Emotion Recognition Music Player

Tristan Perez
University of San Francisco
San Francisco, CA 94117, USA
Tperez2@dons.usfca.edu

Abstract
The paper presents an affective Android music player that gives song selections based on user emotions. By analyzing user facial expressions, the music player is play music appropriate for the mood. The paper discusses the implementation of the application and user studies and results.

Author Keywords
Emotions; Facial Expression; Detect; Mood; Music;

Introduction
The idea behind the project was influenced by my frequent use of my music player. I like to listen to music, whether that is walking to class or studying. It became a habit for me to skip through my playlists in order to play a song that really matches how I feel. At times, I would feel annoyed when I couldn't find the right song and would have to settle for a song closest to what I feel. I always felt that there should be a better way to find that certain song without having to press skip multiple times. This is the motivation behind the project, to make it easier to find and play music that matches emotion.

The project is an emotion detecting music player that uses the Affectiva and Spotify SDK programmed on Android Studio. The music player captures people's

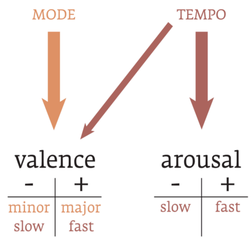

Figure 1: The double effect of mode and tempo on valence and arousal. Tempo has a major impact on arousal and a minor impact on valence.

Musical characteristics and emotional expression of music
(adapted from Bruner II, 1990, p. 100)

musical	Emotional Expression							
element	Serious	Sad	Fear	Serene	Humorous	Happy	Exiting	Majestic
Mode	Major	Minor	Minor	Major	Major	Major	Major	Minor
Tempo	Slow	Slow	Slow	Slow	Fast	Fast	Fast	Medium
Pitch	Low	Low	Low	Medium	High	High	Medium	Medium
Rhythm	Firm	Firm	Low	Flowing	Flowing	Uneven	Firm	Firm
Harmony	Cons	Diss	Diss	Cons	Cons	Cons	Diss	Diss
Loudness	Medium	Varied	Soft	Soft	Medium	Medium	Loud	Loud

Figure 2: The model shows keys characteristics in determining emotion expression of music

emotions while listening to music and gives feedback by playing music based on the person's emotions. The objective is to create an emotionally intelligent music player by differentiating between the emotions felt while listening to music. The emotion detector detects people's emotions and plays appropriate music to calm or boost their mood.

My contributions include:

- Capturing people's emotion while listening to music
- Giving real-time feedback of different songs that match people's emotions

Related Works

There have been many related works on music recommendation through emotion recognition, but none of them give real-time feedback through the use of a music player.

Machine Recognition of Music Emotion: A Review by Yi-Hsuan Yang and Homer H. Chen focuses on recognizing the emotion of a music through arousal and valence [1]. Affective Content Analysis of Music Video Clips by Ashkan Yazdani, Krista Kappeler, and Touradj Ebrahimi focus on analyzing and characterizing multimedia contents by the emotions it may elicit in viewers [3]. Both focus on recognizing and analyzing music emotions. My project will be different as it provides the analysis as feedback using the user's emotions.

Personalized Music Recommendation in a Mobile Environment by Claus Schabetsberger and Markus Schedl generates a list of track recommendations for a user based on music features extracted from the audio content directly on the user's device and on social features inferred from YouTube data, such as the type of music videos the user searches [2]. Collaborating with an Autonomous Agent to Generate Affective Music by Fabio Morreale and Antonella De Angeli explains the use of musical characteristics in order to make music easier to untrained users. In addition, their work explains the musical influence on the expressiveness of a composition [4]. MoodMusic: A Method for Cooperative, Generative Music Playlist Creation by created a method to generate music playlists that suit everyone's tastes and the mood of the group without managing playlists [5]. These works on music recommendations helps create the appropriate playlists used in the application.

Methodology

The facial recognition was detected using Affectiva's SDK. The data for the visualizations were taken from the Affectiva SDK and add to a simple moving average (SMA) of 100 data points for each emotion. For every face recognized gave values for different emotions such as joy, anger, sadness, and calm. The data was filtered to take out any emotions whose total values came out to zero, and as each new data point comes in, the oldest data point in the window is discarded. Values that were not zero were added to appropriate array of emotion. For each sliding window, the mean was calculated. If the window is not full, nothing was calculated. When the threshold of 40 was reached for the average of data, the next appropriate song is queued and the data points in the array is cleared.

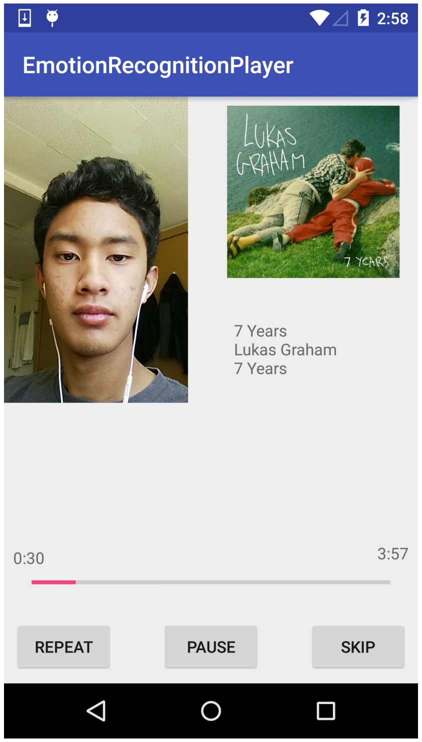

Figure 4: The application begins to detect your facial expressions and queues music based on the user's desired target emotion.

Each data point was written to a CSV file for data visualization and statistical analysis.

Playlists were constructed using Spotify API. With the audio features, I was able to construct different playlists consisting of 10 songs each. For the calm playlists, I measured for serene elements in songs. So, I added songs with low levels of danceability, energy, loudness, and valence. For the joyful playlists, I measured for happy elements in songs. So, I added high levels of danceability, energy, loudness, and valence. In addition to the Spotify API, I followed the Musical characteristics and emotional expression of music model (Figure 2) with emphasis on mode and tempo (Figure 1). I used the serene and happy music elements to create the playlist.

There were two questionnaires given before and after the experiment. First questionnaire asked the user to specify what types of music players he/she uses, how many hours they listen to music, and if they listen to their "Discover" playlist. Getting the number of hours spent listening to music gives a better understanding of their experience with music players. Asking about their "Discover" playlist use helps understand their impressions of it. The Second questionnaire asks for the user's and experience and inputs for improvements.

For the user study, there were two experiments with different conditions. The first experiment ran a music that tried to boost the user's mood. If the user was angry or sad, calm music would play in order to get the user to calm down first. If the user was calm or happy, the music player would play joyful music to boost or keep the user's mood up.

For the control experiment, the user was asked to choose the option of being in a calm or joyful mood. For the joyful mood, the music player played music that was meant to boost the user's emotion. If the user was angry, the music player would calm down the user first. Then, if the user was calm, joyful playlist B, which played 80s and 90s music, would queue. If the user was happy, then joyful playlist A, which played current popular music, would queue. The test was supposed to see if the music player was able to play appropriate music based on the mood. For the calm mood, the music player played music that was meant to calm the user down. If the user was calm, then the music player would keep playing from the current playlist that is queued (Figure 4). If the user was happy, angry, or sad, the music player would try calm him/her down by playing songs with lower energy and valence. Playlist queues are switched when facial expressions do not match with the target emotion. When facial expressions do match with the target emotion, the music player keeps queuing from the current playlist.

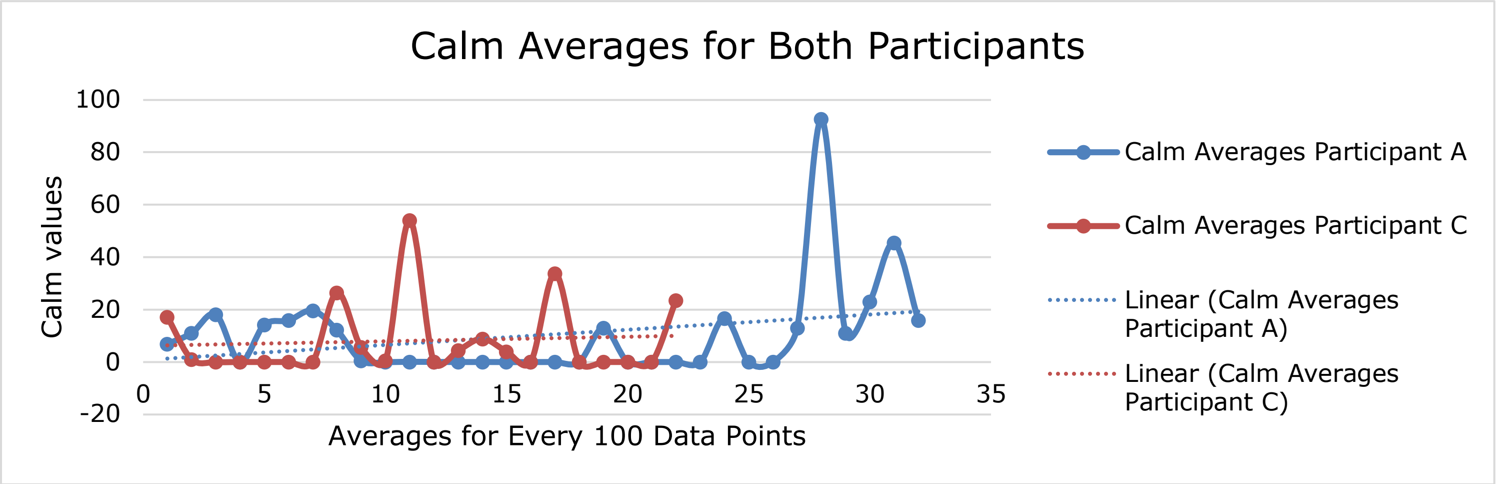

Figure 1: The graph shows similar linear progression of both participants' emotion. Both participants chose their target emotion as calm.

Results

I analyzed participants' emotions when they chose a specific mood. I implemented a t-test to figure out the similarities between the two means for the calm emotion. The test calculated a t-Stat value of 3.145 which is greater than both one-tail and two-tail (Figure 5). The test means that the results were not significantly different when user's chose the calm mood. The study was conducted on three participants, so the results can't be very accurate unless more studies are conducted. However, future user studies look promising based on the current results.

t-Test: Two-Sample Assuming Unequal Variances		
	Participant A	*Participant C*
Mean	10.18161758	7.754254112
Variance	900.9140535	656.4265406
Observations	3126	2136
Hypothesized Mean Difference	0	
df	5008	
t Stat	3.145488264	
P(T<=t) one-tail	0.000833906	
t Critical one-tail	1.645157951	
P(T<=t) two-tail	0.001667811	
t Critical two-tail	1.960437793	

Figure 5: T-test for calm mood

Based on the questionnaires given to the participants, participants rated their experience five or above on a scale from one to seven and also said that they would use the app again if given the chance. Participants indicated that the “Discover” playlists were terrible and most felt that the app played better music selections than the current music player that they used. In terms of how accurate the music matched their mood, participants rate the accuracy no greater than five on the same scale. In addition, the participants gave feedback for improvement. They felt that music should be queued faster, and that the layout of the music player should look better.

Discussion

One learning outcome was the positive feedback from using the app. The participant indicated that he/she had never used a music player that gave real-time feedback. Usually feedback playlists only give music selections based on what the person listens to for a whole week.

After creating the music player, I realize that there are many possible future works that could be done with it. Future works could be implementing a heart rate monitor with the music player to get a more accurate recognition of the user’s emotion. To further the apps functionality, the music player could be used as a therapeutic tool that could possible reduce the user’s stress whenever the he/she is angry or sad.

Conclusion

The paper presented an emotion recognition music player that gives real-time feedback by playing appropriate music based on the target mood. From the study, we can see that adding emotion inputs to music players increases its functionality and improve user experience.

Acknowledgements

I would like to thank all the volunteers for the user and pilot studies and to thank Professor Yuksel for help in developing the music player.

References

1. Bauer, Jared S., Alex Jansen, and Jesse Cirimele. MoodMusic: a method for cooperative, generative music playlist creation. In *Proceedings of the 24th annual ACM symposium adjunct on User interface software and technology*. ACM, (2011).
2. Morreale, Fabio, and Antonella De Angeli. "Collaborating with an Autonomous Agent to Generate Affective Music." *Computers in Entertainment (CIE)* 14.3 (2016).
3. Schabetsberger, Claus, and Markus Schedl. Personalized Music Recommendation in a Mobile Environment. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*. ACM, (2013).
4. Yang, Hsuan, and Homer H. Chen. Machine Recognition of Music Emotion: A Review. In *ACM Transactions on Intelligent Systems and Technology (TIST)* 3.3 (2012).