UDC 004.421, DOI: 10.2298/CSIS1001211Q

A Novel Hierarchical Speech Emotion Recognition Method Based on Improved DDAGSVM*

Qi-rong Mao and Yong-zhao Zhan

Department of Computer Science and Communication Engineering of Jiangsu University, zhenjiang, Jiangsu Province, 212013, China
{mao_qr,yzzhan}@ujs.edu.cn

Abstract. In order to improve the recognition accuracy of speech emotion recognition, in this paper, a novel hierarchical method based on improved Decision Directed Acyclic Graph SVM (improved DDAGSVM) is proposed for speech emotion recognition. The improved DDAGSVM is constructed according to the confusion degrees of emotion pairs. In addition, a geodesic distance-based testing algorithm is proposed for the improved DDAGSVM to give the test samples differently distinguished many decision chances. Informative features and SVM optimized parameters used in each node of the improved DDAGSVM are gotten by Genetic Algorithm (GA) synchronously. On the Chinese Speech Emotion Database (CSED) and the Audio-Video Emotion Database (AVED) recorded by our workgroup, the recognition experiment results reveal that, compared with multi-SVM, binary decision tree and traditional DDAGSVM, the improved DDAGSVM has the higher recognition accuracy with few selected informative features and moderate time for 7 emotions.

Keywords: Speech Emotion Recognition, Improved DDAGSVM, Hierarchical Recognition Method, Confusion Degree, Geodesic Distance.

1. Introduction

Speech emotion recognition plays an important role in affective computing. Up to now, many pattern recognition methods have been used in the speech emotion recognition. For example, ANN[1], HMM[2-3], GMM[4], Gaussian supervector with SVM [5].In paper [6] and [7], visual and audio signals were fused to detect emotion by using many recognition methods. Despite many promising results have been gotten in the past, it is not satisfying. Research results have shown that contributions of different features to an emotion are different[8], and the separabilities between different emotions are different[9]. This finding implies that it maybe improve the recognition accuracy of speech emotion to use effective features to recognize different emotions by the means

* Supported by the National Natural Science Foundation of China under Grant No.60673190, and the University Science Research Project of Jiangsu Province No.09KJB520002.

of hierarchical classification method. But up to now, in many of speech emotion recognition methods, classifiers are used to classify all emotions at the same layer by using the same feature subset.

SVM has good classification ability, and it is a kind of 2-class classifier. So it is very easy to use SVM to construct the hierarchical classification method. Binary decision tree is one kind of hierarchical classification method by using SVM[10]. It only needs *N*-1 SVMs when *N* emotions need to be classified. But in each node of SVM-based binary decision tree, classifier is 1-vs-r SVM, and the training samples of it are unbalance in most cases. Furthermore, the training of 1-vs-r SVM is more time-consuming than 1-vs-1 SVM.

SVM-based Decision Directed Acyclic Graph (DDAGSVM) is another kind of hierarchical classification method by using SVM[11]. In this method, 1-vs-1 SVM is used as the classification function of each node. A 1-vs-1 SVM can only exclude one class from consideration classes. For a problem with *N* classes, *N*-1 decision nodes will be evaluated in order to derive an answer, and there are *N* layers in the DDAG. And *N**(*N*-1)/2 1-vs-1 SVMs are trained in DDAG. The path taken through the DDAG SVM from root node to the last node is known as the evaluation path. Research result in paper [11] shows that DDAGSVM is amenable to a VC-style bound of generalization error, and that it yields comparable accuracy and memory usage to the standard 1-v-r SVM. But in the traditional DDAGSVM, the choice of the class order is arbitrary. The unclassifiable regions are assigned to the classes associated with the leaf nodes. In paper [12], an optimizing DDAGSVM is proposed to improve the generalization ability of DDAG. This research work enlightens us.

In another aspect, there are many feature selection methods proposed for the speech emotion recognition recently, which can be broadly classified as random or nonrandom selection[13-16]. The selected features by the random selection methods can get higher recognition accuracy on the classifier employed. Moreover, informative features selection and SVM parameter optimization can be finished synchronously by Genetic Algorithm (GA). Thus, GA is designed to select features and optimize parameters in this paper.

In this study, an improved DDAGSVM is proposed to reduce error classification cumulation of DDAGSVM, and then a novel hierarchical speech emotion recognition method is put forward by using the improved DDAGSVM. In addition, GA is used to select informative features for each class pair and optimize SVM parameters in each node.

Section 2 describes the confusion degrees of emotion pairs. Hierarchical speech emotion recognition method based on improved DDAGSVM is presented in section 3. Experiment results on the Chinese Speech Emotion Database (CSED) and the Audio-Video Emotion Database (AVED) for 7 emotions are illustrated in section 4. Section 5 draws a conclusion.

2. Confusion degrees of emotion pairs

Definition 1 (Confusion degree between two emotions): It is assumed that E_i and E_j denote the i^{th} emotion and the j^{th} emotion respectively. Then, confusion degree between E_i and E_j is the mean of the probability that the samples labeled E_i are classified into E_j and the probability that the samples labeled E_j are classified into E_i. It is denoted by I_{ij}, and it is defined as Eq. (1).

$$I_{ij} = \frac{p(r = i \mid x \in E_j) + p(r = j \mid x \in E_i)}{2} \quad . \tag{1}$$

In Eq.(1), x is a test sample, and r is the classification result. We can see clearly that, the greater the confusion degree between two emotions is, more difficultly the two emotions are differentiated. The confusion degrees of emotion pairs can be gotten by confusion matrix.

3. Improved DDAGSVM

In this paper, to improve the error classification cumulation of the DDAGSVM, an improved DDAGSVM is proposed. Compared with the traditional DDAGSVM, improved DDAGSVM has two improvements. One is the constructing method of DDAGSVM, and the other is the geodesic distance-based testing algorithm for the test samples differently distinguished. In the improved DDAGSVM, emotion pairs having smaller confusion degrees are classified in the upper nodes, and emotion pairs with greater confusion degrees are differentiated in the lower nodes. The improved DDAGSVM is constructed with many evaluation paths. According to confusion degrees of emotion pairs, an improved DDAGSVM can be set up.

In DDAGSVM, if the recognition accuracy of the test sample differently distinguished is increased, the classification ability of DDAGSVM can be improved. In the following, 1-vs-1 SVM is analyzed to find the samples which may be classified in error.

Eq. (2) is the equation of the hyperplane of SVM[17]. It makes the margin between two classes largest when $f(x)$=0. The hyperplane is denoted as H_0. In Eq.(2), $K(x_i,x)$ denotes kernel function, and x_i (i=1,2,3…m) denotes m-dimensional inputs. y_i denotes the class label, and y_i=1 for class 1 and y_i=-1 for class 2. Two parallel plane of H_0 are called as insulation plane, and they are denoted as H_1 and H_2 respectively. The equation of plane H_1 is $f(x)$=1, and the equation of plane H_2 is $f(x)$= -1. They are shown in Fig.1. For the position of a test sample x_0, there are three possible situations.

$$f(x) = \sum_{i=1}^{m} y_i \alpha_i^* k(x_i, x) + b^* \quad . \tag{2}$$

Situation 1: x_0 locates on or above the plane H_1, namely, $f(x_0) \geq 1$;
Situation 2: x_0 locates on or under the plane H_2, namely, $f(x_0) \leq -1$;

Situation 3: x_0 locates in the margin between the plane H_1 and the plane H_2, namely,$-1<f(x_0)<1$;

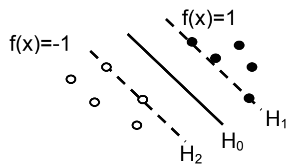

Fig. 1. Largest margin of SVM

If situation 1 or situation 2 happens, x_0 is classified as class 1 or class 2 correctly. But if situation 3 happens, x_0 may be classified by error in all probability. Therefore, the key to reduce error classification cumulation of DDAGSVM is whether the test samples in situation 3 can be classified correctly. In this paper, except the leaf nodes in the improved DDAGSVM, a new testing algorithm for each node is proposed as follows.

Algorithm 1. Geodesic distance-based testing algorithm for each node
It is assumed that T_{sv1} and T_{sv2} denote the support vector set of class 1 and class 2 respectively. It is also assumed that there are n_1 support vectors in T_{sv1} and n_2 support vectors in T_{sv2}. x_0 denotes a test sample. $f(x_0)$ denotes the distance between x_0 and the hyperplane H_0 of SVM. $GD(x_0,s_i)^1$ denotes the geodesic distance between x_0 and the i^{th} support vector in T_{sv1}, and $GD(x_0,s_i)^2$ denotes the geodesic distance between x_0 and i^{th} support vector in T_{sv2}. GD_{mean}^1 and GD_{mean}^2 denote the average distances between x_0 and support vectors of class 1 and class 2 respectively.
x_0 is classified by SVM, and $f(x_0)$ can be gotten;
The label of x_0 recognized by SVM is saved as L_0;
If ($f(x_0)\le-1$ or $f(x_0)\ge1$)
{ x_0 is classified as the class denoted by L_0;}
else if ($-1<f(x_0)<1$)
{ Compute $GD(x_0,s_i)^1$ and $GD(x_0,s_i)^2$ according to Algorithm 2;

$$GD_{mean}^1=\frac{1}{n_1}\sum_{i=1}^{n_1}GD(x_0,s_i)^1;$$

$$GD_{mean}^2=\frac{1}{n_2}\sum_{i=1}^{n_2}GD(x_0,s_i)^2;$$

If (((GD_{mean}^1- GD_{mean}^2)<0) and (L_0 is class 1))
{ x_0 is classified as class 1;}
else if (((GD_{mean}^1- GD_{mean}^2)>0) and (L_0 is class 2))
{ x_0 is classified as class 2;}
else if (((GD_{mean}^1- GD_{mean}^2)<0) and (L_0 is class 2))
{ if (| GD_{mean}^1- GD_{mean}^2|>δ)
{ x_0 is classified as class 1;}
else if (| GD_{mean}^1- GD_{mean}^2|$\leqslant\delta$)
{ x_0 is sent to both the left and the right node to be classified in further. The

---

class label of x_0 is decided by the membership degree sequences of evaluation path, which is defined in the last paragraph of this section.}}
else if (((GD_{mean}^1- GD_{mean}^2)>0) and (L_0 is class 1))
{ if (| GD_{mean}^1- GD_{mean}^2|>δ)
{ x_0 is classified as class 2;}
else if (| GD_{mean}^1- GD_{mean}^2|$\leqslant\delta$)
{ x_0 is sent to both left and right node to be classified in further;}}}

In Algorithm1, there are two reasons to use the average geodesic distance between a test sample and support vectors of one class to reflect the distance between the test sample and the class. One reason is that the geodesic distance between samples can reflect the factual distance between samples in feature space[18]. The other reason is that support vectors of one class can be regarded as the representative samples of this class. It consumes less time than calculating distance between a test sample and all the training samples of one class. The calculating method of geodesic distance between sample x_0 and a support vector s_i is shown as Algorithm 2.

Algorithm 2. Calculating method of the geodesic distance $GD(x_0,s_i)$ between the sample x_0 and the support vector s_i.
Input: x_0 and s_i;
Output: the geodesic distance $GD(x_0,s_i)$ between the sample x_0 and the support vector s_i.
1) T_{sv1} and T_{sv2} are merged into one set T, and x_0 is inserted into T. The elements in T are denoted as v.
2) Calculate the euclidean distances of every two samples in T.
3) Calculate k neighbors of each sample in T. The k neighbors of v_i is denoted as $\Omega(v_i)$.
4) Create the neighbor matrix G according to Eq.(3) in the following.

$$G(i,j)=\begin{cases} d(v_i,v_j) & v_j\in\Omega(v_i) \\ \infty & v_j\notin\Omega(v_i) \end{cases} \qquad (3)$$

Where, $G(i,j)$ denotes the element of i^{th} line and j^{th} column in matrix G. $d(v_i,v_j)$ denotes the euclidean distance between v_i and v_j.
5) Estimate the shortest distance $GD(v_i,v_j)$ between v_i and v_j according to Floyd Algorithm.

$$GD(v_i,v_j)=min(GD(v_i,v_j),\ GD(v_i,v_k)+GD(v_k,v_j)) \ . \qquad (4)$$

When v_i is x_0, and v_i is s_i, $GD(v_i,v_j)$ is the geodesic distance between x_0 and s_i. In Algorithm 1, when a test sample is sent to both the left and the right node to be classified in further, which evaluation path of the test sample is reliable? In this paper, the membership degree sequence of evaluation path is defined to judge the most reliable evaluation path. Here, if it happens in l^{th} node that the test sample is sent to both the left and the right node according to Algorithm 1, the evaluation path is checked from $(l+1)^{th}$ node to the last node. Since each node eliminates one class from the list in the evaluation path of improved DDAGSVM, the membership degrees of the test sample x_0 not belonging to emotion class 1 and emotion class 2 in the node l are defined as

Eq.(5). They are denoted by λ_l^1 and λ_l^2 respectively.

$$\lambda_l^1 = \begin{cases} 0 & f(x_0) \geq 1 \\ \frac{1}{GD_{mean}^2} & -1 < f(x_0) < 1 \\ 1 & f(x_0) \leq -1 \end{cases} \qquad \lambda_l^2 = \begin{cases} 1 & f(x_0) \geq 1 \\ \frac{1}{GD_{mean}^1} & -1 < f(x_0) < 1 \\ 0 & f(x_0) \leq -1 \end{cases} . \tag{5}$$

When the membership degrees of x_0 in each node traveled are gotten, these membership degrees make up of membership degree sequences for evaluation paths. It is assumed that there are two sequences of two evaluation paths P_j and P_k, and they are denoted as $q_j=(\lambda_{(l+1,j)}^i, \lambda_{(l+2,j)}^i, \lambda_{(l+3,j)}^i, \ldots, \lambda_{(N-1,j)}^i)$ and $q_k= (\lambda_{(l+1,k)}^i, \lambda_{(l+2,k)}^i, \lambda_{(l+3,k)}^i, \ldots, \lambda_{(N-1,k)}^i)$ $(i \in \{E_1, E_2, \ldots, E_N\})$. Lexicographic order is used as the rule to compare two membership degree sequences. The test sample x_0 is classified as the emotion class of the leaf node of P_j if and only if $q_j > q_k$ according to the lexicographic order.

4. Experiment Results and Discussions

4.1. Selecting informative features and optimizing SVM parameters

In this paper, two emotion speech corpora are used for experiments. One corpus is CSED [10], and the other is AVED [19]. These two emotion corpora are recorded by our workgroup. Each database includes short utterances covering 7 emotions. In this paper, 132 dimensional original features are extracted from the aspects of the voice of quality, energy, pitch, formant frequency, MFCCs (Mel Frequency Cepstral Coefficients), Mel Frequency energy Dynamic Coefficient and multi-fractal [19-21].

By using GA, we can select informative features for each emotion pair and optimize parameters for the corresponding SVM. Both the recognition accuracy and the number of selected features are taken into consideration in the fitness function, which is defined by Eq.(6).

$$fitness = w_r * SVM\_accuracy + 10 * w_n * features\_num^{-1} . \tag{6}$$

In Eq.(6), *SVM_accuracy* denotes SVM recognition accuracy, w_r denotes the SVM recognition accuracy weight, w_n denotes the features number weight, and *features_num* denotes the number of the features selected. The constant 10 in the second item of Eq.(6) is to make the two items keep balance. According to the fitness function of Eq. (6), w_r and w_n can influence the experiment result. We defined w_r =0.8 and w_n =0.2 in this paper according to experiment. The best chromosome is obtained when the termination criteria is satisfying. Both the selected informative features and the pair of (C, σ) for each SVM can be gotten by GA. Here, $C \in [2^{-5}, 500]$, $\sigma \in [2^{-10}, 100]$. The number of the united informative features of all emotion pairs is 46. That is to say, only 46 features need to be extracted from samples before testing. It improves the efficiency of the improved DDAGSVM.

4.2. Constructing Improved DDAGSVM

Before confusion degrees of emotion pairs are calculated according to Eq.(1), confusion matrix of 7 emotions must be gotten. It is gotten by using multi-SVMs. Feature subset and SVM parameters used in this step are gotten by GA. The confusion matrix on CESD is shown in Table 1. According to the confusion matrix, confusion degrees of emotion pairs can be gotten, and then the improved DDAGSVM is set up. The part of it is shown in Fig.2.

Table 1. Confusion matrix by multi-SVMs on CESD

%	Happiness	Sadness	Surprise	Anger	Fear	Disgust	Neutral
happiness	83.3	0	6.7	6.7	3.3	0	0
Sadness	0	80	0	0	0	13.3	6.7
Surprise	6.7	0	83.3	3.3	3.3	0	3.3
Anger	0	0	10	90	0	0	0
Fear	0	3.3	0	0	86.7	3.3	6.7
Disgust	3.3	0	0	0	0	90	6.7
Neutral	3.3	6.7	0	0	3.3	20	66.7

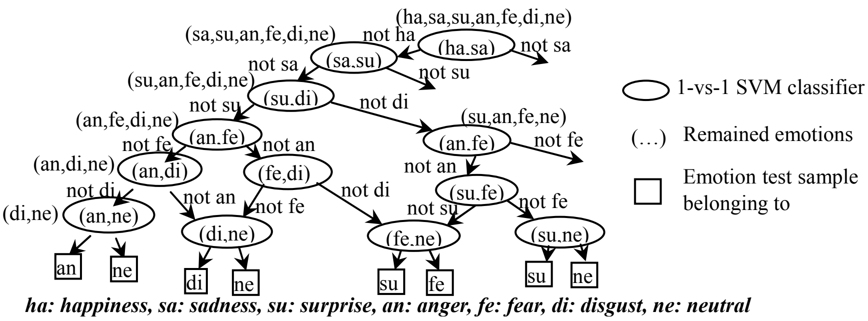

Fig.2. Part of the improved DDAGSVM for 7 emotions

4.3. Recognition experiments

For the improved DDAGSVM, the value of δ in Algorithm 1 and the value of k in Algorithm 2 are important, and they are difficult to choose. If the value of δ is adjusted to greater, the recognition accuracy may be improved, but the number of the test samples that need to be classified in many evaluation paths will increase. It will consume much time to test. On the contrary, if the value of δ is adjusted to smaller, the recognition accuracy may be decrease. In this paper, we compromised between the recognition accuracy and the test time.

50 is chosen as the value of δ according to the experiment. Moreover, according to the experiment, 6 is chosen as the value of k.

The hierarchical speech emotion recognition method based on improved DDAGSVM proposed in this paper was evaluated on two different emotion speech databases CSED and AVED. On each database, we evaluated five methods for 7 emotions:

- Multi-SVM. 21 1-vs-1 SVMs of it are executed in the same layer.
- SVM-based Binary Decision Tree(SVM-based BDT). In this decision tree, the emotions having the better clustering performance are distinguished in the upper layer, and the emotions having the worse clustering performance are distinguished in the lower layer [10].
- Traditional DDAGSVM. The choice of the class order of this method is arbitrary.
- Improved DDAGSVM not with the Geodesic distance-based testing algorithm. In this method, the DDAGSVM is constructed according to the confusion degrees of emotion pairs. The training and testing processes of this method are the same with those of the traditional DDAGSVM.
- Improved DDAGSVM with the Geodesic distance-based testing algorithm proposed in this paper (improved DDAGSVM with GD test).

In the method mentioned above, both the features and the parameters used by each SVM are selected or optimized by GA. The recognition accuracies of these five methods on the two speech emotion databases are listed in Table 2. Table 3 list the number of SVMs needed to train and to test. The mean training time and the testing time of one sample needed by the five classification methods are also listed.

Table 2. Average recognition accuracy on the two databases

Database	Classification methods	Recognition accuracy(%)
CSED	Multi-SVM	77.5
	SVM-based BDT	78.8
	Traditional DDAGSVM	78.9
	Improved DDAGSVM not with GD test	80
	Improved DDAGSVM with GD test	82.7
BDESAVED	Multi-SVM	74.8
	SVM-based BDT	76.9
	Traditional DDAGSVM	76.8
	Improved DDAGSVM not with GD test	78.9
	Improved DDAGSVM with GD test	80.5

From the results in Table 2, we can see clearly that the hierarchical methods have better recognition accuracy than multi-SVM. It proves that the hierarchical methods can improve the recognition accuracy of speech emotion recognition. Compared with the other four methods, the improved DDAGSVM with GD test proposed in this paper has the highest recognition accuracy on the two databases. And the improved DDAGSVM not with GD test has better recognition than the traditional DDAGSVM, SVM-based BDT and multi-SVM. These experiment results indicate that the construction method of DDAGSVM and the geodesic distance-based testing algorithm proposed in this paper can improve the classification ability of DDAGSVM. In the aspect of time consumed, the results in Table 3 illustrate that SVM-based BDT needs the longest time to train. It is because that the classifier in binary decision tree is 1-vs-r SVM, the training of which will consume much time. The testing time consumed by the improved DDAGSVM not with GD test is the same with that of the traditional DDAGSVM. But the improved DDAGSVM with GD test has longer testing time than the two methods, because the calculation of geodesic distance and the membership degree sequence will consume much time when the test sample is distinguished differently. Although the improved DDAGSVM with GD test will consume much time in testing, the increasing testing time is little. Compared with the least testing time 0.026s, the average testing time of the improved DDAGSVM with GD test only increases 0.006s.

Table 3. Consumed time by the five methods

Classification method	The number of SVM needed to train	The number of SVM needed to test	Training time(s)	Average testing time of one test sample(s)	Maximum testing time of one sample (s)
Multi-SVM	N*(N-1)/2 (1-vs-1)	N*(N-1)/2 (1-vs-1)	15	0.035	0.035
SVM-based BDT	N-1 (1-vs-r)	<(N-1) (1-vs-r)	26.6	0.03	0.038
Traditional DDAGSVM	N*(N-1)/2 (1-vs-1)	N-1 (1-vs-1)	15	0.026	0.026
Improved DDAGSVM not with GD test	N*(N-1)/2 (1-vs-1)	N-1 (1-vs-1)	15	0.026	0.026
Improved DDAGSVM with GD test	N*(N-1)/2 (1-vs-1)	>N-1 (1-vs-1)	15	0.032	0.041

5. Conclusions

In this paper, a novel hierarchical method based on improved DDAGSVM was proposed for the speech emotion recognition. Improved DDAGSVM is set up according to the confusion degrees of emotion pairs. In addition, a geodesic distance-based testing algorithm is proposed to give the test samples differently distinguished many decision chances. The informative features for each emotion pair are selected by GA while the SVM parameters are optimized by GA. Before testing, only 46 features are extracted from testing samples. This makes training and testing of the improved DDAGSVM more quickly. The classification experiments have been done on two speech emotion databases CSED and AVED. The experiment results reveal that, compared with the other four methods, the improved DDAGSVM with the Geodesic distance-based testing algorithm has the highest recognition accuracy. The disadvantage of it is that the testing time is longer. But the increasing extent of it is small. It is still fit for the real-time applications. This is worth to be studied in the further work. The method proposed in this paper can be applied in other pattern recognition fields.

6. References

1. Bhatti M. W., Wang Y., Guan L.: A Neural Network Approach for Human Emotion Recognition in Speech. In Conf. ISCAS. Vancouver, BC, Canada, 181-184. (2004)
2. Nwe T. L., Foo S. W., Silva L. C. D.: Speech Emotion Recognition Using Hidden Markov Models. Speech Communication, Vol. 41, No. 4, 603-623. (2003)
3. Schuller B., Rigoll G., Lang M.: Hidden Markov Model-Based Speech Emotion Recognition. In IEEE Int. Conf. ICASSP. New York, USA, 1-4. (2003)
4. Moataz M. H., Mohamed S., Fakhri K.: Speech Emotion Recognition Using Gaussian Mixture Vector Autoregressive Models. In IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP). Beijing, China, IV-957-960. (2007)
5. Hu H., Xu M. X., Wu W.: GMM Supervector Based SVM with Spectral Features for Speech Emotion Recognition. In Conf. of ICASSP. Beijing, China, 413-416. (2007)
6. Wang Y. J., Guan L.. Recognizing Human Emotional State From Audiovisual Signals. IEEE TRANSACTIONS on Multimedia,Vol.10, No.4, 659-667. (2008)
7. Zeng Z. H., Tu J. L., Brian M. P., Thomas S. H.: Audio–Visual Affective Expression Recognition Through Multistream Fused HMM. IEEE Transaction on Multimedia, Vol.10, No.4, 570-577. (2008)
8. Lin Y. L., Wei G., Yang K. C.: A Survey of Emotion Recognition in Speech. Journal of Circuits and Systems, Vol.12, No.1, 90-98 (in Chinese). (2007)
9. Dimitrios V., Constantine K.: Emotional Speech Recognition: Resources, Features, and Methods. Speech Communication, Vol. 48, 1162-1181. (2006)
10. Mao Q. R., Wang X. J., Zhan Y. Z., Pan Z. G.: Speech Emotion Recognition Method Based on Selective Features and Decision Binary Tree. Journal of Computational Information Systems, Vol. 4, No.4, 1795-1801. (2008)
11. John C. P., Nello C., John S. T.: Large Margin DAGs for Multiclass Classification, S.A. Solla, T.K. Leen and K.-R. M¨uller (eds.), MIT Press, 547–553. (2000)
12. Takahashi, F., Abe, S.: Optimizing Directed Acyclic Graph Support Vector Machines. In Proc. of ANN in Pattern Recognition, 166–170. (September 2003)
13. Huang C. H., Wang C. J.: A GA-based Feature Selection and Parameters Optimization for Support Vector Machines. Expert Systems with Applications, Vol. 31, No. 2, 231-240. (2006)
14. Ververidis D., Kotropoulos C.: Fast and Accurate Sequential Floating Forward Feature Selection with the Bayes Classifier Applied to Speech Emotion Recognition. Signal Processing, Vol.88, 2956-2970. (2008)
15. Altun H., Polat G.: Boosting Selection of Speech Related Features to Improve Performance of Multi-class SVMs in Emotion Detection. Expert Systems with Applications, Vol. 36, No. 4, 8197- 8203. (2009)
16. Wang X. J., Mao Q. R., Zhan Y. Z.: Speech emotion feature selection method based on contribution analysis algorithm of neural network. AIP Conference proceedings, Vol.1060, 336-339. (2007)
17. Hsu C. W. and Lin C. J.: A Comparison of Methods for Multiclass Support Vector Machines. IEEE Transaction on Neural Networks, Vol. 13, No. 2, 415-425.(2002)
18. You M. Y., Chun C., Bu J. J., Tao J. H.: Analysis on Nonlinear Manifold. In proc. of ICPR'06, Vol.3, 91-94. (2006)
19. Mao Q. R.: Research on the feature extraction and the recognition metod of speech emotion. Ph. D. thesis. Zhenjiang, Jiangsu University, 27-28 (in Chineese). (2009)
20. Roddy C., Randolph R. C.: Describing the Emotional Classes That are Expressed in Speech. Speech Commuication, Vol. 40, 5-32. (2003).

21. Zhan Y. Z., Cao P.: Research and Implementation of Emotional Feature Extraction and Recognition in Speech Signal, Journal of Jiangsu University (National Science Edition), Vol. 26, No.1, 72-75(in Chinese). (2005)
22. Lin Y. L.: Research on Speech Emotion Recognition. Ph.D. Thesis. Guangzhou, South China University of Technology, 51-57, (in Chinese). (2006)

Qirong Mao got her M.S. degree in Computer Science from Jiangsu University in China, in 2002. She is currently an associate professor at the Department of Computer Science and Communication Engineering in Jiangsu University at Zhenjiang (China). Now she is working on the affective computing for the Ph.D.. Her current research interests include Human-Computer Interaction (HCI) and Computer Supported Collaborative Work (CSCW).

Yongzhao Zhan is a Professor of Computer Science at Jiangsu University, Zhenjiang, China. He got his Ph.D. in Computer Science from Nanjing University, China, in 2000. His current research focuses on Human-Computer Interaction (HCI), Computer Supported Collaborative Work(CSCW) and the security of wireless network.

Received: May 05, 2009; Accepted: September 29, 2009.