

Application of Music Emotion Recognition for Design of Audio Content in terms of Affective Gaming

GAD Ganepola¹ and LP Kalansooriya²

^{1,2}General Sir John Kotelawala Defence University, Sri Lanka

^{#1} 33-is-008@kdu.ac.lk, ^{#2} pradeepkalansooriya@kdu.ac.lk

Abstract

Music is a medium of communicating human emotion. In fact, a research area for studying on musicality features and variations of perceived emotions in accordance has been unfolded due to this proven fact. It is referred to as Music Emotion Recognition (MER). Various researches are being conducted on determining the relationships among music structural features and perceived emotions. This research in fact, will be focusing on a practical implementation based on MER. The paper interestingly, describes on deploying a basic prototype of applying MER when playing car racing video games. This is performed with the aim of increasing the user interactivity of game players and improving the state of current video game industry to "Affective Gaming" state. The emotions experienced during racing gameplay was analyzed through player annotations. Results depicted that players frequently experience various degrees of Arousal values when playing video games. Hence, Music structural features that contribute to Arousal factor of emotions were selected. Music features are to be extracted from Electronic/Rock music genre and a machine learning model is to be built to predict emotions that will be perceived. The emotions utilized to classify music are "Happy", "Sad", "Anger", "Tension" and stored in a database. This will be connected to the integrated software that will play music accordingly afterwards. The integrated software consists of the classified database, the game engine and a hardware component that captures EEG signals and analyze according to Arousal & Valence dimensions of emotions.

Keywords - MER, Video Games, Affective Gaming, Human Emotions.

I. INTRODUCTION

Music Emotion Recognition (MER) is an emerging research area combining the fields of Music, psychology & computational analysis. Numerous

researches had been conducted on various genre of music to distinguish music features' patterns relating with perceived emotions. However, research that utilizes MER into other application areas had not yet been popularized. Playlist generation is the most common appliance of using MER still. Considering this problem context, this research will focus on design of audio content in car racing video games based on MER to bring about "Affective Gaming" experience to players.

The term "affect" in psychology is defined as a feeling or an emotion expressed for some event (Laura et al.,2017). In the psychology perspective, video games consist of higher amounts of "affect" because video game environments are built to change according to the user interactivity. The interactions occur when the game environment

receives inputs from the player's side by executing actions via the game controller.

"Affective" gaming can be considered as detecting the real-time emotional state of a player during various stages in game play and then enhancing the user interactivity accordingly to emotional state. An instance of tracking the real state of emotions is by monitoring the physiological responses such as variations in EEG signals, pulse signals etc. during user interactivity within the game environment (Gilleade, 2005), (Kalansooriya, 2016). Examples of such games include arousal-driven appearance of non-player characters (NPCs) in *Left 4 Dead 2* the fearful combat skills of the opponent NPCs in *F.E.A.R.*, the avatars' emotion expression in the *Sims series*, the emotional responses of game characters in *Prom Week and Façade*, the emotion-driven narrative building system in *Storybricks* etc.

The basic principle of affective gaming lies within the communication between human to human, human to computer and human to human through computer. According to Sundstrom, design of an affective loops in video games makes possible all mentioned forms of communication. The affective

loop treats emotions as processes. The affective loop in gaming can be viewed as a closed loop with three sequential key phases.

- 1) the player expresses his emotions through the interaction with a game;
- 2) the game then detects the emotional reactions of the player, and interprets those reactions according to the context of the game;
- 3) based on that interpretation, the game makes adjustments that can be achieved via emotional modelling and expression of NPCs or via affect-driven content generation adapting the game to the player. This in turn affects the player (both her mind and body) making her respond through game actions and emotional reactions (step 1 again) (Sundstrom,2005).

The player expresses his emotions when she reacts to the game content. Game content includes mainly game mechanics, game environment, audio content etc. Game mechanics are the methods and rules specific for any game type that makes the game player interact with the game environment. Game environment consist of the visual settings such as graphics, background lighting levels, saturation levels etc. Audio content includes music clips played in the background.

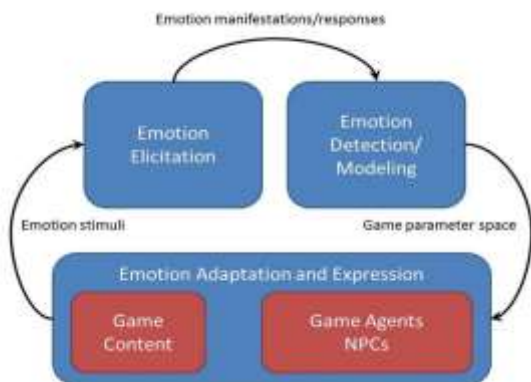


Figure 1- Affective Game Loop
Source - (Sundstrom,2005).

At present, game mechanics are maximumly being utilized for emotion elicitation. However, the impact of music for emotion awakening is not being weighted much by the game developers. Music Emotion Recognition (MER) is a widely discussed research topic in the field of psychology. MER is a cross disciplinary field of auditory perception, psychology and music theory. With technology

mixing in, MER is now a field consisting of signal processing and machine learning (Kim et al.,2010).

MER does have a potential for a high emotion elicitation but is still in its early stages and is yet to be developed and utilized in day-to-day activities. Reasons may include music being extremely subjective for humans. In fact, it is a timely and an exasperating task to identify to various emotions perceived. In the context of gaming industry, MER is minimally used because of the low budget allocated for audio content when compared with the budget allocated to that of graphics [8]. Nevertheless, composing music based on MER costs much more than composing graphic content. Also, various existing software that classify emotions in music are not customized for the gaming industry. This in turn, makes game developers for non-utilizing of MER in gaming environment.

The aim of this paper is illustrating a design of an audio-based software prototype for car racing video game with the application of MER. That way, the concept of MER can be popularized and let the general public experience the effectiveness of MER. In conclusion, the aim of this research is to transform MER into a Technology rather than interpreting as a Science.

II. DESIGN APPROACH

The audio-based software prototype is an integrated software, designed to play music excerpts automatically, in accordance with the user's current emotions, during gameplay. The objective of this concept is to enhance the user experience by improving the desirability during gameplay. This integrated software to be designed will consist of

1. Real-time emotions felt by the player which are detected and analyzed through physiological signals of the player. (Emotion Elicitation) – The emotion detection/ modelling area would be done through a software which is being developed as part of the wholesome project. The initial dataset for emotion trials were taken from the Mahnob HCI Tagging Database (Soleymani,2012). The EEG feature extraction and preprocessing were done using EEGLab in MATLAB. They were mapped to arousal valence scale and then classified into emotion labels as per

the same labels classified for music excerpts.(This will be a separate research under the development of the integrated software.)

2. Game mechanics in a specific time frame will be fed as another form of an input.
3. A rule base system that will compare the two inputs and decides on which music excerpt to be selected from database.
4. A database consisting of music excerpts classified according to emotions.

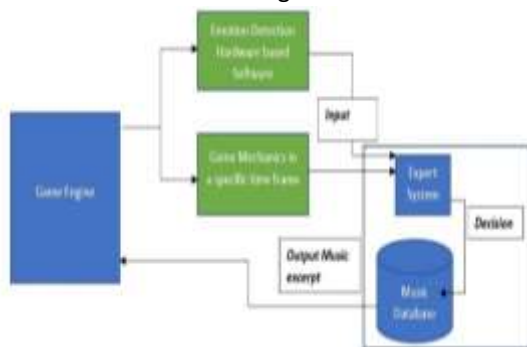


Fig 2 – High level user diagram of the integrated software prototype

Design of this integrated software will occur in stages and this research paper is centered on the work of the music database design based on MER. This data base consists of music that are classified according to a predefined set of emotions. The emotion label set can be change with the game genre because present video game industry is very much diverse that a specific genre explicit certain emotion. For Instance, Zombie games evoke Fear in high amounts. In order to narrow the scope of this research, Car Racing games were selected for designing this MER based application. Racing Games are considered to evoke strong emotional arousal where players are put in a mode of ‘Fight or Flight’. In fact, Racing games can be taken into account as a computer simulator of real- time racing sports. In fact, Arousal is considered as the key to many sporting games like racing. Arousal will make players experience a change in physiological responses such as increase in Blood Pressure, increased Heartbeat, etc. (Juslin et al.,2010).

III. METHODOLOGY

This research is performed under four stages.

1. Selection of a Standard Data Set
2. Customization of Standard Data Set
3. Music Information Retrieval
4. Music Emotion Recognition

The figure below illustrates the detailed view of this research.

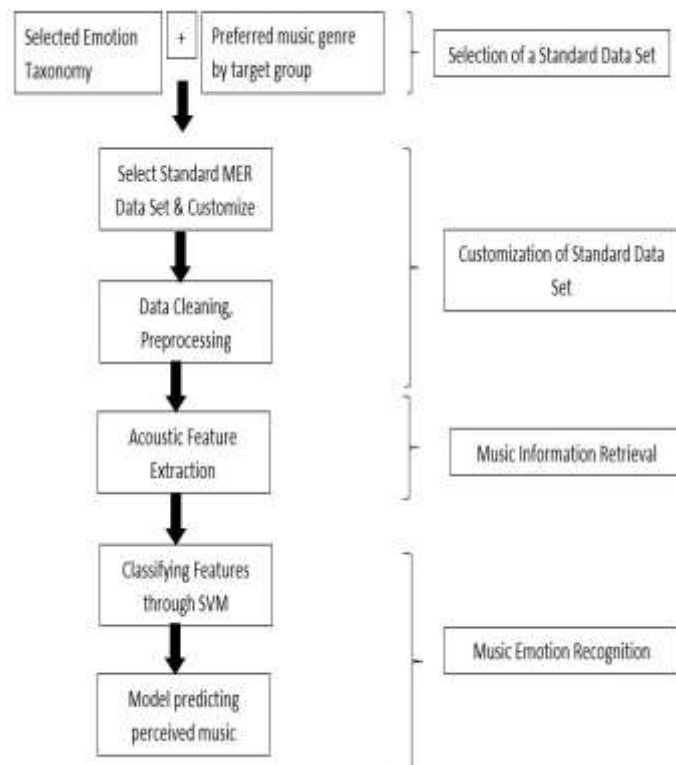


Fig 3 – Methodology (detailed diagram)

A. Selection of a standard data set

‘Emotify’ Data Set (Alijanaki,2015) was selected for this research. Emotify is a standard database listed by the International Society of Music Information Retrieval (ISMIR) which consist of dataset induced on musical emotion from a game “Emotify”. The dataset consists of 400 excerpts of genres (pop, classical, rock, electronic). The selection was based on mainly two factors, Emotion taxonomy pertaining to Racing Games and preferred Music Genre by Racing Game players.

1.Selected Emotion Taxonomy for Mood Detection in Racing Games

A subjective test consisting of 16 game players playing car racing games were closely examined during their game play. Their behavioral changes, language tonal changes and gesture movements were closely examined. A questionnaire consisting of questions were put in forth to a group of car racing game players. This consist of questions related to game mechanics & inquiring them their

emotions felt when they meet certain game mechanics during the game.

On analysis of the above input sources, it was recognized that emotion labels 'Happy', 'Sadness', 'Anger' and 'Tension' were the most frequent emotions experienced during car racing games. The degree of arousal levels correlating to the above emotions was determined by Soleymani et al.'s research work.

Arousal Class	Emotional Keywords
Calm	Sadness, Disgust, Neutral
Medium Arousal	Joy & happiness, amusement
Excited/Activated	Surprise, Fear, Anger, Anxiety

Table 1: Source – (Soleymani, et al.,2012)

However, Soleymani had not defined the Arousal Class for Tension. Based on the psychological theory – "Drive Reduction", we define the arousal class for Tension as Increasing Arousal Rate.

2 Selected Music Genre for Racing Games Audio Content

Present racing video games consist of Electronic music & Rock music with vocals played in the background. Also, it was observed that most of the players listen to their own playlists comprised of electronic & Rock music when playing games. Hence, the above-mentioned music genre was selected. Also, a survey was conducted among the same player group who participated in the subjective test to gather suggestions on playing music based on their emotions. Their suggestions were as follows.

Happy, positive, motivational background music played at the beginning of the game and when the payer has successfully completed a challenge

Disappointing music played when a player loses a racing game.

Music that increases arousal will be ideal to build Tension in an instance where the player is being prepared for a challenge.

Music that evokes anger/ high levels of arousals played during the time of player facing the challenge.

B. Customization of "Emotify" Data Set

Electronic & Rock music genre of "Emotify" Data Set was selected. The Emotify" data set's annotations had been collected using GEMS (Geneva Emotional Music Scales). The emotions that were identified through the survey was matched with the GEMS Scale to extract music excerpts for the training purpose of algorithm.

Identified Arousal Class	GEMS Scale
Calm	Sadness
Medium Arousal	Joy Activation
Excited/Activated	Power
Increasing Arousal Rate	Tension

Table 2: Matched Arousal Classes with GEMS Scale Source – (Soleymani, et al.,2012)

This dataset consists of multiple emotion annotations for a single music excerpt. Hence, the data set needs to be cleaned. This is done by computing Fleiss's kappa, which is a statistical measure designed to estimate agreement, when the answers are binary or categorical. For each song, a number of people (players of the game) rate some items (emotions) on a binary scale (emotion is either present or not).

C. Music Information Retrieval

The music excerpts from the customized version of "Emotify" data set were preprocessed to 44 000Hz mono channel and to 30 second duration.

The preprocessed music excerpts were then subjected to Acoustic feature extraction. Features were extracted using Mirtoolbox. According to previous research, MER researches generally comment music structural features to be as follows:

Joy: loud, fast tempo, regular rhythm

Sad: slow tempo, minor mode

Anger: Loud tempo, dissonance, irregular tempo.

Tension: dissonance, irregular tempo, large dynamical changes, lack of tonal stability.

Considering the above, acoustic features relevant for Arousal detection were selected. The selection of features to this research were mostly inspired by Tan's research (Tan et al.,2019). They were Root Mean Square Energy, Entropy of Energy, Spectral Energy, Spectral Flux, Spectral Roll-off, Beats per Minute, and their standard deviations.

Factors selected for Arousal Detection

1. Root Mean Square Energy
2. Entropy of Energy,
3. Spectral Energy,
4. Spectral Flux,
5. Spectral Roll-off,
6. Beats per Minute.
7. Standard deviations of the above factors.

D. Music Emotion Recognition

The acoustic features that contribute to Low, Medium and High Arousal levels are identified and classified using Support Vector Machine(SVM Classifier), thereby drawing into conclusion the list of factors and their respective value ranges which produces the emotions "Anger", "Joy" & "Sadness". "The rate of change of Arousal rates from medium to high is concluded to produce emotion "Tension".

Afterwards, a model that will predict arousal values given a music excerpt, will be trained using the dataset.

The Support Vector Machine classifier will be deployed using Python's sklearn library (Pedregosa *et al.*, 2011).

IV. DISCUSSION

MER is a rapid emerging research area where many researchers are focusing on processing, searching, organizing and accessing music-related data in terms of perceived music emotions.

However, it was evident that although many researches had been conducted, they were merely based on finding the relations between the acoustic features and perceived music emotions. Various researchers had put forward the numerous acoustic features that contribute for emotion generation within humans. There were minimal numbers of research conducted to develop applications using MER. An instance of an application of MER was a music by Paiva *et al* where the players can select songs from a playlist that were previously classified from MER (Panda *et al.*, 2010). It was observed during literature survey, that even the MER based applications were only music players.

Applications using the technology of MER should not be limited to music players because MER has the capacity to build a unique user centered design experience. Perceived music emotion can be applied into many more user interacting application area domains for better user desirability like E -

Gaming, Advertising, Music therapy sessions etc. In fact, MER should be added as a component for an enhanced User Experience design.

V. CONCLUSION

The sole purpose of this research is to develop a basic platform where games are utilizing the benefits of MER. This paper presents on developing a MER based music excerpts' database exclusive for car racing games so that users will be able to experience the assets of Affective Gaming.

This research paper above describes the conceptual design view of the project. Currently the MER audio classification software is being developed using latest technologies to deploy as a web application.

REFERENCES

- A. Aljanaki, F. Wiering, R. C. Veltkamp 2015. Studying emotion induced by music through a crowdsourcing game. *Information Processing & Management*.
- Kalansooriya P., Mitsui Y, Takaki S., Kinefuchi M., Takahashi A., Nomura S., 2016. The Effect of a Bio-Signal Interactive Virtual Reality on Subjective Impression – A Preliminary Study. *International Journal of Affective Engineering Vol 15 No.3*
- Kim *et al.*, "Music Emotion Recognition: State of Art Review," 11th International Society for Music Information Retrieval Conference (ISMIR 2010) MUSIC.
- K. M. Gilleade, A. Dix, and J. Allanson, "Affective Videogames and Modes of Affective Gaming: Assist Me, Challenge Me, Emote Me."
- K R Tan *et al* 2019 IOP Conf. Ser.: Mater. Sci. Eng. 482 012019
- M. C. Laura and A. M. Andrea, "Emotion recognition techniques using physiological signals and video games – Systematic review – Técnicas de reconocimiento emocional utilizando señales fisiológicas y," vol. 26, no. 46, pp. 19–28, 2017.
- Pedregosa *et al.*, 2011. *JMLR* 12, pp. 2825-2830.
- P. Sundstrom, 2005. "Exploring the affective loop,"
- R. E. S. Panda, 2010. "Automatic Mood Tracking in Audio Music," *Electr. Eng.*, p. 100
- S. LEKAMGE, A. MARASINGHE, P. KALANSOORIYA, and S. NOMURA, 2016. "A Visual Interface for Emotion based Music Navigation using Subjective and Objective Measures of Emotion Perception," *Int. J. Affect. Eng.*, vol. 15, no. 2, pp. 205–211,

Soleymani M., Pantic M & Pun T., 2012. Multimodal Emotion Recognition. OEE TRANSACTIONSON AFFECTIVE COMPUTING,3(2)