>
<tr>
<td>Polynomial</td>
<td>d=4</td>
<td>97.33</td>
</tr>
<tr>
<td>RBF</td>
<td>σ = 0.5</td>
<td>96.67</td>
</tr>
<tr>
<td>RBF</td>
<td>σ = 0.1</td>
<td>95.33</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>b=0.5, c=0</td>
<td>91.33</td>
</tr>
</tbody>
</table>

According to the above results, the polynomial kernel function is chosen in our experiment, its nuclear parameters chosen as $d = 3$. The results of the samples classification by SVM with polynomial kernel function are shown in Fig. 7. In order to facilitate drawing, the six patterns of motion are respectively denoted as: EF(elbow flexion), EE(elbow extension), WIR(wrist internal rotation), WER(wrist external rotation), HC(hand close) and HO(hand open).

Due to there has large variances for each BP training, the best result from the repeated experiment is demonstrated in this paper.

![Figure 7](image1.png)

![Figure 8](image2.png)

![Figure 9](image3.png)
From the Fig. 8 and Fig. 9, the training process of the BP neural network achieves a faster speed than SVM method, but the result of the classification is not very stable. Especially, the effect of classification is not satisfying when BP neural network fall in the overfitting case. According repeatedly comparing and analyzing the classification results between the BP neural network and the SVM, the following conclusions can be drawn:

(1) The training speed of the BP neural network is faster than the SVM. The classification results of the BP neural network are different in each time while the classification results of SVM are more stable.

(2) More parameters need to be set for the BP neural network method, but the SVM just needs to adjust the parameters of the kernel function.

(3) In the training process, the BP classifier always focuses on minimizing the training error, but the good classification performance is not guaranteed for the test data. Furthermore, when the target error is set too small or the complexity of the network increases, the risk of the overfitting phenomena will correspondingly increase. It may lead to the worse classification performance.

In last experimental, the training time and testing accuracy of the SVM and incremental SVM are compared. In our experiment, the samples are divided to 6 groups according to the different motion. The samples number of the each group increased from 30 to 300. 15 support vectors machine can be obtained by the training process. The total training time shown in the fig. 10 exhibits the different performance between SVM and incremental SVM algorithm. In this experiment the training time did not includes the time of the feature extraction. Because the testing correct rates for each motion are different, only the average correct rates are compared in case of different samples number. Furthermore, the testing shows that whether SVM or Incremental SVM is selected, the similar correct rates will be found when the total samples number exceeds 300. So the correct rates are tested and compared in case samples number is smaller than 300. The samples number of the each group increased from 10 to 50.

The experimental results of the training time and correct rate are respectively shown in Fig. 10 and Fig.11. From Fig.11 it can be seen that the two methods has similar training time when the number of the samples is smaller than 600. But when the number of the samples is over 600, the training time of the basic SVM will increase rapidly. The training time of the incremental SVM is more flat. In generally, to the recognition of the EMG, the basic SVM has higher accurate than the incremental SVM. But it is also demonstrated that the correct rate of the incremental SVM continues to improve with the increase of the sample number. It is obvious that the incremental SVM is suitable for the online training of the EMG recognition.

V. CONCLUSION

In this work the feature vectors of the EMG signals are extracted by the wavelet transform method and 6 patterns of motion have been identified by the SVM. The experimental results show that the wavelet decomposition is an effective tool of the feature extraction and the EMG signals can be well characterized by a few wavelet coefficients. SVM exhibits the similar classification accuracy under the different kernel function, which illustrates the effectiveness and robustness of the SVM classifier. Comparing the classification result by the BP neural network and the SVM, it is shown the SVM classifier has better classification and generalization performance than the traditional BP neural network. The incremental SVM and the basic SVM are also compared in the correct rate and training time. The experimental results illustrate the incremental SVM is effective and suitable for the online training of the EMG recognition. This study has established a proper platform for further research and experiment. For the purpose of application in the future, the EMG data obtained from amputees should be bench-tested. Technically the robustness against shifts of the electrodes and the behavior in non-static contractions will be thoroughly investigated.

ACKNOWLEDGMENT

This work has been supported by project of Zhejiang Education Department under Grant Y201121841, the
Natural Science Foundation of China under Grants 61202203 and in part by Zhejiang Provincial Natural Science Foundation of China under Grant LY12F01023.

REFERENCES


