The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

Music Recommendation System Aimed at Improving Recognition Ability

H. Konishi[1] and Y. Yoshitomi[2]

1: Department of Environmental Information, Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan,
2: Graduate School of Life and Environmental Sciences, Kyoto Prefectural University,
1-5 Nakaragi-cho, Shimogamo, Sakyo-ku, Kyoto 606-8522, Japan,
E-mail: yoshitomi@kpu.ac.jp

Abstract: Recently, music therapy has been used for improving the recognition ability of people. Music therapy may be more effective when the favorite music of each person is adopted. We propose a music recommendation method that fuses content-based music recommendation and collaborative filtering. For characterizing the music in our content-based music recommendation system, we use three kinds of feature parameters for characterizing music: time series of wavelet transform coefficients, time series based on Mel-frequency cepstral coefficients, and parameters characterizing the rhythmic content. Several strategies for music recommendation are prepared. Each strategy is composed of a feature vector and a decision rule. The system can determine a good strategy for music recommendation for each user by not only exploiting the user's music evaluation history up to the present time but also the music evaluations by other users. In the experiments, 12 users rated 52 songs coming from a textbook database of songs for elementary schools. The number of recommended songs by the proposed method was 6.75 per user, and the number by collaborative filtering was 5.17 per user. The recommendation accuracy of the proposed method was 81.8%, and that of collaborative filtering was 74.1%.

Keywords: Recognition ability, Music therapy, and Music recommendation**.**

I. INTRODUCTION

In Japan, the average age of the population has been increasing, and this trend is expected to continue. Recently, music therapy has been used for improving the recognition ability of people, particularly older people. Music therapy may be more effective when the favorite music of each person is adopted. To the best of our knowledge, no research reports exist on the technology of music recommendation aimed at improving recognition ability. In the present study, we propose a music recommendation method that combines content-based music recommendation [1] and collaborative filtering to improve recognition ability. We evaluate the proposed method with children's songs, which tend to be familiar to older people.

II. FEATURE PARAMETERS

In the content-based music recommendation system, we use three kinds of feature parameters for characterizing music: time series of wavelet transform coefficients, time series based on Mel-frequency cepstral coefficients (MFCCs) [2], [3], and parameters describing the rhythmic content [4]. These parameters are described in the following subsections.

1. Wavelet transform coefficients

Original audio data $s_k^{(0)}$, where k denotes the element number in the data, are used as the level-0 wavelet decomposition coefficient sequence. The $s_k^{(0)}$ data are decomposed into the multi-resolution representation (MRR) and the coarsest approximation by repeatedly applying the discrete wavelet transform (DWT). The wavelet decomposition coefficient sequence $s_k^{(j)}$ at level j is decomposed into two wavelet decomposition coefficient sequences at level $j+1$, as shown below in (1) and (2):

$$s_k^{(j+1)} = \sum_n \overline{p_{n-2k}} s_n^{(j)}, \qquad (1)$$

$$w_k^{(j+1)} = \sum_n \overline{q_{n-2k}} s_n^{(j)}, \qquad (2)$$

where p_k and q_k denote the scaling and wavelet sequences, respectively, and $w_k^{(j+1)}$ denotes the development coefficient at level $j+1$. The development coefficients at level J are obtained by using (1) and (2) iteratively from $j=0$ to $j=J-1$.

In the present study, we use Daubechies wavelet for the DWT. As a result, we obtain the following relation between p_k and q_k:

$$q_k = (-1)^k p_{1-k}. \qquad (3)$$

It is known that the histogram of the wavelet coefficients of each domain of MRR sequences has a distribution centered at approximately 0 when the DWT is performed on music [5]. We found out that the standard deviation of the wavelet coefficients of each domain of the MRR sequences of music changed in its time series. In addition, the change depended on the music. Therefore, we use the time series of the wavelet coefficients of each domain of the MRR sequences as elements of the feature vector.

2. MFCC features

MFCCs are obtained for each frame of a sound signal by the conventional method [3]. The following are used as elements of the feature vector: the time series of the mean values and the standard deviations of 12-dimensional MFCCs, the logarithmic power of 12-dimensional MFCCs, the 12-dimensional MFCC difference between frames, and the 12-dimensional MFCC logarithmic power difference between frames.

3. Rhythm content features

Rhythmic content feature parameters are obtained by using the techniques described in Ref. [4]. A set of feature parameters based on a beat histogram are calculated. These are as follows:

- *A0*, *A1*: the relative amplitude (divided by the sum of the amplitudes) of the first and second histogram peaks, respectively [4];
- *RA*: the ratio of the amplitude of the second peak divided by the amplitude of the first peak [4];
- *P1*, *P2*: the period of the first and second peaks, respectively, in beats per minute [4];
- *SUM1*, *SUM2*, *SUM3*: the sum of beat strength in the histogram in the range of 40–90, 90–140, and 140–250, respectively, in beats per minute.

Each of the three kinds of feature parameters, which are *a*: [*SUM1*], *b*: [*A0*, *A1*, *SUM1*, *SUM2*, *SUM3*], c: [*A0*, *A1*, *P1*, *P2*, *RA*, *SUM1*, *SUM2*, *SUM3*], is used as a feature parameter in the method described in Section III.

III. CONTENT-BASED METHOD

To explain some of the methods for music recommendation, we describe a set of music indices as $M = \{m \mid 1, \cdots, N_M\}$, an evaluation by a user as a score $s (1 \le s \le 5)$, where 1, 2, 3, 4, and 5 mean "dislike," "slightly dislike," "neutral," "slightly favorite," and "favorite," respectively, and a set of evaluated music indices as $M_s = \{m_s \mid 1, \cdots, N_{Ms}\}$.

1. Methods using one kind of feature parameter

After principal component analysis (PCA) on the feature vectors obtained from unevaluated music m' file and the number N_{Ms} of evaluated music files of a user, the principal components up to the l th component are selected under the condition that the accumulated contribution ratio first exceeds 80% at the l th component. The score s for music m^* file having the maximum value of similarity to music m' file among the number N_{Ms} of evaluated music files is assigned to the score of music m' file. The similarity is calculated as the inverse value of the Euclid distance in the l dimensional feature vector space obtained by the above PCA. When the assigned score is 4 or 5 ("slightly favorite" or "favorite"), the unevaluated music m' file is recommended for the user.

2. Methods using two or three kinds of feature vectors

Table 1 shows the conditions for music recommendation by methods using two or three kinds of feature parameters.

Table 1. Conditions for music recommendation by methods using two or three kinds of feature parameters

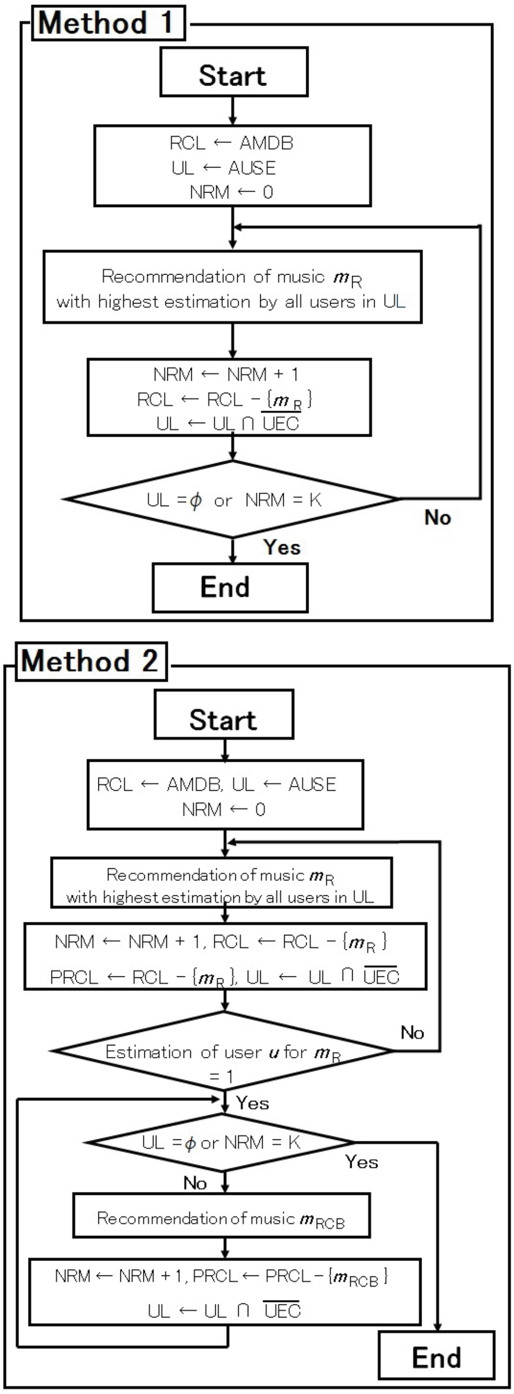
Method A	Score (1 kind of feature parameters) = 4 or 5 Score (another) = 3～5
B	Score (1 kind of feature parameters) = 4 or 5 Similarity order (another) : within top 50 %
C	Score (1 kind of feature parameters) = 4 or 5 Similarity (another) : average or higher than average

When using 3 kinds of feature parameters, 'another' is replaced by 'other two', and then naming replacement (A=>D, B=>E, C=>F) is performed.

IV. PROPOSED METHOD

Fig. 1 shows two music recommendation methods. Method 1 is collaborative filtering only and it is used for comparison with Method 2, which is our proposed method that combines the content-based method and collaborative filtering. Our system recommends music stored in a database to user u, as described in Fig. 1.

In the flowchart of Method 2 described in Fig. 1, the estimation of user u for music m_R is set as 1 when the score of user u for music m_R is 4 or 5 ("slightly favorite" or "favorite"), and it is set as 0 when the score is 1 to 3 ("dislike," "slightly dislike," or "neutral"). Moreover, m_{RCB} is decided for user u by using the most suitable recommendation method and feature parameter(s) selected among 74 combinations [1], [6] of method and feature parameter(s) by our previously reported method [1], in which we used the content-based recommendation method described in



Abbreviations ;

AMDB : set of all music in data base

RCL : set of music in recommendation candidate list

AUSE : set of all users with subjective estimation

UL : set of user in reference user list

NRM : number of recommended music

UEC : set of user(s) who estimate(s) recommended music m_R contrary to user u

PRCL : set of music in recommendation candidate list made of music having 3.5 or higher than 3.5 of average score by users in UL

m_{RCB} : music having the most similarity to the currently recommended music having score of 4 or 5 by user u among music in PRCL

Fig. 1. Flow chart of music recommendation methods

Section III and the feature parameter(s) described in Section II. In selecting the most suitable recommendation method and feature parameter(s) in Method 2, we use the scores of user u for music that has already been recommended for and evaluated by user u and the scores of other users in a reference user list (UL) for music not yet recommended for user u.

V. PERFORMANCE EVALUATION

1. Conditions

Because older people tend to prefer children's songs [7], we selected a CD described as an anthology of good older songs enjoyed by older people with dementia [8], and then we selected 52 songs on the CD that were also included in a music textbook database for elementary schools [9]. For evaluating the music recommendation methods, all 52 of selected songs in the database were assigned scores $s\,(1 \le s \le 5)$ by 12 users (teens: 1, twenties: 6, fifties: 5). Using the same conditions as used in our previous research [1], the feature parameters were obtained by the method described in Section II. We used 10 as the value of K in Fig. 1. For evaluating the two music recommendation methods described in Section IV, we chose each of the 12 users as user u and put the remaining users in the reference user list UL described in Fig. 1. Then, we obtained the result of the music recommendation for each user for each method described in Section IV.

2. Results and discussions

Table 2 shows the process of the music recommendation process for user 8. As shown in Table 2, Method 2 tended to recommend more music and have a higher accuracy of music recommendation than did Method 1. Fig. 2 shows the performance of the two methods. The number of recommended songs by the proposed method (Method 2) was 6.75 per user, whereas that of collaborative filtering (Method 1) was 5.17 per user. The recommendation accuracy of the proposed method was 81.8%, whereas that of collaborative filtering was 74.1%. For both the recommendation accuracy and the number of recommended songs, the proposed method was better than collaborative filtering. In both methods, the recommendation process was terminated at the rate of 5/6, when the number of users staying in the UL became zero. Accordingly, an increase in the number of users in UL might contribute to an increased number of recommended songs.

Table 2. Music recommendation process for user 8

[Method 1]

Order	Recommended music No.	Acceptance	User No. in UL
1	52	○	1,2,3,4,5,6,7,9,10,11,12
2	41	○	1, 3,4,5,6,7,9,10,11,12
3	50	○	1, 3,4,5, 7,9,10, 12
4	21	○	1, 3, 5, 7,9, 12
5	37	×	9
6	49	×	none

[Method 2]

Order	Recommended music No.	Acceptance	User No. in UL
1	52	○	1,2,3,4,5,6,7,9,10,11,12
2	16	○	1,2, 3,4,5, 7,9, 12
3	3	○	1, 3,4,5, 7,9
4	26	○	1, 3,4,5, 7,9
5	21	○	1, 3, 5, 7,9
6	5	○	1, 5, 7
7	17	○	1, 7
8	4	○	1
9	13	×	none

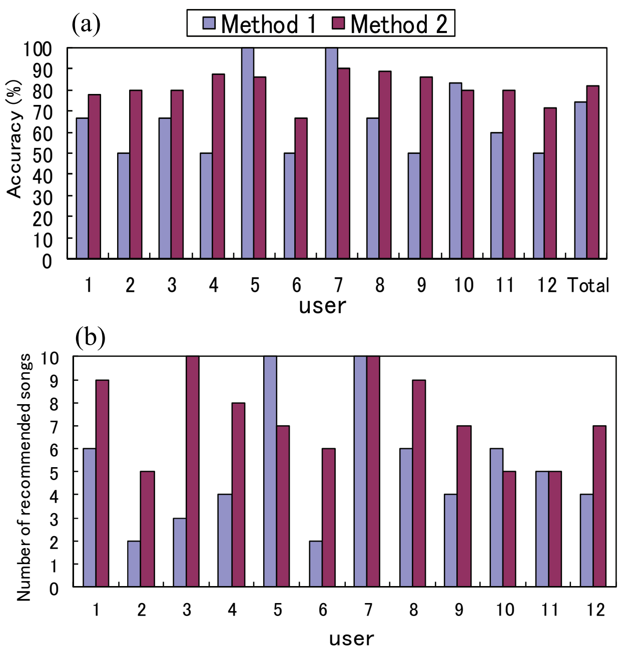


Fig. 2. Performance of music recommendation methods;
(a) recommendation accuracy,
(b) number of recommended songs

VI. CONCLUSION

We propose a music recommendation method combining our previously reported method based on music features and collaborative filtering. We showed that the proposed method is more effective for music recommendation than using only collaborative filtering when used on a music database composed of children's songs. In future work, we will increase the number of users who evaluate the music in the database and apply the proposed method to people who are much older and/or have a cognitive impairment.

Acknowledgments

We would like to thank Associate Professor M. Tabuse of Kyoto Prefectural University for his valuable advice and support on this research. We would also like to thank all the participants who cooperated with us in the experiments.

REFERENCES

[1] Tokuyama T, Yoshitomi Y, Tanaka M and Tabuse M (2009), A study on music recommendation system using the time-series discrete wavelet transform (in Japanese). Proc. of the 8th annual convention of forum on information technology. 2:325-328
[2] Davis S and Mermelstein P (1980), Experiments in syllable-based recognition of continuous speech. IEEE Trans. Acoust., Speech Signal Processing. 28:357-366
[3] Shikano K, Itou K, Kawahara T, Takeda K, and Yamamoto M (2001), Speech recognition system. Ohmsha (in Japanese)
[4] Tzanetakis G and Cook P (2002), Musical genre classification of audio signals. IEEE trans. speech audio process. 10(5):293-302
[5] Murata S, Yoshitomi Y and Ishii H (2007), Optimization of embedding position in an audio watermarking method using wavelet transform (in Japanese). Abstracts of autumn research presentation forums of ORSJ 210-211
[6] Horiike K, Yoshitomi Y, Tokuyama T, and Tabuse M (2011), Music recommendation system using the time-series discrete wavelet transform and the FastICA. Proc. of 16th Int. Symposium on Artificial Life and Robotics. in press
[7] Takahashi T (1997), Research report on songs familiar to people advanced in years (in Japanese) Journal of Japanese music therapy associate 15(1):68-75
[8] Akahoshi T (2009), Good old anthology enjoyable for people advanced in years and troubled with dementia. Kirara shobo (in Japanese)
[9] Music textbook database for elementary school by Kanagawa prefectural education center (in Japanese); http://kjd.edu-ctr.pref.kanagawa.jp/daizai_music/