

Music Mood Classification Model Based on Arousal-Valence Values

JungHyun Kim, Seungjae Lee, SungMin Kim, WonYoung Yoo

Content Research Division, ETRI, South Korea

{bonobono, seungjee, morethannow, zero2}@etri.re.kr

Abstract— This paper presents a music mood classification model based on arousal-valence (AV) values for music recommendation system. We analyse the collected music mood tags and AV values from 10 subjects and classify music mood on the AV plane into 8 regions by using k-means clustering algorithm. For each region, we propose representative mood tags by using statistics. We find that some regions can be identified by representative mood tags like previous models but some mood tags are overlapped in almost regions. The proposed model can express a region with several representative moods and a mood distributed uniformly in almost regions.

Keywords—mood classification, music emotion, arousal-valence

I. INTRODUCTION

Various music recommendation services generating music playlists automatically and searching desired music easily have been popular. Last.fm [1] collects playback logs of users by scrobbling and generates recommendation playlists by analysing playback logs and user's information with similar tastes. Pandora [2], the internet radio service recommends similar music using more than four hundred audio features, "music genome" extracted by experts.

A Genre, an age, an artist, playback history, and search keywords have been used for classifying and searching music. Music emotion or mood is also used for music classification and recommendation, some study [3],[4] had been proposed. Most of previous works focused on feature extraction and classification method, a few works studied music emotion and mood model for music recommendation system. To improve performance, we need well-defined music emotion or mood model.

In this paper, we propose music emotion or mood model based on AV values. By using k-means algorithm, AV planes are classified into 8 regions and each region is analysed. We present representative mood tags for each region and statistics on mood tags.

This paper organized as follows. Section 2 reviews existing music mood models and explains experimental setup and results. Section 3 presents proposed model and compare with previous models. Section 4 concludes this paper.

II. MUSIC MOOD CLASSIFICATION

Section 2 reviews existing music mood classification models and describes experimental procedure and presents experimental results.

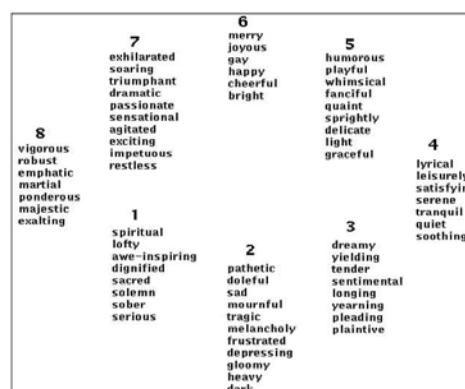


Figure 1. Hevner's adjective circle

A. Music Emotion Model

Generally, emotion model have been studied by two approaches: categorical and dimensional. A categorical approach introduces distinct classes that can be basis for all other possible emotional variations. A dimensional approach classifies emotions along several axes, such as valence (pleasure), arousal (activity), and potency (dominance). The dimensional approach to emotional modelling is most commonly used in musical application. Such approaches include the followings.

Kate Hevner [5],[6] examined the affective value of six musical features such as mode, tempo, pitch, rhythm, harmony, and melody and studied how they relate to mood. The study categorised 67 adjectives into eight different emotional groups with similar emotions and compared musical characteristics in 8 affective states. Figure 1 shows emotional groups in a circle.

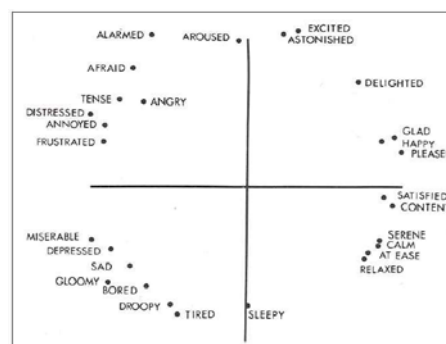


Figure 2. Russell's circumflex model of emotion

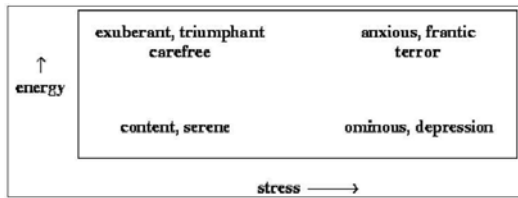


Figure 3. Thayer's energy-stress model of emotion

James Russell [7] presented circumflex model of emotion arranging 28 adjectives in a circle on two-dimensional bipolar space(arousal-valence) keeping away the opposite emotions. Figure 2 shows Russell's circumflex model of emotion.

Robert Thayer [8] proposed energy-stress model in two-dimensional spaces as shown in Figure 3, where contentment is defined as low energy/low stress, depression as low energy/high stress, exuberance as high energy/low stress, and anxious as high energy/high stress.

The dimensional approach such as Russell's and Thayer's model can represent characteristics of emotion simply and visualize emotion and relationship between emotions effectively. In this work, we use two-dimensional space (arousal-valence) and emotion is located by arousal and valence level.

B. Experiment Method

For 446 constant music clips, we gathered arousal/valence values and mood tags from 10 paid persons. In the experiments, the subject entered a mood tag and clicked a point on AV planes after listening to a song by the program shown in Figure 4.

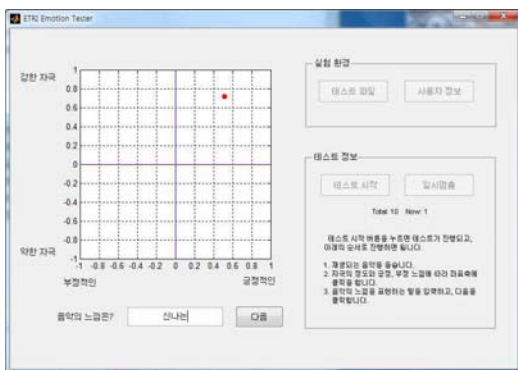


Figure 4. Program used in experiment

The music database is made up of 446 songs from various genres such as jazz, R&B, country, rock, hip-hop, metal, dance, newage, and electronica. From the music database, we extract emotionally constant 20-second segments by three experts. Table 1 shows genre distribution of music database.

TABLE 1. GENRE DISTRIBUTION OF MUSIC DATABASE

Genre	#	Genre	#	Genre	#
Jazz	61	Rock	48	Dance	41
R&B	53	Hip-Hop	32	Newage	69
Country	38	Metal	58	Electronica	46

C. Analysis of Relation Between Mood and AV values

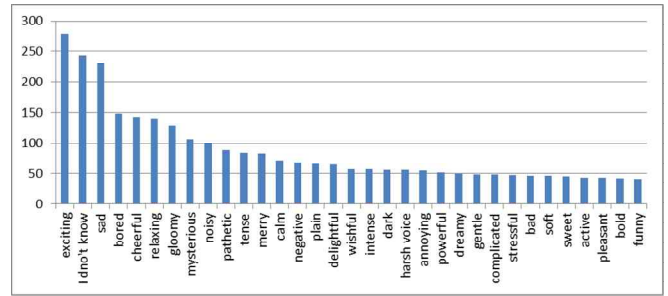


Figure 5. Histogram of mood tags more than the frequency 40

We collected 4460 mood tags and AV values through the experiment. Among the collected tags, we refine and clarify ambiguous and duplicated expressions and obtain 115 independent mood tags. To classify those mood tags, we use top 32 mood tags (excluding 'I don't know') in the frequency histogram as shown in Figure 5 which presents histogram of mood tags more than 40.

Based on similarities of meaning and usage, top 32 mood tags are classified into eight mood groups shown in Table 2. The representative mood is the highest frequency expression in each group.

TABLE 2. MOOD GROUP BASED ON SIMILAR MEANING AND FEELING

	Representatives	Mood tags
1	exciting('신나는' in Korean)	exciting, cheerful(경쾌한), merry(흥겨운), delightful(즐거운), wishful(희망적인), pleasant(기분좋은), funny(재미있는)
2	intense(강렬한)	intense, powerful(힘찬), active(활기찬), bold(도전적인)
3	tense(긴장되는)	tense, dark(어두운)
4	relaxing(편안한)	relaxing, calm(차분한), gentle(잔잔한), soft(부드러운), sweet(감미로운)
5	mysterious(신비한)	mysterious, dreamy(몽환적인)
6	sad(슬픈)	sad, gloomy(우울한), pathetic(애절한)
7	bored(지루한)	bored, plain(평범한)
8	noisy(시끄러운)	noisy, negative(부정적인), harsh-voice(듣기싫은), annoying(짜증나는), stressful(정신없는), uneasy(불편한), complicated(복잡한)

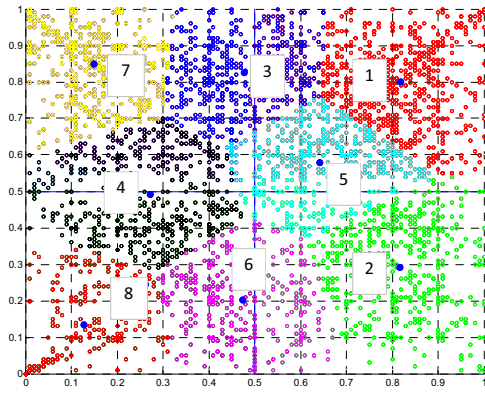


Figure 6. Classified AV values by k-means clustering

To know the relation between mood tags and AV values, we classify AV values through k-means clustering. We partition AV values into 8 clusters using classification of mood tags in Table 2. Figure 6 shows the classified regions on AV planes by k-means clustering.

Each region has the frequency histogram of mood tags shown in Figure 7. At this time, we discard the mood tags less than 10 times. The region 1, 2, 7 and 8 are placed in diagonal line on AV planes and can be expressed one or two representative moods. The region 3, 4, 5, and 6 are placed on arousal-valence axes. These regions have common mood tag, 'I don't know' with high frequency and relatively several representative moods. This shows that people feels vaguely in axes area.

Also, we find that such moods as 'mysterious/dreamy' and 'tense' in Table 2 are distributed uniformly in most of regions.

An existing model classifies AV planes and each classified region is mapped to one representative mood. It is difficult to express these moods by existing model. Therefore, we propose music mood classification model considering the analyzing results.

III. PROPOSED MODEL

To express a mood on AV planes, we use frequency histogram of the mood classes (in Table 2) for each region (in Figure 6) shown in Figure 8.

The region 1 consists of such moods as 'exciting', 'intense', 'peaceful', 'tense', and 'mysterious'. The portion of each mood in the region 1 is respectively 73, 16, 2, 3, and 6 percent and we know that representative mood is 'exciting'. The region 2 is comprised of 5 moods as, 'exciting', 'relaxing', 'mysterious', 'sad', and 'bored'. The portion of each mood in the region 2 is respectively 14, 72, 7, 4, and 2 percent and we know that representative mood is 'relaxing'.

Likewise, a region on AV planes can be expressed with probability distribution of moods. The region 1 and 2 can be expressed as followings:

$$R1=[P_{M1}(0.72), P_{M2}(0.16), P_{M4}(0.02), P_{M3}(0.03), P_{M5}(0.06)] \quad (1)$$

$$R2=[P_{M1}(0.14), P_{M4}(0.72), P_{M5}(0.07), P_{M6}(0.04), P_{M7}(0.02)] \quad (2)$$

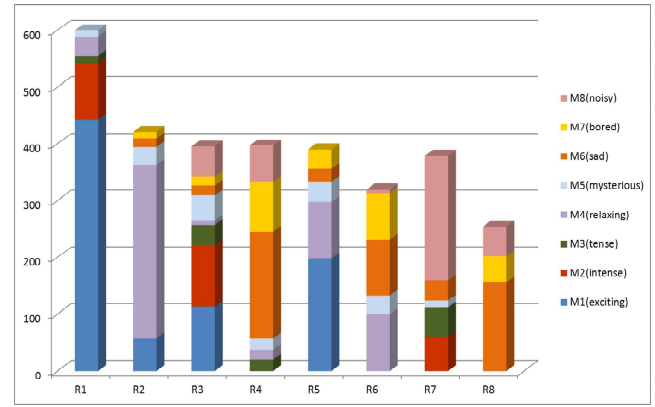


Figure 8. Frequency histogram of the mood in each region

In the same way, a mood can be presented using probability distribution of regions including the mood. The mood class 1 and 2 in Table 2 can be expressed as followings:

$$M1=[P_{R1}(0.55), P_{R2}(0.07), P_{R3}(0.14), P_{R5}(0.24)] \quad (3)$$

$$M2=[P_{R1}(0.38), P_{R3}(0.40), P_{R7}(0.22)] \quad (4)$$

The expression (5) and (6) shows respectively how to express a classified region on AV plane and a mood by probability distribution of mood and regions. R_n and M_n are defined as a classified region on AV planes and mood class, respectively.

$$R_n=[P(M1), P(M2), \dots, P(M_n)] \quad (5)$$

$$M_n=[P(R1), P(R2), \dots, P(R_n)] \quad (6)$$

IV. CONCLUSIONS

In this paper, we propose mood classification model for music recommendation system. We analyze the collected music mood tags and AV values from 20 subjects and classify music mood on the AV plane into 8 regions by using k-means clustering algorithm. We show that some regions can be identified by representative mood tags like previous models but some mood tags are overlapped in almost regions. The proposed model can express a region with several representative moods and a mood distributed uniformly in almost regions. In the future, we plan to add study of emotion classification based on similarity of mood tags and apply it to classify music.

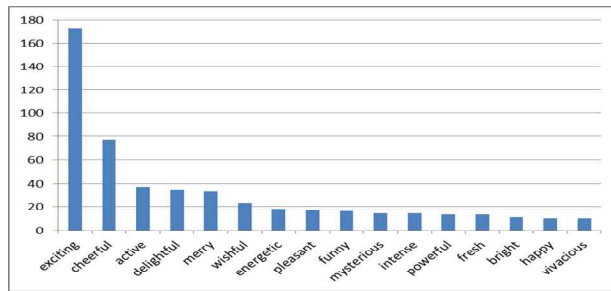
ACKNOWLEDGMENT

This work was supported by the IT R&D program of MSCT/KEIT. [10035229, Development of contents protection and distribution technology on DRM-free environment]

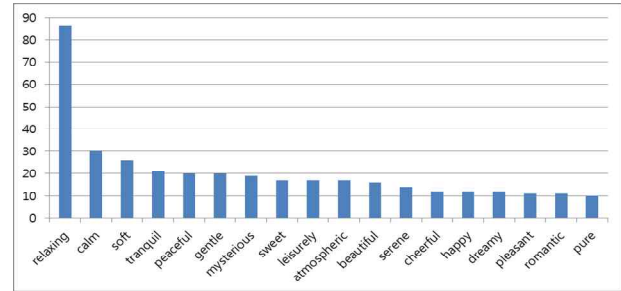
REFERENCES

- [1] <http://www.lasf.fm>
- [2] <http://www.pandora.com>
- [3] Maria M.Ruxanda, Bee Yong Chua, Alexandros Nanopoulos, Christian S.Jensen, Emotion-based Music Retrieval On Well-reduced Audio Feature Sapce, ICASSP Proceeding, 2009.

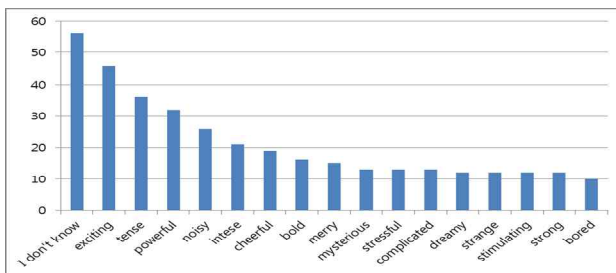
- [4] Laurier C., Herrera P., Audio music mood classification using support vector machine, ISMIR Proceeding, 2007.
- [5] K.Hevner, The affective character of the major and minor modes in music, The American Journal of Psychology, 47(1), 1935.
- [6] K.Hevner, The affective value of pitch and tempo in music, The American Journal of Psychology, 49(4), 1937.
- [7] J.A.Russell, A circumflex model of affect, Journal of Personality and Social Psychology, 37(3), 1980.
- [8] R.E. Thayer, The Biopsychology of Mood and Arousal, Oxford University Press, Oxford University, 1989.



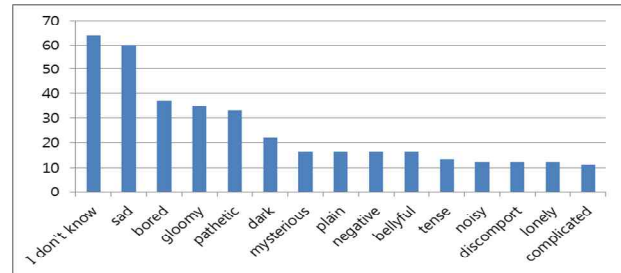
(a) Region 1



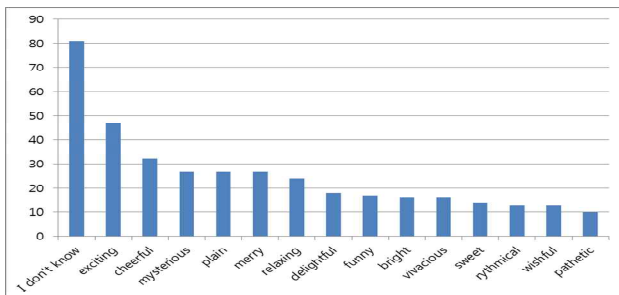
(b) Region 2



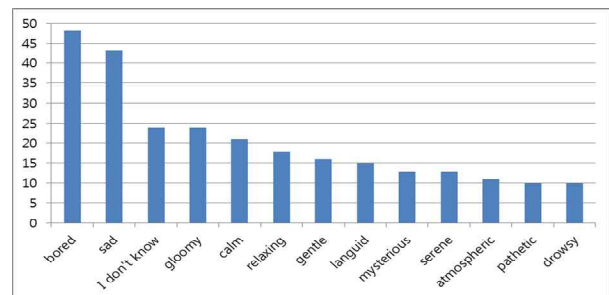
(c) Region 3



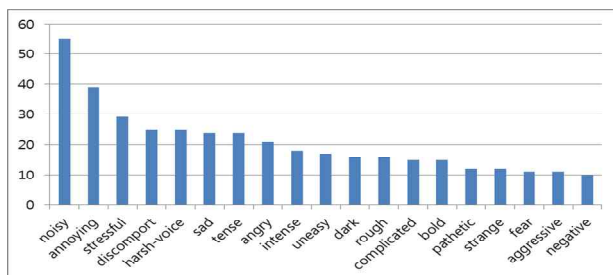
(d) Region 4



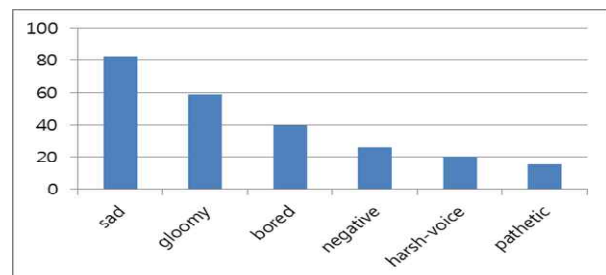
(e) Region 5



(f) Region 6



(g) Region 7



(h) Region 8

Figure 7. Frequency distribution of mood tags in each region