



Contents list available at JMCS

Journal of Mathematics and Computer Science

Journal Homepage: www.tjmcs.com



Speech Emotion Recognition by Using Combinations of C5.0, Neural Network (NN), and Support Vector Machines (SVM) Classification Methods

Mohammad Masoud Javidi and Ebrahim Fazlizadaeh Roshan

Department of Computer Science, Shahid Bahonar University of Kerman, Kerman, Iran.

javidi@mail.uk.ac.ir

Article history:

Received December 2012

Accepted March 2013

Available online April 2013

Abstract

Speech is the fastest and most natural method for human to communicate. This has led several researches to be done in the field of the interaction effects between human and machine. Hence, it is necessary to design machines which can intelligently recognize the emotion of a human voice. However, we are still far from having a natural interaction between the human and machine because machines cannot distinguish the emotion of the speaker. This has established a new field in the literature, namely the speech emotion recognition systems. The accuracy of these systems depends on various factors such as the number and type of the emotion manners as well as the feature selection and the classifier sort. In this paper, classification methods of the Neural Network (NN), Support Vector Machine (SVM), the combination of NN and SVM (NN-SVM), NN and SVM (NN-SVM), NN and C5.0 (NN-C5.0), C5.0 and SVM (SVM-C5.0), and finally the combination of NN, SVM, and C5.0 (NN-SVM-C5.0) have been verified, and their efficiencies in speech emotion recognition have been compared. The utilized features in this research include energy, power, Zero Crossing Rate (ZCR), pitch, and Mel-scale Frequency Cepstral Coefficients (MFCC). The presented results in this paper demonstrate that using the proposed NN-C5.0 classification method is more efficient in recognizing the emotion states-to the extent of 6%- to 30% depending on the number of emotions states-than SVM, NN, and other aforementioned combinations of classification methods.

Keywords: Keywords: Emotion recognition; Feature extraction; Mel-scale Frequency Cepstral Coefficients; Neural Network; Support Vector Machines; C5.0;

1. Introduction

Dealing with speech is one of the latest achievements of the today's technology. Numerous aspects of speech processing have been recognized so far including the speech emotion recognition-which is one of the most important matters in our daily lives. By the speech emotion recognition, we mean recognizing the speaker's emotions by using their speech samples. The speech emotion recognition

has several applications in, for example, web movies, computer tutorial, call center, in-car board system, and etc. [6].

In the latest three decades, several attempts have been made in order to recognize the speech emotion whose most important ones in the recent years include [1] -[5]. The voice and prosodic features, the speaking style, speaker's characteristics, and the linguistic features can affect the emotional [5].

Several models such as hidden Markov models (HMM) in [9], Gaussian mixture model (GMM) in [5], [14], [15], NN in [13], [8], and SVM in [5], [16] have been utilized in order to recognize the speech emotion. It is well investigated in [7] that the SVM and HMM lead to the most and least recognitions, respectively.

We propose in this paper the NN-C5.0 method in order to model the speech emotion recognition systems more precisely. Regarding the existing features in the speech processing such as energy, power, ZCR, pitch, MFCC, and etc., we extracted 68 of such statistical features from an emotion database-which has been implemented in German language by professional individuals in Berlin University.

Emotion states of anger, happiness, fear, sadness, disgust, boredom, and neutral have been considered. We applied these recognizing emotion states to different classification models of SVM, NN, NN-SVM, NN-C5.0, and NN-SVM-C5.0 by using the CLEMENTINE software and verified the results. The presented results in this paper demonstrate that using the proposed NN-C5.0 classification method is more efficient in recognizing the emotion states-to the extent of 6% to 30% depending on the number of emotions states-than SVM, NN, and other aforementioned combinations of classification methods.

This is how the rest of the paper is organized: in section II, we briefly review the most recently proposed speech emotion recognition systems. In the following of this section, we verify the database of Berlin. Section IV and V hold the feature extraction and utilized models, respectively. In section VI, we present the experimental results and also compare different methods. In the last section, the conclusion of the paper is presented.

2. Related work

Speech emotion recognition systems are constructed of three components: 1. Feature extraction; 2. Feature selection; and 3. Classification. The latter component of the system, i.e. classification, predicts the emotional state.

In the 1990's, most recognizing models were proposed based on Linear Recognize Classification (LDC) [17] and Maximum Likelihood Bays (MLB) [10], [27]. In the recent years, on the other hand, GMM [5], [12], [14], NN [13], [8], Multilayer Perceptron (MLP) [5], K-Nearest Neighbor (KNN) [11], [14], HMM [9], and SVM [5], [16] were used to recognize speech emotion.

Table I shows the utilized models in the recent years along with a number of emotions and their recognition rates. In [27], Haq et al. used the MLB classification model and the following seven emotions: the anger, disgust, fear, happiness, neutral, sadness, and surprise. They could attain the recognition rate of 53%. In [10], Ververidis and Kotropoulos used the MLB classification model and the anger, happiness, neutral, and sadness emotions and could achieve the recognition rate of 53.7%.

In [8], Yu et al. used the SVM and ANN classification models and the following four emotions: anger, happiness, neutral, and sadness, and could reach the recognition rates of 71% and 42% respectively for the SVM and ANN models, which demonstrate that the SVM classification model efficiently outperforms that of the ANN. In [9], Ayadi et al. by using the HMM and ANN classification models and seven emotions could gain the recognition rates of 71% and 55% for the HMM and ANN models, respectively, which reveals that the HMM model performs better the one of ANN. In [11], Petrushin used the KNN model and the anger, happiness, sadness, fear, and neutral emotions and reached the recognition rate of 70%.

Table I. The most recent utilized models and their associated results

Paper	Emotion number	Classifier	Recognition rate
PetrushinVA (2000) [11]	5	KNN	70%
Yu F, Chang e, Xu y, Shum h (2001) [8]	4	SVM , ANN	71%,42%
Ververidis D, Kotropoulos c (2006) [10]	5	MLB	53.7%
Ayadi M, Kamel S, Karray F (2007) [9]	6	ANN, HMM	55%,71%
Haq S, Jackson PJB, Edge J (2008) [27]	7	MLB	53%
Gharavian D, Sheikhan M, Pezhmanpour M(2011)[12]	4	GMM	65.1%
He L, Lech M, Maddage NC, Allen NB (2011) [14]	5	GMM, KNN	77%,77%
Sheikhan M, Bejamin M, Gharavian D (2012)[5]	3	GMM, C5.0, MLP, MODULAR NEURAL-SVM	65.9%,56.3%,68.3%,76%
Hamidi M, Mansorizadeh M (2012)[13]	5	NN	78%
FersiniE, Messina E, Archetti F (2012)[16]	6	SVM	70.83%

In [12], Gharavian et al. used the GMM model and four emotions and could reach the recognition rate of 65.1%. In one of the most recent researches in the literature in [5], Sheikhan et al. used three emotions of happiness, anger, and neutral and two classification models, namely modular neural-SVM and C5.0. Their findings showed the recognition rates of 76.3% and 56.3% for the former and latter models, respectively.

3. Berlin database of emotional speech

The Berlin's database of the emotional speech is used to classify discrete emotion states. This database is one of the most applicable ones used to recognize the speech emotion state [7] based on which several works have been presented (for example, as two of the most recent researches, we refer to [16] and [18] proposed in 2011).

This database has been implemented in the Technical University of Berlin. There are seven emotion states utilized in this database: the anger, happiness, boredom, sadness, fear, disgust, and neutral. Ten artists, including 5 men and 5 women, implemented the database in German language by saying 10 sentences, including 5 short and 5 long ones with time duration between 1.5 to 4 seconds. Samplings were done as the single channel and by 16 kHz frequency. From these obtained speeches, it is possible to recognize seven real emotions of human. Number of speech files are approximately 800. Audio files associated with 7 emotions were classified as follows: anger (127), boredom (81), disgust (46), fear (69), happiness (71), and neutral (79).

4. Feature extraction

One of the most important parts of speech emotion recognition systems is the feature extraction. Features that we have extracted to do this research are: energy, power, ZCR, pitch, MFCC which is described below [21], [19].

Pitch: is a very sensitive factor responds to the auditory sense, also called fundamental frequency, refers to the periodic time of a wave pulse generated by air compressed through the glottis from the lungs. In our work, we employed Autocorrelation Function (ACF) method, which is one of the popular methods in pitch tracking process to extract pitch-related features. It helps us to find out the similarity between the signal and a shifted version of itself to have peaked at multiples of the fundamental frequency.

Energy: also known as power or energy, the intensity of a voice can be physically detected through the pressure of sounds or a subjective level of noisiness. Normally, the simple intensity is the sum of the absolute values for each data frame.

Power: the relative intensity value, usually used in speech signal processing for its familiarity with the normal sense of hearing. The range of sound that human's ear can hear is from the ratio of the sound pressure that causes permanent damage from short exposure to the limit that (undamaged) ears can hear is more than a million. In order to simplify the representation of a large range, logarithmic units are used.

ZCR: is considered as one of the duration-related features to be used in our experiment, which represents the number of times that the speech signals crossing the zero point. It can be easily calculated by counting the times that the wave touches the level zero reference. Instead of speech rate mentioned in the psychology, ZCR is more appropriate for language-independent speech recognition.

MFCC: MFCC, a very common groups of features used in automatic speech recognition (ASR), which convert the basic features (pitch, power Db, and phase features) into a 12 MFCC features. They are synthesis features may have no physical meaning.

We extracted 68 features, 56 of which are obtained from 7 statistical data (minimum, maximum, mean, median, standard deviation, range, and variance) of energy, power, pitch, and ZCR. Discrete Fourier Transform (DFT) is applied to these features and their statistical data have been extracted. 12 features of the MFCC have also been calculated.

Table II. Extracted features of each statement

Feature	Normal	DFT
Pitch	Pitch_max	fft_Pitch_max
	Pitch_mean	fft_Pitch_mean
	Pitch_min	fft_Pitch_min

	Pitch_median Pitch_range Pitch_std Pitch_var	fft_Pitch_median fft_Pitch_range fft_Pitch-std fft_Pitch_var
Energy	energy_max energy_mean energy_min energy_median energy_range energy_std energy_var	fft_energy_max fft_energy_mean fft_energy_min fft_energy_median fft_energy_range fft_energy_std fft_energy_var
Power	power_max power_mean power_min power_median power_range power_std power_var	fft_power_max fft_power_mean fft_power_min fft_power_median fft_power_range fft_power_std fft_power_var
ZCR	zcr_max zcr_mean zcr_min zcr_median zcr_range zcr_std zcr_var	fft_zcr_max fft_zcr_mean fft_zcr_min fft_zcr_median fft_zcr_range fft_zcr_std fft_zcr_var
MFCC	mfcc1 mfcc2 mfcc3 mfcc4 mfcc5 mfcc6	mfcc7 mfcc8 mfcc9 mfcc10 mfcc11 mfcc12

All the extracted features are shown in Table 2. In order to extract the feature, each voice is divided into windows of length 320 sections, and the overlapping size is chosen to be 20. Features have been extracted by using the MATLAB software. The calculation methods of these features are presented in the following (in all equations, N and $s(n)$ stand for the window length and the values of samples in the time domain, respectively) [20].

(A) The energy (E) of a signal frame of length N is obtained by $E = \sum_{n=0}^{N-1} s(n) * s(n)$

(B) The power (P) of a signal frame of length N is obtained by $P = \frac{1}{N} \sum_{n=0}^{N-1} (s(n) * s(n))$

(C) The ZCR of a signal frame of length N is obtained by $ZCR = \frac{1}{2} \sum_{n=0}^{N-1} |Sign(s(n)) - Sign(s(n - 1))|$

(D) The pitch feature extraction method of a speech signal frame of length N whose n th sample is denoted by $s(n)$ is obtained by (the autocorrelation method):

In this method, we draw the alternation of the autocorrelation function ($r(\mu)$) versus the frame samples (μ). The distance of the first peak is the period of pitch [20], which is the maximum value of $r(\mu)$ associated with the μ th sample. The following equation shows how $r(\mu)$ is calculated for each μ .

D- For each speech frame, the Fourier transform and the power spectrum was calculated. The power spectrum was then mapped from the linear scale (with frequency in Hz) scale onto the mel-frequency scale using:

$$f_{mel}(f) = 2595 \log_{10} \left(1 + \frac{f_{Hz}}{700} \right)$$

Finally, the discrete cosine transform (DCT) of the mel log powers was calculated, and the first 12 DCT coefficients provided the MFCC values corresponding to a given frame.

5. Classifiers

Classification is another component of a speech emotion recognition system. In this research, we used three classification methods: NN, SVM, and C.5 as well as their combinations. In the following of this section, we briefly explain the three major ones:

5.1. Neural Network (NN)

NN is a method used for the pattern recognition. It is, in fact, a simulation model of the human brain's performance by which a feature vector is classified. Supposing that the feature vector is as $X=[X_1, X_2, \dots, X_n]$, then a neuron unit-which is the main core of a NN-can be presented as Figure 1.

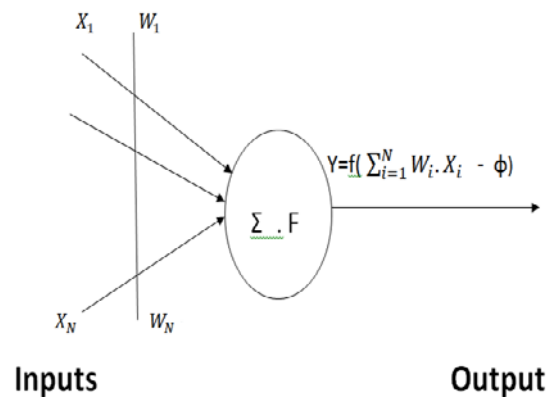


Figure 1. A neuron of a neural network

A neuron unit, indeed, is a super plane in the feature space separating two areas. The value of ϕ denotes the threshold of the neuron unit; and W_i and W represent the weight vectors and the superplane, respectively [26]. NNs have numerous applications in, for example, the simulation of the flying path and vehicle scheduling [22]. They also are applicable in the process of speech emotion recognition [7], [8], [13].

5.2. SVM classifier

The SVM is one of the learning with observation methods which has been used in the classification and regression. In the recent years, this comparatively new method is shown to outperform other older classification ones such as NN.

This method is used to recognize speech emotion state which had resulted in a very good performance [5], [7], [17]. The base of the SVM method is to linearly classify data; and in the linear division of data, that line which has the highest safety margin is selected.

By using the quadratic programming (QP) methods-which are well-known in solving problems with constraints-the optimum line for data is obtained. Before the linear division, in order that the machine can classify data with high complexity, data are transformed into a space with very higher dimensions by using the “phi” function.

In order to solve the problem with very high dimensions by using these methods, the Lagrangian dual theorem is utilized to transform the intended minimization problem in its dual form. In the dual form, instead of the complicated high-dimension “phi” function, a simpler one called the core function is appeared. It is possible to use such various core functions as exponential, polynomial, and sigmoid [23].

5.3. C5.0 classifier

In the recent years, C5.0 has been used as a popular method in data mining. However, it has not been used to recognize emotion till 2012. During this year, Sheikhan used the method, but the results were not satisfying enough [5]. This algorithm is the developed version of the C4.5 and ID3 [24], and it is performed by either constructing the decision tree or utilizing a set of rules.

A C5.0 model works by analyzing the sample based on the quality which has the maximum received information. Each sub-sample defined by the first analysis will be analyzed again. This analysis is mainly done based on various fields. The analysis process continues until the sub-samples can no longer be analyzed. Finally, the analyses of the lowest stage are examined again and those which add no cost to the model are either omitted or trimmed [25].

6. Implementation and results evaluation

In this research, we used the SVM, NN, and C5.0 classifiers as well as their combinations (i.e. NN-SVM, SVM-C5.0, NN-C5.0, NN-SVM-C5.0) in the CLEMENTINE environment; then we stored the data in the EXCEL environment and mined them by the CLEMENTINE software. There were 68 features as inputs which were obtained by programming in the MATLAB environment from statements in the Berlin’s database. Our output was a set of emotion states (anger, happiness, boredom, sadness, fear, disgust, and neutral).

Our statements were 535, 20% of which were randomly chosen for the test while another 80% were used in the learning process. We implemented each set of data 10 times by using the SVM and NN classifiers and their combinations with C5.0, and averaged over the obtained values of recognitions. In our first experiment, we tried to recognize two emotional states of neutral and anger.

Our total statements were 206, 127 of which were associated with the anger state and the other 79 statements were related to the neutral state. The results of each test were obtained by 10 times implementations for the following classification methods: SVM, NN, NN-SVM, SVM-C5.0, NN-C5.0, and NN-SVM-C5.0. The values of the average recognitions associated with these experiments are shown in Table III. As this table demonstrates, for the NN-C5.0 and SVM-C5.0 methods, the obtained recognition rates for the neutral and anger states are approximately the same, whilst they are better than others as much as 3% to 8%.

In our second experiment, we tried to recognize three emotional states of anger, happiness, and neutral. Like the former case, we used 68 features. Our total statements were 272: 128, 78, and 66 statements were accordingly associated with the anger, neutral, and happiness states. The results of each classification test are shown in Table IV. As this table shows, the recognition rate of the NN-SVM-C5.0 is 2% better than that of NN-C5.0. Also, these two classification methods perform much better than others.

In the following, we examined 4, 5, 6, and 7 emotion states whose results are presented in Table VI. This table demonstrates that for four emotion states, NN-C5.0 classification method outperforms SVM-C5.0, NN-SVM-C5.0, NN-SVM, SVM, and NN methods to the extent of approximately 3%,

11%, 14%, 14%, and 26%, correspondingly. Also, for five emotion states, it outperforms SVM-C5.0, NN-SVM-C5.0, NN-SVM, SVM, and NN methods as much as about 1%, 7%, 13%, 13%, and 11%, respectively. Furthermore, for six emotional states, it outperforms SVM-C5.0, NN-SVM-C5.0, NN-SVM, SVM, and NN methods to the extent of approximately 7%, 6%, 13%, 16%, and 32%, correspondingly. Finally, for seven emotional states, it outperforms SVM-C5.0, NN-SVM-C5.0, NN-SVM, SVM, and NN methods as much as about 10%, 11%, 18%, 20%, and 32%, respectively.

Table III. The average recognition rates for the anger and neutral states

Counter Execute	1	2	3	4	5	6	7	8	9	10	Accuracy Mean
NN	84.21	91.49	83.33	88.37	86.36	88.64	85.29	87.1	90.7	74.29	85.978
SVM	86.36	80.95	83.78	95.45	93.48	94.12	79.55	86.84	86.05	89.19	87.577
NN-SVM	84.62	91.49	93.02	91.43	93.88	88.64	87.5	94.87	81.58	91.89	89.892
NN-C5.0	87.5	92.68	92.31	95	89.74	95.12	94.87	91.11	91.43	86.49	91.625
SVM-C5.0	93.18	94.12	90.7	91.3	93.48	94.74	88.89	86.84	87.1	100	92
NN-SVM-C5.0	91.67	96.88	94.23	95	87.5	97.67	92.5	95	90.7	90.32	84.11

Table IV. The average recognition rates for the anger, happiness, and neutral states

Counter Execute	1	2	3	4	5	6	7	8	9	10	Accuracy Mean
NN	79.55	76	71.93	69.49	74.07	67.31	75.81	58.93	71.88	70	71.49
SVM	78.33	71.64	68.52	75	71.19	77.55	64.44	74.55	78.95	81.36	74.153
NN-SVM	84	81.67	84.09	72.92	68.63	77.36	84.62	79.07	77.68	76.27	78.63
NN-C5.0	73.21	81.13	79.25	79.25	81.13	83.33	87.8	80.39	85.25	90.48	82.122
SVM-C5.0	75.44	78.57	86.27	79.59	78.33	79.03	78.12	85.71	73.33	83.33	79.772
NN-SVM-C5.0	88.89	83.93	85.96	86.79	85.71	78.57	77.08	85.96	80.36	89.47	84.272

Table VI. Average recognition rates for 4, 5, 6, and 7 emotion states

Models	4 Emotions Anger, sadness, happiness, neutral	5 Emotions Anger, happiness, disgust, sadness, fear	6 Emotions Anger, happiness, fear, sadness, disgust, boredom	7 Emotions Anger, happiness, fear, sadness, disgust, boredom, neutral
NN	56.272	51.368	43.183	39.41
SVM	68.272	62.561	59.732	53.215
NN-SVM	68.331	62.573	62.047	55.01
NN-C5.0	82.214	75.807	74.946	72.621
SVM-C5.0	79.397	74.842	67.381	61.221
NN-SVM- C5.0	71.701	68.936	68	62.246

7. Conclusion and discussion

In this paper, we used pitch, energy, ZCR, power, and MFCC to recognize speech emotion states. We analyzed 535 emotional speeches from Berlin's database. A set with 68 features were calculated for each statement. Three classification methods: SVM, NN, and their combinations with C5.0 were applied.

According to our findings, for neutral and anger states the recognition rate of NN-C5.0 and SVM-C5.0 models was approximately 92% and they slightly outperform NN, SVM, NN-SVM, and NN-SVM-C5.0 models. For anger, happiness, and neutral emotional states, the recognition rate of NN-SVM-C5.0 was 2%, 5%, 6%, 20%, and 13% better than those of NN-C5.0, SVM-C5.0, NN-SVM, SVM, and NN models, respectively.

For more emotion states, we obtained the following results: For four, five, six, and seven emotional states, the proposed NN-C5.0 classifier performed better than the other ones as much as 3% to 26%, 1% to 24%, 6% to 31%, and 10% to 33%, correspondingly. It is evident that the proposed NN-C5.0 classification method is more accurate than the other ones used in this paper.

8. References

- [1] T. Pao and Y. Chen, J. Yeh, Y. Chang. Emotion recognition and evaluation of mandarin speech using weighted D-KNN classification. *Int. Innov. Comput. Info. Control.* 4, 1695- 1709(2008).
- [2] H. Altunand, G. Pollat., Boosting selection of speech related features to improve performance of multi-class SVMs in emotion detection. *Expert Syst. Appl.* 36, 1897-8203(2009).
- [3] M. L. Yang. Emotion recognition from speech single using new harmony feature. *Single Process.* 90, 1415-1423(2010).
- [4] L. He and M. Lech, N.C. Maddage, N.B. Allen. Study of empirical mode decomposition and spectral analysis for stress and emotion classification in natural speech. *Biomed Signal Process Control.* 6, 139-146 (2011).
- [5] M. Sheikhan and M. Bejamin, D. Gharavian. Modular neural-SVM scheme for speech emotion recognition using ANOVA feature for method. *Neural Comput & Applic*(2012).

- [6] B. Schuller and G. Rigoll, M. Lang. Speech emotion recognizing combining acoustic features and linguistic information in a hybrid support vector machine-belief network architecture, in proceeding of the ICASSP, 1, 397-401(2004)
- [7] M. Ayadi and M.S. Kamel, F. Karray. Survey on speech emotion recognition: features, classification schemes, and databases. Pattern Recognition. 44, 572–587(2011).
- [8] F. Yu and E. Chang, Y. Xu, H. Shum. Emotion detection from speech to enrich multimedia content. In proceedings of the IEEE Pacific Rim conference on multimedia. Advances in multimedia information processing, 550-557(2001).
- [9] M. Ayadi and S. Kamel, F. Karray. Speech emotion recognition using Gaussian mixture vector autoregressive models. In proceeding of the international conference on acoustics, speech, and signal processing, 5, 957-960(2007).
- [10] D. Ververidis and C. Kotropoulos. Fast sequential floating forward selection applied to emotional speech features estimated on DES and SUSAS data collection, In proceeding of the European signal processing conference, 1-5(2006).
- [11] V.A. Petrushin. Emotion recognition in speech signal: experimental study, development, and application. In proceedings of international conference on spoken language processing, 222-225(2000).
- [12] D. Gharavian and M. Sheikhan, M. Pezhmanpour. GMM-based emotion recognition in Farsi language using feature selection algorithms, World ApplSci J 14:626-638 (2011).
- [13] M. Hamidi and M. Mansorizadeh. Emotion recognition from Persian speech with NEURAL NETWORK, International Journal of Artificial Intelligence & Applications (IJAI), 3, No.5(2012).
- [14] L. He and M. Lech, N.C. Maddage, N.B. Allen. Study of empirical mode decomposition and spectral analysis for stress and emotion classification in natural speech. Biomed Signal Process Control 6: 139–146(2011).
- [15] R. Kohavi and G.H. John. Wrappers for feature subset selection. ArtifIntell 97:273–324(1997).
- [16] E. Fersini and E. Messina, F. Archetti. Emotional states in judicial courtrooms: an experimental investigation. Speech Commun 54:11–22(2012).
- [17] F. Dellaert and T. Polzin, A. Waibel. Recognizing emotion in speech. In proceedings of international conference on spoken language processing, 3, 1970–1973(1996).
- [18] W.U. Siqing and H. Tiao, C. Wai-yip. Automatic speech emotion recognition using modulation features. Speech Communication 53:768–785(2011).
- [19] L.R. Rabiner and R.W. Scheafer. Introduction to Digital Speech Processing. The essence of knowledge, Boston- Delft(2007).
- [20] J.R. Deller and J.G. Proakis. Hanson discrete-TIME Processing of Speech Signals. Macmilan, New York(1993).
- [21] J. Rong and G. Li, Y.P. Chen. Acoustic feature selection for automatic emotion recognition from speech. Information Processing and Management 45 :315–328(2009).
- [22] M.T. Hagan and H.B. Demuth, M. Beale. Neural networks design. Boston: PWS(1996).
- [23] J. Christopher and C. Burges. A Tutorial on Support Vector Machines for Pattern Recognition. Data Mining and Knowledge Discovery 2, 121 – 167(1998).
- [24] J. Quinlan. Programs for machine learning. Morgan Kaufmann, San Francisco, CA, (1993).
- [25] Clementine®12.0 Clementine Modeling Nodes(Chapter 6), (2007).
- [26] L. Fausett. Fundamentals of Neural Networks, Prentice-Hall(1994).
- [27] S. Haq and P.J.B. Jackson, J. Edge. Audio-visual feature selection and education for emotion classification. In proceedings of international conference on auditory-visual speech processing, 185–190(2008).