퍼지 및 지능시스템학회 논문지 2001, Vol. 11, No. 7, pp. 611-614

Design of Intelligent Emotion Recognition Model

Yi-gon Kim

Dept. of Electrical Engineering. Yosu National University
Yosu 554-749, Korea

Abstract

Voice is one of the most efficient communication media and it includes several kinds of factors about speaker, context emotion and so on. Human emotion is expressed in the speech, the gesture, the physiological phenomena(the breath, the beating of the pulse, etc). In this paper, the emotion recognition method model using neuro-fuzzy in order to have cognizance of emotion from voice signal is presented and simulated.

Key words : Emotion recognition, Voice signal, Wavelet transform, Neuro-Fuzzy model, Emotion Engineering

1. Introduction

In recent years, we have been interested in artificial Intelligence, according to the rapid development of life environment, to solve the problems for human being in various fields of human society. Moreover, with the progress of the study about human intelligence, the study of emotion is required little by little too. Emotion is a sense and thought caused by external physical stimulation, that is psychological experience, joy, hate, anger, etc. This study for emotion is of practical value in various field of industry, being representative of that is the active interface between the human and the machinery. The active interface doesn't mean that user make an offer with menu to a machinery, then the machinery makes a response to the request. Rather this means that the machine grasp the user's motion and psychology when user make an offer to machinery, and then the machine respond suitably to that situation. For this reason, to grasp the user's motion and psychology(emotion) and estimate user's emotion is needed for this active bidirectional interface technology.

New field of study that is spotlighted now is emotion engineering that is to apply a quantitative and qualitative analysis of human's emotion for design of products to be safe and convenient for the human life.. In our country, electric home appliance company is in charge of organizing the emotional engineering study group for an emotional design of electric home appliance.

Japan carried out a project "development of human sensation measurement application" in connection with emotion engineering from 1990. Also, in USA MIT leads others in this field.

There are some kinds of methods to recognize basic emotion from expressed language, facial expression and gesture, physiological phenomena, and voice signals [1-6]. Each method has a special feature respectively. First is passive and difficult, because used language must express an emotional mean and languages are very much. Second is visual method that uses a camera and microprocessor to recognize emotion. So this recognition method is very complex. Third is more accurate than others. But it is not natural because to wear. Last is not only simple and natural, but also it has not limit in application.

In past studies, researches use prosodic parameters (intensity, rhythm and speed) for a recognition of emotion using voice signals in speech[4]. In this study, a new clarification of the definition of emotional factors in voice signal is proposed. Proposed recognition model which interferes the emotion with an analyzed physical parameters of the voice signals is shown. The recognition model is realized by traditional neuro-fuzzy model. This model is investigated, which is able to recognize emotion in speech.

2. Extraction of emotional parameters in voice signals

There are many ways to recognize emotion, but in this study, we proposed the emotion recognition model that recognize speaker emotion from his voice signals.

It is very difficult to extract effective information data that express emotion in his voice signals. Yanaru proposed the mathematical model of emotion expression[1,2]. Yanaru's study that human emotion can be divided in eight of basic emotions is based on P.Plutchik and P.T.Young's. R. Plutchick and P.T. Young stands on their opinion that human has some of basic emotion and expressed human emotion is mixed with those in their studies. But this study has a defect that this is complex and difficult. Because the past and the present emotional states of speaker are necessary to express one's emotion.

In this study, we propose a simpler a recognition method, that inferences a present emotional state from only speaker's

Manuscript received September 18, 2001 ; revised November 10, 2001.

voice signals. For this purpose, the extraction of emotional parameters is based on expression method of joy and anger together with sorrow and pleasure with rhythm of Korean classical music. Emotional parameters are the magnitude and the variation of a energy distribution, the duration of "A", wavelet factors of one's a tone in "A, G-RUB-SSUM -NI-GGA?" is used for analyzing the quality.

We used those parameters to recognize the eight emotional states of a speaker.

3. Construction of Experiment device and Data acquisition

Fig.1 shows data acquisition system that is constructed with a microphone (Brüel Kjæl Type 2671) and a LPF(NF Electronic Instrument FV-664(fc is 15KHz)), A/D converter (This is AD2838 that is product of Sam-Hyung Data System Co.) And we acquired experiment data with sampling Time 40KHz. As is shown in Fig.1, microphone is to transfer voice signals to electric signals, next this signal is amplified in pre-amplifier, last amplified signal becomes an input of A/D converter. And then, in PC this digital signals is analyzed with analyzing algorithms.

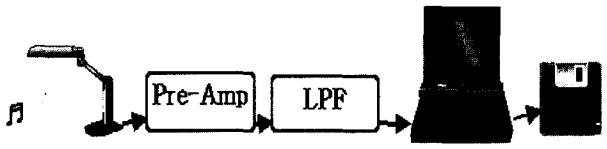

Fig. 1 Structure of signal acquisition system

In experiment, we asked speaker to say "A, G-RUB-SSUM-NI-GGA?" repeatedly three times on each eight of basic emotions. In this study, eight of basic emotions are joy, sadness, expectancy, surprise, anger, fear, hate, acceptance. Speakers are fives. For various expression of emotion, we allowed the speaker to express their own emotion with their voice tones and intonations.

4. Analyzing algorithms and experiment

Experiment system is as shown in Fig.1. we ask five speakers to say "A, G-RUB-SSUM -NI-GGA?" in three times that cotains respective basic emotions. and then it is analyzed by proposed algorithms. This algorithms is as follows.

1 step. start algorithms
2 step. measure the duration of voice.
3 step. measure the energy distribution of voice signals (voice intensity)
4 step. analyze the voice signals with wavelet transforms.-that is multi-filter
5 step. infer emotion with neuro-fuzzy model.
6 step. stop algorithms

If you are interested in wavelet, then you can refer the reference papers[14,15]. Fig.2 is a flow diagram of emotion recognition algorithms that show the flow of analyzing algorithms of voice signals and inference model.

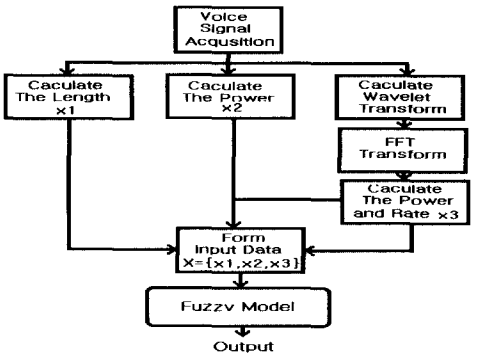

Fig.2 Flow diagram of recognition algorithms

we get the Input vector that is X([x1,x2, x3]) in quantitative analysis, at the conclusion we learned the recognition model with the input vectors. Fig. 3 is sample data that is acquired from emotional speaker's speech in time domain.

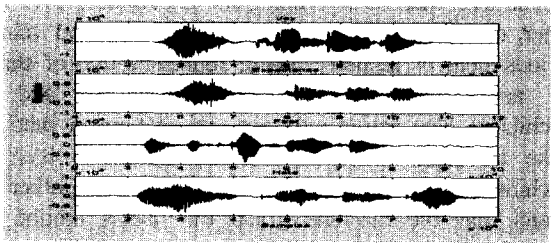

Fig. 3. Samples of Voice Signals

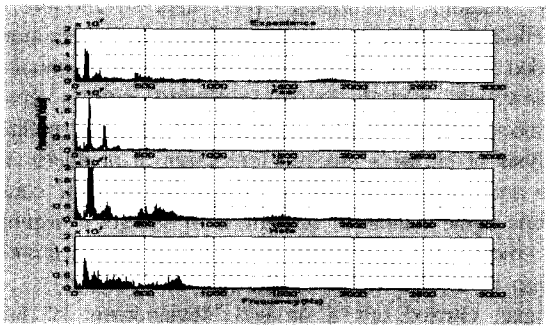

Fig. 4. FFT spectrums of sample's voices

Fig.4 is the frequency spectrums of voice signals that are analyzed by using FFT. In Fig.4, we can distinguish each spectrum that has a characteristic features. But it is dificult to extract a characteristic data from signals. For resolving this problem, wavelet transform is used to extract pattern data from voice signals. Fig.5 is result of Wavelet transform analysis.

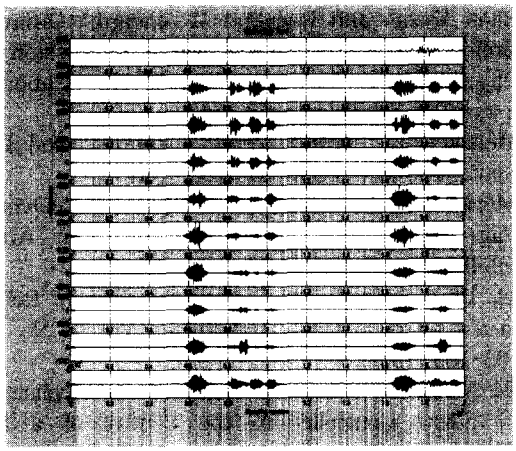

Fig. 5. Result signals of Wavelet Transform(Fear)

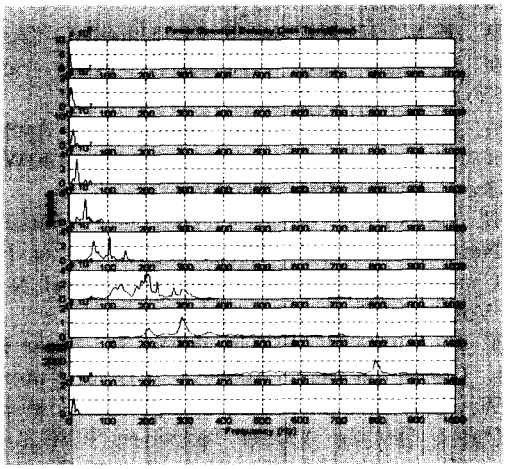

Fig. 6. FFT spectrums of result signal of Wavelet transform

As is shown in Fig. 4., we know that characteristic feature of voice signals appeared in frequency band according to each kind of emotions that differs from emotions to emotions. Fig. 6 shows the spectrum distribution of wavelet transform that is analyzed by using FFT.

Table 1. Characteristic data of voice signals

Characteristics / Emotions	Intensity(x1)	Period of tones (x2) (Samples,40KHz)
Surprise	284	7000
Anger	442	6000
Sadness	203	11000
Expectancy	166	10000
Acceptance	308	17000
Joy	312	13000
Hate	200	12000
Fear	162	8000

Table 2. Characteristic datas of x3

Channel / Emotions	S	a8	d1	d2	d3	d4	d5	d6	d7	d8
Surprise	2.7e9	37.2	0.05	0.25	0.79	5.25	27.6	40.6	24.5	28.8
Anger	6.1e9	0.07	0.06	0.3	1.6	8.4	32.2	41.0	16.34	0.17
Sadness	6.7e8	5.8	0.05	0.16	0.4	1.6	3.88	22.6	48.8	13.5
Exp.	3.58e8	10.7	0.21	1.27	3.94	6.0	8.32	9.4	32	28.6
Accept.	1.9e9	0.93	0.04	0.38	2.03	10.43	25.76	15.6	31.5	13.6
Joy	2.1e9	4.65	0.11	0.51	1.76	4.86	8.51	19.43	48.9	8.59
Hate	6.2e8	1.93	0.06	0.56	3.28	13.39	19.44	19.77	26.33	15.75
Fear	5.14e8	3.95	0.02	0.09	0.3	0.99	5.55	30.63	47.95	11.08

We analyzed the rate of low and high frequency energy distribution by means of analyzed result of frequency spectrums.

The analyzed result of voice signals that expressed eight basic emotions is represented in Table.1 and Table.2, in Fig7.

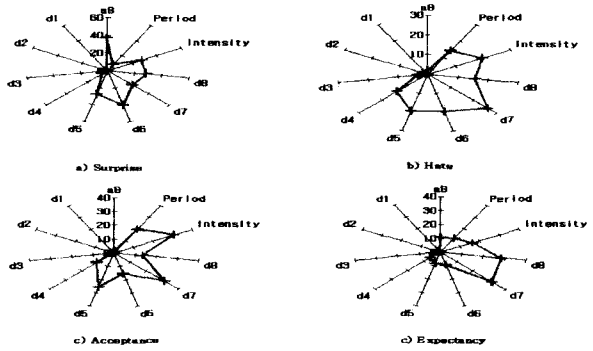

Fig.7. Characteristic data

5. Design of emotion recognition model

Fig.7 is emotion recognition model that is constructed by using ANFIS(Adaptive Neuro Fuzzy Inference System) for recognizing emotion from voice signals[8-13].

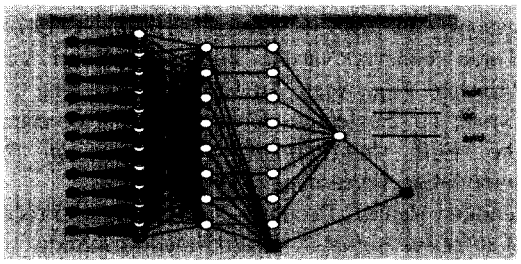

Fig. 8. Structure of Nero-Fuzzy Model

Fig.9 is the surface of learned nero-fuzzy Model with learning data, This model is evaluated with experiment data, and this result is shown in Fig.9. As a result, inference error is about 10%.

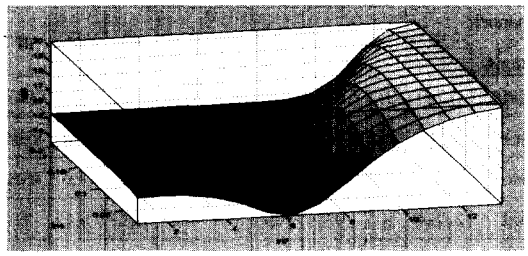

Fig. 10. Surface of learned Nero-Fuzzy Model

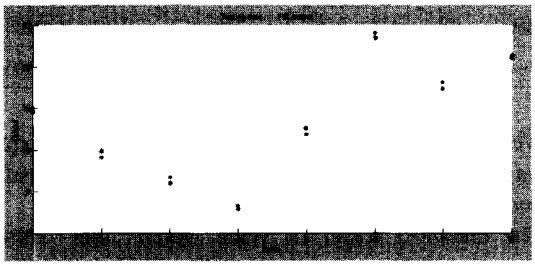

Fig. 11. Result of Simulation

6. Conclusion

In this paper, we proposed emotion recognition model that recognizes one's emotion using voice signals only. The input vectors of recognition model are an intensity, a period of tones and an energy distribution shape of frequency spectrums on voice signals.

We design the recognition model using acquired input vectors, and estimate this model with real data. Fig.9 is result of this simulation in which win rate is about 85%. But this experiment is restricted within five speakers. Henceforth, we will study to apply this for making intelligent toy.

References

[1] Torao Yanaru, Naruki Shirahama, Kaori Yoshida, Masahiro Nagamatsu, "An Emotion Processing System Based on Fuzzy Interference and Subjective Observations." Information Sciences 101(3-4): 217-247 (1997)

[2] Torao Yanaru, Naruki Shirahama, Kaori Yoshida, Masahiro Nagamatsu,"An Emotion Processing System Based on Subjective Observation," Proceedings of the Fifth International Conference on Neural Information Processing (1998) pp.475-477

[3] Jennifer Healey, Rosalind Picard,"Digital Processing of affective signals", picard@media.mite.edu

[4] Tsuyoshi Moriyama 'and Shinji Ozawa,"A Mweasurement of Human Vocal Emotion Using Fuzzy Control."

[5] L.W.Cambell, D.A.Becker, A.Azarbayjani, A.F.Bobick, and A.Pentland, "Invariant feature for 3-D gesture recognition", in proceedings, International Conference on Automatic face and Gesture Recognition, Pp 157-162, Killington,VT, 1996, IEEE

[6] Elias Vyzas and Rosalind W. Picard," Offline and Online Recognition of Emotion Expression from Physiological Data", MIT Media Laboratory, e-mail:Picard@Media.mit. edu

[7] Medsker, L., Hybrid Neural Networks and Expert Systems. New York: Kluwer, 1994

[8] Mozetic, I., Model-Based Diagnosis: A Overview. Page 419-430 of: Advanced Topics in Artificial Intelligence. Springer-Verlag, 1992

[9] M. Braee and D. A. Rutherford, "Fuzzy Relations in a Control Setting", Kybernetes Vol. 7, No. 3, pp. 185-188, 1978.

[10] E. M. Scharf and N. J. Mandic, "The Application of a Fuzzy Controller to the Control of a Multi-Degree-Freedom Robot Arm", in Industrial Applications of Fuzzy Control, M. Sugeno, Ed. Amsterdam: North-Holland, pp. 41-62, 1985.

[11] W. J. M.. Kicker and E. H. Mamdani, "Analysis of a Fuzzy Logic controller", Fuzzy Sets an Systems Vol. 1, No. 1, pp. 29-44, 1978.

[12] W. J. M.. Kickert, "Futher Analysis and Application of Fuzzy Logic Control", Interal Rep. F/WK2/75, Queen Mary College, London, 1975.

[13] R. Parekh, J. Yang, V. Honavar," Constructive Neural-Network Learning Algorithms for Pattern Classification", IEEE Trans. on Neural-Network ,Vol.11, No.2, pp. 436-451,March,2000

[14] T. K. Sarkar, C. Su, R. Adve, M.Salazar-Palma, L. Garcia-Castillo, Rafael R. Boix,"A Tutorial on Wavelets from an Electrical Engineering Perspective, Part 1: Discrete Wavelet Techniques."IEEE Antennas and Propagation Mag.,Vol.40, No.5, pp49-69, Oct.1998

[15] F. X. Canning, J. F. School," Diagonal Preconditioners for The EFIE Using a Wavelet Basis," IEEE Trans. on Ant. and Prop.,Vol. 44,No.9,pp.1239-1246, 1996

Yi-gon Kim

received the MS degree in avionic electric engineering from Hankuk aviation university, Seoul Korea, in 1986 and 1988 respectively. He recieved Ph.D. degree in electrical engineering from Chonnam National university in 1993. He performed research at Tokyo Institute of Technology by research member in 1991 and at Iowa State University by visiting professor in 2000-2001. He is an associate professor in the School of ECC at Yosu National University. His interests fuzzy-Neuro Modeling and its application to diagnosis of industrial systems and control of intelligent systems