New dataset of emotional and color responses to music

Primož Godec[1], Matevž Pesek[1]

[1]*University of Ljubljana, Faculty for computer and information science, Ljubljana, Slovenia*
E-mail: p.godec9@gmail.com, matevz.pesek@fri.uni-lj.si

Abstract

We present new dataset containing perceived and induced emotions for 200 audio clips. Dataset also provide users association of color for each clip and other data that explain how users perceive emotions, what is users emotion state, what genre users prefers and some other data important for mood information retrieval. We collected more than 7000 responses with online survey for dataset of 200 audio clips. That mean that dataset provides average 37 responses per clip. We used new methodology for gathering user perception of emotions in 2D space named MoodGraph and MoodStripe used for gathering presence of emotions by user. Our dataset was used to train simple regression algorithm and compared with some other datasets.

1 Introduction

What kind of music you listen when you are happy? And what when you are sad? It is known that there is strong connection between emotions and music. Your current mood influences on your music choice. When you are happy your music choice would be different than that when you are sad or frustrated. There is also inverse connection. Music can affect on your current mood.

But can we good enough determine relationship between music and emotions? Many researchers have worked on this, but there is a lot of undiscovered. Music mood from year to year has become very important field in music information retrieval. There are many algorithm for estimate mood from audio music files. In [5] they uses regression for mood classification. Other use support vector machine [1] an some d other approaches. MIREX organises mood classification task from 2007. Also some other fields such as psychology, musicology, sociology works on relationship between music and emotions. One of the important goals of all these filed is to make better user experience of system such us music recommendation. That goal is also important to us, so we decided to make research of music mood. The beginning of our research was to gather new dataset on which we have building our research.

Several dataset was made in past and some of them were available online. Sound track for music and emotions provides single mean rating with label and values in three-dimensional model. Dataset contains values for 361 film music clips [3]. The Mood Swings Turk Dataset contains average 17 valence-arousal ratings for 240 audio clips [6]. Clips in this dataset are of popular music. Cal500 provides mood labels for 500 western popular songs [8]. They have around 3 annotations per song. MTV Music Dataset contains 5 bipolar valence-arousal ratings for 192 popular songs [7]. Songs were selected between songs MTV has played on his channel.

This collections of datasets we want to add our new dataset for 200 audio clips. For each clip we gathered emotions that song induced in user and emotions that are perceived in song. Together with this we also capture color that users things best describe song. With this dataset we also provide some other data, that can help us in music mood classification. That are user emotional state, users genre preferences and their perception of emotions.

But what is the quality of this captured data? We wanted to try this on the algorithms for music classification and compare this with data on other datasets. Hence we select algorithm that was presented in [5]. They used regression to estimate valence and arousal value for audio clip using MFCC and chroma. We implemented similar algorithm and test it on our dataset and the Mood Swings Turk dataset. Predictions on our dataset are better than predictions on dataset that author provided.

[se o svn ki ga bomo naredili]

[struktura paperja]

2 Online survey

As mentioned, we gathered data with online survey. Before our survey was made, decision about choosing mood labels had made. Some basic emotion labels exist [2], but there is not standard set music and mood research. Some author choose label sets intuitively, without any explanations [9]. Because we wanted to have good dataset, we decided to prepare preliminary survey.

This survey has two parts. In first we prepared some basic demographical questions and questions about user perception of mood. We test basic structure and questions type of main survey. But the most important was second part. There user has to describe his emotional state on scale from 1 to 7 for each of 48 emotional labels. In this part we also test response on continuous color wheel used to describe connection between mood and colors.

Depending on results of this questionnaire we selected 17 basic emotion labels which strongly correlate to three basic components that explain 64% of the variance to the dataset. Depending on user responses and results we also decided to restrict continuous color wheel on that with 49 colors.

2.1 The survey

We structured our survey in three sections. We captured user personal characteristic in first part. User was asked to answer on some demographical questions: age, area of living and native language. Some data about users music education, listening to the music and genre preference was captured. We did not want to be too long with this part, but we think that it is important to have that type of data, because this help us to understand users mood perception. For example we think that users who like metal, marked that that type of music is more positive, than other users.

Second part contains questions about user perceptions of mood and connection between color and mood. First we asked for user current mood state with three different tasks. User had place his current mood in valence arousal space. This is 2D plane to describe pleasantness and activeness of the mood. Then we wanted to know how user describes his current mood by selecting color. User also had to mark how strong one emotional state is expressed in him. We capture this with new element we described in [4] named MoodStripe (Figure 1). Second part has also two tasks that directly capture music mood perception. First users have to place 10 emotional labels in valence-arousal space described before. For this part we used self designed element named MoodGraph (Figure 2). Then they were asked to choose color that best match to labels used before. That part was important because it tells us user perception of mood labels, that was also used in third part of our survey.

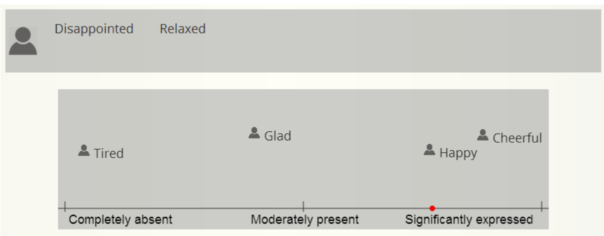

Figure 1: New element used in our survey named MoodStripe. With drag and drop technique user place emotion labels in plane depending on how is emotion expressed. When place on left this mean that emotion is totally absent, when right this mean that emotion is significantly expressed.

In last part of our survey user was asked to make two actions for each of 10 audio clips. Clips are 15 second long and was randomly selected from base of 200 clips. All selected audio clips are for most of users unknown, to avoid bias because of familiarity of song. We gathered music from four sources. Eighty songs was chosen from the free online music service Jamendo. We selected songs from as many genre as possible. Next 80 clips was from

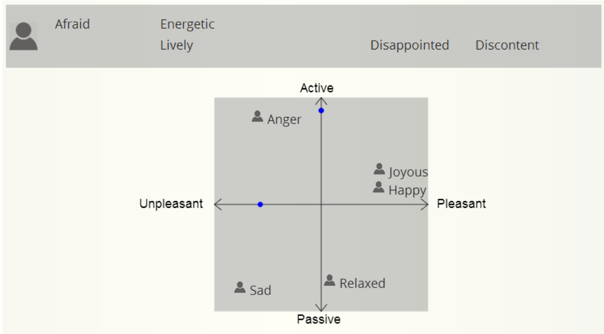

Figure 2: New element used in our survey named MoodGraph. With it we capture pleasantness on x axis and activeness on y axis.

film music dataset and is described here [3]. We also provide 20 clips from collection of slovenian folk music collection and 20 of the contemporary electro-acoustic music collection.

For each clip we asked user to select color, that he think best describes music clip. User was also asked select emotion labels that best describe clip and place them in valence-arousal space. We provide two group of labels. From first group user select emotions that was induced by song in music and from second emotions perceived in song.

References

[1] Asa Ben-Hur and Jason Weston. A user's guide to support vector machines. In *Data mining techniques for the life sciences*, pages 223–239. Springer, 2010.

[2] Tim Dalgleish and Michael J Power. *Handbook of cognition and emotion*. Wiley Online Library, 1999.

[3] Tuomas Eerola and Jonna K Vuoskoski. A comparison of the discrete and dimensional models of emotion in music. *Psychology of Music*, 2010.

[4] Matevz Pesek, Primoz Godec, Mojca Poredoš, Gregor Strle, Joze Guna, Emilija Stojmenova, Matevz Pogacnik, and Matija Marolt. Gathering a dataset of multi-modal mood-dependent perceptual responses to music. In *Proceedings of the 2nd Workshop Emotions and Personality in Personalized Services (EMPIRE 2014), Aalborg, Denmark*, pages 1613–0073, 2014.

[5] Erik M Schmidt and Youngmoo E Kim. Projection of acoustic features to continuous valence-arousal mood labels via regression. In *10th International Society for Music Information Retrieval Conference. ISMIR*, 2009.

[6] Erik M Schmidt and Youngmoo E Kim. Modeling musical emotion dynamics with conditional random fields. In *ISMIR*, pages 777–782, 2011.

[7] Björn Schuller, Clemens Hage, Dagmar Schuller, and Gerhard Rigoll. 'mister dj, cheer me up!': Musical

and textual features for automatic mood classification. *Journal of New Music Research*, 39(1):13–34, 2010.

[8] Douglas Turnbull, Luke Barrington, David Torres, and Gert Lanckriet. Semantic annotation and retrieval of music and sound effects. *Audio, Speech, and Language Processing, IEEE Transactions on*, 16(2):467–476, 2008.

[9] Bin Wu, Simon Wun, Chung Lee, and Andrew Horner. Spectral correlates in emotion labeling of sustained musical instrument tones. In *ISMIR*, pages 415–420, 2013.