Emotion Recognition of EMG Based on Improved L-M BP Neural Network and SVM

Shanxiao Yang, Guangying Yang
Department of Electronics Engineering, Taizhou University, Taizhou, China
Email: ysxtzc@126.com

***Abstract*— This paper compares the emotional pattern recognition method between standard BP neural network classifier and BP neural network classifier improved by the L-M algorithm. Then we compare the method Support Vector Machine (SVM) to them. Experiment analyzes wavelet transform of surface Electromyography (EMG) to extract the maximum and minimum wavelet coefficients of multi-scale firstly. We then input the two kinds of classifier of the structural feature vector for emotion recognition. The experimental result shows that the standard BP neural network classifier, L-M improved BP neural network classifier and support vector machine's overall pattern recognition rate is 62.5%, 83.33% and 91.67 respectively. Experimental result shows that feature vector extracted by the wavelet transform can characterize emotional patterns through the comparison with the BP neural network classifier and Support Vector Machine, indicating that the Support Vector Machine have a stronger emotional recognition effect.**

***Index Terms*— Surface Electromyography (EMG) Signal; Emotional Pattern Recognition; Support Vector Machine (SVM); Wavelet Transform; L-M algorithm**

I. INTRODUCTION

Surface EMG(surface electromyography)signals, also known as EMG, is the one-dimensional time series signal obtained(guide, zoom, display and records) by bioelectrical changes in the skin surface with the activity of neuromuscular system, its detection has non-invasive, real-time, multi-target measurement and other advantages. It reflects the functional status of nerves and muscles and has been widely used in clinical medicine, sports medicine and other fields [1].

The study of emotion recognition has an important significance in understanding human emotions in the role of human intelligence. Emotion recognition is one of the key steps towards emotional intelligence in advanced human-machine interaction. Although many efforts have been taken recently to recognize emotions using facial expressions, speech and physiological signals [2, 3, 4, 5], current recognition systems are not yet advanced enough to be used in realistic applications.

In daily life, human intelligence not only shows in the normal rational thinking and logical reasoning ability, but also in the normal emotional capabilities. Computer science, this ability to promote the establishment of a friendly man-machine interface is of great significance. As the deepening of affective computing, the request for emotion recognition technology will be correspondingly enhanced.

This article uses the surface EMG signal with objective data for six-scale decomposition of surface EMG with the method of wavelet transform and extract the maximum and minimum of multi-scale wavelet coefficients, constructing 14-dimensional feature vector, respectively, input into the standard BP neural network, the L-M algorithm improved BP neural network classifier and Support Vector Machine (SVM) for emotion recognition.

II. RECOGNITION METHODS

A. L-M algorithm based on improved BP neural network

BP neural network is fully named as the Back-Propagation Network, that is, back-propagation network. It is a forward multi-layer network, which uses the error back-propagation algorithm to train the network. BP algorithm was proposed by Rumelhart et al [6] in 1986, and since then, due to simple structure, multi-adjustable parameters, much training algorithm and good operational performance, BP neural network got a wide range of practical application.

The network structure of the three-layer BP neural network is shown in Figure 1, from which we can see that, BP neural network contains an input layer, a middle layer (hidden layer) and an output layer. There is a full connectivity between the upper and lower layers and no connections between neurons in each layer. For the input signal, it needs to spread towards to hidden layer nodes and transformed by the function, then transmit the input signal of hidden layer nodes to the output layer nodes. Usually, the transfer function of BP neural network is Sigmoid Type differentiable function, which can achieve arbitrary non-linear mapping between the input and output, so BP network has been widely applied in pattern recognition, function approximation and other areas [6].

The three nodes of the BP network is represented as: input node x_j, hidden node y_i, output node σ_j, network weight of input node w_{ij}, network weight of hidden node

This work is supported by education department Program of Zhejiang Province in University (2010).

doi:10.4304/jsw.6.8.1529-1536

and output node T_{li}, the expectation output the output node t_l, Figure 1 shows the BP neural network structure.

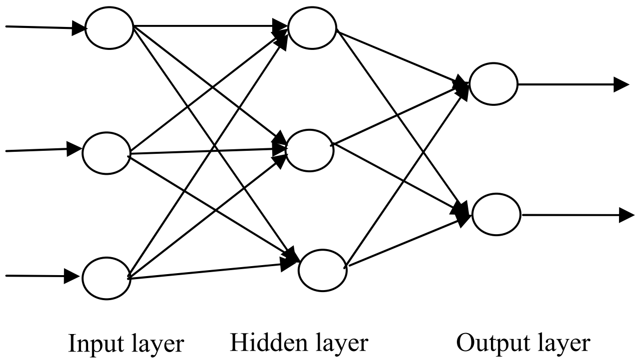


Figure 1. BP neural network structure

The basic learning algorithm of BP neural network:

l) Determine the various learning parameters on the basis of the known network structure, including neuron number of input layer, hidden layer, output layer, learning rate and error parameters.

2) Initialize the network weights and thresholds.

3) Provide learning sample: input vector and target vector.

4) Start to learn, and do the following for each sample:

a. Forward-calculation of the j unit in the l layer:

$$v_j^{(l)}(n)=\sum_{i=0}^{T} w_{ji}^{l}(n) y_i^{l-1}(n) \tag{1}$$

Equation (7) $y_i^{l-1}(n)$ is the signal transmitted from unit i(i=0, set $y_0^{l-1}=-1$, $w_{j0}^{l}(n)=\theta_j^{l}(n)$) of the $(l-1)$ layer.

If the activation function of j unit is sigmoid function, then:

$$y_j^{(l)}(n)=\frac{1}{1+\exp\left(-v_j^{(l)}(n)\right)} \tag{2}$$

And we can get:

$$f'\left(v_j(n)\right)=\frac{\partial y_j^{(l)}(n)}{\partial v_j(n)}=y_j(n)\left[1-y_j(n)\right] \tag{3}$$

If the j unit belongs to the first hidden layer $(l=1)$, then

$$y_j^{(0)}=x_j(n) \tag{4}$$

If the j unit belongs to the output layer $(l=L)$, then

$$y_j^{(L)}(n)=O_j(n) \tag{5}$$

$$e_j(n)=d_j(n)-O_j(n) \tag{6}$$

b. back-calculation of δ:

For the output units,

$$\delta_j^{(l)}(n)=e_j^{(L)}(n)O_j(n)\left[1-O_j(n)\right] \tag{7}$$

And for the hidden layer units,

$$\delta_j^{(l)}(n)=y_j^{(l)}(n)\left[1-y_j^{(l)}(n)\right]\sum_k \delta_k^{(l+1)}(n) w_{kj}^{(l+1)}(n) \tag{8}$$

c. Fix the right values according to the following:

$$w_{jk}^{(l)}(n+1)=w_{ji}^{(l)}(n)+\eta\delta_j^{(l)}(n)y_i^{(l-1)}(n) \tag{9}$$

5) Enter a new sample until it reaches the error requirement, and the input order of each cycle in training samples needs a re-random order.

The specific program flow chart of training network using BP algorithm shows as fig.2.

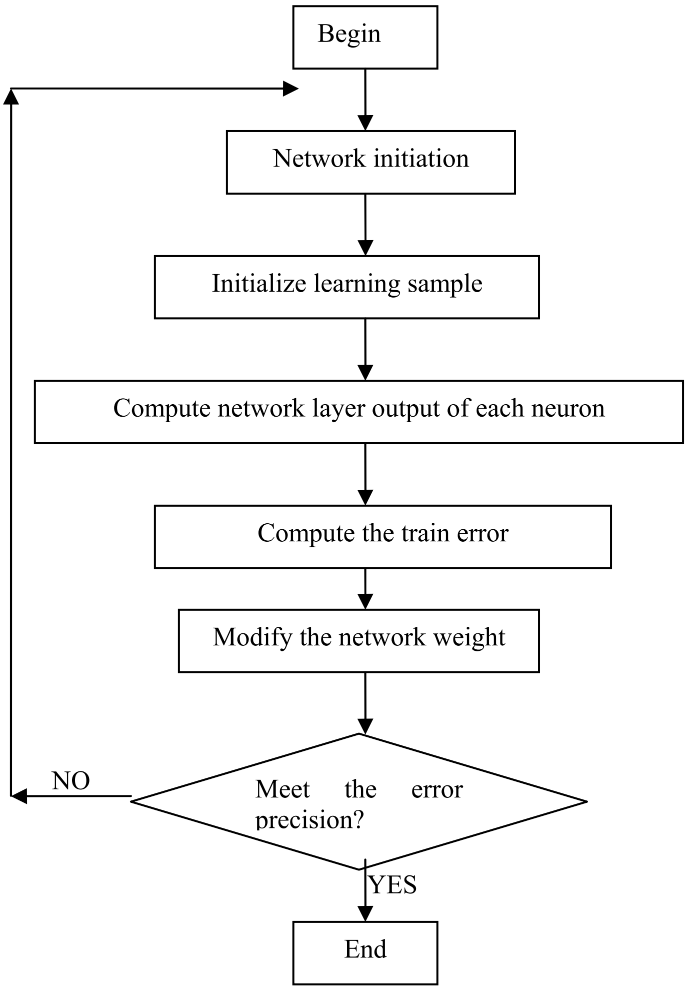


Figure 2. BP Algorithm Flow Chart

At present, a number of improved algorithm is put forward for so many shortcomings of the standard BP algorithm, such as the second-order fast BP algorithm, momentum factor method, role function adjust method and so on. These improved methods improve obviously in convergence speed and approximation accuracy than the gradient descent method. But for the problem with many variables and large sample input emotional identification, there is low approximation precision and even paralysis of training problems. To this end, L-M algorithm is needed for BP network training. L-M algorithm is an effective method of nonlinear least squares problem. It is designed to minimize the error sum of squares and is superior to conventional BP algorithm with fast convergence and high approximation accuracy. The network weights, the threshold adjustment is essentially least squares problems of dealing with an non-linear error function E, and therefore L-M algorithm can be used to achieve adjustment of the weight value and the threshold[7, 8].

Specific steps of L-M algorithm are as follows:

(1) Give the allowed training error ε, constants β and μ0; as well as the initial weight value and threshold vector, giving that k = 0, μ = μ0;

(2) Compute network output and error objective function E (xk);

(3) Calculate Jacobian matrix J (x);

(4) Calculate △x;

(5) If E<ε, go to step (6), else, compute error objective function E (xk). If E(x k+1) < E(x k), then k=k+1, μ=μ/β, and go to step (2), else go to step (4);

(6) End.

B. Pattern classification using support vector machine

Support vector machine method is based on the statistical learning VC dimension theory and the structural risk minimization principle. Here, we briefly describe the principle of pattern classification using the SVM. A two-class classification problem is assumed for simplification. The problem of finding a linear classifier for given data points with a known class label can be described as a problem of finding a separating hyper plane $\omega^T x + b$ that satisfies:

$$y_i(\omega^T x_i + b) \ge 1,\ for\ i = 1, 2, ..., N \tag{10}$$

Where xi and $y_i \in \{+1, -1\}$ denote a feature vector and its given correct class label, respectively. If it is not possible to classify them with a linear classifier, as is the case with most practical problems, the problem can be described in a less strict form as follows:

$$y_i(\omega^T x_i + b) \ge 1 - \xi_i,\ for\ i = 1, 2, ..., N \tag{11}$$

Here, ξ_i is called a slack variable, and it represents deviation from the ideal condition of its linear reparability. We can pose a problem of finding the optimum one among the separating hyper planes by minimization of the cost function $\frac{1}{2}\omega^T \omega + C\sum_{i=1}^{N} \xi_i$ subject to the constraints:

$$y_i(\omega^T x_i + b) \ge 1 - \xi_i,\ for\ i = 1, 2, ..., N$$

$$\xi_i \ge 0,\ for\ i = 1, 2, ..., N \tag{12}$$

The above cost function is defined so that its minimization coincides with the maximization of margin and minimization of classification error under the constraint of eq. (12).

This constrained optimization problem can be solved by using the Lagrange multiplier method. From the theory of the Lagrange multiplier method, it can be shown that the above problem can be expressed as a problem of finding Lagrange multipliers $\alpha_i s$ as follows: given the training set $\{(x_i, y_i),\ i = 1, 2, ..., N\}$, find $\alpha_i s$ that maximize the objective function $Q(\alpha_1, \alpha_2, ...\alpha_N)$, i.e.

We maximize Q:

$$Q(\alpha_1, \alpha_2, ...\alpha_N) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{i=1}^{N} \alpha_i \alpha_j y_i y_j x_i^T x_i \tag{13}$$

It subjects to:

$$\sum_{i=1}^{N} \alpha_i y_i \ge 0\ and\ 0 \le \alpha_i \le C\ for\ i = 1, 2, ..., N \tag{14}$$

This is called the dual problem of the original problem that seeks to find the optimum separating hyper plane. After finding $\alpha_i s$, the classification of a data point $\mathrm{x}_{\mathrm{new}}$ is performed as

$$f(\mathrm{x}_{\mathrm{new}}) = sign(\sum_{i=1}^{L} y_i a_i \mathrm{x}_{\mathrm{new}}{}^T x_i + b) \tag{15}$$

Here L is the number of support vectors obtained from the maximization. The relationship between the parameter of hyper plane ω and the Lagrange multiplier $\alpha_i s$ is given as $\omega = \sum_{i=1}^{L} y_i a_i x_i$. Actually, the SVM algorithm adopts a preliminary non-linear mapping to higher-dimensional feature space before the linear discrimination. The feature space is hidden from both input and output. The rationale for the non-linear mapping is taken from the Cover theorem. It states that a non-linear mapping to high dimension increases the likelihood of linear separation. The decision function is now expressed as follows, instead of eq. (15):

$$f(\mathrm{x}_{\mathrm{new}}) = sign(\sum_{i=1}^{L} y_i a_i \{\varphi(\mathrm{x}_{\mathrm{new}})\}^T . \varphi(x_i) + b) \tag{16}$$

Here $\varphi(x)$ denotes the non-linear mapping to high dimension. Previous expressions involve computation of an inner product in high dimensional space. Not every mapping $\varphi(x)$ can be expressed in this fashion, and the criteria are stated by Mercer's theorem. Here, we used a Gaussian kernel. Finally, the decision rule of a given data point is

$$f(\mathrm{x}_{\mathrm{new}}) = sign(\sum_{i=1}^{L} y_i a_i \mathrm{K}(\mathrm{x}_{\mathrm{new}}, x_i) + b) \tag{17}$$

For the general multi class classification problem where the number of classes is larger than 2, 'one-against-one' and 'one against-all' approaches can be used. The best class label for a specific input vector is determined from voting. It is generally accepted that the 'one-against-one approach' gives a better result. Fivefold cross-validation is used to determine the final parameters of the classifier.

III. EXPERIMENT PROCESS

Experiment was carried out in the Matlab7.6 environment. Acquisition of a high-quality database of physiological signals is vital for the emotion recognition algorithm development. An important concern is the selection of signals that are to be used as input to the emotion recognition system. It is desirable that the influence of emotion on the activity of the nervous system is effectively reflected in the physiological signals employed. Unlike the case of speech recognition or facial expression recognition, where knowledge of the correct class label of a given data point is self-evident, the acquisition of a high-quality physiological signal database with confidence in the underlying emotional status is an intricate task.

It is not at all easy to judge whether the targeted emotional status is properly induced. Even if it is properly induced, the variation in physiological responses among individuals is expected to be enormous. Moreover, it is generally hard to determine whether the phenomenological changes in the physiological signals are from emotional status change or other factors, such as cognition, thought and sensory stimuli [9].

The physiological signal data of EMG is from the Augsburg University in Germany, it is four kinds of emotions, joy, anger, sadness and pleasure, generated by a subject's conduct of music by Johannes Wagner and others through the selective emotional music, with a total of Record a 25-day EMG physiological signals whose signal sampling frequency is 32Hz [10].

This paper uses a quadratural compact Daubechies 5 wavelet as the base function for six-scale decomposition of the EMG physiological signal data each day. And extract the maximum and minimum values composition vector of each layer in wavelet decomposition as the feature vector of the surface EMG signal vector, constituting a 14-dimensional feature vector. The EMG signals Waveform of sadness and wavelet transform coefficients Wf (a, b) in different scales are shown in fig.3.

Then we extract wavelet coefficients of the signal's maximum value and minimum value. We classify and save four kinds of models of the wavelet coefficients. Figure 4 shows feature extraction procedure of sadness sample of s3 completely. Figures 4(a), (b), (c) and (d) are statistical analysis of samples 1a corresponding to approximation weight a6, detail weight d6, weight d5, weight d4, detail weight d3, detail weight d2 respectively. The features of the EMG can be extracted from the statistical analysis figure, and this paper extracts the maximum and minimum value as the feature vector of different pattern of emotional.

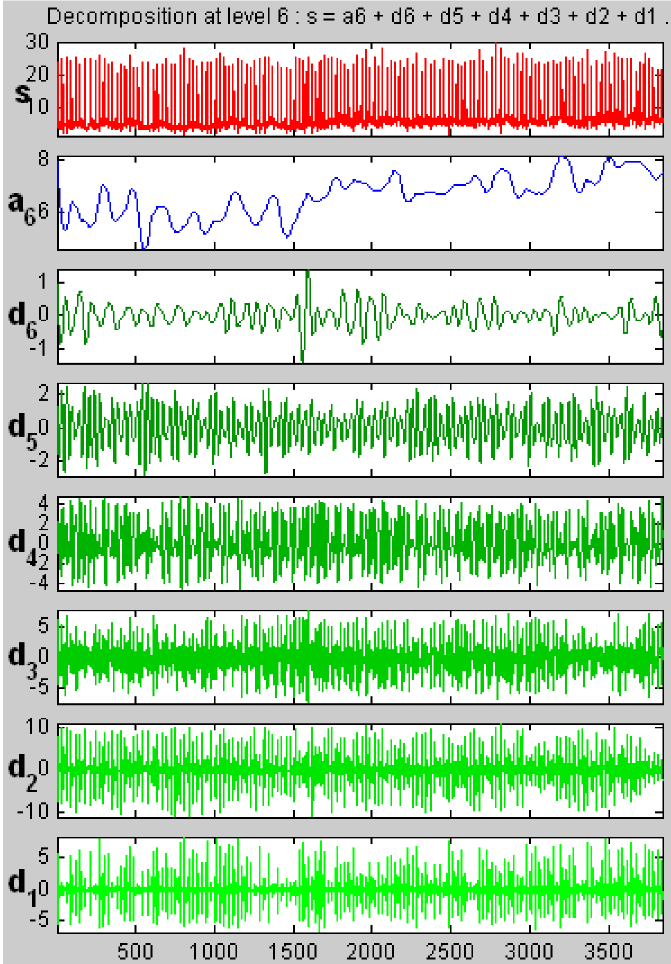
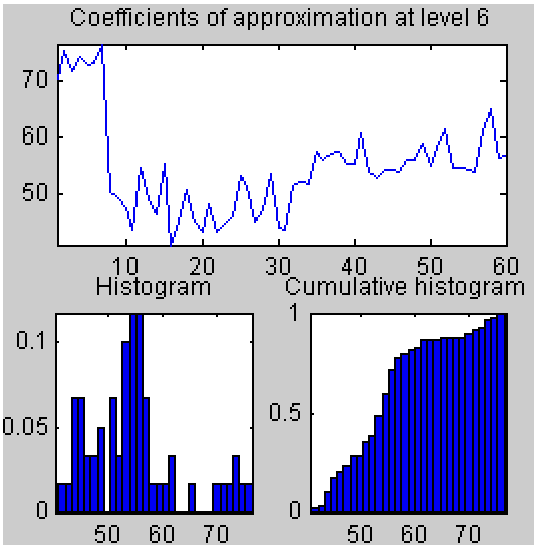
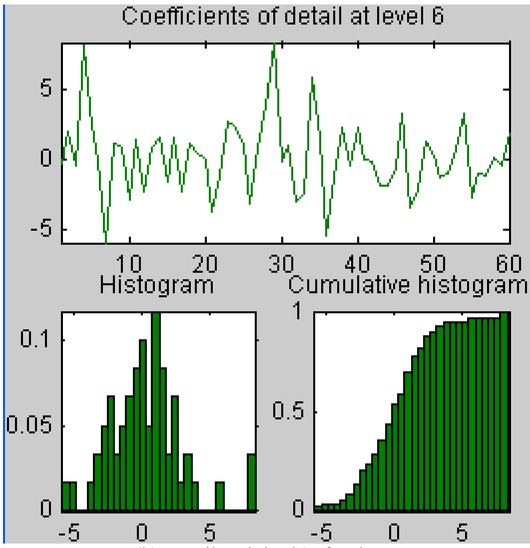


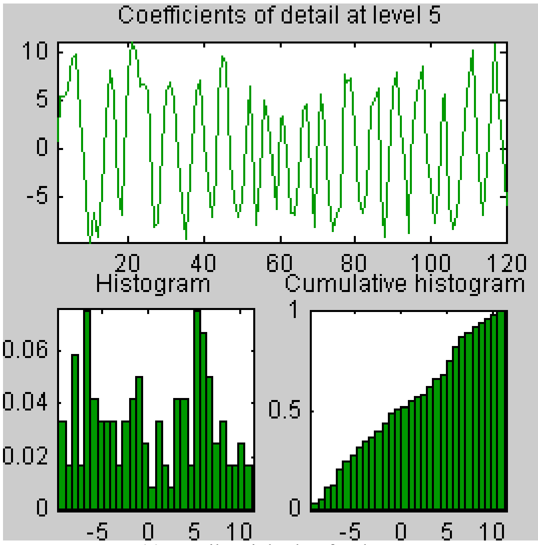
Figure 3 .The EMG signals of sadness and wavelet transform in six scales



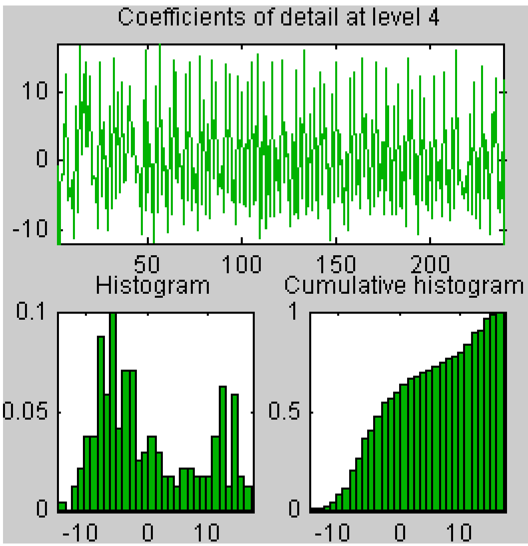
(a) Approximation weight a6 of sadness



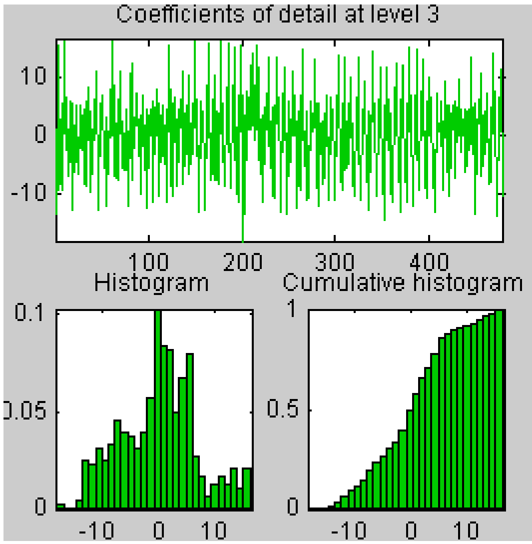
(b) Detail weight d6 of sadness



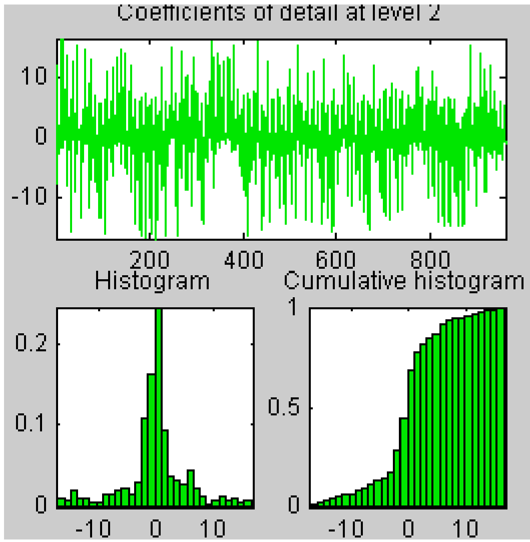
(c) Detail weight d5 of sadness



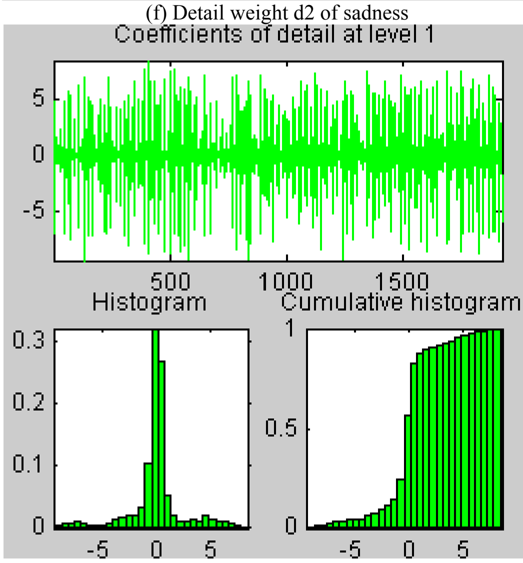
(d) Detail weight d4 of sadness



(e) Detail weight d3 of sadness



(f) Detail weight d2 of sadness



(g) Detail weight d1 of sadness

Figure 4. Complete feature extraction of sample sadness from s3

Wavelet Transform analyzed EMG parameters are statistically classified as emotional pattern joy, anger, sadness and pleasure. These parameters are then applied to ANN as training and testing data. Also, these parameters are considered as neurons in ANN. The neurons in a feed forward neural network are organized as a layered structure and connected in a strictly feed forward manner. The structure of a basic feed forward neural network is presented in Fig.1. The feed forward neural network is one of the most widely used ANN s. A great number of successful applications of this type of network have been reported [11].

So, after the six-scale decomposition procedure, we get wavelet coefficients of maximum value and minimum value of a typical emotion pattern of sadness shows as Table I and Table II. There are only six representative samples given. Each emotion classification pattern signal is composed by seven coefficients. Then decompose the six-scale wavelet to get the 14-dimensional feature vector as the input feature vector of EMG for pattern recognition of emotion classification with method of BP neural network, improved LM-BP neural network and Support vector machine.

TABLE I
THE MAXIMUM VALUE OF REPRESENTATIVE SADNESS EMG SIGNAL

	sadness						
	a6	d6	d5	d4	d3	d2	d1
1	82.51	5.792	9.752	21.09	15.56	13.77	6.406
2	56.66	4.453	7.254	11.55	11.69	10.54	5.287
3	77.68	5.62	11.88	15.22	16.54	18.02	9.093
4	89.72	9.779	15.09	18.54	19.54	17.39	8.798
5	59.27	6.432	10.22	15.93	15.49	15.86	7.524
6	60.03	6.782	8.033	12.1	10.35	9.607	4.82

TABLE II
THE MINIMUM VALUE OF REPRESENTATIVE SADNESS EMG SIGNAL

	sadness						
	a6	d6	d5	d4	d3	d2	d1
1	37.34	-6.993	-8.262	-11.41	-18.02	-14.71	-7.219
2	29.48	-5.432	-6.654	-8.705	-10.54	-11.23	-5.864
3	45.8	-7.404	-13.5	-17.24	-15.39	-16.83	-9.787
4	41.32	-4.566	-10.5	-16.3	-17.11	-18.86	-10.02
5	22.38	-7.003	-11.19	-14.46	-14.72	-15.45	-8.408
6	32.37	-7.193	-8.506	-7.885	-10.19	-11.59	-5.564

Then a three-layer BP neural network is used, input nodes is 14 and the output nodes is 4,which represents four kinds of emotional states, respectively, joy(1000), anger(0100), sadness(0010) and pleasure(0001). For the nodes selection in the middle hidden layer, experiments show that the hidden layer nodes have a significant impact for the performance of neural networks. If too few nodes, each category can not be separated by the network, and if too many nodes, the operation is too big, there maybe "over learning", therefore system performance and efficiency must be taken into comprehensive consideration to determine the hidden layer nodes. In this study, after many experiments comparison, finally select 14 as the hidden layer nodes and the effect is quite good. The training sample set is closely related to network performance. To design a good set of training samples, it is necessary to note that the sample size and also the quality of the sample, that is, the determination of the samples number and sample selection and organization. In this paper, 19 days of 25 days is selected as the training set through experimental selection and comparison and the remaining six days data of data as the test set. Learning rate on network performance impact will be greater. The experiments show that the learning rate is 0.01 and the precision control parameter is 0.01. We take Tansig function as the hidden layer activation function and the logsig function as the output layer activation function.

In this paper, the trained training function in the network training, the function is a standard gradient descent BP algorithm. Figure 5 shows a curve changed by the training process with the number of the error training. As shown in Figure 3, the network training process convergent slowly after 1000 times of training, and also vary greatly with the network training goal error.

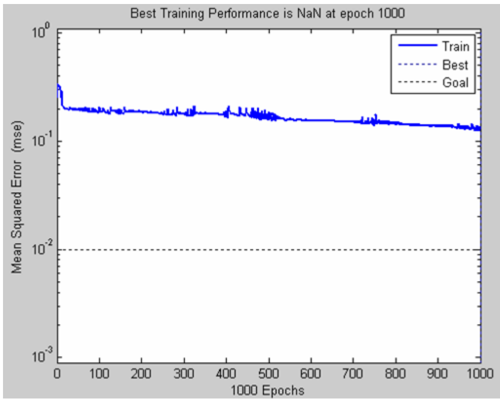


Figure 5 The training results of the standard BP algorithm (training function: trained)

Then L-M algorithm is used to train improved BP neural network classifier and the training function is trainlm. Figure 6 is a curve of error changing with the train number in the training process. As shown by Figure 6, after 273 times of training, the improved BP neural network has reached the requirements of the goal error and the training process convergence fast.

After comprehensive analysis of the above two kinds of training process, we can see that convergence speed of the L-M algorithm improved BP neural network training is fast and the network training error is small. But, at the same time, we find that identification effects of joy and pleasure by BP neural network is not very good, so we use method of Support Vector Machine (SVM) classifier for pattern recognition. The specific model, select a group and led a group of closely related variables instrumental variable, as the SVM input variables as the leading export.

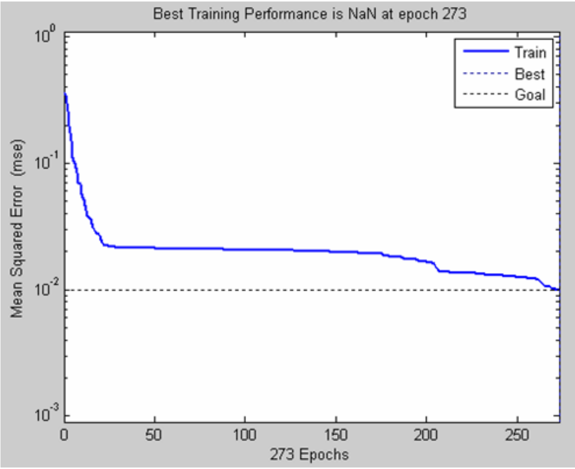


Figure 6 Training results of the LM improved BP algorithm (training function: trainlm)

Among them, the output (decision rules) is described as follows: kernal function $k(x_i, \ x) = \exp\{{-|x - x_i|^2}/{2\sigma_2}\}$ is , the right value is $a_i y_i$. K (Xn, X) is the nonlinear transformation based on n-support vector $X = (X_1, X_2, \ldots X_n)$, the input vector in this paper is $X = (X_1, X_2, \ldots X_s)$.

Enter SVM classifier training set and test set are same as BP neural network. This paper selects radial basis function as kernel function, that is, set kernel-type = 'RBF-kernel'. When Setting gam and sig2, the value of these two parameters are input parameters of the trainls svm function, among which gam is the adjustable parameters to control the degree of punishment if it is right or wrong sub-sample, sig2 is the parameters of RBF kernel function. Through many experiments in this paper the two values are selected as 1. This is the realization of support vector machine algorithms.

In the training process by using the SVM toolbox, the worst impact on the training effect is the selection of related parameters. Here we set on the gam, sig2 the value through several times of choice and setting, thus choose a group of parameters with the largest correct recognition rate as the optimal gam and sig2. Experiment result shows the SVM recognition rate recorded in the experiment through multiple choice of setting different gam and sig2.Experimental result shows that when setting the gam to 1, sig2 to 0.5, we received the best recognition rate.

The result of emotion recognition results by the use of SVM and two types of BP classifier the same set of data shows in Table III. The discrimination efficiency toward the four kinds of emotional of the standard BP neural network classifier is 62.5%, while the discrimination efficiency of the L-M improved BP neural network classifier is 83.33%. We can see that the emotion recognition using the SVM toolbox is more than 87.5% and it is feasible. This Proves that the wavelet coefficients extract from the wavelet transform can characterize emotional patterns. Support vector machine for solving linear equations used to achieve the classification makes the training time greatly reduced.

TABLE III
RECOGNITION EFFECT IN TWO KINDS OF CLASSIFIERS

	Joy	Anger	Sadness	Pleasure	Total
BP	50%	66.7%	83.3%	50%	62.5%
L-M BP	66.7%	83.3%	83.3%	100%	83.3%
SVM	100%	83.3%	83.3%	83.3%	87.5%

IV. CONCLUTION

Identifying the person's emotional state through the physiological signal has drawn increasing attention [12]. This experiment multi-scale decomposition wavelet of EMG signals by wavelet transform and extract the maximum and minimum of wavelet decomposition coefficients to constitute a signal feature vector. Then enter it into the standard BP neural network classifier and the L-M improved BP neural network classifier for emotion recognition method. These two kinds of classifiers are able to detect and identify the surface EMG of four kinds of emotions, joy, anger, sadness and pleasure. And the L-M improved BP neural network has a better recognition effect than the standard BP neural network classifier. The improved BP Network has improved the system response speed and identification accuracy and can effectively eliminate the phenomenon of over-fitting with a good, generalization ability, overcoming some shortcomings of standard BP algorithm. Thus, the L-M improved BP neural network has a great prospect in some areas such as emotion recognition.

Surface EMG analysis, however, is still in stage of the initial exploration. Particularly, not only the surface EMG is with human body movement and its physiological state, also the shape and placement location of the guide electrode have a great influence on the results of the test and analysis, thus strengthening the integration of multiple information technology and integrating a wide range of information, such as ECG, pulse, body temperature, etc., to guess people's inner feelings more accurately, it will be a further study [13,14].

In this article, emotions are classified into four categories, joy, anger, sadness and pleasure. But the actual life emotions of human are extremely rich. It is far short of expectations that sum up with four categories of emotion, and thus, how to define the type of emotion and how to categorize these categories will become a research direction in the future.

ACKNOWLEDGMENT

We wish to thank our colleagues at Taizhou University for many suggestions and stimulating discussions. This work is supported by education department Program of Zhejiang Province in University (2010).

REFERENCES

[1] FrigoC, FerrarinM, FrassonW, et al. EMG signals detection and processing for on-line control of functional electrical stimulation [J]. Electromyogram & Kinesiology, 2000 , 10 (5): 351–360.

[2] J. A. Healey: Wearable and Automotive Systems for Affect Recognition from Physiology, PhD thesis, MIT, Cambridge, MA, May 2000

[3] R. W. Picard, E. Vyzas, and J. Healey: Toward Machine Emotional Intelligence: Analysis of Affective Physiological State, IEEE Transactions Pattern Analysis and Machine Intelligence, Vol.23, No.10, pp.1175-1191, Oct. 2001

[4] A. Haag, S. Goronzy, P. Schaich, J. Williams: Emotion Recognition Using Bio-Sensors: First Step Towards an Automatic System, Affective Dialogue Systems, Tutorial and ResearchWorkshop, Kloster Irsee, Germany, June 14-16, 2004

[5] F. Nasoz, K. Alvarez, C. L. Lisetti, N. Finkelstein: Emotion Recognition from Physiological Signals for Presence Technologies, International Journal of Cognition, Technology and Work, Special Issue on Presence, Vol 6(1), 2003

[6] Xinliang Zhang, Yonghong Tan. The adaptive control using BP neural networks for a nonlinear servo-motor [J]. Journal of Control Theory and Applications, 2008, 6(3): 273-276.

[7] WU Fang-liang, SHI Zhong-kun, YANG Xiang-hui, et al. Submarine Sonar Self-Noise Forecast Based on BP Neural Network and Levenberg-Marquart Algorithm [J].Shipbuilding of China, 2006, 47(3): 45- 50.

[8] ZHANG Kun WANG Zhi-zhong. The Application of BP Neural Network Improved with LM Algorithm in Surface EMG Signal Classification[J]. Chinese Journal of Medical Instrumentation, 2005, 29(6): 399-401.

[9] K. H. Kim, S. W. Bang, S. R. Kim. Emotion recognition system using short-term monitoring of physiological signals [J]. Med. Biol. Eng. Comput., 2004, 42, 419–427

[10] YANG RuiQing, LIU Guang Yuan. Emotion Recognition Using Four Physiological Signals Based on BPSO[J]. Computer Science,, 2008, (03), 137-138.

[11] Pao, Y. H., Adaptive Pattern Recognition and Neural Networks, Addison-Wesley, Reading, MA, 1989.

[12] Wagner J, Kim J, André E, et al. From physiological signals to emotions: implementing and comparing selected methods for feature extraction and classification [C], IEEE International Conference on Multimedia & Expo, New York, 2005, 940-943.

[13] Federica Cavicchio, Massimo Poesio. Annotation of Emotion in Dialogue: The Emotion in Cooperation Project [J]. Lecture Notes in Computer Science, 2008, 5078: 233-239.

[14] Joyce H.D.M. Westerink, Egon L. van den Broek, Marleen H. Schut, et al. Computing Emotion Awareness Through Galvanic Skin Response And Facial Electromyography[J]. Philips Research, 2008, 8(2): 149-162.

Shanxiao Yang (1957-): An associate professor of taizhou university, major in electronic and information engineering, research on signal processing, intelligent robot.

Guangying Yang (1980-): A lecturer of taizhou university, major in control theory and control engineering, research on biomedical signal processing, multi sensor integration, data fusion and intelligent robot.