Brainwave-based Mood Classification Using Regularized Common Spatial Pattern Filter

Saim Shin[1], Sei-Jin Jang[2] , Donghyun Lee[1], Unsang Park[1], and Ji-Hwan Kim[1]
[1] Department of Computer Science and Engineering, Sogang University
35 Baekbeom-ro, Mapo-gu, Seoul 121-742, Republic of Korea
[e-mail: {miror, redizard, unsangpark, kimjihwan}@sogang.ac.uk]
[2] Smart Media Center, Korea Electronics Technology Institute
410-837 9FL, M-city Bldg, 195, Baegma-ro, Ilsan-Dongu, Goyang-si, Gyeonggi-do, Republic of Korea
[e-mail: sjjang@keti.re.kr]
*Corresponding author: Ji-Hwan Kim

Abstract

In this paper, a method of mood classification based on user brainwaves is proposed for real-time application in commercial services. Unlike conventional mood analyzing systems, the proposed method focuses on classifying real-time user moods by analyzing the user's brainwaves. Applying brainwave-related research in commercial services requires two elements - robust performance and comfortable fit of.

This paper proposes a filter based on Regularized Common Spatial Patterns (RCSP) and presents its use in the implementation of mood classification for a music service via a wireless consumer electroencephalography (EEG) device that has only 14 pins. Despite the use of fewer pins, the proposed system demonstrates approximately 10% point higher accuracy in mood classification, using the same dataset, compared to one of the best EEG-based mood-classification systems using a skullcap with 32 pins (EU FP7 PetaMedia project). This paper confirms the commercial viability of brainwave-based mood-classification technology.

To analyze the improvements of the system, the changes of feature variations after applying RCSP filters and performance variations between users are also investigated. Furthermore, as a prototype service, this paper introduces a mood-based music list management system called MyMusicShuffler based on the proposed mood-classification method.

1. Introduction

Human moods do not follow structured patterns. Moods vary temporally, and are inherently ambiguous. Mood classification results vary quite substantially between human annotators. Moods are expressed and perceived over multiple modalities. As a result, EEG-based mood classification systems have not achieved high accuracies despite involving lesser mood classes.

In the mood related studies, the Audio Music Mood Classification (AMC) tasks are annually held in Music Information Retrieval Evaluation eXchange (MIREX) [1]. AMC task evaluates music audio classification algorithms. MIREX 2013 published that the best accuracy obtained was 66.33% for this task involving five mood clusters (cluster 1: passionate, rousing, confident, boisterous, rowdy. cluster 2: rollicking, cheerful, fun, sweet, amiable/good natured, cluster 3: literate, poignant, wistful, bittersweet, autumnal, brooding. cluster 4: humorous, silly, campy, quirky, whimsical, witty, wry. cluster 5: aggressive, fiery, tense/anxious, intense, volatile, visceral) [2]. Mood classification was evaluated using speech. In this evaluation, a speech mood classification system detected moods from speech signals. An HMM based speech mood classification system showed a 54.3% accuracy across five mood classes - excitement, frustration, happiness, neutral, and sadness [3].

A Brain-Computer Interface (BCI) system is aimed at directly controlling a machine using a human brain signal. A typical BCI system calibrates the neural activities of the brain and builds a classification model to generate a command. Electroencephalography (EEG) is the recording of electrical activity along the scalp and is widely used as the measurement of the neural activities in a noninvasive BCI system. Owing to its usability, numerous research and competitions have been conducted to build a robust classification model for BCI tasks [4].

The non-stationary nature in EEG data has been studied in related researches. Even if a subject performs the same task under the same conditions, the EEG data may indicate a significant variation compared to those acquired during a different session owing to the subject's cerebral condition including stress, emotion, hormones, subject attention, and fatigue [5]. EEG data analyzing techniques consider defects arising from noise and undesirable signals.

In this paper, we propose an EEG data classification technique to determine a user's real-time moods to overcome this limitation. This paper also proposes a preprocessing technique for EEG-based mood classification. Because of the low spatial resolution of EEG data, spatial filtering is often used to facilitate the extraction of useful features for BCI. Spatial filtering applies a linear combination to the channels. The system proposed in this paper utilizes Regularized Common Spatial Patterns (RCSP) as a preprocessing filter prior to mood classifications using EEG data [6]. The experimental results demonstrate that the RCSP filter successfully alleviates the non-stationarities of brainwave signals, especially in the extraction of long-term patterns.

In this work, performance variations between pin positions and users are further investigated to analyze improvements of the proposed system. These discussions of the experimental results also reflect the status of BMI technology for mood related services.

The proposed service, MyMusicShuffler, uses personalized user-mood models to recognize positive and negative moods based on the real-time brain signals received from the user. The objective of MyMusicShuffler is to maximize user convenience by allowing a user

to enjoy music without interruption or stress [7]. After analyzing the user's mood using these personalized mood models, this system automatically updates the music playlist. The music entries in the music list in this system are classified to a positive mood based on brainwaves. As such, it precludes the use of heavy devices for the acquisition of brainwaves; instead, wireless devices are utilized. EEG is the recording of electrical activity along the scalp. Compared to the devices used in previous mood analysis systems, consumer EEG devices are wireless and convenient, thus enhancing the applicability of the proposed EEG-based system for commercial purposes.

The remainder of this paper is organized as follows. In Section 2, previous studies on mood classification based on brainwaves are discussed. In Section 3, the proposed mood classification system is introduced. In Section 4, data acquisition and other experimental conditions are outlined and the mood-classification results are analyzed. Section 5 introduces a prototype service, MyMusicShuffler, which applies the proposed mood-classification techniques to a music service application. Finally, Section 6 concludes this paper.

2. Previous Works

In the fields of medical science and psychology, numerous brain functions have been investigated using neuroimaging technologies including functional magnetic resonance imaging [8], positron emission tomography [9], and magnetoencephalography [10]. These technologies require expensive equipments and experts who collect and analyze the information. Consequently, applications for brain-information systems have been developed primarily in hospitals and research centers.

2.1 Brain Computer Interfaces Based on EEG Data

An important advance is the evolution of EEG acquisition devices. Although EEG skullcaps guarantee relatively stable and sufficient information with multi-channel signals, they are not appropriate for applications in commercial services. Because wearing a skullcap is time consuming and users should endure the uncomfortable design, user-friendly, and low cost EEG devices (e.g., headsets or headband-style devices) have recently become available on the market [11].

Many relatively lower cost devices have been introduced using EEG data [12-14]. EEG represents the voltage fluctuation resulting from ionic current flows within the neurons in the brain. EEG resolution is lower than that of the above-mentioned methods. The EEG sensors are located on the scalp of the user. Although EEG data suffer from this relatively poor sensitivity compared to other brain information, these devices remains popular because its hardware costs are significantly less than those of the above-mentioned techniques [15]. To obtain accurate EEG data, the sensors must be local positions and the pressure between the sensors and the scalp should be maintained at a constant level. Many devices, called EEG skullcaps, have been designed in a skullcap style.

Performance comparisons between a skullcap and a headset-style EEG device were presented in a P300 speller domain. A P300 wave occurs when a desired target is observed. The P300 speller presents an *n*-by-*n* matrix on a monitor. One of the letters flashes randomly. If it is a letter desired to be written, a P300 wave appears. The difference in accuracy between an EEG skullcap and EEG headset with the same EEG classifier was measured as 4.77% and 6.12% in sitting and walking conditions, respectively [16]. However, the BMI techniques for the P300 speller involve recognition of the instantaneously evoked potential signals in an EEG.

Analyzing long and continuous patterns is necessary in the process of classifying moods. Therefore, various discussions on the experimental results are necessary for their application to a mood-based mood related systems.

Another important advance in EEG classification is using spatial filters. EEG classifiers extract features from the multiple signals based on the neighboring pins simultaneously. In this process, there exist redundant and correlated semantics between the signals from neighboring pins. To solve this problem, multiple spatial filtering techniques are applied in EEG classifications. Bipolar [17] and Laplacian filters [18] have reported that these techniques can emphasize the localized activity and reduce the diffusion of spatial activity. Recently, Common Spatial Pattern (CSP) filters have been applied in the Motor Imaginary (MI) domain [19]. These are designed to identify spatial projections that maximize the power/variance ratios of the filtered signals for two classes [20]. In particular, CSP-based filters have proven their superiority over other filters, having recorded 7 to 9% accuracy improvements in Motor Imaginary classifications and P300 spellers.

2.2 Mood Classification Based on EEG Data

In the case of EEG-based mood-classification systems, many advances have been made using different machine-learning techniques [10-12], [21]. However, their performance varies significantly according to the mood classes and EEG data acquisition environments. Mood-classification performance demonstrated a 93% precision when 60 pins of EEG acquisition points on a skullcap were used; however, the system recorded only a 35% precision in four pins for the mood classes defined as 'asleep', 'awake', and 'other' [10], [22].

Mood-classification performance using an EEG skullcap was evaluated using neural network [11]. In this research, EEG data were collected from 68 EEG sensors while participants listened to original sound tracks with four different moods ('joy', 'anger', 'sorrow', and 'relaxation'). The mood-classification accuracies were measured as 54.5%, 67.7%, 59.0%, and 62.9% for 'joy', 'anger', 'sorrow', and 'relaxation', respectively. The Bayes classifier was applied to classify three levels of arousal [23]. In this research, an accuracy of 58% was recorded when 37 signals were used: 32-channel EEG data from a skullcap, galvanic skin response [24], blood pressure, heart rate, respiration, and temperature. However, in this study, only the arousal level of the user in the mood space was measured and not the actual emotional state. Therefore, the system could not measure more complex responses.

Binary moods are abstract and comprise combination of many emotions [25]. For example, the emotions such as anger and fear, popularly referred to as 'negative', have 'negative valence' [26-27]. Therefore, binary mood classification systems cannot always be expected to yield high accuracies. An SVM-based binary mood classification has been performed, which involved measuring EEG signals while listening to Chinese pop songs; its highest accuracy was 72.21% [28]. Another representative mood-classification research using an EEG skullcap was the EU FP7 PetaMedia project [15]. In this research, the user preference for a given music video clip was identified in binary-mood classes (positive or negative) using 32-channel EEG data and additional features. This project utilized a Dataset for Emotion Analysis using EEG, Physiological, and video signals (DEAP) [22], which acquires EEG and peripheral physiological signals when the user views music video clips. The binary-mood classification achieved an accuracy rate of 70.25% when a Support Vector Machine (SVM) classifier with radial basis function (RBF) kernels using 32-channel EEG data.

3. Mood Classification Using EEG Signals

3.1 Mood-based services and EEG

The proposed EEG-based mood-classification system identifies the user's emotional state by analyzing the received EEG signals as the user listens to music. The mood classifier shows the positive or negative mood of a user while listening to a music clip. By applying this technique, this paper proposes a mood-based music list management service. Music clips indicating a positive mood are recommended to the users. Music clips classified with a negative mood in the EEG-based mood classifier are changed to different music items according to user's policy: Random selection or music items within the user's favorite list. If a new item is not contained in a user's favorite music list, the newly determined positive music is added to the user's favorite list. The proposed EEG-based music service has several advantages compared to other music services. First, the proposed service can respond to a user's real-time emotional state because the EEG data reflect the real-time emotional response of the user. Second, BMI is one of the best methods for analyzing man-machine interaction in multi-tasking environments. Users typically listen to music as they perform other activities such as working, exercising, or reading. The proposed system is based on a stress-free and hands-free interface that requires no direct user interaction.

3.2 Mood Classification by Analyzing EEG Data

Fig. 1 illustrates the process of the proposed mood classification using EEG. The brain feature extractor receives the time-domain EEG data from the wireless EEG acquisition device. In this experiment, the Emotiv EPOC is used [29]. This device was verified as the best EEG headset among 13 available EEG headsets in the usability test [21].
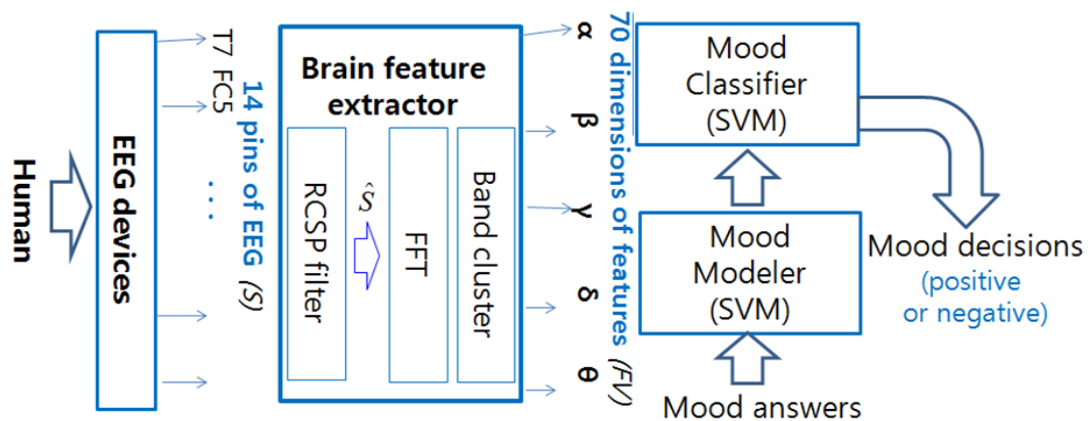

Fig. 1. Process of mood classification using EEG

This study uses EEG data for the real-time detection of the user's mood. The 10–20 international system is the standardized electrode placement of the American Electroencephalographic Society and an internationally recognized method for describing and applying the locations of scalp electrodes in the context of an EEG test or experiment [30]. Each site has a letter that identifies the lobe and a number that identifies the hemisphere. The

'10' and '20' refer to the fact that the actual distance between adjacent electrodes is either 10% or 20% of the total front–back or right–left distance of the skull [30]. This system ensures acquisition of the stereotypical signals based on the relationship between the location of an electrode and the underlying area of the cerebral cortex.

The 14 EEG acquisition points in EPOC are illustrated in **Fig. 2**. The EEG collector gathers the continuous multi-channel signals and then transfers these to the brain feature extractor through a wireless connection. As indicated in **Fig. 2**, the system continually acquires signals from the 14 pins including two references as the user listens to the music items and categorizes them based on the pin positions.

The brain feature extractor receives the power spectra from these bands after employing a RCSP preprocessor and a Fast Fourier Transform (FFT) of the 14 EEG data in a window of eight seconds. FFT is a popularly simple and fast method for signal processing and the most widespread method in EEG data processing. FFT is not a separate transform; rather, it is an efficient algorithm for computing the Discrete Fourier Transform (DFT) of a sequence. It is useful in signal processing of EEG data, where it uses range from filtering and frequency analysis to power spectrum estimation [31].
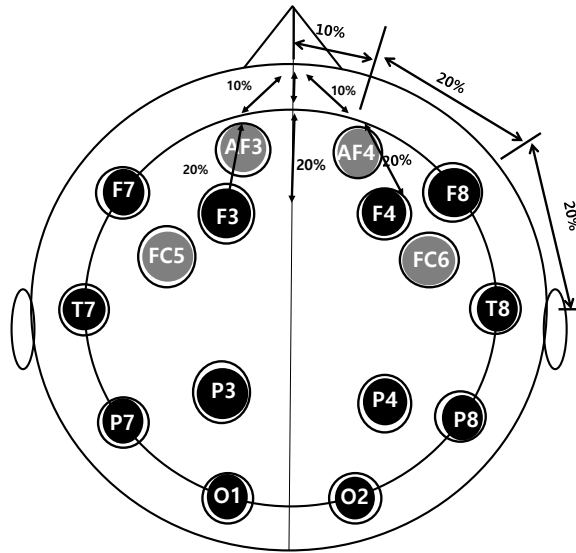

Fig. 2. Map of EEG electrode sites in EPOC device based on the 10-20 international system. The 10-20 international system was standardized by the American Electroencephalographic Society [30]
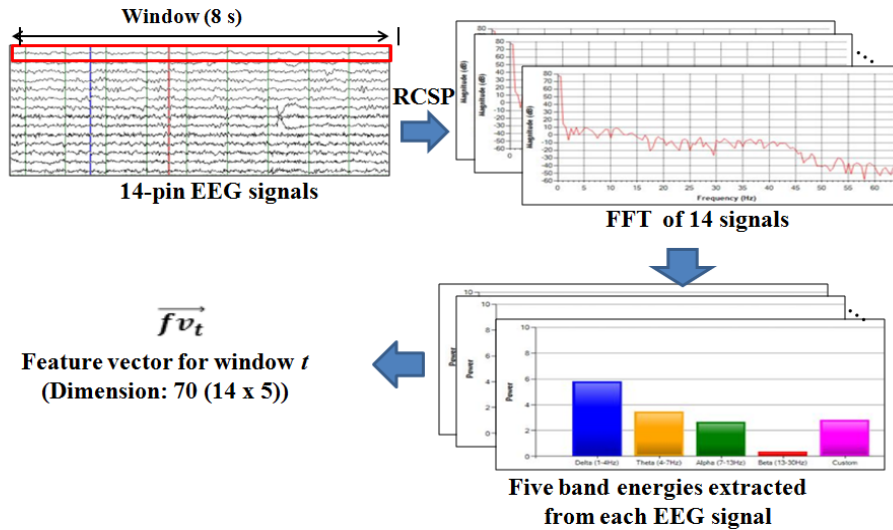
In the RCSP filter, the EEG signals are estimated to reduce the diffusion because of spatial activity among the signals. The applied RCSP filter encodes the most discriminative information from multiple signals [6]. Due to the smearing effect of the skull, the underlying source signal is spread over several channels. RCSP is useful in recovering the original source signal.

The brain feature extractor in **Fig. 1** analyzes the signals to extract the necessary features for the mood classification. EEG signals are typically described in terms of rhythmic activity. The rhythmic activity is divided into frequency bands. The representitive of the cerebral signals observed in a scalp EEG is in the range of 1-20 Hz when its clinical recording techniques are used [32].

Table 1. Representative cerebral frequency bands and their corresponding frequency ranges

Frequency Band	Frequency Range (Hz)
Delta (δ)	0-4
Theta (θ)	4-7
Alpha (α)	7-13
Beta (β)	13-30
Gamma (γ)	30-50

The EEG sampling rate of the Emotive EPOC is 128 Hz [29]. The window length for the 512-point FFT of the received EEG data is eight seconds, and the shift period is 8.7 ms [10], [33]. In the band cluster of the brain feature extractor, the power spectra of the EEG data are assessed in five frequency bands, as indicated in **Table 1** [15] [26-27], [33-34].

Fig. 3. Process of feature extraction from EEG data. Five band energies extracted from each EEG signal constitute the feature vector for their corresponding window.

The process of feature extraction is illustrated in **Fig. 3**. p and t denote a pin index ($1 \leq p \leq 14$) and a window index ($1 \leq t \leq T$. T: the total number of windows in an acquired EEG signal), respectively. $E_{\delta(t, p)}$ is defined as the δ band energy in the t-th window from the p-th pin. In a similar manner, $E_{\theta(t, p)}$, $E_{\alpha(t, p)}$, $E_{\beta(t, p)}$, and $E_{\gamma(t, p)}$ are defined. fv_t' is defined as the set of the five band energies from the 14 pins in the t-th window. Consequently, the dimension of fv_t' is 70. The set of feature vectors FV for mood classification in MyMusicShuffler is presented as:

$$FV = \left\{ fv_1', \cdots, fv_t', \cdots, fv_T' \right\}$$
$$fv_t' = (E_{\theta(t,1)}, E_{\alpha(t,1)}, E_{\beta(t,1)}, E_{\gamma(t,1)}, E_{\delta(t,1)}, \cdots, E_{\theta(t,14)}, E_{\alpha(t,14)}, E_{\beta(t,14)}, E_{\gamma(t,14)}, E_{\delta(t,14)}) \quad (1)$$

The proposed EEG mood classifier gathers fv_t' for 15 seconds, while a user listens to a music clip. The set of FV for a song is separated with 1,724 fv_t', which are based on an eight second window and 8.7 ms shift period. Each fv_t' is classified as a 'positive' or 'negative' mood class using the user's personal mood model. Mood-classification results for a music clip are counted and the EEG music recommender determines the dominant emotional mood as the

most frequently counted mood for the music clip.

MyMusicShuffler constructs a personalized mood model for a single user. To construct the personal mood model, EEG data while listening to music are collected. 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' mood class and the other levels are categorized as the 'positive' mood class. fv_t' for each music clip is scattered to 70 dimensional space with the corresponding mood classes. An SVM with an RBF kernel is applied to classify these two mood classes [35].

3.3 RCSP Filter based EEG Signal Estimation

To extract the normalized spectral features from 14-pin EEGs, we propose the feature estimation applying a spatial filter technology. Because of the low spatial resolution of EEGs, spatial filtering can be used to facilitate the extraction of the useful features for BCI classification. Common Spatial Pattern (CSP) is a supervised method to find spatial filters. This technique finds a direction vector which maximizes the variance of band-passed EEG signals from one class while minimizing the signal variance of another class [6]. CSP based filters still show the sensitivity to noise and over-fitting to the applied samples. To overcome this sensitivity, this paper applies Regularized CSP (RCSP) among the several CSP variants [6].

S_i is an EEG signal matrix for class *i*, with EEG channels as columns and time samples as rows. An average spatial covariance matrix of S_i, is represented by C_i, $i \in \{1, 2\}$. In this paper, classes are aligned with a positive and a negative emotion. To find CSP based filters, a spatial filter $\hat{w}$ is determined by Equation (2) such that the variance of the filtered signal becimes the maximum for one class and the minimum for the other class.

$$\hat{w} = \arg \max_w \left(\frac{w^T C_1 w}{w^T C_2 w} \right) \qquad (2)$$

The maximization problem of Equation (2) is solved as a generalized eigenvalue problem by Equation (3). The eigenvector *w* corresponding to the maximum eigenvalue is the direction vector. This direction vector *w* is a spatial filter for S_i and w^{-1} is called a common spatial pattern for S_i.

$$C_1 w = \lambda C_2 w \qquad (3)$$

RCSP uses a regularization framework to penalize undesired solutions by two user-defined regularization parameters. This paper applies Tikhonov Regularization CSP algorithm [36], which is expressed as:

$$\hat{w}_I = \arg \max_w \left(\frac{w^T \tilde{C}_1 w}{w^T \tilde{C}_2 w + \alpha w^T I w} \right) \quad with \ \tilde{C}_i = (1 - \beta)C_i + \beta I \qquad (4)$$

Where $\tilde{C}_i$ is the regularized C_i , and *I* is a identity matrix whose size is *N* by *N*. In this process, α and β are user-defined regularization parameters ($0 \leq \alpha, \beta \leq 1$). $\hat{w}_I^{-1}$ is a common spatial pattern for S_i. This regularization is able to constrain the influence of outliers in the

solution of filters. In this RCSP feature extraction, input EEG signal S_i is projected as in Equation (5). $\hat{S}_i$ is the input of FFT module in brain feature extractor in **Fig. 1**.

$$\hat{S}_i = \hat{w}_I^{-1} S_i \tag{5}$$

The RCSP is a suitable filter for reducing the variance associated with sample-based estimates. Noisy EEG data from the wireless EEG acquisition headsets can also be normalized by using RCSP filter.

4. Experiments

4.1 Experimental Setup

To evaluate the proposed method, two kinds of EEG evaluation datasets for mood classification, DEAP and KETI EEG datasets, were used. The age and gender distribution of the participants is listed in **Table 2**.

Table 2. Participant age and gender distribution in the applied EEG data sets

Age	Number of Participants			
	DEAP		KETI EEG	
	Male	Female	Male	Female
10s	-	1	-	-
20s	10	12	3	2
30s	7	2	4	3
40s	-	-	1	-
Total	17	15	8	5

4.1.1 Construction of DEAP

The DEAP was developed in the EU FP7 PetaMedia project [22]. Thirty-two people participated in the preparation of the DEAP. While each participant watched 40 music video excerpts of one-minute duration each, 32-dimensional EEG signals using 10-20 international system [30] and various peripheral physiological signals - Galvanic skin response (GSR) [24], respiration amplitude, skin temperature, electrocardiogram (ECG) [37], blood volume by plethysmograph, electromyograms of Zygomaticus and Trapezius muscles (EMG) [37], and electrooculogram (EOG) [17], skin temperature and respiration - were collected. The annotated results for binary mood are contained in this dataset.

4.1.2 Construction of KETI EEG Dataset

The KETI EEG dataset was constructed for mood classification in this research. The EEG data and 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 [38]. Thirteen subjects participated in this construction. Each participant reviewed the music list by listening to clips of 30-second duration and selected his or her ten favorite clips and ten least favorite clips from the KETI AFA2000 corpus. Because the objective of this experiment was

to demonstrate the effect of audio stimuli on EEG data without visual stimuli, the participants only listen to music clips in this experiment. The mood-modeling module gathered the EEG responses while the participants listen to the selected clips. One-minute EEG data were then extracted from the twenty clips. These 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 their ten least favorite music clips. The extracted dataset consisted of approximately 280 EEG signals of one-minute duration from the 13 participants.

For the EEG acquisition, an Emotiv 14-pin, wireless EEG headset was used [29]. This headset is designed to use the 14 specific sensor positions identified 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 data wirelessly, users felt more comfortable during the EEG acquisition process compared to a situation where wired devices were used.

The signal acquisition space was restricted to a single room to maintain the same speaker systems and illumination conditions. The EEG acquisition room was equipped with soundproof walls, a light control system, quality stereo speakers, and a comfortable chair. **Fig. 4** illustrates the room and the EEG device setup in this room.

Fig. 4. EEG data acquisition room and EEG device setup

4.2 Experimental Results

This section presents the performance evaluation of the EEG mood classifier by measuring the mood-classification accuracies. The accuracy is defined in the ratio of the correctly detected results by the evaluated system to the total number of samples examined [39]. In the experiments, the accuracies were measured using a ten-fold cross validation method. **Table 3** summarizes the results. Accuracy is used by a performance measure in this paper.

The proposed system achieved an accuracy rate of 83.01% when the classifier used all data from the KETI EEG dataset from the 14 pins. In particular, when using the RCSP spatial filter before the SVM classifier is applied, there was a remarkable improvement in the performance. The performance comparison with the EU FP7 PetaMedia project using the DEAP is also described in **Table 3**. The EU FP7 PetaMedia project reported an accuracy of 70.25% [22]. In the experiments, two mood classifiers were constructed using all of the 32 EEG signals and 14 signals corresponding to the pin positions indicated in **Fig. 2**. The accuracy of the mood

classifier for the proposed system by applying the SVM only was 65.63% with 14 pins and 68.89% with 32 pins. These performances were slightly lower than the system introduced in the EU FP7 PetaMedia project, because the proposed method uses only a subset of the EEG features in the DEAP. In **Table 3**, the performance of the proposed classifier by applying the SVM and RCSP filter was shown to be 80.55% with 14-pin EEG data and 82.4% with 32-pin EEG data, respectively. The proposed system showed superior performance compared to the classifier in the EU FP7 PetaMedia project [17]. Because the performance from EU FP7 PetaMedia project system did not apply the RCSP filters, the results confirm that the classifier with an RCSP filter is suitable for mood classification even though the applied EEG data is constructed with lesser number of channels.

Table 3. Evaluation results

Data	System	Feature	Accuracy (%)	Remarks
D E A P	SVM	14-pin EEG data	65.63	The pin positions are the same as in **Fig. 2**.
		32-pin EEG data	68.89	Part of 32 pins in 64-pin of 10-20 international system are used [29].
	RCSP-SVM	14-pin EEG data	80.00	The pin positions are the same as in **Fig. 2**.
		32-pin EEG data	82.4	Part of thirty two pins in 64-pin of 10-20 international system are used [29].
	EU FP7 PetaMedia project system [17]	Thirty two-pin + physiological features + audio-visual features	70.25	All features in DEAP are used.
KETI EEG	SVM	14-pin EEG data	81.07	
	SVM-RCSP	14-pin EEG data after RCSP filtering is applied.	83.02	

The EEG data in the DEAP were captured by applying both audio and visual stimuli while watching music video clips, but the KETI EEG dataset did not use the visual stimuli; the subjects only listened to the audio clips. The experimental results indicate that the proposed system can achieve robust performance only with audio stimuli.

To analyze the effectiveness of the proposed method, the analysis of effect by using RCSP filtering in EEG signals is explained in Section 4.2.1 and the experimental results of performance variation among users is discussed in Section 4.2.2.

4.2.1 Effect analysis of RCSP filtering in EEG signals

The results of the feature distribution patterns after applying RCSP filters in EEG classifications are shown in **Fig. 5**. Sample EEG signals from three pins for a participant in each evaluation dataset are selected. The participants recording the best increase in accuracy when applying RCSP filters from two datasets are assigned to the sample participants. A sample in these datasets contains 14 signals, however, to represent the dimensions in 3D, three dimensions are selected (DEAP: F7, F4, and T8; KETI EEG dataset: AF3, F7, and F8 in **Fig. 2**) per sample participant as dimensions in **Fig. 5**. The selected three dimensions show the best

accuracy improvements after RCSP filtering, in the experimental results, for each participant when constructing single-pin based mood classifiers. From the sample signals, the sums of δ-band energies are extracted and scattered in each graph; the selected pins from each signal decide the dimensions in each graph. The scattered band in this figure is decided based on the feature value that is the lowest band among the introduced five frequency bands in **Table. 1**, in order to reduce the unit of the graphs considering expressiveness of the graphs. The changes of feature areas after applying RCSP filtering are represented in these feature dimensions. The graphs in (a) are the analyzed results for DEAP and the graphs in (b) are sampled feature dimensions from KETI EEG dataset. The coordinate axes of the two graphs in **Fig. 5** are expressed using different scales because the EEG signals in two datasets were acquired by using two different EEG devices that show different responsiveness. The colored points in each graph represent the distributions of δ-band energies of each time point from the signals. The blue and green points mean the sampled δ-band energies of the features obtained from the signals using RCSP filters and the features of the signals obtained without RCSP filters, respectively. The effects of RCSP filters were demonstrated in **Fig. 5** by comparing the distributed areas of δ-band energies.
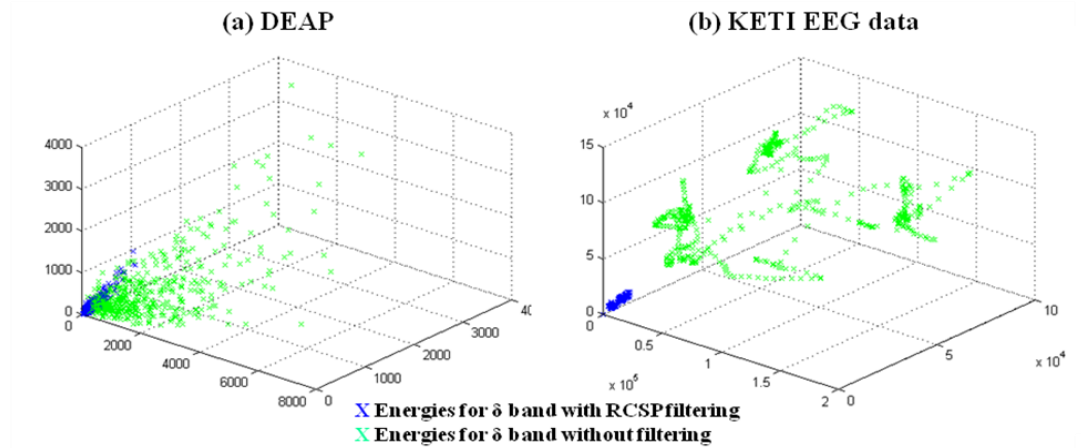

Fig. 5. Distributions of energies in δ-band after applying RCSP filtering in feature dimensions

In these results, the blue points show lower variance of feature distributions than the green points. By referencing in Equation (2), RCSP based filters are selected in order to maximize the variance of the filtered signals. Based on the assumption of this filter, in the experimental results, the significant decrease of the variances in feature distributions are proven in **Fig. 5**. The sparse areas need a more complex feature space with frequent regression cases compared to the dense area in the machine learning stage of SVM. This implies that the changed feature areas can be successfully detected in the machine learning stages. Therefore, this result demonstrates that the proposed RCSP based filter effectively retrieves the hidden semantics from EEG signals to the feature dimensions for EEG mood classification. Moreover, RCSP filtered feature areas tend to change linearly. These patterned feature dimensions by RCSP filters support the successful classifications with the kernel functions in SVM. Therefore, the performance improvement in the proposed mood classification methods based on RCSP filters and SVM is observed. This result also implies that RCSP filtering provides a robust increase in the performance of the mood classification systems by effectively reducing the feature dimensions as a preprocess module.

4.2.2 Analysis of performance variation between users

To analyze the performances by applying RCSP filters between users, the mood-classification results for all the participants from the two evaluation datasets are shown in **Fig. 6**.
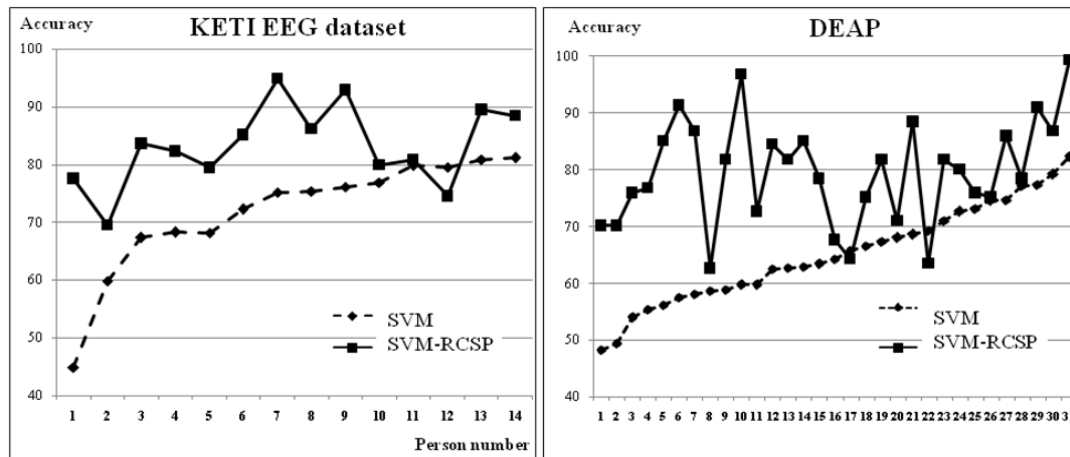

Fig. 6. Mood-classification accuracies for each participant. The total number of participants depicted in these graphs is 45. 14 dots in the left graph represent the accuracies for the KETI EEG dataset, and the 31 dots in the right graph present the accuracies for the DEAP. All results were measured with the classifier using all signals from the 14 pins.

The two graphs in **Fig. 6** show the performance changes after applying the RCSP filter. In each graph in **Fig. 6**, the dotted lines represent the accuracies of the mood classifications while applying only the SVM. The solid lines show the distributions of the performances when applying the SVM after the RCSP filter. In **Fig. 6**, when using RCSP filters, the more the number of recorded improvements, the lower the accuracies are shown when only applying SVM. By comparing the differences in these changes, it can be seen that applying RCSP filter improves the performances for cases with low accuracies as compared to the cases with high accuracies. The differences indicate that the RCSP filter is an effective preprocessor for extracting hidden semantic layers for mood classification. Therefore, RCSP filter-based preprocessing modules can guarantee robust performances by reducing the performance gap that exists between the results for various participants in EEG based mood classification systems.

5. MyMusicShuffler: A Prototype Service Based on Mood Analysis

This paper introduces a new prototype automatic music shuffling service by applying the proposed mood-classification technique. The BMI developed for MyMusicShuffler is designed for user comfort. **Fig. 7** displays the system architecture for the prototype. The proposed service, MyMusicShuffler, uses personalized user-mood models, which recognize positive and negative moods based on the real-time EEG data received from the user. The objective of MyMusicShuffler is to maximize user convenience by allowing the users to enjoy music without interruption or stress [4]. After analyzing the user's mood using these personal mood models, the system automatically updates the music playlist. The music items on the music list in the system are classified by mood based on the EEG. Because signal acquisition is

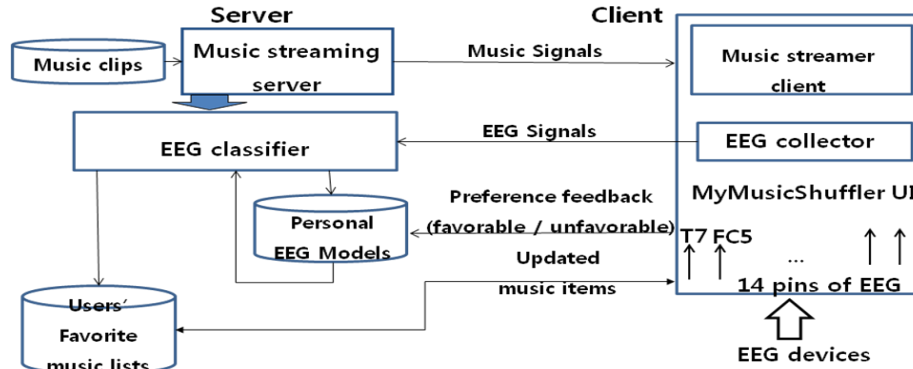based on wireless devices, the use of heavy devices is avoided.

Fig. 7. System architecture of MyMusicShuffler. T7 and FC5 pins are the pin positions in the 10-20 international system [30].

The implementation of the proposed music service are displayed in **Fig. 8**. The proposed UI supports the following scenario. Based on the dominant mood determined by the binary-mood classifier, the EEG mood classifier manages the user's playlist without the requirement for user interaction. Users can change the settings to allow the music to be limited to the items in a given playlist or to be selected randomly from the entire library. If the classification result from the user's EEG is negative when the user listens to a particular music item, that item is removed from the playlist. If the result is positive, the system saves the item in the playlist. The mode of 'starting music item' in this interface can be randomly selected from all music items or each user's previously positive music items based on each user's choice.

The prototype interfaces in **Fig. 8** suggest music video clips related to user's favorite music items. However, in the process of generating the models, the video stimuli are not applied. Therefore, the proposed method can be applied to different audio-based services.

MyMusicShuffler was constructed for a client/server system. The client module gathers the EEG data and delivers the signals to the server. The client module indicates the classification results graphically to the user and manages the playlist. The mood modeler and the EEG music recommender on the server train the mood models, classify the mood, and deliver the updated list to the client.

Fig. 8. Implementation results of MyMusicShuffler.

6. Conclusions

In this paper, mood classification using user brainwaves is proposed. The proposed RCSP based algorithms achieved an overall accuracy of 83.4% in the binary mood classification for the KETI AFA2000 music corpus. The performance of MyMusicShuffler was higher than one of the best EEG-based mood-classification systems (the proposed: 83.4%, EU FP7 PetaMedia project: 70.25%), despite the use of only audio stimuli and simple features.

For further analysis, effect of RCSP filters in feature dimensions and performance variations between different users were investigated. In these experimental results, RCSP filter is a successful preprocessing module to increase the accuracies of mood classification by decreasing the feature variations including same class. RCSP filter-based preprocessing modules can also guarantee the robust performances by reducing the performance gap existing between the results for various participants. The prototype music list management system, MyMusicShuffler, was implemented based on the proposed mood-classification technique using EEG data. Conventional music services primarily consider metadata generated in the past, whereas the proposed system focuses on real-time user moods based on user brainwave analysis for managing music lists.

References

[1] D. Stephen, K. West, A. Ehmann et al., "The 2005 music information retrieval evaluation exchange (MIREX 2005): preliminary overview," in *Proc. of the Sixth International Conference on Music Information Retrieval (ISMIR 2005)*, London, UK, pp. 320-323, Sep. 2005. Article (CrossRef Link).

[2] R.Panda, B. Rocha, R. Paiva, "Music emotion recognition with standard and melodic audio features," *Internatioinal Journal of Applied Artificial Intelligence*, vol. 29, no. 4, 2015. Article (CrossRef Link).

[3] K. Han. D. Yu, I. Tashev, "Speech emotion recognition using deep neural network and extreme learning machine," *INTERSPEECH*, Dresden, pp.223-227, Sep. 2015. Article (CrossRef Link).

[4] R. Krepki, B. Blankertz, G. Curio, et al., "The Berlin brain-computer interface (BBCI) – towards a new communication channel for online control in gaming applications," in *Proc. of International Conference of Multimedia Tools and Applications*, pp.73-90, Feb. 2007. Article (CrossRef Link).

[5] S. Sun, J. Xhou, "A review of adaptive feature extraction and classification methods for EEG-based brain-computer interfaces," *International Joint Conference on Neural Networks (IJCNN)*, pp. 1746-1754, Beijing, China, Jul. 2014. Article (CrossRef Link).

[6] H. Lu, H. Eng, C. Guan, et al., "Regularized common spatial pattern with aggregation for EEG classification in small-sample setting," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 12, pp. 2936-2946, 2010. Article (CrossRef Link).

[7] F. Lotte, C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE Transactions on biomedical engineering*, vol. 58, no. 2, pp. 355-362, 2011. Article (CrossRef Link).

[8] S. Shin, D. Jang, J. Lee, et al., "MyMusicShuffler: mood-based music recommendation with the practical usage of brainwave signals," in *Proc. of IEEE International Conference on Consumer Electronics*, Las Vegas, USA, pp. 355-356, Jan. 2014. Article (CrossRef Link).

[9] D. Bailey, D. Townsend, P. Valk, et al., "Positron emission tomography: basic sciences," *Springer-Verlag*, 2005. Article (CrossRef Link).

[10] S. Huettel, A. Song, G. McCarthy, "Functional magnetic resonance imaging," *Sinauer*, 2004. Article (CrossRef Link).

[11] N. Carlson, "Physiology of behavior," *Pearson Education*, 2013. Article (CrossRef Link).

[12] D. McFarland, L. McCane, S. David, et al., "Spatial filter selection for EEG-based communication," *Journal of Electroencephalography and Clinical Neurophysiology*, vol. 103, no. 3, pp. 386–394, 1997. Article (CrossRef Link).

[13] Y. Wang, Z. Zhang, Y. Li, et al., "BCI competition data set IV: an algorithm based on CSSD and FDA for classifying single trial EEG," *IEEE Transactions on Biomedical Engineering*, vol.51, no. 6, pp. 1081-1086, 2004. Article (CrossRef Link).

[14] Z. Wei, "Utilizing EEG signal in music information retrieval" *Master Thesis*, Department of Computer Science, National University of Singapore, 2010. Article (CrossRef Link).

[15] K. Ishino, M. Hagiwara, "A feeling estimation system using a simple electroencephalograph," *IEEE International Conference on Systems, Man and Cybernetics*, Washington, D.C., pp. 4204-4209, Oct. 2003. Article (CrossRef Link).

[16] P. Vespa, V. Nenov, M. Nuwer, "Continuous EEG monitoring in the intensive care unit: early findings and clinical efficacy," *Journal of Clinical Neurophysiology*, vol. 16, no. 1, pp. 1-13, 1999. . Article (CrossRef Link).

[17] D. Matthieu, C. Thierry, P. Mathieu, et al., "A P300-based quantitative comparison between the Emotiv EPOC headset and a medical EEG device," *International Journal of Biomedical Engineering*, vol. 12, no. 56, 2013. Article (CrossRef Link).

[18] J. Wolpaw, D. McFarland, G. Neat, et al., "An EEG based brain-computer interface for cursor control," *Journal of Electroencephalography and Clinical Neurophysiology*, vol. 78, pp. 252-259, 1991. Article (CrossRef Link).

[19] Y. Wang, S. Gao, X. Gao, "Common spatial pattern method for channel selection in motor imagery based brain-computer interface," *Engineering in Medicine and Biology Society*, pp. 5392-5393, Jan. 2005. Article (CrossRef Link).

[20] S. Koelstra, C. Muehl, M. Soleymani, et al., "DEAP: a database for emotion analysis; using physiological ignals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, 2011. Article (CrossRef Link).

[21] K. Stamps, Y. Hamam, "Towards inexpensive BCI control for wheelchair navigation in the enabled environment - a hardware survey," *Brain Informatics*, *Lecture Notes in Computer Science, Springer*, vol. 6334, pp. 336-345, 2010. Article (CrossRef Link).

[22] S. Koelstra, C. Muehl, M. Soleymani, et al., "DEAP: a database for emotion analysis; using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, 2011. Article (CrossRef Link).

[23] H. Kim, C. Seon, J. Seo, "Review of Korean speech act classification: machine learning methods," *Journal of Computing Science and Engineering*, vol. 5, No. 4, pp. 288-293, 2011. Article (CrossRef Link).

[24] C. Metz, "Basic principles of ROC analysis," in *Proc. of Seminars in Nuclear Medicine,* vol. 8, no. 4, pp. 283–298, 1978. Article (CrossRef Link).

[25] K. Trohidis, G. Tsoumakas, G. Kalliris, et al., "Multi-label classification of music by emotion," *European Association for Signal Processing Journal on Audio, Speech, and Music Processing*, vol. 1, no. 4, 2011. Article (CrossRef Link).

[26] N. H. Frijda, "The emotions," *Cambridge University Press*, p. 207, 1986.

[27] G. Chanel, J. Kronegg, D. Grandjean, et al., "Emotion assessment: arousal evaluation using EEG's and peripheral physiological signals," *Multimedia Content Representation, Classification and Security, Lecture Notes in Computer Science, Springer*, vol. 4105, pp. 530-537, Sep. 2006. Article (CrossRef Link).

[28] R. Panda, R. Paiva, "Music emotion classification: analysis of a classifier ensemble approach," in *Proc. of International Workshop on Music and Machine Learning in conjunction with the International Conference on Machine Learning, Edinburgh*, Scotland, Jun. 2012. Article (CrossRef Link).

[29] "Emotiv software development kit - user manual for release 1.0.0.4," *Emotiv*, 2011.

[30] R. Gilmore, "American electroencephalographic society guidelines in electroencephalography, evoked potentials, and polysomnography," *Journal of Clinical Neurophysioogyl*, vol. 11, no. 1, pp. 1-147, 1994. Article (CrossRef Link).

[31] M. Shaker, "EEG waves classifier using wavelet transform and Fourier transform," *International Journal of Medical, Health, Biomedical and Pharmaceutical Engineering*, vol. 1, no. 3, pp. 163-168, 2007. Article (CrossRef Link).

[32] N. Jatupaiboon, S. Pan-ngum, P. Israsena, "Real-time EEG-based happiness detection system," *The Scientific World Journal*, vol. 2013. Article (CrossRef Link).

[33] P. Lang, M. Greenwald, M. Bradely, et al., "Looking at pictures - affective, facial, visceral, and behavioral reactions," *Psychophysiology*, vol. 30, no. 3, pp. 261–273, 1993. Article (CrossRef Link).

[34] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biological Psychology*, vol. 84, no. 3, pp. 394–421, 2010. Article (CrossRef Link).

[35] G. Nam, B. Kang, K. Park, "Robustness of face recognition to variations of illumination on mobile devices based on SVM," *Korea Society for Internet Information Trans. on Internet and Information Systems*, vol.4, no.1, pp. 25-44, 2010. Article (CrossRef Link).

[36] Q. Zhang, M. Lee, "Analysis of positive and negative emotions in natural scene using brain activity and GIST," *International Journal of Neurocomputing*, vol. 72, no. 4-6, pp. 1302-1306, 2009. Article (CrossRef Link).

[37] Y. Lin, C. Wang, T. Jung, et al., "EEG-based emotion recognition in music listening," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1798-1806, 2010. Article (CrossRef Link).

[38] C. Song, H. Park, C. Yang, et al., "Implementation of a practical query-by-singing/humming (QbSH) system and its commercial applications," *IEEE Transactions on Consumer Electronics*, vol. 59, no. 2, pp. 407-414, 2013. Article (CrossRef Link).

[39] R. Fan, P. Chen, C. Lin, "Working set selection using second order information for training support vector machines," *Journal of Machine Learning Research,* vol. 6, pp.1889-1918, 2005. Article (CrossRef Link).

Saim Shin received her B.E. degree in Computer Science from Sookmyung Women's University *in 2000 a*nd M.E. degrees in Computer Science from KAIST (Kore*a Advance*d Institute of Science and Technology) *in 2002. She is currently pursuing a* Ph.D. degree in Computer Science and Engineering at Sogang University. Since 2013, she has worked as a senior research engineer at the Next Generation Sound Supporting Center of Korea Electronics Technology Institute (KETI), Seoul, Korea. Her research interests include speech and audio signal processing, mood analysis for contents recommendation and retrieval.

Sei-Jin Jang received the B.S. and M.S. degrees in electronics engineering from Kyungpook National University, Korea, in 1995 and 1997, respectively. From 1997 to 2002, he has been working as a research engineer in Digital TV Research Center of Daewoo Electronics Co., Korea. Since 2002, he has been working at Korea Electronics Technology Institute (KETI) and he is presently in charge of the Next Generation Sound Supporting Center as a Director. His research interests include the areas of multimedia communication and multichannel audio system.

Donghyun Lee received his B.E. degree in Computer Science and Engineering from Sogang University in 2013. He is currently pursuing a Ph.D. degree in Computer Science and Engineering at Sogang University. His research interests include speech recognition and spoken multimedia content search.

Unsang Park received the B.S. and M.S. degrees from the Department of Materials Engineering, Hanyang University, Seoul, Korea, in 1998 and 2000, respectively. He received the M.S. and Ph.D. degrees from the Department of Computer Science and Engineering, Michigan State University, MI, USA in 2004 and 2009, respectively. He has been an assistant professor in the Department of Computer Science and Engineering at Sogang University since 2012. His research interests include pattern recognition, image processing, computer vision, and machine learning.

Ji-Hwan Kim received the B.E. and M.E. degrees in Computer Science from KAIST (Korea Advanced Institute of Science and Technology) in 1996 and 1998 respectively and Ph.D. degree in Engineering from the University of Cambridge in 2001. From 2001 to 2007, he was a chief research engineer and a senior research engineer in LG Electronics Institute of Technology, where he was engaged in development of speech recognizers for mobile devices. In 2004, he was a visiting scientist in MIT Media Lab. Since 2007, he has been a faculty member in the Department of Computer Science and Engineering, Sogang University. Currently, he is an associate professor. His research interests include spoken multimedia content search, speech recognition for embedded systems and dialogue understanding.