AFFECTIVE JUKEBOX: A CONFIRMATORY STUDY OF EEG EMOTIONAL CORRELATES IN RESPONSE TO MUSICAL STIMULI

Joel Eaton, Duncan Williams, Eduardo Miranda
Interdisciplinary Centre for Computer Music research, Plymouth University, Devon, PL4 8AA, UK
joel.eaton@postgrad.plymouth.ac.uk, duncan.williams@plymouth.ac.uk

ABSTRACT

This paper presents a proof-of-concept pilot study investigating whether 2-dimensional arousal-valence correlates determined from electroencephalogram (EEG) readings can be used to select music based on the affective state of a user. Self-reported emotional states are used to evaluate a system for estimating arousal and valance from EEG by means of music selection from a real-time jukebox, with stimuli that have strong emotional connotations determined by a perceptual scaling analysis.

Statistical analysis of participant responses suggests that this approach can provide a feasible platform for further experimentation in future work. This could include using affective correlations to EEG measurements in order to control real-time systems for musical applications such as arrangement, re-composition, re-mixing, and generative composition via a neurofeedback mechanism which responds to listener affective states.

1. INTRODUCTION

In this paper we propose a neurofeedback system that performs real-time control over music selection, based on an individual's affective state. Akin to the control one has over song selection with a traditional music jukebox we have designed and built a system that selects musical clips based on a user's mood, as measured via electrical information in the brain. Clips with suggested emotional qualities based on crowd-sourced metadata were used in the study after being subjected to a perceptual scaling experiment in order to confirm their suggested emotional descriptors. Each clip is played back to a user over loudspeakers, generating a constant playback of musical excerpts in response to the affective state measured during the previous excerpt. Thus, a real-time prototype *affective jukebox* is realised. An experiment was conducted in order to validate the electroencephalogram (EEG) readings of the affective state reflected in, and interpreted via, the corresponding musical clips.

In recent years the field of brain-computer interfacing (BCI) has developed to the point of offering practical control mechanisms that can be used in music making activities. We aim to exploit this further by integrating affective control of BCI systems with novel creative applications. If on-going work in this field proves successful, one exciting prospect lies in the possibility of inducting specific affective states through music, or further, plotting affective trajectories using responsive musical stimuli. This could have potential in designing shared affective experiences in collaborative settings with multiple individuals, in real-time 'affectively-driven' algorithmic composition systems, or real-time arrangement with EEG as a 'hands free' control device. This offers advantages to non-musically trained end-users and to, for example, those with severe disabilities.

In recent years computer music systems have been adapted to respond to brain-control, for example in generating real-time notation [1], controlling performance systems [2] and working towards achieving therapeutic goals for users with physical disabilities [3]. Whereas these examples utilise explicit cognitive control over brainwaves, and subsequent musical parameters, this paper presents a passive approach to control that is determined by unconscious recognition of emotion. It remains to be seen whether mood is actively controllable in a neurofeedback system of this kind, however initial results suggest that a hybrid method of active/passive control may be achievable using this method. For example, affective states might change over time in response to the *desired* emotional content from listening to a style of music that reflects a previous affective quality. By choosing to feel more positive in *expectance* of more positive music, the system shows early signs of a measure of active control.

Studies where passive control, such as emotion recognition, is utilised can be difficult to quantify if success is measured against intention. This challenge is also increased as responses to music can often be unpredictable, depending on a range of factors such as cultural and social interpretations, personal taste, prior experience, memory and so forth. These factors highlight a need for experimentation with real-time systems for rapid feedback against a set of carefully selected stimuli.

Affective correlations to EEG have been suggested when mapping affective states to musical parameters [4]. In other studies listeners have been asked to self-report emotions that are compared against such EEG readings [5]. Our approach combines the two to see whether self-report confirms the suggested EEG affective correlation. Stimuli were selected from a pool with suggested emotional characteristics established by crowd-sourcing and corroborated by a pre-experiment perceptual scaling exercise. In simple terms the goal of the affective jukebox is to select music based on an individual's affective response as measured via EEG.

1.1 Affective model

Music psychology typically documents three types of emotional responses to music: mood, affect, and emotion [6]. Russell's circumplex model of affect [7] provides a way of parameterising emotional responses to musical stimuli in two dimensions: valence (positivity) and arousal (energy or activation), as shown in Figure 1. This model maps neatly on to Hevner's adjective cycle [8] and is corroborated by other studies of music and emotion in 2-dimensions [9]. Interested readers can find more exhaustive reviews on the link between music and emotion in [10] and the recent special issue in *Musciae Scientiae* [11]. Other models of music and emotion have been used but the 2-dimensional model was implemented in this work as it has been well documented in respect to music and to neurophysical measurement by means of EEG [12, 13, 4].

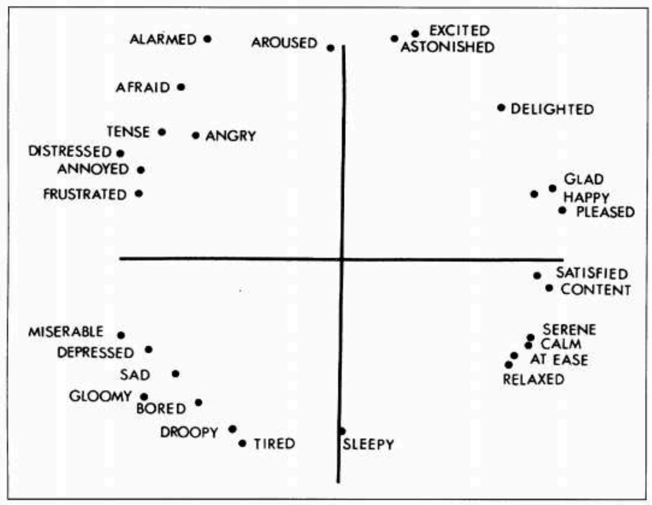

Figure 1. Circumplex model of affect, from Russell, p1168. Adjectives have been scaled in two dimensions, with valence on the horizontal axis and arousal on the vertical axis.

This system divides the circumplex model into quadrants which are then indexed via a Cartesian co-ordinate such that 12 discrete co-ordinates can be referenced corresponding to individual affective states, as shown in Figure 2. 2 axis give four quadrants, each of which is sub-divided by 3 to give 12 approximate affective states. 12 adjectives were selected from the circumplex model such that 'basic' emotions (*sad*, *calm*, *angry*, and *happy*), which have been well-documented in almost all music and emotion studies, are represented in the centre of the quadrants, with adjectives for lower and higher arousal levels spaced as evenly as possible from these adjectives in each quadrant. In this manner, a co-ordinate of (v1, a1) would refer to *tired*. Two adjectives were deliberately avoided in the selection process: *Sleepy* and *aroused*, as they were both placed near to the centre of the circumplex model of affect (shown in Figure 1) and might therefore be considered somewhat ambiguous as to their valence. However, there is no reason why additional adjectives could not be incorporated in the future to such a model, providing they are adopted from an appropriately scaled dimensional model.

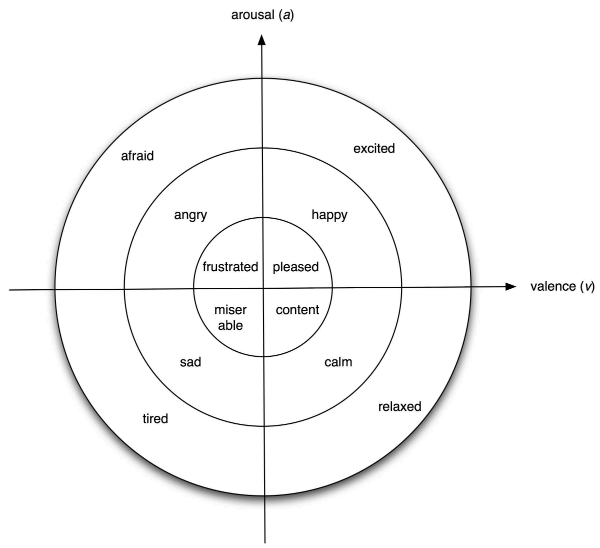

Figure 2. Quadrants with 12 discrete affective adjectives from the circumplex model (afraid, angry, frustrated, excited, happy, pleased, tired, sad, miserable, relaxed, calm, and content).

The main emotions (*sad*, *calm*, *angry*, and *happy*) can all be seen in the second level (v1, a2), (v2, a2), (v1, a5), and (v2, a5) respectively. An emotional trajectory moving from *pleased*, via *happy*, to *excited*, can be represented by a vector which gradually increases in arousal whilst maintaining positive valence: (v2, a4), (v2, a5), (v2, a6). Values for these co-ordinates can then be assessed from listeners and used to select music in real-time in response to, and by means of meta-information corresponding to, the selected emotional adjective.

2. SYSTEM DESIGN

EEG is a common choice for measuring electrical brain activity due to its non-invasive nature and relatively affordable and customisable hardware. However, interpreting meaningful information within EEG is a challenging task. Problems with noise inherent in complex signals that overlap into frequency or time-based ranges are commonly faced in BCI research. The quality of hardware components can make a significant difference in improving the signal-to-noise ratio and the ultimate success of mapping brainwaves to task related functions. The system presented and evaluated in this paper was designed using the g.tec Sahara dry electrode and pre-amplifier system with the g.MOBIlab+ amplifier. EEG data was passed into Matlab using the g.tec Simulink API for processing, feature extraction and classification and the resultant data was then passed to MAX/MSP via OSC for music clip selection. The raw EEG was pre-processed using a low-pass filter to accommodate the full range of frequencies measured and a notch filter with a centre frequency of 50Hz to reduce mains interference.

To reduce interfering noise from blinking, muscle activity or movement related artefacts Tenke and Kaysers method of segmenting incoming EEG into epochs of samples (50% overlap; Hanning window) and rejecting those that are clipped above a threshold of +100 μv [14] was adopted. EEG data was passed through butterworth bandpass filters and the spectral power of alpha (8-13Hz) and beta (13 – 30Hz) frequency bands was isolated. Mean values of spectral power are normalised across the last 10 seconds of 30 windows (which equals the length of each clip of music) to gauge response to the previously selected window of music. This method is useful to counter the known effect of diminishing arousal over time as subjects familiarise themselves with the stimuli and the environment [15].

To ascertain levels of valance and arousal electrodes were placed on the front of the scalp, over the prefrontal cortex an area that plays a significant role in emotion handling. Electrodes are positioned across points F3 and F4 (using the international 10–20 system). A subject's arousal can be derived from the ratio between alpha and beta activity. Strong alpha activity is known to indicate a relaxed state of mind, and this combined with increased activity in the beta band can indicate arousal; alertness in mental activity [16]. The balance of activation levels across the left and right hemispheres indicates a difference between a motivated approach or a more negative, withdrawal type of mental state, which is directly related to valence [4].

Real-time classification of emotional EEG characteristics is either crudely conducted using threshold values derived against stimuli presented to users or more accurately through training a classification model with techniques such Linear Discriminant Analysis (LDA) and Support Vector Machines (SVM) [4, 5]. Currently this system adopts the former approach, similar to that of Lui et al [17], and we acknowledge that although results were corroborated offline using LDA, an online classifier would yield much greater real-time accuracy.

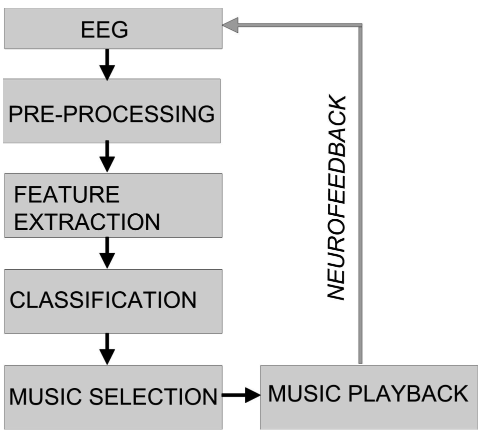

Figure 3. Diagram of the BCI neurofeedback system.

3. EVALUATION

In order to determine whether the EEG affective model was accurate, a perceptual experiment was undertaken whereby listeners were invited to self-report their emotional response to music selected autonomically in direct response to the EEG correlations.

Musical mood classification is a growing field in the realm of musical information retrieval, with various possibilities for stimulus selection including systems that utilise machine learning, crowd-sourcing, and acoustic analysis [18]-[20]. In this case, the stimulus set was selected from music which had already been tagged with emotional descriptors by crowd-sourcing in the *Stereomood* project, an "emotional on-line radio that provides music that best suits users' mood and activities" [21]. The stimuli included material from a range of genres, with a variety of tempos and instrumentation. Material with tags that correspond to the affective adjectives shown in Figure 2 was edited to 30-second clips and subjected to loudness equalisation in order to create the stimulus set shown in Table 1. There was, however, potential for some bias in the crowd-sourced metadata of the musical stimuli. For example, a search for *afraid* yielded *Non, je ne regrette rien* as performed by Edith Piaf. Knowing the lyrical content and delivery of this song, it seemed reasonable to assume that some of the crowd involved in tagging this stimulus did so as *afraid* as they felt that this song might give them the opposite i.e. *courage*. Four sources for each stimulus were initially short-listed, as a mechanism for elimination of erroneously tagged material was required. Therefore, these short-listed stimuli were evaluated in a short perceptual scaling experiment where listeners were asked to confirm how much they agreed with each of the crowd-sourced tags. Only stimuli with the largest amount of universal agreement to their documented correlation were progressed (>66% agreement with a standard deviation <0.20).

Stimulus number / Cartesian co-ordinate	Corresponding affective adjective	Musical stimulus (Title, performer)
1. (a1, v1)	*Tired*	*Dissociation EP*, Gelatinous
2. (a2, v1)	*Sad*	*One day I'll fly away*, Keith Jarret & Charlie Haden
3. (a3,v1)	*Miserable*	*Fade into you*, Chelsea Burgin
4. (a4, v1)	*Frustrated*	*What Kind of Girl*, Kid Moxie
5. (a5, v1)	*Angry*	*Sneak Chamber*, Tsutchie & Force of Nature
6. (a6, v1)	*Afraid*	*Perfect Drug*, Nine Inch Nails

7. (a1, v2)	*Relaxed*	*I'll take the road*, Dave Reachill
8. (a2, v2)	*Calm*	*Jung Greezy*, Snake Oil
9. (a3, v2)	*Content*	*Get Lucky*, Daft Punk
10. (a4, v2)	*Pleased*	*All around*, Tahiti 80
11. (a5, v2)	*Happy*	*Theo*, Apples
12. (a6, v2)	*Excited*	*Tropp'Cazz'Pa'Capa*, Smania Uagliuns

Table 1. Stimulus set used in evaluation experiment. Cartesian co-ordinates for arousal and valence are determined by EEG analysis. The corresponding affective adjective is then used to select a musical stimulus.

3.1 Listening panel and experimental procedure

Six listeners participated in the experiment. Each participant had some experience of critical listening. All participants were male, aged between 22-35. The experiment was conducted via a Max/MSP graphical user interface. Listening tests were conducted via full range loudspeakers in a quiet room with dry acoustics. Participants were allowed to adjust volume levels according to their own preference during a familiarisation exercise. Instead of integrating an external library of stimuli, such as the Affective Digitized Sounds database (IADS) of audio stimuli for emotion induction [22] the affective jukebox system was calibrated using excerpts from the full stimulus set in order to ensure that the EEG recordings were directly relevant to the stimulus set.

3.1.1 Familiarisation and calibration stage

The familiarisation exercise allowed listeners to hear the following stimuli in order to calibrate the EEG response: tired (v1, a1), relaxed (v2, a1), afraid (v1, a6), and excited (v2, a6), meaning that listeners had experienced one stimulus from each quadrant at the maximal and minimal arousal levels before undertaking the experiment. AV levels were recorded to determine the maximum and minimum levels on a participant-by-participant basis. These values were used to normalise the individual's responses in the main experiment.

3.1.2 Main experiment

After an initial measure of a user's resting affective state to select the first musical clip, the self-report/listening task was repeated over a series of trials, as illustrated in figure 4. EEG is collected from the last 10 seconds of each listening session (one complete stimulus) and the corresponding affective co-ordinate is used to select the next clip.

Once playback of the selected stimulus has finished, listeners are asked to scale their agreement with the intended emotion:

"Thinking about the music you just heard would you agree that it reflected the way you felt whilst the previous piece of music was playing (or in the first test, before any music was playing at) all?"

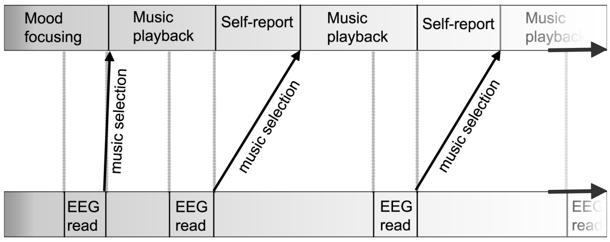

Figure 4. Design of the experimental paradigm. A staggered method of selection based playback and self-reporting was used over trials for each participant.

In response, listeners were presented with a slider using a hidden 100-point scale with end points marked "totally agree" to "totally disagree". It is acknowledged that in repeatedly answering this question users were required to recall previous responses to music in every instance, which is by no means an easy feat, but is necessary in order to reflect the paradigm of a jukebox playback system. Each experiment was conducted twice, the first in order to acquaint the participant with the experiment paradigm (see Figure 4.) and user interface. Data was only recorded and analysed from the second run of each participant, in order to ensure that participants understood the task.

4. RESULTS

Listener responses were analysed in the SPSS package. Mean agreement across all participants was 0.85 with a standard deviation of 0.14 and a standard error of mean of 0.07, as shown in Table 2. Although only a limited number of listeners took part in the evaluation, mean agreement was relatively high suggesting that there was a good degree of corroboration between the EEG measurement and the mood meta-tagging that was used to select the stimuli. However, the overall standard deviation (σ) was also quite high. The low sample size implies that an improvement in the margin of error could be achieved by using a larger number of participants for such an evaluation in future – to improve the σ to a confidence level of > 92% (α=<0.07) would mean at least halving the margin of error, requiring quadruple the number of participants (to achieve a confidence level of > 95% would require ~ 8 times the number of participants). The relatively low standard error of mean suggests that this hypothesis might be borne out by a larger scale evaluation, but such testing was beyond the scope of this pilot study.

Mean	Std. Deviation (σ)	Std. Error of Mean	Variance
0.85	0.14	0.07	2.14

Table 2. Mean agreement, standard deviation, standard error of mean, and variance across all participants (shown to 2 decimal places)

The mean agreement and standard deviation across each individual participant, as shown in Table 3, was then examined. Although mean agreement was good, the individual participants' standard deviation remained high with α=>0.05 in all participants, and particularly high standard deviation in two participants (participant 2 and 5).

Participant number	Mean	Std. Deviation (σ)
1	0.90	0.06
2	0.79	0.25
3	0.88	0.11
4	0.77	0.14
5	0.90	0.12
6	0.85	0.13

Table 3. Mean agreement and standard deviation from each individual participant (shown to 2 decimal places).

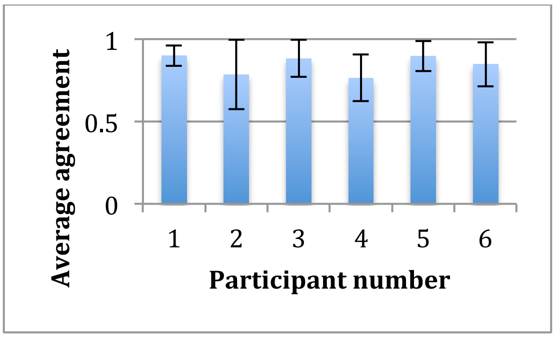

Figure 5. Mean agreement across individual participants with standard deviation shown in error bars. Note that the participants with the highest mean agreement also exhibit the lowest standard deviation in their responses.

Three participants (participants 1, 3, and 5) had additional training in EEG use and showed both a higher mean agreement and lower standard deviation in their responses, as shown in Figure 3, than the three participants who had not previously been trained in the system (participant number 2 and 4). However, there was only a relatively small variation in standard deviation between participants 3-6, regardless of previous training. This suggests that a robust evaluation of a training system might be useful in further work to determine whether any significant improvement in mean agreement and can be generated in previously untrained participants.

5. DISCUSSION

5.1 Performance and limitations

It is important to acknowledge that this was a pilot study using only a limited panel of participants of similar age and gender. The small number of musical stimuli was specified in order to reduce unmanageable variations in listener responses, however an improved system could increase the number of stimuli tailored to the individual; and thereby increase the performance of the system.

Moreover the tests highlighted that individual musical preferences, even within this small subset of possible users, caused a wide variation of responses. However, we hypothesise that within these caveats the system should be scalable to larger numbers of users.

One possible method to tackle the problem of individual musical preferences would be to carry out a pre-screening exercise (for example the STOMP system [23]) to improve the affective resolution by reducing individual bias to specific genres. Similarly a wider pool of musical stimuli could be used to address this.

5.2 Observations

Whilst conducting the experimental evaluation, a particular neurofeedback loop was observed. Listeners tended to select the same clip of music repeatedly once they had reached a settled affective state. This suggests an intuitive affective state, which is perhaps to be expected. When a listener actively engages with music that reflects their mood, their mood is unlikely to change for a period of time as they enjoy the listening experience associated to their affective state. Over time it could be seen that valence decreased after repeated exposure to the same clip. It is easy to imagine becoming annoyed after listening to the same 30 seconds repeatedly. Anecdotally, users most readily understood the corresponding states of stimuli in the extremes of the arousal dimension, but this would require a more rigorous evaluation to confirm.

It was also noted that listeners with more training found it possible to *actively* select different music clips, a function which we hope to investigate further in future experiments. This may hold potential for future applications with active control, such as selecting music to improve mood or to aid relaxation.

6. CONCLUSIONS

This paper reports on the successful pilot implementation of a real-time system for passive music selection and playback based on affective states derived from EEG to determine co-ordinates for arousal and valence. The pilot suggests that it is possible for a jukebox style affective music playback system to be controlled via EEG. Listener self-report confirmed that there was a good deal of corroboration between the EEG co-ordinates and individual affective state whilst engaging with the music selections. Both the affective jukebox system and the experimental methodology are in a pilot stage but these early results are promising and suggest that this is an appropriate methodology to employ when conducting further experiments with a larger number of participants and more complicated EEG derived affective correlations. Possible applications using this approach might include a more advanced affective jukebox taking into account personal music preferences, affectively driven composition engines and performance systems.

Acknowledgments

The authors gratefully acknowledge the support of EPRSC grants EP/J003077/1 and a doctoral scholarship awarded by Plymouth University.

7. REFERENCES

[1] J. Eaton and E. Miranda, "Real-Time Notation Using Brainwave Control," in Proc. Sound and Music Computing Conference, Stockholm, Sweden, 2013.

[2] J. Eaton and E. Miranda, "New Approaches in Brain-Computer Music Interfacing: Mapping EEG for Real-Time Musical Control," in Proc. Music, Mind, and Invention Workshop, The College of New Jersey, USA, 2012.

[3] E. R. Miranda, W. L. Magee, J. J. Wilson, J. Eaton, and R. Palaniappan, "Brain-Computer Music Interfacing (BCMI): From Basic Research to the Real World of Special Needs," Music and Medicine, vol. 3, pp. 134-140, 2011.

[4] R. Ramirez and Z. Vamvakousis, Detecting Emotion from EEG Signals Using the Emotive Epoc Device. Brain Informatics Lecture Notes in Computer Science. Springer, pp 175-184, 2012.

[5] Yuan-Pin Lin, Chi-Hong Wang, Tzyy-Ping Jung, Tien-Lin Wu, Shyh-Kang Jeng, Jeng-Ren Duann, and Jyh-Horng Chen, "EEG-Based Emotion Recognition in Music Listening," IEEE Trans. Biomed. Eng., vol. 57, no. 7, pp. 1798–1806, Jul. 2010.

[6] J. A. Russell and L. F. Barrett, "Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant." Journal of Pers. Soc. Psychol., vol. 76, no. 5, p. 805, 1999.

[7] J. A. Russell, "A circumplex model of affect." Journal of Pers. Soc. Psychol., vol. 39, no. 6, p. 1161, 1980.

[8] K. Hevner, "Experimental studies of the elements of expression in music," Am. J. Psychol., vol. 48, no. 2, pp. 246–268, 1936.

[9] E. Schubert, "Measuring Emotion Continuously: Validity and Reliability of the Two-Dimensional Emotion-Space," Aust. J. Psychol., vol. 51, no. 3, pp. 154–165, Dec. 1999.

[10] K. R. Scherer, "Which Emotions Can be Induced by Music? What Are the Underlying Mechanisms? And How Can We Measure Them?," J. New Music Res., vol. 33, no. 3, pp. 239–251, Sep. 2004.

[11] A. Lamont and T. Eerola, "Music and emotion: Themes and development," Music. Sci., vol. 15, no. 2, pp. 139–145, Jul. 2011.

[12] G. Chanel, J. Kronegg, D. Grandjean, and T. Pun, "Emotion assessment: Arousal evaluation using EEG's and peripheral physiological signals," Multimed. Content Represent. Classif. Secur., pp. 530–537, 2006.

[13] L. A. Schmidt and L. J. Trainor, "Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions," Cogn. Emot., vol. 15, no. 4, pp. 487–500, 2001.

[14] C. E. Tenke and J. Kayser, "Reference-free quantification of EEG spectra: combining current source density (CSD) and frequency principal components analysis (fPCA)," Clin Neurophysiol, vol. 116, pp. 2826-46, Dec 2005.

[15] F. R. Dillman Carpentier and R. F. Potter, "Effects of music on physiological arousal: Explorations into tempo and genre," Media Psychology, vol. 10, pp. 339-363, 2007.

[16] A. Choppin. "EEG-based human interface for disabled individuals: Emotion expression with neural networks". Masters thesis, Tokyo Institute of Technology, Yokohama, Japan, 2000.

[17] Lui Y, Sourina O, Nguyen M.K, "Real-time EEG-based human emotion recognition and visualization", In Proc. 2010 International Conference on Cyberworlds (CW), pp 262-269, 2010

[18] X. Hu and J. S. Downie, "Improving mood classification in music digital libraries by combining lyrics and audio," in Proc. of the 10th annual joint conference on Digital libraries, pp. 159–168, 2010.

[19] S. Davies, P. Allen, M. Mann, and T. J. Cox, "Musical Moods: A Mass Participation Experiment for Affective Classification of Music," in ISMIR, pp. 741–746, 2011.

[20] S. Knobloch and D. Zillmann, "Mood management via the digital jukebox," J. Commun., vol. 52, no. 2, pp. 351–366, 2002.

[21] E. Cambria, A. Hussain, and C. Eckl, "Taking refuge in your personal sentic corner," in Proc. Workshop on Sentiment Analysis where AI meets Psychology, pp. 35–43, 2011,

[22] M. Bradley and P. Lang, "The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. Technical report B-3," University of Florida, Gainesville, Fl., 2007.

[23] Rentfrow, P. J., & Gosling, S. D. "The do re mi's of everyday life: The structure and personality correlates of music preferences". Journal of Personality and Social Psychology, 84(6), 1236-1256. 2003.