Music Recommendation System Based on Preference Classification of Real-time User Brainwave

Abstract—In order to predict user-favorite songs, managing user preferences information and genre classification are necessary. In this paper, we propose a preference classification about content based on real-time user brainwave and a music recommendation system based on it. We focused on classifying real-time user preferences by analyzing the user’s brainwaves. The brainwaves are acquired using a wireless consumer Electroencephalography (EEG) device with small-sized pins in order to enhance the system’s usability for mobile devices. The performance of preference classification accuracy is nearly equal as that of one of the best EEG-based preference analyzer, despite the use of a comparatively lesser number of feature dimensions. In our study, a very short feature vector, obtained from low dimensional projection and already developed audio features, is used for music genre classification problem.

Keywords—Preference classification; Real time; EEG (Electroencephalography); Music recommendation component;

I. INTRODUCTION

It is very important to classify user’s preferences in a recommendation engine. For that, many explicit or implicit ways such as usage patterns analyzing and thumb up/down are used. In the past research, we proposed the MusicRecom system based on user’s music usage patterns. The preference classification through usage patterns is an implicit analyzing. This system has a cold-start problem because when a user starts recommendation system at first, the system doesn’t have his usage history. In this paper we propose another implied user’s preference analyzing based on user’s brainwaves. We didn’t use traditional heavy devices for EEG recording. The Brain Machine Interface (BMI) developed for MusicRecom is designed for the user’s comfort.

II. SYSTEM ARCHITECTURE

We propose the MusicRecom system which consists of four main functional modules. The preference classification module collects user’s brainwave signals and classifies them. The genre classification module extracts the features and classifies genres of music. We utilize the advanced algorithms for the proposed genre classification [9, 10]. The music information generator collects music information such as Singer, Title, Artist, Composer, and Album Name. The last one is the recommendation engine selecting the music that has the most similar features to user’s favorite songs.

![System architecture MusicRecom system based on EEG](image1)

![Feature extraction process](image2)

The feature extraction process of our system is shown in Fig. 2. First, input audio is pre-processed with decoding, down-sampling, and mono-conversion. The pre-processed audio is framed using hamming window of about 23ms with 50% overlap. From each window, raw features are obtained. In our system, mel-frequency cepstral coefficients (MFCC), Decorrelated filter bank (DFB), and Octave-based spectral contrast (OSC) are used. MFCCs represent the spectral characteristics based on Mel-frequency scaling, and they are also used in various music classification systems. The DFB...
considers variation of amplitudes between neighboring bands. It is extracted from subtraction of log spectrum in neighboring mel-scale band. The OSC considers the spectral peak, spectral valley, and spectral contrast in each octave-based subbands. After raw feature extraction, length of feature vector is 42: 13 for MFCCs, 13 for DFB, and 16 for OSC. After extracting 3 features, their statistical values are computed in order to represent temporal variation.

We used mean, variance, feature-based modulation spectral flatness measure, and feature-based modulation spectral crest measure [9]. Length of the feature vector is quadrupled, and we get a feature vector of dimension 168. This 168-dim vector can be used as it is, but sometimes, the feature vector is used after dimension reduction depending on system design. For an application using low-computational power, short feature vector is necessary. To reduce feature dimension without performance degradation, distance metric learning is applied in our system [10]. The feature vector of length from 5 to 168 is used after dimension reduction. In this study, we use 5 and 10-dim feature vector.

B. PREFERENCE CLASSIFICATION BASED ON BRAINWAVES

The proposed music recommendation system identifies the user’s emotional state by analyzing the EEG signals received as the user listens to music. The EEG-based music recommendation system has various advantages as compared with other recommendation techniques. First, the proposed service is able to respond to the user’s real-time emotional state, since EEG reflects the real-time emotional response of the user. Second, the BMI is one of the best methods for analyzing man-machine interaction in multi-tasking environments.

We use the Emotiv EPOC to collect EEG signals [3]. This device was verified as the best EEG headset among 13 outstanding EEG headsets in the usability test [4]. In order to extract the spectral features from 14-pin EEGs, the brain feature extractor receives the power spectra from these bands after taking the Fast Fourier Transform (FFT) of the 14 EEG signals in a window of 8 seconds. The EEG sampling rate of the emotive EPOC is 128 Hz [6], the window length for the 512-point FFT of the received EEG signals is 8 seconds, and the shift period is 8.7 ms [1], [4]. In the band cluster of the brain feature extractor, the power spectra of the EEGs are assessed in five frequency bands, as shown in Table I [2] [5], [7-8].

<table>
<thead>
<tr>
<th>TABLE I. REPRESENTATIVE CEREBRAL FREQUENCY BANDS AND THEIR CORRESPONDING FREQUENCY RANGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency Band</td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Delta (δ)</td>
</tr>
<tr>
<td>Theta (θ)</td>
</tr>
<tr>
<td>Alpha (α)</td>
</tr>
<tr>
<td>Beta (β)</td>
</tr>
<tr>
<td>Gamma (γ)</td>
</tr>
</tbody>
</table>

The process of feature extraction is illustrated in Fig. 4, \( p \) and \( t \) denote a pin index \( (1 \leq p \leq 14) \) and a window index \( (1 \leq t \leq T) \). The total number of windows in an acquired EEG signal, respectively. \( \text{pow}_{\delta}(t, p) \) is defined as the \( \delta \) band energy in \( t \)-th window from \( p \)-th pin. In a similar way, \( \text{pow}_{\theta}(t, p) \), \( \text{pow}_{\alpha}(t, p) \), \( \text{pow}_{\beta}(t, p) \), \( \text{pow}_{\gamma}(t, p) \) are defined. \( \text{fv}_t \)' is defined as the set of the five band energies from 14 pins in the \( t \)-th window. As a result, the dimension of \( \text{fv}_t \) is 70. The set of feature vectors \( \text{FV} \) for preference classification in MyMusicShuffler is presented as:

\[
\text{FV} = \{\text{fv}_1', \ldots, \text{fv}_t', \ldots, \text{fv}_N'\} \quad 1 \leq t \leq T
\]

\[
\text{fv}_t' = (\text{pow}_{\delta}(t,1), \text{pow}_{\alpha}(t,1), \text{pow}_{\beta}(t,3), \text{pow}_{\gamma}(t,3), \text{pow}_{\gamma}(t,34), \ldots, \text{pow}_{\delta}(t,14), \text{pow}_{\alpha}(t,14), \text{pow}_{\beta}(t,14), \text{pow}_{\gamma}(t,14), \text{pow}_{\gamma}(t,34))
\]

(1)

MusicRecom constructs a personalized preference model for a single user. A user listens to music clips and annotates each music clip with his or her preference on four levels (1-4, where four is a strong preference). Level 1 and 2 are classified as the ‘negative’ preference class, and the other levels are categorized as the ‘positive’ preference class. \( \text{fv}_t' \) in each music clip are scattered to 80 dimensional space with their corresponding preference classes. An SVM with an RBF kernel is applied to classify these two preference classes [9].

In testing, the proposed EEG music recommender gathers \( \text{fv}_t' \) for 8 seconds while a user listens to a music clip. Each \( \text{fv}_t' \) is classified as the ‘positive’ or ‘negative’ preference class using the user’s personal preference model. Preference classification results for a music clip are counted, and the EEG music recommender determines the dominant emotional preference as the most frequently counted preference for the music clips.

![Fig. 3. Process of feature extraction from EEG signals.](image)

C. RECOMMENDATION ENGINE

After user preference modeling using brainwaves, the system can know which content is suitable or not. The recommendation engine suggests songs similar with user’s preferred songs. We use the cosine coefficient to compare user’s preferred songs and music datasets. The recommendation engine extracts user preference features and dataset features. In Equation (2), we can control recommendation focus through preference weight. If user doesn’t want recommendation by singers, then singer preference weight can be set 0. At last we sort candidate songs by preference value and recommend top 15 songs.

\[
\text{PrefVal}_i = \sum_{k=0}^{w} \text{Sim}_k(D_i, T_j)
\]

(2)
where $D_i$ is the music feature vector lists of user preference music and $T_i$ is the music feature vector lists of dataset.

### III. EVALUATION AND IMPLEMENTATION

#### A. USERPREFERENCE CLASSIFICATION ACCURACY

To evaluate the proposed system, an EEG dataset is constructed for music recommendation. EEG data and the feedback related to the selected music items were collected. A music corpus, KETI AFA2000, which contains approximately 2,400 Korean pop mp3 clips, was used [10]. Thirteen participants took part in the experiment. Each participant reviewed the music list by listening to thirty second clips, and selected his or her ten favorite clips and ten least favorite clips from the AFA2000 corpus. The preference modeling module gathered the EEG responses to the selected clips. One-minute EEG signals were extracted for each of the twenty clips. One-minute clips of the selected items were played from the beginning of the song to the one minute mark. First, the participants listened to their ten favorite music clips; then, they listened to the ten least favorite music clips. The extracted dataset consisted of approximately 260 EEG signals of one minute in duration from the thirteen participants. For EEG acquisition, an Emotiv 14-pin and wireless EEG headset was used [11]. This headset is designed to use 14 specific sensor positions shown in Fig. 2: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4. Because the device acquires multi-channel EEG signals wirelessly, users feel more comfortable during the EEG acquisition process compared to a situation where wired devices are used. The signal acquisition space is restricted to a single room in order to maintain the same speaker systems and illumination conditions. The EEG acquisition room is equipped with soundproof walls, a light control system, quality stereo speakers, and a comfortable chair. These accuracies were measured by using ten-fold cross validation. The proposed system achieved an 81.07% accuracy rate when the classifier uses all the EEG signals from the fourteen pin.

#### B. RECOMMENDATION SATISFACTION

The MusicRecom suggests 10 songs to participant after learning about their brainwaves patterns. We had gathered usage history for 5 days in the MusicRecom system. In the past research [12], we used usage history for analyzing user preference. The assumption in the evaluation is that music which each user played or commented is their answers, and we compared these answers and the recommended music by the MusicRecom system. In order to evaluate the user’s satisfaction about the MusicRecom system, we proposed MRR-based evaluation method given Equation (3).

$$MRR = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{r_i}$$  

(3)

In Equation (3), $r_i$ means the ranking in the recommended list in by the proposed user recommendation engine. Table II shows the evaluation results of the MusicRecom recommended contents.

### IV. CONCLUSION

In this paper, we proposed a recommendation system based on a preference classification using real-time user brainwaves and genre feature classification. Proposed user’s preference classifier achieved an overall accuracy of 81.07% in the binary preference classification for the KETI AFA2000 music corpus. And we could recognize the user’s satisfaction when we use brainwaves. This system can be applied to various audio devices, apps and services.

### ACKNOWLEDGMENT

This research is supported by Ministry of Culture, Sports and Tourism(MCST) and Korea Creative Content Agency(KOCCA) in the Culture Technology(CT) Research & Development Program 2015.

### REFERENCES


---

**TABLE II. EVALUATION RESULTS WITH MRR**

<table>
<thead>
<tr>
<th>User preference</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on EEG</td>
<td>0.66</td>
</tr>
</tbody>
</table>

---