Expert Systems with Applications 36 (2009) 10035–10044

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A novel method for personalized music recommendation

Cheng-Che Lu, Vincent S. Tseng *

Department of Computer Science and Information Engineering, National Cheng Kung University, Tainan 701, Taiwan, ROC

ARTICLE INFO

Keywords:
Personalized recommendation
Feature extraction
Music representations
Music emotion
Weighting

ABSTRACT

With the development of digital music technologies, it is an interesting and useful issue to recommend the 'favored music' from large amounts of digital music. Some Web-based music stores can recommend popular music which has been rated by many people. However, three problems that need to be resolved in the current methods are: (a) how to recommend the 'favored music' which has not been rated by anyone, (b) how to avoid repeatedly recommending the 'disfavored music' for users, and (c) how to recommend more interesting music for users besides the ones users have been used to listen. To achieve these goals, we proposed a novel method called *personalized hybrid music recommendation*, which combines the content-based, collaboration-based and emotion-based methods by computing the weights of the methods according to users' interests. Furthermore, to evaluate the recommendation accuracy, we constructed a system that can recommend the music to users after mining users' logs on music listening records. By the feedback of the user's options, the proposed methods accommodate the variations of the users' musical interests and then promptly recommend the favored and more interesting music via consecutive recommendations. Experimental results show that the recommendation accuracy achieved by our method is as good as 90%. Hence, it is helpful for recommending the 'favored music' to users, provided that each music object is annotated with the related music emotions. The framework in this paper could serve as a useful basis for studies on music recommendation.

1. Introduction

In recent years digital music like MIDI Manufacturers Association has rapidly increased due to its convenience for people. Although people can access the music via internet, it is very difficult to search out the 'favored music' from music databases. Generally speaking, shopping webs can enable users to search music by inputting singer and title, but users often forget the information of a song. To solve the problem Themefinder and Meldex can let users search for similar music by comparing the musical melodies, without the necessity of inputting the singer and title. The searching method based on a comparison of melodies is called music information retrieval. On the other hand, the music shopping web developed the recommendation system which can encourage people to buy the music such as from Amazon and iTunes, especially the hot music often recommended for users after computing the music ratings. For example, if many people buy both music A and music B, then music B is suggested when someone buys music A.

In fact, these recommendation methods incur two problems: they cannot recommend the music which is not rated by anyone and users may be not interested in highly rated music. Shardan and Maes (1995) implemented a network system called Ringo which is based on similarities between the interest profile of a user and those of other users. The method can recommend music for users by computing the social information, so it is collaborative recommendation. Furthermore, Kuo and Shan (2002) used melodic style to represent a user's music preference and built a personalized content-based filtering system. The melodic style involves musical contents, so the method can be regarded as content-based recommendation. Dixon, Bainbridge, and Typke (2007) thought that music recommendation systems concentrated on content-based analysis and collaborative filtering, and after carrying out a survey, discovered that little attention has been paid to the reasons why these techniques have been effective. Moreover, to further cater to users' listening, the musical emotion is important for the recommendation of works of music. Yeh, Tseng, Tsai, and Weng (2006) proposed a personalized music emotion prediction system to predict users' different emotions concerning the music they preferred.

In the past few years, numerous studies have been devoted to multimedia analysis and personalized recommendations, such as: web image annotation (Tseng, Su, Wang, & Lin, 2007), video annotation (Tseng, Su, & Huang, 2006), music classification (Lu et al., 2006) and recommendation systems (Lee, Liu, & Lu, 2002; Lee & Lu, 2003). In this paper, we propose a novel method called personalized hybrid

* Corresponding author. Tel.: +886 6 2757575x62536; fax: +886 6 2747076.
E-mail addresses: cclu@idb.csie.ncku.edu.tw (C.-C. Lu), tsengsm@mail.ncku.edu.tw (V.S. Tseng).

0957-4174/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2009.01.074

music recommendation, to achieve three goals: to recommend music which has not been rated by anyone, to avoid repeatedly recommending some 'disfavored music', and to recommend the more interesting music besides what the music users are used to listen to. To evaluate the recommendation accuracy, we constructed a system that, after mining users' hobbies, can recommend music to users. By the feedback of the user's options, the proposed methods accommodate the variations of the users' musical interests and then recommend the favored and more interesting music, with prompt consecutive recommendations. Experimental results show that the recommendation accuracy achieved by our method is as good as 90%. The rest of the paper is organized as follows. The second section of the paper is a review of the literature, addressing both empirical and theoretical aspects of the features in musical research. The third section clarifies how to obtain the features from a score, and describes the method of personalized hybrid music recommendation. The results for the various analyses are presented following each of these descriptive sections. Finally, conclusions are presented and suggestions are made for further research.

2. Related work

With the development of digital music technology, it is important to determine how to search out the 'favored music' for users from a music database. Users usually input some musical metadata in music searches, for example, singer or title. Furthermore, users can search music without inputting the singer and title, such as with Meldex based on musical contents. Musical contents have been extensively applied in a wide variety of applications, such as music recommendation and music information retrieval. In the section on music information retrieval, the method is usually based on a comparison of melodies and Liu, Hsu, and Chen (2001) proposed that the repeated and longest patterns were usually the main melody which is used to retrieve similar music. To reduce retrieval errors, Tseng et al. (1999) considered the problem of extracting the melody and used the N-index technique. Weyde et al. (2005) segmented melody into motifs and phrases, which were then encoded according to variations. Typke, Wiering, and Veltkamp (2005) provided an overview of content-based music information retrieval (CBMIR), both for audio and for symbolic music notation; and thought that content-based music search engines can be useful for finding musical scores similar to a given query and Query-by-Humming. To retrieve the musical data easily, Chai and Vercoe (2000) used an XML-like language in a multimedia information retrieval system called MusicCat. Furthermore, Karydis, Nanopoulos, Papadopoulos, Katsaros, and Manolopoulos (2005) proposed query representation with reducing length and selective policy for the routing of answers, which are helpful in reducing the response times on wireless mobile ad hoc network.

As discussed above, we thought that the information of users' music interests not only is overlooked in the process of music information retrieval but is also more suitable for discovering the 'favored music' for users. However, up to now, there has been relatively little research conducted on the process of music recommendation. In general, the techniques of music recommendation can be divided into content-based and collaborative methods. Kuo and Shan (2002) used melodic style to represent user's music preference and built a personalized content-based filtering system to test the precision of recommendation; the results achieved an average precision of 63.28%. Furthermore, Chen and Chen (2001) employed users' transaction data to predict the music users like. The results show that an average of 62% precision for content-based method and 23% precision for collaborative method can be achieved. To improve the precision of music recommendation, Uitdenbogerd and Schyndel (2002) learned that demographic and personality factors have been shown to be factors influencing music preference, and discovered that tempo, pitch, tonality and rhythm are the main factors. To reduce the prediction error, Tiemann, Pauws, and Vignoli (2007) used the ensemble learning methods for hybrid music recommenders, combining the social and the content-based recommender algorithm in an initial experiment by applying a simple combination rule to merge recommender results.

Furthermore, In order to further understand the reason why users like the specific musical genres, the research has been extended to discover the relationships between musical emotion and users' habit. Krumhansl (2002) and Sadie and Latham (1990) thought that music can convey emotions. Although most music recommendation systems are based on the users' preference, sometimes recommending music by the emotion is needed. For example, users need a peaceful music when studying. Discovering the musical emotions can prove helpful in the development of music recommendation. Lu, Liu, and Zhang (2006) used the emotional model, psychologist suggested, to classify the musical genre and discovered the relationships between emotions and three musical features (intensity, timbre and rhythm). For example, the intensity of contentment and depression in music is usually quite weak, but that of exuberance and anxiety is usually powerful. Furthermore, Yang et al. (2004) used the emotional words in lyrics to improve accuracy of emotion-based music classification, and Wang, Zhang, and Zhu (2004) discovered that the music can enrich listener's moods by the variations of tonality, beat, tempo and interval. On the other hand, to avoid the laborious work of emotion labeling, Kuo, Chiang, Shan, and Lee (2005) employed an affinity graph to discover the relationships between music features and emotions from film music, whereby the results can achieve 85% accuracy, on average.

3. Proposed method

In this section, we describe our method in detail by presenting a conceptual framework to carry out personalized hybrid music recommendation. The workflow of our method is shown in Fig. 1. For the objectives to be achieved, we first illustrate how to obtain the features from a score in Section 3.1. In Section 3.2, we state how to perform the proposed methods which involve: content-based, collaboration-based and emotion-based music recommendations, and show how to achieve hybrid music recommendation, by computing the weights of three recommendation methods, according to users' interests.

3.1. Feature extraction

In this section, we describe our method in detail with the focus on the extraction of features. The workflow of our method is shown

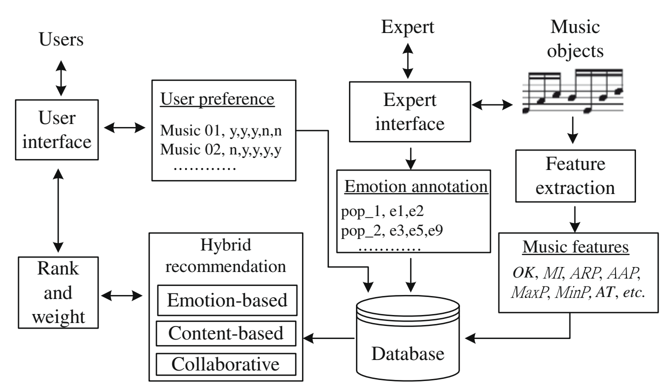

Fig. 1. Workflow for proposed method.

in Fig. 2. The main idea is to extract the features that can be employed to represent polyphonic music. The features are obtained from XML files of the score by the proposed approach that mainly includes the following steps: parsing the MusicXML files, extracting the qualitative features, representing polyphonic music as the music representations we proposed, encoding the pitches into codes and computing the quantitative features.

3.1.1. Music representations

A score consists of notes and musical notations, which can be used in composing music. To discover interesting features, the scores must first be translated into XML files by the music software (Recordare, xxx). The polyphonic music can be translated as some melodies using the tags of XML, such as ⟨score-part id=melody1⟩, ⟨score-part id= melody2⟩, ⟨score-part id=melody3⟩ and ⟨score-part id=melody4⟩, as shown in Fig. 2a. It is worth noting that the qualitative features have been assigned in XML files after translating the scores. We can directly extract a qualitative feature by parsing the XML files and then encoding the *original key* into the codes we suggested according to Table 1, as shown from Fig. 2a and b.

However, in order to obtain other quantitative features, the music has to be represented as music representations we proposed and then the quantitative features can be obtained by formulations in Section 3.1.2. In this paper we developed the program to parse XML files and automatically achieved the following works for music representations. The process is shown in Fig. 2a and b. We proposed that the melodies in score can be represented as a sequence of symbols by integrating the three points below (A)–(C) before extracting the quantitative features. The representations are convenient for analyzing the variations of the pitch:

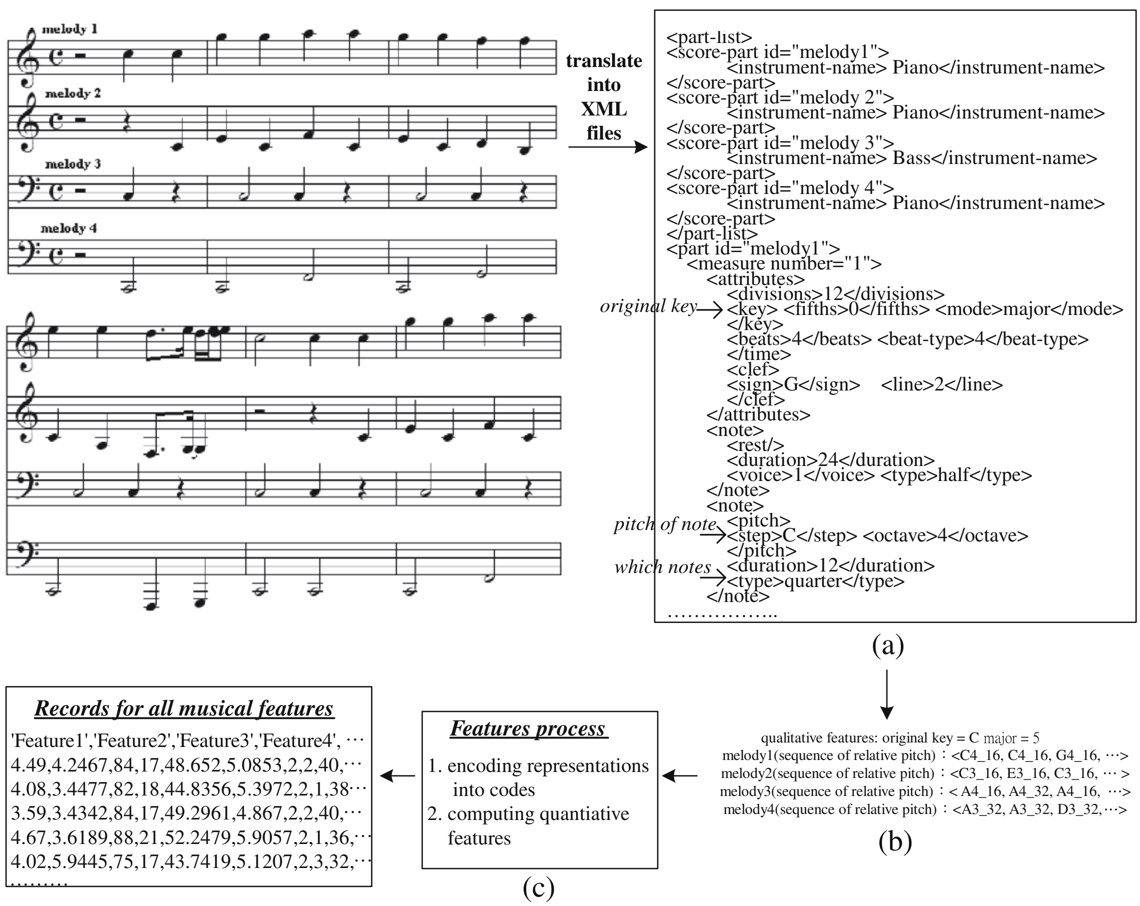

Fig. 2. An example illustrating how to obtain features from polyphonic music.

Table 1
Transforming the key signature into the suggested value.

Value	Term	Tag in MusicXML files
5	C major, A minor	⟨fifths⟩0⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩0⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
6	G major, E minor	⟨fifths⟩1⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩1⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
7	D major, B minor	⟨fifths⟩2⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩2⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
8	A major, F$^{\#}$minor	⟨fifths⟩3⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩3⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
9	E major, C$^{\#}$minor	⟨fifths⟩4⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩4⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
10	B major, A^{b} minor	⟨fifths⟩5⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩5⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
11	F$^{\#}$ major, E^{b} minor	⟨fifths⟩6⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩6⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
4	C$^{\#}$ major, B^{b} minor	⟨fifths⟩7⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩7⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
3	A^{b} major, F minor	⟨fifths⟩-1⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩-1⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
2	E^{b} major, C minor	⟨fifths⟩-2⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩-2⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
1	B^{b} major, G major	⟨fifths⟩-3⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩-3⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩
0	F major, D major	⟨fifths⟩-4⟨/fifths⟩ ⟨mode⟩major⟨/mode⟩, ⟨fifths⟩-4⟨/fifths⟩ ⟨mode⟩minor⟨/mode⟩

Relative pitch : <C3_16&C4_16, G3_16, E3_16&C4_16, G3_16, C3_16&G4_16, G3_16, E3_16&G4_16, G3_16,..>

Fig. 3. Music represented as a symbolic sequence.

(A) The pitch of a note can be represented as relative pitch. Relative pitch means that the notes with the same chroma in different octaves are regarded as different pitches. Absolute pitch means that the notes with the same chroma in different octaves are regarded as same pitch. In our work, we can compute two features: average relative pitch and average absolute pitch, because composers may compose a same piece with different octaves.
(B) When more than one note is played simultaneously, the result is a chord (Miller & Williams, 1991). We propose that a chord can be represented as the highest and lowest pitches if a chord consists of two or more notes. For example, the relative pitch sequence in Fig. 3, the first chord (C3_16& C4_16) with two eighth-notes is represented as a relative highest pitch C3_16 and a relative lowest pitch C4_16, where C denotes tone 'Do'; 3 denotes middle tone; _16 denotes duration with four units; and and denotes the combination of pitches. The rest may be deduced by analogy.

Table 2
The proposed features related to the reasons for choosing the music.

Term	Reason of choice (as shown in Fig. 4)
Maximum Interval (*MI*)	Pitch
Maximum Pitch (*MaxP*)	
Minimum Pitch (*MinP*)	
Average Relative Pitch (*ARP*)	
Average Absolute Pitch (*AAP*)	
Average Tempo (*AT*)	Tempo
Rhythmic Speed (*RS*)	Rhythm
Original Key (*OK*)	Tonality
Chord Density (*CD*)	Other
Key Density (*KD*)	
Meter Density (*MD*)	

(C) Note with duration means that the note sounded for a specific time such as whole note, half-note and quarter-note. Composers use notes with duration to enrich the rhythm. In this paper we treat 128th note as a base unit because the 128th note is minimum. For example, a quarter-note equals 32 128th notes so a melody can be converted into a sequence of pitches with duration, as shown in Fig. 3. For example, _16 denotes duration with sixteen units in the figure.

3.1.2. Feature extraction

In this paper the proposed features consist of quantitative and qualitative features. As Section 3.1.1 described, we can directly extract qualitative features when parsing the XML files except for quantitative features. For example, *original key* = *C* major = 5. In this section, we explain the proposed features in Table 2 and describe how to obtain quantitative features in detail. We can compute the quantitative features by music representations, as shown in Fig. 2b and c. All pitches in music representations are first encoded into the codes, for example, $C4 = 48$, $C4^{\#} = 49$, $D4^{b} = 49$ and $D4 = 50$. In Fig. 4, maximum interval (*MaxI*), maximum pitch (*MaxP*), minimum pitch (*MinP*), average relative pitch (*ARP*) and average absolute pitch (*AAP*) can be obtained by computing the variations of pitches. The other qualitative features are described as follows:

$$Average\ Tempo\ (AT) = number\ of\ beats/m \tag{1}$$

where the *number of beats* denotes the number of beats of music and *m* denotes the time of music in minutes. For example, the music for two minutes comprises 240 beats, so the *Average Tempo* can be determined by the formula (1) ($AT = 240/2 = 120$).

$$Rhythmic\ Speed(RS) = maximum\ number\ of\ pitches/m \tag{2}$$

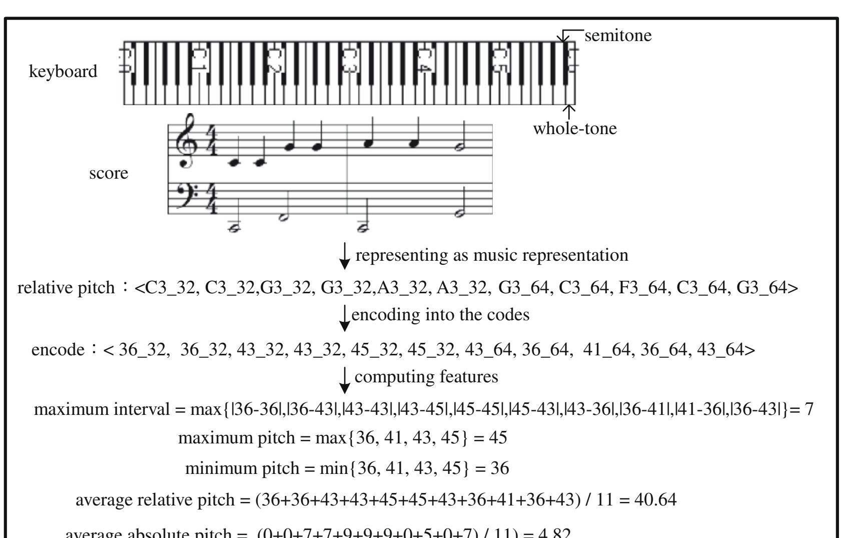

Fig. 4. An example illustrating how to compute pitch-related features.

where the *maximum number of pitches* denotes the maximum number of pitches among all melodies (both a note and a chord are viewed as a pitch); and *m* denotes the time of music for minute. For example, if the music for two minutes consists of two melodies that are played simultaneously, the main melody has 300 pitches and the minor melody has 250 pitches, then the *Rhythmic Speed* can be determined by formula (2) ($RS = 300/2 = 150$).

$$\textit{Chord Density}\ (CD) = \frac{NC}{100} \tag{3}$$

where *NC* is the number of distinct chords in the tracks. For example, if the music consists of ten distinct chords, so the *Chord Density* can be determined by the formula (4) ($CD = 7/100 = 0.07$).

$$\textit{Key Density}(KD) = \frac{NK}{12} \tag{4}$$

where *NK* is the number of distinct keys in the tracks. For example, if the music consists of three distinct keys (C major, D major and E minor), so the *Key Density* can be determined by the formula (4) ($KD = 3/12 = 0.25$).

$$\textit{Meter Density}(MD) = \frac{NM}{18} \tag{5}$$

where *NM* is the number of distinct meters in the tracks. For example, if the music consists of two distinct meters (3/4 and 4/4), so the *Meter Density* can be determined by the formula (5) ($MD = 2/18 = 0.1111$).

3.2. Personalized hybrid music recommendation

In recent years, despite the fact that the recommendation systems have been able to recommend some hot music by other users' ratings, the system still cannot recommend music which has not been rated by anyone, such as the latest or non-popular music. Furthermore, users may be not interested in the rating of music but the system always recommends the highly rated music according to other users' ratings. We proposed that the content-based, collaboration-based and emotion-based recommendations can be combined into the proposed recommendation method with feedback, called personalized hybrid music recommendation. The proposed method not only can recommend the 'favored music' not rated by anyone, but can also recommend more interesting music besides the music that the users have been used to listening to. By the feedback of the user's options, the proposed methods can adapt the variations of the users' interests and then promptly avoid recommending some 'disfavored music' to users in the recommendation process.

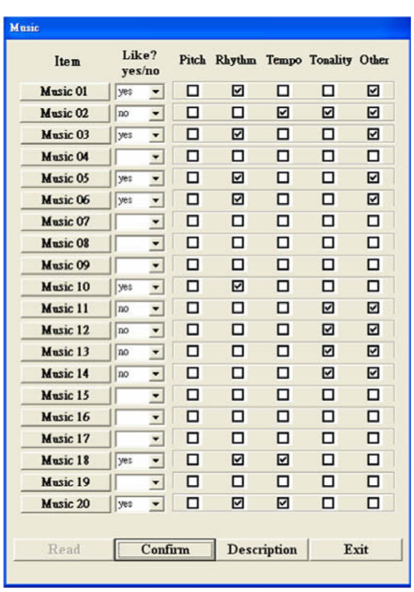

Fig. 5. The questionnaire for users.

To achieve our objective, we first construct a user's music questionnaire to discover user's interests, as shown in Fig. 5. The 20 music objects in the questionnaire are assigned by the expert. After registering, the user cannot only choose the music as favored (yes) or disfavored (no) or space (neutral), but also can mark the reasons why the user chooses the preferred music. Hence, we used the users' questionnaire records and feedback in the process of recommending interesting music more effectively to users. The rest of this section is organized as follows. The content-based, emotion-based and collaboration-based recommendation methods are used to demonstrate how to perform the recommendation work in Section 3.2.1, 3.2.2 and 3.2.3, respectively. Section 3.2.4 presents how the three above-mentioned recommendations are combined according to the weights computed for the three recommendations, based on users' interests.

3.2.1. Content-based recommendation

The content-based recommendation method is able to recommend music that has not been rated by anyone and approaches the practical musical contents that users like. To achieve the desired results, the proposed features in Section 3.1 are treated as the musical contents which are used in the process of the content-based recommendation. The details are indicated in Fig. 6b and the process of the proposed methods are as shown in the following steps:

Stage 1: Transform the qualitative feature into the value shown in Table 1.
Stage 2: The values of all features are normalized.
Stage 3: Compute the weights of eleven features according to the 'favored music' and the reasons of choice which all relate to the eleven features shown in Table 2, so each $\boldsymbol{w}_i$ is between 0 and 1 after computing.
Stage 4: Compute the similarity value between each music object in the database and the 'favored music' by using formula (6). Then, each music object in the database can obtain a similarity value, for example, by computing the similarity values between the music object (M_1) in database and the 'favored music' (A_1 and A_2), and then assuming: f (M_1, A_1) = 0.12 and f (M_1, A_2) = 0.25, so similarity value of M1 = Max (f (M_1, A_1), f (M_1, A_2)) = 0.25.
Stage 5: All music objects in the database are ranked by descending.

$$\mathrm{f}\ (\mathrm{M}, \mathrm{A}) = \sum_{i=1}^{11} w_i \times (1 - |m_i - a_i|) \tag{6}$$

where **M** denotes one of the music objects from the database; **A** denotes one of the music objects assigned as yes by the user; $\boldsymbol{w}_i$ denotes the weights of the features for 'favored music'; and $\boldsymbol{m}_i$, $\boldsymbol{a}_i$ denotes the attribute values of the music objects.

3.2.2. Emotion-based recommendation

To further cater to users' listening, musical emotions are important for the process of music recommendation. We proposed that the systems determine the music of interest for users by computing the differences between users' interests and musical emotions. In practice, Reilly (1996) and Thayer (1989) had already clarified the emotional types and Kuo et al. (2005) also employed an affinity graph to discover the relationships between music features and emotions from film music. To achieve the emotion-based recom-

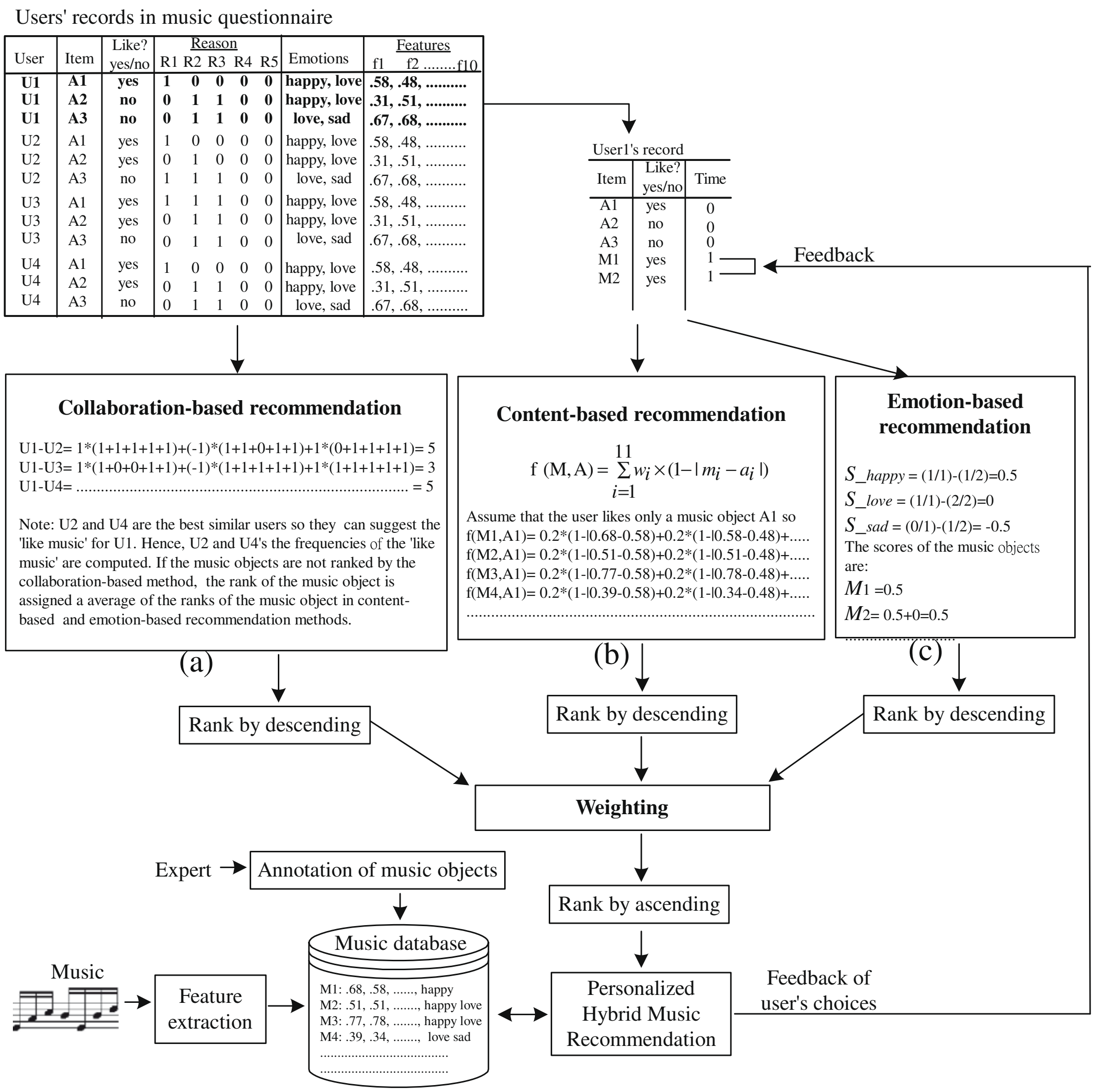

Fig. 6. Personalized hybrid music recommendation for user1 in details.

mendation method we proposed, the expert first has to annotate one or more emotions for the music objects by using the form shown in Fig. 7. As shown in Table 3, all of the emotions are divided into positive and negative. Then, the scores of the favored and disfavored emotions for the user are computed by formula (7) and the scores of the emotions of each music in the database also computed, by formula (8) and (9). Finally, all of the music objects in the database are ranked by descending. The details of this method are shown in Fig. 6c.

$$s_i = \frac{n_{i\text{–favored}}}{N_{\text{favored}}} - \frac{n_{i\text{– disfavored}}}{N_{\text{disfavored}}} \tag{7}$$

where the s_i denotes the score of the emotion e_i for the user; and $n_{i\text{–favored}}$ denotes the number of the emotion e_i in whole 'favored music'; and N_{favored} denotes the number of 'favored music'; and $n_{i\text{–disfavored}}$ denotes the number of the emotion e_i in whole 'disfavored music'; and $N_{\text{disfavored}}$ denotes the number of 'disfavored music'.

$$m(e_j) = \begin{cases} s_i, & \text{if } e_j = e_i \\ p, & \text{if } e_j \neq e_i \text{ and } e_j \text{ is positive emotion} \\ q, & \text{if } e_j \neq e_i \text{ and } e_j \text{ is negative emotion} \end{cases} \tag{8}$$

where the $m(e_j)$ denotes the score for the music objects with emotion e_j; p denotes the average for the positive emotions; and q denotes the average for the negative emotions; e_i denotes the user's favored and disfavored emotions; e_j denotes the music objects with emotions.

$$M_i = \frac{\sum_{j=1}^{n} m(e_j)}{n} \tag{9}$$

where the M_i denotes the score of the music object i in the database and n denotes the music object with the number of emotions.

3.2.3. Collaboration-based recommendation

To find out the more interesting music for user, the recommendation system has to depend on other users' suggestions. We proposed the collaboration-based recommendation method that can compute the differences between users' interests and further discover the interesting music objects suggested by other users with the most similar interests. First, the system must obtain each user's music interests by the form shown in Fig. 5, and then the difference value between users can be computed by using the proposed algorithm shown in Fig. 8. In Fig. 8, the difference value is between 100 and −100, representing the difference between the user and that of

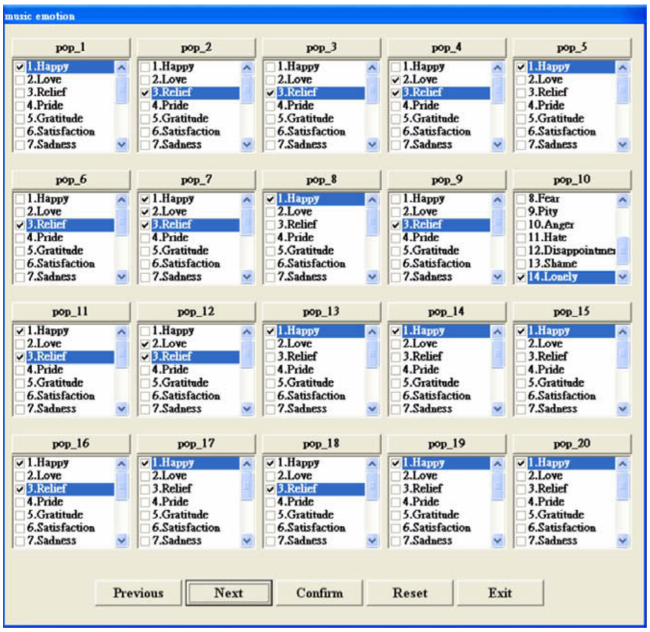

Fig. 7. The emotions of each music object annotated by the expert.

Table 3
Emotions used in this paper.

Cluster of emotion	Affection
1. Happiness, 2. Love, 3. Relief, 4. Pride, 5. Gratitude, 6. Satisfaction	Positive
7. Sadness, 8. Fear, 9. Pity, 10. Anger, 11. Hatred, 12. Disappointment, 13. Shame, 14. Loneliness, 15. Anxiety	Negative

the other users. After discovering the users with most similar interests (one or more users), the frequencies of the 'favored music' objects are computed and then the 'favored music' objects are ranked by descending. However, the other music objects in the database may not be ranked after using the method above, so the ranks are forced to refer to the ranks of the music objects in content-based and emotion-based recommendation methods. For example, assume that music object m is not ranked in the works of collaboration-based recommendation; if the ranks of music object m are 9 for content-based recommendation and 1 for emotion-based recommendation respectively; the rank of music object m for collaboration-based recommendation are 5 $((9+1)/2 = 5)$. The details of this method are shown in Fig. 6a.

3.2.4. Weighting for personalized hybrid music recommendation

The goal of the weighting is to combine three proposed recommendation methods to cater to users' needs which involve recommendations of the latest music, the recommendations of musical emotions and the recommendations of more interesting music. As indicated in Section 3.2.1, 3.2.2 and 3.2.3, the music objects can be assigned their ranks after performing the three recommendation methods. In this section, the system can compute the weights of the three recommendation methods using formula (10), respectively, according to the ranks of the music objects chosen by the user, as shown in Fig. 9 (stages 1–4). In Fig. 9, because the ranks of the music objects represent the accuracy of the recommendation methods, we proposed that the sum of the ranks of the music objects chosen by the user for three proposed recommendation methods, respectively, can be used to evaluate the weights of the recommendation methods for the user. In other words, for the single recommendation method, the weight is bigger if the sum of the ranks is smaller. The weights represent the importance of the recommendations for the user. Furthermore, in the processes of making the recommendations, the system can gather the favored and disfavored music chosen by the user every time, called the user's feedback, which shows the user's variations of interests, so the proposed method can adjust the weights of three recommendations to adapt to the user's realistic needs according to the user's record of interests before recommending music to the user. All of the details of the proposed method are presented in Fig. 6.

$$w_i = \frac{1 - \frac{\sum_{j=1}^{n} r_j}{R}}{2} \tag{10}$$

The **Collaboration-based** function computes the difference between two users' choices in the initial music questionnaire.

Input: the interests in questionnaire involving the user and the one of other users' choices.
Output: the difference is between -100 and 100.
Variable:
$U1[i,0]$, $U2[i,0]$: the choice of i-th music, such as favored(yes) or disfavored(no) or space (neutral).
$U1[i,1]$, .., $U1[i,5]$: the five reasons why the user chooses the music.
$U2[i,1]$, .., $U2[i,5]$: the five reasons why the one of the other users chooses the music.
P: positive reward or negative penalty.
D: the difference between two users for the same music.
S: sum of the difference

Collaboration-based (U1, U2):

```
S = 0
for i from 1 to 20
    P = -1
    if U1[i, 0] = U2[i, 0]  then
        P = 1
    end
    for j from 1 to 5
        if U1[i, j] = U2[i, j]  then
            D = D + 1
        end
    S = S + (D × P)
Collaborative-based (U1, U2) = S
```

Fig. 8. The Collaboration-based algorithm.

Stage 1:

User has already chosen the 'favored music' : a1,a2,a3,a4 — $\sum_{j=1}^{n} r_j$

Rank for emotion-based $\longrightarrow$	a1:1, a2:2, a3:10, a4:7	1+2+10+7 = 20
Rank for content-based $\longrightarrow$	a1:23, a2:6, a3:11, a4:20	23+6+11+20 = 60
Rank for collaboration-based $\longrightarrow$	a1:1, a2:2, a3:4, a4:13	1+2+4+13 = 20

R = 20+60+20 = 100

Stage 2:

Computing the $$w_i = \frac{1 - \frac{\sum_{j=1}^{n} r_j}{R}}{2} \rightarrow$$

$w_{\text{emotion-based}}$ = (1-(20/100))/2=0.4

$w_{\text{content-based}}$ = (1-(60/100))/2=0.2

$w_{\text{collaboration-based}}$ = (1-(20/100))/2=0.4

Stage 3:

ready to recommend the music from database: m1,m2,m3,m4, ...

Rank for emotion-based $\longrightarrow$	m1:2, m2:5, m3:1, m4:10 ...
Rank for content-based $\longrightarrow$	m1:20, m2:6, m3:15, m4:30 ...
Rank for collaboration-based $\longrightarrow$	m1:3, m2:2, m3:8, m4:20 ...

Computing the score of rank of music objects = $r_{\text{emotion-based}} * w_{\text{emotion-based}} + r_{\text{content-based}} * w_{\text{content-based}} + r_{\text{collaboration-based}} * w_{\text{collaboration-based}}$

for example:

m1: 2*0.4+20*0.2+3*0.4 = 6
m2: 5*0.4+6*0.2+2*0.4 = 4
m3: 1*0.4+15*0.2+8*0.4 = 6.6
m4: 10*0.4+30*0.2+20*0.4 = 18
m5: ....
....

Stage 4:

To recommend user the favored music, the music objects in the database are ranked by ascending after stage 3.

Fig. 9. The weighting method in details.

where the w_i denotes the importance of the recommendations for the user, n denotes the number of 'favored music', r_j denotes the ranks of the 'favored music' for one of three recommendation methods, and R denotes the sum of the ranks of the 'favored music' for the three recommendation methods.

4. Experimental evaluation

To evaluate the proposed method, we collected 680 MIDI files from public Web sources (DownWithTheLoads; Folksongs; KIDiddles; MIDI Database). We constructed a system that can recommend the music to users after mining the records of users' musical interests. In Fig. 10, the recommendation system can show the information involving music items, the choices (yes or no or space), the reasons of the choice (pitch, rhythm, tempo, tonality and other) and the 'favored music' objects have chosen by the user (favorite). The system can recommend ten music selections to users every time and then the users can listen to the music by pressing the buttons of the music items, as shown in Fig. 11. After listening, the user cannot only designate the choices for music items but can also mark the reasons. In addition, provided that the user wants to understand what the meanings of the reasons are, the user can press the button of the description and then the explanations of the reasons are presented (see Fig. 12). In the experiments, the 27 volunteers used the recommendation system. We tested whether the proposed methods can accurately find the music that users like, and then compared the weights (importance) of the recommendation methods for users.

4.1. Evaluation for personalized hybrid music recommendation

To evaluate the recommendation accuracy, the system can recommend 10 music items to each user every time. In Fig. 13, the vertical lines represent the accumulative number of the 'favored

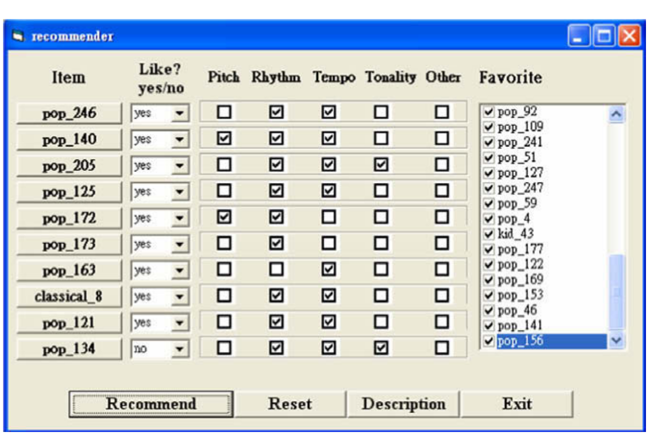

Fig. 10. The system recommends ten music selections to the user every time.

Fig. 11. User's listening to the music.

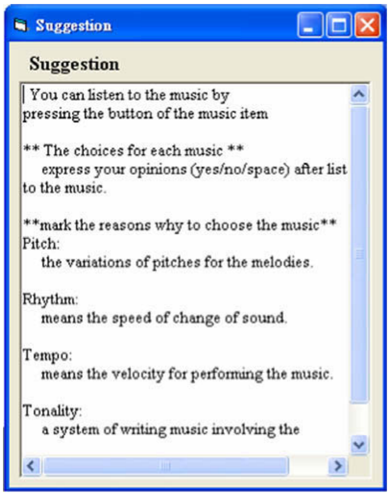

Fig. 12. Descriptions of the reasons.

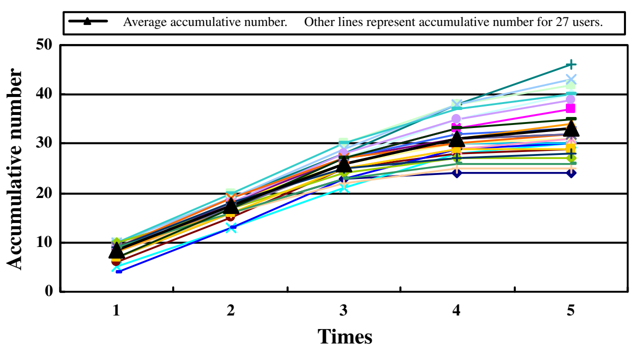

Fig. 13. The accumulative number of 'favored music' during five consecutive recommendations.

music' consecutive recommendations, and the wavy lines represent the times of recommendations. Then, the accumulative numbers of the 'favored music' chosen by 27 users are depicted in Fig. 13 and the average of the accumulative number is represented as the rough and black line with triangular symbols. Fig. 13 shows that the numbers of 'favored music' are raised during consecutive recommendations, and shows that an average of 90% for best recommendation accuracy can be obtained for a second recommend. However, for the users' interests, the number of 'favored music' in the database is limited. The experimental results show that the system may not find music conforming to users' interests if the music users like has been completely obtained by consecutive recommendations; hence, the average accumulative number is speedily decreased. As shown in Fig. 13, after three consecutive recommendations, we discovered that the numbers of 'favored music' between users obviously differ because the selections were obtained from the music database; this is to say, the number of 'favored music' will be greater if the users' interests are extensive, and fewer if the users' interests are limited. Whether these interests are limited or extensive, the results presented that the proposed methods still can effectively and accurately determine the 'favored music' for users during consecutive recommendations, with the average accuracy for the preceding three times being: 84.5%, 90% and 86.5%, respectively. Afterwards, the average accuracy of the recommendations is speedily decreased because the users have nearly obtained the entire 'favored music' from the music database.

4.2. Comparison of weights for recommendations

In this paper, to cater to users' needs involving the recommendations of latest music, recommendations of musical emotions and recommendations of more interesting music, three recommendation methods (content-based, collaboration-based and emotion-based methods) are combined by the weighting method. In this section, we tested the variations of weights for three recommendation methods to understand the relationships between recommendation methods and users' needs more clearly. In Fig. 14, the vertical lines represent the weight of the recommendation method, and the wavy lines represent the times of recommendations. Then, the variations of weights for 27 users are depicted in Fig. 14 and the average of the weight is represented as the rough and black line with circular symbols. On the whole, the variations of weights for the collaboration-based method are fewer than those for other methods, showing that the parts of the music users like are similar. At first, the content-based method computes the ranks of the 'favored music' in users' questionnaires, for the content-based methods we proposed; all the ranks of the 'favored music' equal to 1 so the weight of the content-based method is bigger than the weights of the other methods. Afterwards, the weight of the emotion-based

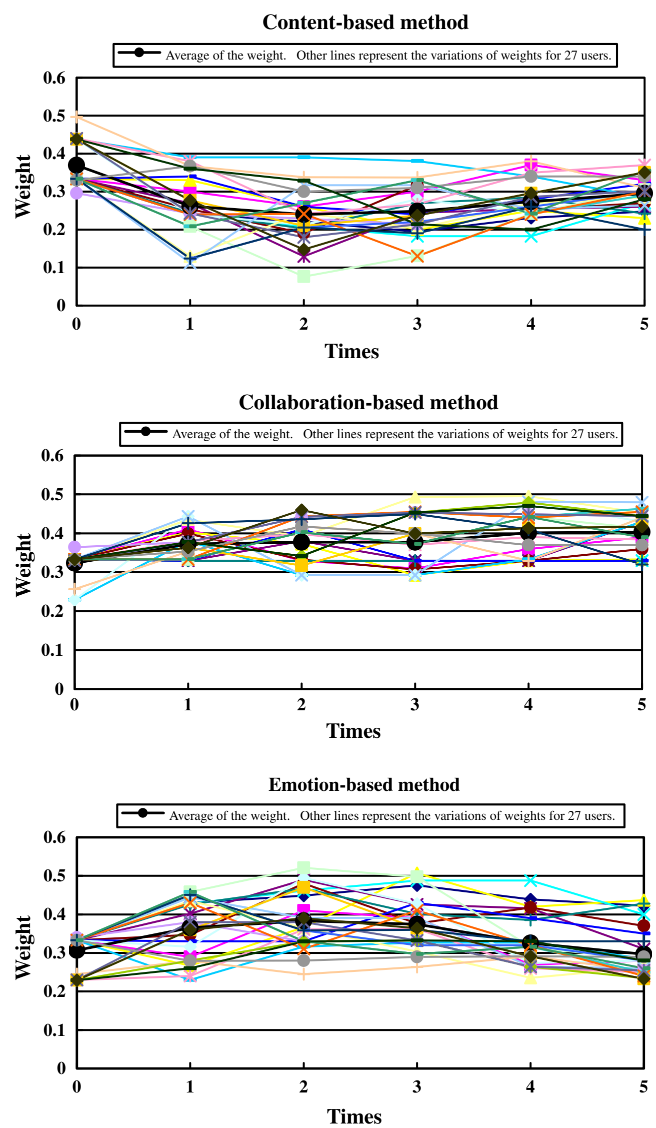

Fig. 14. The variations of weights for 27 users' needs.

method is bigger than that of the content-based method during consecutive recommendations, showing that the recommendation accuracy for emotion-based method is better than the recommendation accuracy for content-based method; in other words, it is helpful for recommending interesting music if the music can be annotated via music emotions. Furthermore, as the results in Section 4.1 indicate, the users have nearly obtained the entire 'favored music' from the music database before the preceding three times, and Fig. 14 presents that the average weights for collaboration-based method are gradually raised, showing that the collaboration-based method can help users to obtain more interesting music.

5. Conclusions and future work

In this paper, we proposed a novel method called *personalized hybrid music recommendation* to achieve three goals: to recommend the 'favored music' which has not been rated by anyone, to avoid repeatedly recommending some 'disfavored music', and to recommend more interesting music for users in addition to the music which users have been used to listen to. In fact, Uitdenbogerd and Schyndel (2002) proposed that tempo, pitch, tonality and rhythm are the key music preferences for determining users' interests. The features we proposed can be extracted from musical scores and then treated as the musical contents to find out the 'favored music'. With the duration of consecutive recommendations, based on the feedback of the user's options, the proposed methods can automatically adjust the weights of the recommendation methods to accommodate the variations of the users' musical interests, and then promptly recommend the favored and more interesting music. Experimental results show that the recommendation accuracy achieved by our method is as good as 90%. Hence, it is obviously helpful for recommending the interesting music if the music objects are annotated with the music emotions. For future work, we will apply our method to the research of music emotion annotations and test more categories of music with our method, in order to extend the applications. Moreover, the other features and recommendation methods will be also investigated.

Acknowledgements

This research was supported by Ministry of Economic Affairs, Taiwan, ROC under Grant No. 97-EC-17-A-02-s1-024 and by National Science Council, Taiwan, ROC under Grant No. NSC 97-3114-E-006-001.

References

Amazon. <http://www.amazon.com/>.
Chai, W., & Vercoe, B. (2000). Using user models in music information retrieval systems. In *Proceedings of international symposium on music information retrieval*.
Chen, H. C., & Chen, A. L. P. (2001). Music recommendation system based on music data grouping and user interests. In *Proceedings of the ACM international conference on information and knowledge management*.
Dixon, S., Bainbridge, D., & Typke, R. (2007). Sociology and music recommendation systems. In *Proceedings of the international symposium on music information retrieval*.
DownWithTheLoads. <http://www.downwiththeloads.com/>.
Folksongs. <http://ingeb.org/>.
iTunes. <http://www.apple.com/itunes/>.
Karydis, I., Nanopoulos, A., Papadopoulos, A., Katsaros, D., & Manolopoulos, Y. (2005). Content-based music information retrieval in wireless ad-hoc networks. In *Proceedings of the 6th symposium on music information retrieval*.
KIDiddles. <http://www.kididdles.com/>.
Krumhansl, C. L. (2002). Music: A link between cognition and emotion. *Current Directions in Psychological Science, 11*(2), 45–50.
Kuo, F. F., & Shan, M. K. (2002). A personalized music filtering system based on melody style classification. In *Proceedings of the IEEE international conference on data mining*.
Kuo, F. F., Chiang, M. F., Shan, M. K., & Lee, S. Y. (2005). Emotion-based music recommendation based on association discovery from film music. In *Proceedings of the ACM international conference on multimedia*.
Lee, W. P., Liu, C. H., & Lu, C. C. (2002). Intelligent agent-based systems for personalized recommendations in internet commerce. *Expert Systems with Applications, 22*(4), 275–284.
Lee, W. P., & Lu, C. C. (2003). Customising WAP-based information services on mobile networks. *Personal and Ubiquitous Computing, 7*(6), 321–330.
Liu, C. C., Hsu, J. L., & Chen, A. L. P. (2001). Discovering nontrivial repeating patterns in music data. *IEEE Transactions on Multimedia, 3*(3), 311–325.
Lu, C. C., & Tseng, V. S. (2006). Music classification by parsing main melody. In *Proceedings of the international computer symposium* (pp. 281–285).
Lu, L., Liu, D., & Zhang, H. J. (2006). Automatic mood detection and tracking of music audio signals. *IEEE Transactions on Audio, Speech, and Language Processing, 14*(1), 5–18.
Meldex. <http://www.nzdl.org/fast-cgi-bin/music/musiclibrary/>.
MIDI Database. <http://www.mididb.com/>.
MIDI Manufacturers Association. <http://www.midi.org/>.
Miller, H. M., & Williams, E. (1991). *Introduction to music*. Perennial Press.
Recordare. <http://www.recordare.com/software.html>.
Reilly, W. S. N. (1996). *Believable social and emotion agents*. Ph.D. dissertation, Carnegie Mellon University.
Sadie, S., & Latham, A. (1990). *The cambridge music guide*. Cambridge University Press.
Shardan, U., & Maes, P. (1995). Social information filtering: Algorithms for automating word of mouth. In *Proceedings of the conference on human factors in computing systems*.
Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford University Press.
Themefinder. <http://essen.themefinder.org/>.
Tiemann, M., Pauws, S., & Vignoli, F. (2007). Ensemble learning for hybrid music recommendation. In *Proceedings of the international symposium on music information retrieval*.
Tseng, Y. H. (1999). Content-based retrieval for music collections. In *Proceedings of the 22nd international ACM SIGIR conference on research and development in information retrieval* (pp. 176–182).
Tseng, V. S., Su, J. H., & Huang, J. H. (2006). A novel video annotation method by integrating visual features and frequent patterns. In *ACM KDD workshop on multimedia data mining, Philadelphia, USA*.
Tseng, V. S., Su, J. H., Wang, B. W., & Lin, Y. M. (2007). Web image annotation by fusing visual features and textual information. In *ACM symposium on applied computing (SAC), Korea*.
Typke, R., Wiering, F., & Veltkamp, R. C. (2005). A survey of music information retrieval systems. In *Proceedings of the 6th symposium on music information retrieval* (pp. 153–160).
Uitdenbogerd, A., & Schyndel, R. (2002). A review of factors affecting music recommender success. In Proceedings of the international symposium on music information retrieval.
Wang, M. Y., Zhang, N. Y., & Zhu, H. C. (2004). User-adaptive music emotion recognition. In *Proceedings of the 7th international conference on signal processing* (pp. 1352–1355).
Weyde, T., & Datzko, C. (2005). Efficient melody retrieval with motif contour classes. In *Proceedings of the 6th symposium on music information retrieval* (pp. 686–689).
Yang, D., & Lee, W. S. (2004). Disambiguating music emotion using software agents. In *Proceedings of the international symposium on music information retrieval* (pp. 218–223).
Yeh, C. C., Tseng, S. S., Tsai, P.C., & Weng, J. F. (2006). Building a personalized music emotion prediction system. In *Proceedings of the 7th Pacific rim conference on multimedia* (pp. 730–739).