Sound Quality Indicating System Using EEG and GMDH-type Neural Network

Kiminobu Nishimura and Yasue Mitsukura
Department of System Design Engineering, Faculty of Science and Technology, Keio University, Yokohama, Japan
E-mail: nishimura@mitsu.sd.keio.ac.jp, mitsukura@sd.keio.ac.jp Tel: +8145-566-1718

***Abstract*—In this paper, we propose a sound quality evaluation system using electroencephalogram (EEG) and group method of data handling (GMDH) type neural network. Recently, EEG is used in various applications, and we focus on sound quality evaluation using EEG. We prepared EEG samples to train a GMDH-type neural network to recognise 3 typical types of sound which was used to create the training data. The results showed that using GMDH-type neural network improved recognition rate compared to the other method. Additionally, we repeated simulations by using different parameter of GMDH-type neural network, and the open test results showed the recognition rate variations in different parameter values.**

I. Introduction

Recently, loudspeakers are used in consumer scenes frequently according to the development of digital media compression technologies and other multimedia technologies including online digital media stores which sell music. Additionally, opportunity of listening to music is increasing by integration of music player function to many portable electronic devices containing smart phones and computers. The users select their preferable media and devices such as loudspeakers, headphones and amplifiers to play music to achieve satisfaction or pleasantness [1]. In other words the users select their devices by its appearances and characteristics. The manufacturers of these devices put great number of efforts to design and create such devices having preferable appearances and characteristics determined by a significant number of users. However, it is usually difficult to determine what sound quality of loudspeakers cause user's preferences; therefore various researches were presented to evaluate them using questionnaires [2]. The results obtained by the questionnaires are usually obscure since the valuation basis of the sound quality is determined by human. There are differences among individuals, for example, some people are familiar to care about sound qualities, and some may not. Thus there are significant errors of evaluation if people did not answer questions correctly. This error makes the sound quality evaluating system useless, hence an efficient evaluating methods are needed.

In this research, we define sound quality as a value of pleasantness of the listener when the evaluating sound is played. Therefore, we use biological signals, e.g. electroencephalogram (EEG) to evaluate the sound quality since it is caused by human emotion including pleasantness. There are several researches using biological signal especially EEG which recognizes human emotion of the listeners and being used to musical therapies [3]-[5]. These researches create systems which measure single or multiple channel EEG and outputs estimated emotion types, but it can not be an effective evaluation because of its discrete property. The ideal output of the sound quality can be thought as some numeric values so that the evaluating target can be compared. In this paper, we propose a sound quality evaluating system using single channel EEG to output a numeric value which indicates the sound quality.

II. EEG analysis and the proposed method

Biological signals including EEG can be used to recognize human emotions [6]. There are two types of electroencephalograph to measure EEG signals. One is the multiple channel EEG, which use a head gear type device having multiple channel electrodes positioned referring to the international 10-20 system. The other device measuring EEG is the single channel electroencephalograph which and can be worn immediately compared to the multiple channel ones. According to its compact body, this device can only measure the Fp1 position based on the international 10-20 system. This position is at the left forehead of a human, which will include the electric potential caused by neuron activities in the frontal lobe. Since this part of the human brain relates to emotions, using only Fp1 position is enough to estimate the emotion. The common analysis of EEG is to measure the EEG before the task and after, and calculate the spectrum by applying various frequency analysis method, for example fast Fourier transform (FFT), and divide the amplitude of the mean of the EEG before the task, for each frequency from the after. The next process would be principal component analysis (PCA), followed by the discriminant analysis, for example, Fisher's linear discriminant analysis to discriminate the different EEG caused by types of sound qualities [7]. This conventional method using PCA is not robust against the data which sometimes become irregular. PCA tries to get the information from the combinations of variables having the most variance, so if some irregular data are used to train the system, it will cause errors leading to overfitting. In this paper, we use a group method of data handling (GMDH) type neural network [8] to perform feature extracting which will replace PCA in the conventional method. GMDH type neural network uses what so called "exterior criterion", which will grow a neural network appropriately to prevent over-fitting by using an evaluation function. The arguments of this function are the values outputted when unseen data is set to the input of the network. This will lead to prevent over-fitting.

We used the conventional method of extracting EEG by the following method. First we perform FFT to the entire measured EEG signal. This is resulted by the FFT parameters of 1 second window size with no time shifting, and because the frequency outputted by the transform were decreased to only in the range of 4-22Hz. Since we measured 180 seconds of EEG, we obtain 19 degree vectors in number of 180. The next process is to extract the EEG difference between before and after the task. This process is operated by the following equations. The transformed data $\mathbf{X}$ can be described as below:

$$\mathbf{X} = \begin{bmatrix} x_4(1) & x_5(1) & \cdots & x_{22}(1) \\ x_4(2) & x_5(2) & \cdots & x_{22}(2) \\ \vdots & \vdots & & \vdots \\ x_4(T) & x_5(T) & \cdots & x_{22}(T) \end{bmatrix}. \tag{1}$$

In the equation above, $x_f(\tau)$ indicates the amplitude value of f Hz at τ th second. T is the measured time, 180. The divided EEG after by the average of before the task, $\tilde{\mathbf{X}}$ can be calculated by the follwing equation.

$$\tilde{\mathbf{X}} = \begin{bmatrix} \frac{x_4(T-29)}{\bar{x}_4} & \frac{x_5(T-29)}{\bar{x}_5} & \cdots & \frac{x_{22}(T-29)}{\bar{x}_{22}} \\ \frac{x_4(T-28)}{\bar{x}_4} & \frac{x_5(T-28)}{\bar{x}_5} & \cdots & \frac{x_{22}(T-28)}{\bar{x}_{22}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{x_4(T)}{\bar{x}_4} & \frac{x_5(T)}{\bar{x}_5} & \cdots & \frac{x_{22}(T)}{\bar{x}_{22}} \end{bmatrix} \tag{2}$$

where

$$\bar{x}_j = \frac{1}{30}\sum_{i=1}^{30} x_j(i) \tag{3}$$

Finally, we apply a moving average filter.

$$\tilde{\mathbf{X}}_{\mathrm{SMA}_t} = \frac{1}{5}(\tilde{\mathbf{x}}_t + \tilde{\mathbf{x}}_{t+1} + ... + \tilde{\mathbf{x}}_{t+4}), (t = 1, 2, ..., 26) \tag{4}$$

We now have 26 $\times$ 19 matrix data. As a consequence, there are usable data of 26 for each EEG measuring. We then construct a GMDH type neural network and train this network by using the prepared data. This network is grown automatically by its method, and the polynomial function of the neuron used is set as $f(\omega, \mathbf{u})$ which is shown below.

$$f(\omega, \mathbf{u}) = \omega_1 + \omega_2 u_1 + \omega_3 u_2 + \omega_4 u_1 u_2 + \omega_5 u_1^2 + \omega_6 u_2^2, \tag{5}$$

where ω_i is the weight for the i th term and u_j is the j th input of the neuron $(j = 1, 2)$. We also set that each neuron can take an input from the input layer in the network. The network will only have one output neuron, which will be a numeric value. The target signal were set to -1, 0 and 1 for EQ1, EQ2 and EQ3, respectively. The evaluation function to investigate the neuron is calculated by the least square method using the training data and its target signal. GMDH type neural network is an algorithm which automatically determines the scale of the neural network by creating layers. Accordingly, there will be a parameter that determines when to stop the network grow. Fig. 1 shows the effect of the value of this parameter "delta". Every time GMDH type neural network creates a new neron layer, an evaluation value to each neuron can be calculated by using the exterior data. This value indicates how the created neural network fits to the exterior data, and it will be lower the better. The vertical axis in Fig. 1 is the criterion value which is the least evaluation value of the all neuron in the created layer. The horizontal axis shows the layer number; therefore GMDH type neural network will keep creating a new layer unless the criterion value of the working layer is less than one on the previous layer. However, this also means that the system being more complex leading to over-fitting. To construct neural network being appropriately scaled, parameter "delta" can be set. This is a value which stops the algorithm when the new criterion value did not get better in a certain value than the one on the previous layer. An appropriate setting of this parameter is necessary to avoid over-fit and suppress calculation costs. From Fig. 1 we can see that by setting the parameter higher, the scale of the neural network is kept smaller than when it is lower.

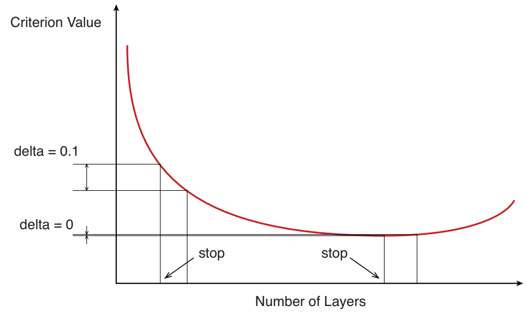

Fig. 1. Effect of the value of parameter delta

We ran two simulations which include closed test and open test. For the closed test, we used 3 types of EEG data, EQ1, EQ2 and EQ3 and recognized them by training a neural network with parameter "delta" set to 0 and compared with a method using PCA. For the open test, we only used 2 types of EEG data, EQ1 and EQ3, and calculated the recognition rates. In this case, "delta" was set to 0.005. This value was selected by numerous simulations which results in highest recognition rates.

III. Experiment

We prepared three types of sound qualities by creating different frequency responses indicated on Fig. 2, and applied those equalizers to a specific pair of loudspeakers so the features other than the frequency response will be the same. Equalizer 1 (EQ1) in Fig. 2 represent a flat response, and EQ3 is a response which cuts below 200Hz, over 5000Hz and boost around 2000Hz. This sound quality refers to a sound

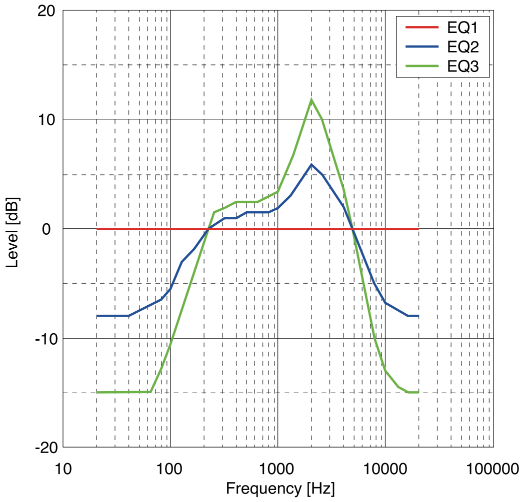

Fig. 2. Frequency response for the three sound qualities

TABLE I
RECOGNITION RATES USING PCA IN CLOSED TESTING

Subject	Recognition Rate[%]	Subject	Recognition Rate[%]
1	44.02	17	44.23
2	41.03	18	39.53
3	42.31	19	47.44
4	40.60	20	44.87
5	40.17	21	39.53
6	40.81	22	41.24
7	41.45	23	49.79
8	49.57	24	44.23
9	39.32	25	42.74
10	42.52	26	41.88
11	34.40	**27**	**50.00**
12	41.45	28	44.02
13	45.09	29	44.87
14	43.16	30	48.29
15	44.87	31	38.25
16	47.44	32	40.38
		Average	43.11 ± 3.59

of an AM radio. EQ2 represents the sound quality between EQ1 and EQ3. Therefore, in this research we define the good sound quality as the sound quality of the loudspeakers itself, and define bad sound quality as a sound quality which is like an AM radio. We gathered 32 subjects, including male and female with the age between 20-50 years old. We ordered each subject a task to close their eyes to prevent artifacts for EEG, and measuring EEG for 30 seconds of rest followed by 120 seconds of music listening with each sound qualities, and an additional 30 seconds of rest. While rest, music are stopped playing. We asked each subject to do the task 6 times for each sound quality in the same day. We used an hair band type electroencephalograph measuring Fp1 position according to the international 10-20 system, and set the reference electrode to the ear tab. The electroencephalograph measures EEG by a sample rate of 128Hz.

TABLE II
RECOGNITION RATES USING GMDH TYPE NEURAL NETWORK IN CLOSED TESTING

Subject	Recognition Rate[%]	Subject	Recognition Rate[%]
1	64.86	17	55.00
2	68.22	18	61.14
3	73.75	19	62.67
4	61.78	20	66.25
5	61.00	21	60.00
6	61.89	22	62.33
7	57.33	23	63.33
8	53.20	24	63.33
9	54.69	25	66.75
10	63.78	26	62.25
11	61.43	**27**	**74.40**
12	64.44	28	63.78
13	66.00	29	55.41
14	55.24	30	68.44
15	68.44	31	58.22
16	56.22	32	64.44
		Average	62.50 ± 5.20

IV. RESULTS AND DISCUSSIONS

We calculated the recognition rate for recognising the 3 sound qualities, by analysing the EEG signals. We used the threshold method and Fischer linear discriminant analysis for the feature extraction by GMDH type neural network and PCA, respectively. 10-fold cross validation was operated for both methods, using PCA and GMDH type neural network, and the results of the final values are indicated by recognition rate on TABLE I and TABLE II, respectively. The bold font items on both tables represents the highest score within all subjects. We can notice that the method using GMDH neural network improved the recognition rate for all subjects. Fig. 3 indicates the plots for each data obtained by the described feature extraction in the previous section and applying PCA, selecting the first two principal components. Both methods used the same data. Red, blue and green plots describes data of EQ1, EQ2 and EQ3 respectively. Fig. 4 indicates the histogram for the output of the neural network. The output values are settled around from -1 to 1, because the target signal were set as it is. The light coloured red, blue and green indicates the data for EQ1, EQ2 and EQ3. The horizontal axis is the output value of the network, and the vertical axis is the count of data within the range of output values. The covered area are shown in mixed colours, indicated in the legend within the figure. We set the threshold to -0.5 and 0.5 for recognising EQ1, EQ2 and EQ3. Both results on Fig. 3, 4 shows results for only subject 3. This subject showed a significant difference between the two methods that we can see large covering area of multiple colour plots in Fig. 3, but the histogram on Fig. 4 is clearly separated. This has resulted in recognition rate difference on TABLE I and TABLE II, which difference is the highest in all subjects. We could not put all plots and histograms on limited paper area, but it was common that the principal component plot had large covered areas, comparing to that the GMDH type neural network method had fewer covered areas which causes miss recognition.

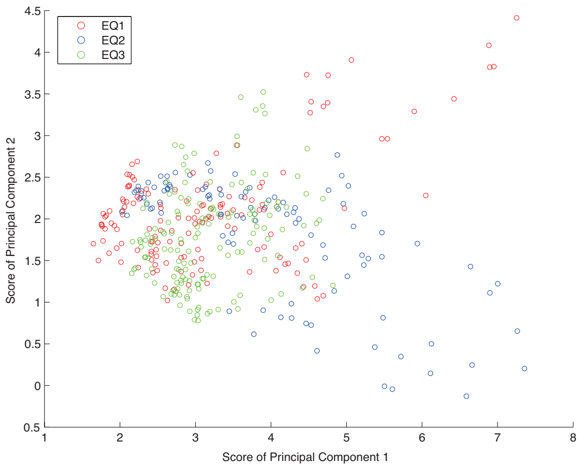

Fig. 3. Principal component plots for subject 3

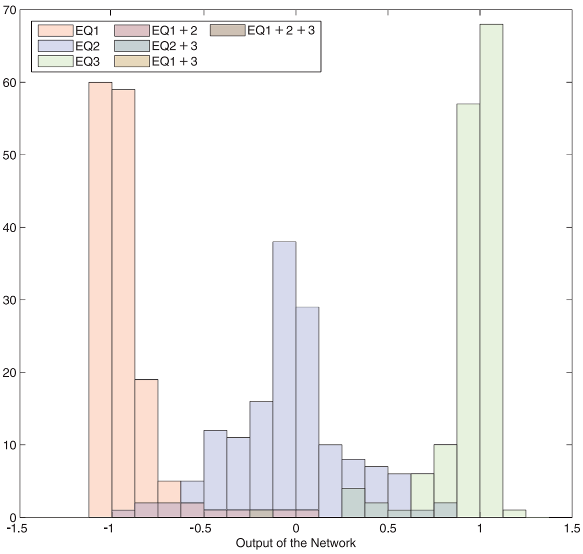

Fig. 4. Histogram using GMDH type neural network for subject 3

TABLE III
RECOGNITION RATE USING GMDH TYPE NEURAL NETWORK IN OPEN TESTING

Subject	Recognition Rate[%]	Subject	Recognition Rate[%]
1	80.00	17	78.52
2	89.00	18	83.20
3	87.33	19	83.00
4	82.33	20	86.00
5	74.81	21	79.00
6	92.40	22	83.00
7	81.18	23	80.00
8	78.82	24	85.67
9	84.09	25	85.93
10	84.00	26	83.33
11	80.37	**27**	**95.29**
12	81.00	28	78.52
13	82.80	29	83.20
14	82.67	30	85.67
15	86.50	31	76.00
16	77.33	32	87.06
		Average	83.06 ± 4.44

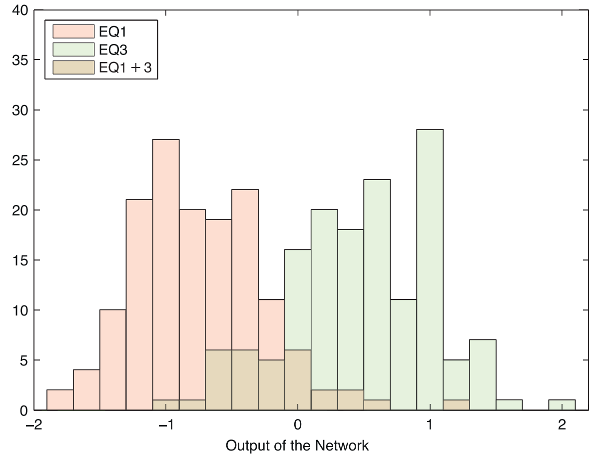

Fig. 5. Histogram using GMDH type neural network for subject 2

The closed test results showed that the method using GMDH type neural network was better than the method using PCA. This can be caused by the appropriate selection of frequency and weights to them to separate the output values. PCA is a method to find out the axis which will have the most variance, but it did not result enough separation to be discriminated by a simple linear analysis. TABLE III indicates the recognition rates for each subjects using GMDH type neural network using delta=0.005. The average was 83.06%, which can be said that it is on high level to recognize 2 classes by analyzing EEG. Histogram on Fig. 5 represents the output of the neural network constructed on subject 2. This diagram show that the output for data of EQ1 and EQ3 have been separated clearly but the skewness of the distibutions are less than the one on Fig. 4 which is delta set to 0. The ideal separation would be all data output becoming the same as the target signal, but to avoid over-fit, this result can be perceived satisfaction. However, we have not yet investigated the actual cause in frequency and also the cause of miss prediction when GMDH type neural network method is used. The proposed method using GMDH type neural network output one dimensional value, compared to two dimensional values when PCA was used. There also is a need to assess what the output value of the neural network actually mean, which should be the pleasantness as we defined the sound quality.

V. CONCLUSIONS

In this paper, we proposed an EEG analysis method using GMDH type neural network to evaluate a sound quality of a pair of loudspeakers. We applied frequency transformation as same as the conventional method, but by creating GMDH type neural network which output one value to extract the EEG difference between such sound quality improved the separation of feature scores. The three types of equalizers followed linearly in the output of the proposed system, as the actual sound qualities were defined linearly also. In the open test results, by setting the parameter delta which is to prevent over construction of neural network layers and the average recognition rate resulted 83.06%. For our future works, we will investigate the system to be actually used to qualify loudspeakers, by testing different types of loudspeakers.

REFERENCES

[1] A. Neuman, M. Bakke, C. Mackersie, S. Hellman, and H. Levitt, "The effect of compression ratio and release time on the categorical rating of sound quality", *J. Acoust. Soc. America*, Vol. 103, No. 5, pp. 2273-2278, 1998.

[2] A. Gabrielsson, and H. Sjogren, "Perceived sound quality of sound reproducing systems", *J. Acoust. Soc. America*, Vol. 65, No. 4, pp. 1019-1033, 1979.

[3] Y. P. Lin, C. H. Wang, T. P. Jung, T. L. Wu, S. K. Jeng, J. R. Duann, and J. H. Chen, "EEG-based emotion recognition in music listening", *IEEE Trans. Bio. Eng.*, Vol. 57, No. 7, pp. 1798-1806, 2010.

[4] Y. P. Lin, J. R. Duann, J. H. Chen, and T. P. Jung, "Electroencephalographic dynamics of musical emotion perception revealed by independent spectral components", *Neuroreport*, Vol. 21, Issue 6, pp. 410-415, 2010.

[5] T. Ogawa, S. Karungaru, Y. Mitsukura, M. Fukumi, and N. Akamatsu, "The EEG analysis by using a neural network in listening to music *IEICE Tech. Report*, pp. 5-9, 2007.

[6] S. Sanei, and J. Chambers, "EEG Signal Processing", John Wiley and Sons, 2009.

[7] M. Nakanishi, Y. Mitsukura, and A. Hara, "EEG analysis for acoustic quality evaluation using principal component analysis", *Proc. 20*$^{\mathrm{th}}$ *IEEE Int. Symp. Robot and Human Interactive Comm.*, pp. 321-324, 2011.

[8] A. G. Ivakhnenko, "Polynomial Theory of Complex Systems", *Sys. Man and Cybernetics, IEEE Trans.*, Vol. SMC-1, No. 4, pp. 364-378, 1971.