International Journal of Future Generation Communication and Networking
Vol. 13, No. 2, (2020), pp. 62 - 69

Brief Survey on Techniques for Intelligent Music Player Applications

Dhaval Pantojee[1], Nikhil Tank[2], Chaitali Wani[3], Vidya Phaphal[4], & Prof. V.S. Khandekar[5]

Department of Information Technology, Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Abstract

Computer vision techniques are utilized in many fields like control , event monitoring, marketing, healthcare field, internal control , military technology, etc. one among the sub-areas of computer vision is countenance recognition. face acts because the main indicator for the behavioral and therefore the spirit of the individual. Facial expressions which may be classified as fear, happiness, joy, sadness, aggressiveness is recognizable with computer vision techniques. Human reacts and responds to music and this music features a high impact on an individual's brain activity. People tend to concentrate to music supported their interests and mood.

This paper gives a survey of various techniques employed by the researchers for mood detection based on humans' face emotions. It also surveys the various sorts of existing applications available and briefly summarizes techniques utilized in each application.

I. INTRODUCTION

Nowadays it has become necessary to identify the facial expression of human which helps the organization as well as individual to recognize the emotions of an individual. It is often applied to all different places where face recognition plays an important role in identifying the emotion.

Emotional aspects have a enormous impact on communication. Emotion recognition is carried out diversely; it may be verbal or non-verbal. Voice (Audible) is a verbal form of communication & Facial expression, action, body postures, and gesture is a non-verbal form of communication. Social intelligence like communication understanding, decision making and also helps in understanding the behavioral aspect of human.

Music is something that enhances an individual's life. Mood detection based on emotion is one of the present topics within the various fields which provides a solution to varied challenges. Most of the music lover's users found themselves in a hectic situation once they don't find songs corresponding to their mood in the situation. Again,

there's increasing advancements within the sector of multimedia and technology, with increasing features like fast forward, reverse, variable playback speed (seek & time compression), local playback, streaming playback with multicast streams and including volume modulation, genre classification. But these features satisfy the user's basic requirements, yet the user has got to face the task of manually browsing through the playlist of songs and choose songs based on his current mood and behavior. So, there's a requirement for the system which will reduce human efforts of manually playing the son based on human mood.

ISSN: 2233-7857 IJFGCN

II. ANALYSIS OF EMOTION RECOGNITION TECHNIQUES

While researching for facial expression recognition, we came across various methods of expression classification. One such method of classification was by classification of emotions using the person's heartbeat as a classification criterion. K Chankuptarat [3] mentioned this method in their paper which was able to distinguish between 4 basic emotions (happy, sad, angry and neutral) using a classifier which used heartbeat as a classifying agent. In the paper, it is stated that neutral emotion has most stable heart rate (around 80 to 100 bpm). Similarly, angry has the highest above 100bpm.

Although this might be one of the easier methods of analyzing expressions due to its relatively lower number of factors involved, it is generally not a good idea to base assumptions about emotions only on one factor (that too on heart rate of a person).

F. Abdat et al [5] in their paper use a real time face detector module which is available in OpenCV library which is an adapted version of the original Viola-Jones Face Detector. This module helped them create an anthropometric model of the human face, since it uses a 68_landmark approach to extract features of the face from the real-time image.

The Anthropometric model of the face had to be localized. For this purpose the paper suggested the following stages of localization. Main Axis localization, facial feature points localization using proportional position, Facial feature points detection using Shi&Thomasi method for more accuracy.

The above mentioned part was just for face detection, when it came to facial expression recognition, their approach was based on a system called Facial Action Coding System(FACS for short). FACS Stated that there were 22 facial muscles which were closely relevant to human expression. Thus the representation of these facial muscles was necessary and was done using a pair of key points, a dynamic one and a fixed one. The dynamic points moved during a real time image of a facial expression while the fixed points were not moved.

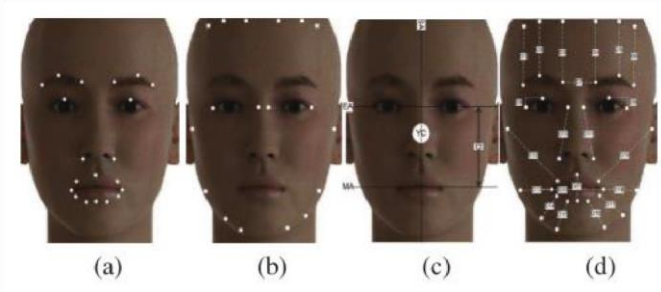


Fig.-1 a:Dynamic points, b:Static Points ,c: Principal axis, d:facial distances[5]

In terms of results regarding the accuracy of emotion classification, here is the confusion matrix for two different databases:

1:Happy 2:Fear 3:Disgust 4:Anger 5:Sadness 6:Surprise 7:Neutral

	1	2	3	4	5	6	7
1	**98.5%**	0%	0%	0.5%	0%	0%	1%
2	0%	**100%**	0%	0%	0%	0%	0%
3	0%	0%	**99%**	0%	1%	0%	0%
4	0%	0 %	0 %	**100%**	0%	0%	0%
5	0 %	0 %	0 %	0%	**97%**	0 %	3%
6	0%	0%	0%	0%	0 %	**100%**	0%
7	0%	0%	0%	0%	2.5%	0%	**97.5%**

ISSN: 2233-7857 IJFGCN

Table 1: Confusion matrix of Emotion classification for Kohn-Kanade Database

	1	2	3	4	5	6	7
1	**98.33%**	0%	0%	0%	0%	0%	1.66%
2	0%	**100%**	0%	0%	0%	0%	0%
3	0%	0%	**98.33%**	0%	0%	0%	1.66%
4	0%	0%	0%	**100%**	0%	0%	0%
5	0%	0%	0%	0%	**91.66%**	0%	8.33%
6	0%	0%	0%	0%	0%	**100%**	0%
7	0%	0%	0%	0%	5%	0%	**95%**

Table 1: Confusion matrix of Emotion classification for FEEDTUM Database

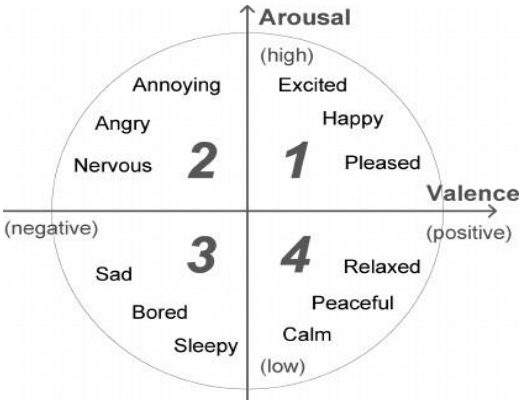
Another method of expression recognition is described by A. Bhardwaj[6] in which EEG methodology was implemented. EEG(Electroencephalography) is a system for measuring electrical signals generated by the brain. The data is acquired through the receptors placed on the scalp of an individual and upon presenting an external stimulus (like an emotion evoking image or video) EEG signals are generated which get captured by the receptors. After capturing the EEG data, SVM is applied to the acquired data to classify the images into different emotions.

The method of using EEG signals was proven to be simpler for the machine to understand but acquisition of such data is bound to be a difficult task since there do not exist any such devices that are able to capture EEG signals without the use of receptors which makes this method of expression recognition less feasible.

F. Abdat et al. [5] implemented an emotion recognition system based on facial expression. Here SVM method (Support Vector Machine) algorithm is used to recognize emotions through facial expression with an emotion recognition rate more than 90% in real-time.

Finally, there is a method called Viola Jones Algorithm which is discussed in “EmoPlayer: An Emotion Based Music Player”
[8] which uses a Haar Cascade Classifier for object detection (from an image) and landmark point extraction after which the features are fed to an SVM for training and testing. The SVM then is able to classify the emotions into 6 categories (happy, sad, fear, disgust, anger and surprise). The emotion is detected using a system known as 68_landmark which extracts positions of 68 different facial features from the image. Then to categorize emotions, displacements from “neutral” face is calculated which help identify and label different facial expressions accurately. With this method in place a proper SVM model can be trained which will be able to classify unseen

ISSN: 2233-7857 IJFGCN



——Fig. 1. Russel 2D Valence Arousal Model [14].

feature displacements in real time.

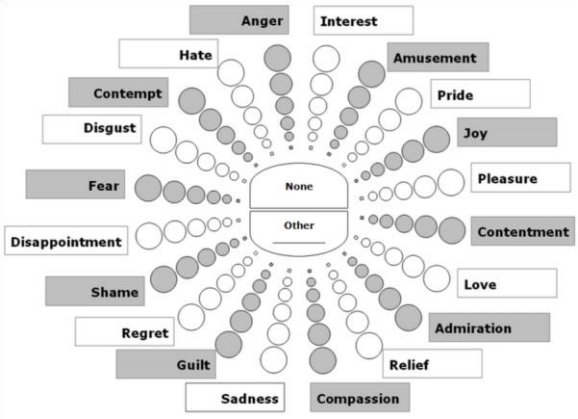


Fig. 2. Geneva Emotion Wheel [15].

An intelligent music player has to not just recognize the person's facial expressions alone to be able to correctly play the music. It should also be intelligent enough that it can categorize the music with accuracy according to the parameters. The following section discusses the different methods of audio feature extraction methods by which classification of the features of music, such as (but not limited to) the pitch, amplitude or the timbre.

III. AUDIO FEATURE EXTRACTION METHODS

Ramya Ramanathan et al. [1] proposes an intelligent system that categorically classifies a music collection based on the emotions conveyed by each song and then recommends an appropriate playlist to the user based on their current mood using k-means clustering.

Shlok G has described a method of Music Classification with the help of features extracted from acoustic feature libraries like LibROSA, aubiopitch, etc. A methodology called Recursive Feature Elimination (RFE in short) is used which works by recursively eliminating features and building a model based on remaining features that best contribute to the accuracy of the model. Based off of these features an artificial neural network is trained which successfully classified the songs into 4 classes. The classified songs are then mapped to the user's mood based on models called the Russel 2D Valence-Arousal Model (Fig. 1) and Geneva Emotion Wheel (Fig. 2).

ISSN: 2233-7857 IJFGCN

In their paper about Evaluation of Musical Features for Emotion Classification [4], Y. Song discusses the difficulties of representing emotion through physical means (like facial expressions). These models only provide an idea of a perceived emotion a song shows.

Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation[3], while preprocessing All the songs were down sampled to a uniform bit-rate of 128 kbps, a mono audio channel and resampled at a sampling frequency of 44100 Hz. The songs are then further divided into smaller clips containing the most meaningful parts of the song. They also discovered many mood sensitive audio features by reading some latest works and results from 2007 MIREX audio mood classification task.

The paper also mentioned the two most popular modular approaches in categorization of music: One is Categorical and the other is a Dimensional model for emotions. The categorical model bounds emotions into limited number of innate categories such as: happy, sad, anger and fear/surprise. While the Dimensional approach considers all the affective terms arising from neurophysiological systems (valence and arousal model). Although dimensional model approach would be more accurate the technologies require us to be able to understand the mindset of the singer/music composer in order to better generate a model for the purpose, which is difficult to achieve currently.

While studying another paper we came across a tool called Mood Cloud [7] which was created for automatic mood prediction from audio content. This tool was able to visualize in real time the prediction probabilities of five mood categories: happy, sad, aggressive, relax, party. In order to make the system aware of the mood concept the system then had to generate something called as a synonym set with the help of wordnet and mapped them with the tags the songs had associated with from a website known as last.fm. Their validation step involved feedback from annotators on the website about the song on why a particular song deserved a certain tag. This way a song database of 1300 songs was developed divided into 5 categories. The internal categorization of songs was done using LibSVM, a popular open source machine learning library written in which implements the sequential minimal optimization algorithm for kernelized support vector machines. Summarizing it, a real-time music mood visualization tool, classifies music emotions into 5 types; namely aggressive, happy, party, relax and sad. It applies the SVM library to analyze the emotion dataset. The result is then presented by using a Flash player.

Along with the methods of extraction of different datapoints required (like facial expressions and audio features), we came across different applications of intelligent music players. We had found six unique and interesting ways in which such a music player was developed. The next section gives a brief summarization of the different styles of implementation of an intelligent music player.

IV. EXISTING MUSIC PLAYER APPLICATIONS

The paper “EmoPlayer: An Emotion Based Music Player (2016) [8]” puts forth the proposed system based on facial expression extracted will generate a playlist automatically there by reducing the effort and time involved in rendering the process manually. Facial expressions are given using inbuilt camera. This paper has proposed Viola-Jones algorithm and multiclass SVM (Support Vector Machine) for face detection and emotion detection respectively. Music Player is created using python. The paper proposed wxpython for creating GUI for Music Player. Music Player has play-pause button, next, previous button, volume slider and playback slider. It also consists of List box to display list of songs according to detected emotion. Emotion is provided by SVM to Music Player for loading and playing songs.

Krittrin Chankuptarat et al. [2] proposes an emotion-based music player, which can suggest songs based on the user's emotions; sad, happy, neutral and angry. The application receives either the user's heart rate

ISSN: 2233-7857 IJFGCN

or facial image from a smart band or mobile camera. It then uses the exact classification to classify the user heart rate to identify the user's emotion. 2 kinds of the classification method; the heart rate-based and the facial image-based methods are identified here. Then, the application returns songs which have the same mood as the user's emotion.

Their paper "Emotion Based Music Player Using Facial Recognition (2017) [2]"puts forth The application is thus developed in such a way that it can manage content accessed by user, analyze the image properties and determine the mood of the user based on mp3 file properties so that they can be added into appropriate play lists according to the mood. The proposed system tries to provide an interactive way for the user to carry out the task of creating a playlist. The working is based on different mechanisms carrying out their function in a predefined order to get the desired output.

The paper "EMO PLAYER: Emotion Based Music Player (2018) [3]" puts forth The Project Emo Player (an emotion based music player) is a novel approach that helps the user to automatically play a song based on the emotions of the users. It recognizes the facial emotions of user and plays the songs according to their emotions. The emotions are recognized using a machine learning method Support Vector Machine Algorithm (SVM) algorithm. The webcam captures the image of user. It then extracts the facial features of user from captured image. The training process involves initializing some random values say smiling and not smiling of our model, predict the output with those values, then compare it with models prediction and then adjust the values in order that they match the predictions that were made previously. Evaluation allows the testing of model against data that has never been seen and used for training and it's meant to be representative of how model might perform when in the real world. According to the emotion, the music will be played from the predefined directions. The paper "Emotion Based Music Player (2017) [4]" puts forth problem, this paper proposes an emotion-based music player, which is able to suggest songs based on the user's emotions; sad, happy, neutral and angry. The application receives either the user's heart rate or facial image from a smart band or mobile camera. It then uses the classification method to identify the user's emotion. This paper presents two kinds of the classification method; the heart rate-based and the facial imagebased methods. Then, the application returns songs which have the same mood as the user's emotion. The experimental results show that the proposed approach is able to precisely classify the happy emotion because the heart rate range of this emotion is wide.

Yading Song, Simon Dixon, Marcus Pearce et al [4] apply an SVM-based approach for classifying the music emotions based on tags of the Last.FM website. There are 4 emotions provided in this research; namely, happy, angry, sad, and relax.

Cyril Laurier and Perfecto Herrera et al. [7], a real-time music mood visualization tool, classifies music emotions into 5 types; namely aggressive, happy, party, relax and sad. It applies the SVM library to analyze the emotion dataset. The result is then presented by using a Flash player.

Maaoui et al. [10] proposed a system fully automatic facial expression and recognition system based on three-step face detection, facial characteristics extraction, and facial expression classification. This system proposed an anthropometric model to detect the face feature point combined to shi-Tomasi method. In this method the variation of 21 distances which describe the facial feature from the neutral face and the classification base on SVM (Support Vector Machine).

Rahul Hirve et al. [8] proposed system based on facial expression extracted will generate a playlist automatically thereby reducing the effort and time involved in rendering the process manually Facial

ISSN: 2233-7857 IJFGCN

expressions are given using an inbuilt camera. We have used the Viola-Jones algorithm and multiclass SVM (Support Vector Machine) [9][10][11] for face detection and emotion detection respectively.

The paper "Intelligent Music Player Based on Human Recognition [5]" This paper proposes an intelligent agent that sorts a music collection based on the emotions conveyed by each song, and then suggests an appropriate playlist to the user supported his/her current mood. The user's local music collection is initially clustered supported the emotion the song conveys, i.e. the mood of the song. This is calculated taking into consideration the lyrics of the song, as well as the melody. Every time the user wishes to generate a mood-based playlist, the user takes a picture of themselves at that instant. This image is subjected to facial detection and emotion recognition techniques, recognizing the emotion of the user. The music that best matches this emotion is then recommended to the user as a playlist.

The paper "Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation (2017) [3]" EMP which recommends music based on the real-time mood of the user. EMP provides smart mood-based music recommendation by incorporating the capabilities of emotion context reasoning within our adaptive music recommendation system. Our music player contains three lash modules: Emotion Module, Music Classification Module and Recommendation

Module. The Emotion Module takes a picture of the user's face as an input and makes use of deep learning algorithms to spot their mood with an accuracy of 90.23%. The Music

Classification Module makes use of audio features to realize an interesting results of 97.69% while classifying songs into 4 different mood classes. The Recommendation Module suggests songs to the user by mapping their emotions to the mood sort of the song, taking into consideration the preferences of the user.

V. CONCLUSION

Identification of human emotion based on human face emotions and has many applications in real life. It avoids the hectic work of choosing the song every time depending on the individual's mood.

From all the papers we were able to study, we found out about many techniques useful for emotion detection and recognition, some of them were accurate but not so feasible (EEG signal processing), while others were feasible to an extent, but not as accurate. We realized there is a need for a system that could both accurately predict the user's emotion while being sufficiently feasible. Implementing a system that is accurate and feasible should be the goal of the future. During our research on this topic, we found many different methods for emotion recognition out of which, 3 of the papers used SVM to classify emotions and one of which has about 90% accuracy rate.

Moreover, we discovered various methods for audio feature extraction where the most feasible method was the method called Recursive Feature Elimination.

Furthermore when researching this topic, we realized that emotion based music applications have existed as a concept for many years, but not in field of research. We found many applications that had their own twist to the application but when it came to technicality, most of the applications were found to be very homogeneous.

REFERENCES

[1] Ramya Ramanathan; Radha Kumaran; R Ram Rohan; Rajat Gupta; Vishalakshi Prabhu "An Intelligent Music Player Based on Emotion Recognition", 2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS).

ISSN: 2233-7857 IJFGCN

[2] Krittrin Chankuptarat; Raphatsak Sriwatanaworachai; Supannada Chotipant,"Emotion-Based Music Player", 2019 5th International Conference on Engineering, Applied Sciences and Technology (ICEAST)

[3] Shlok Gilda,1Husain Zafar, Chintan Soni and Kshitija Waghurdekar"Smart Music Player Integrating FacialEmotion Recognition and MusicMood Recommendation", 2017 IEEE.

[4] Yading Song, Simon Dixon, Marcus Pearce, "EVALUATION OF MUSICAL FEATURES FOR EMOTION CLASSIFICATION", University of London, ISMR, 2012.

[5] F. Abdat; C. Maaoui; A. Pruski,"Human-Computer Interaction Using Emotion Recognition from Facial Expression", 2011 UKSim 5th European Symposium on Computer Modeling and Simulation.

[6] Aayush Bhardwaj; Ankit Gupta; Pallav Jain; Asha Rani; Jyoti Yadav "Classification of human emotions from EEG signals using SVM and LDA Classifiers", 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)

[7] Cyril Laurier and Perfecto Herrera, Mood Cloud: A Real-Time Music Mood Visualization Tool, Music Technology Group Universitat Pompeu Fabra Ocata 1, 08003 Barcelona, Spain.

[8] Rahul Hirve1, Shrigurudev Jagdale2, Rushabh Banthia3, Hilesh Kalal4& K.R. Pathak5, "EmoPlayer: An Emotion Based Music Player", Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-5, 2016

[9] Prof. Vijaykumar R. Ghule, Abhijeet B. Benke, Shubham S. Jadhav, Swapnil A. Joshi,"Emotion Based Music Player UsingFacial Recognition", IJIRCCE2017.

[10] F. Abdat, C. Maaoui and A. Pruski, "Human-computer interaction using emotion recognition from facial expression", 2011 UKSim 5th European Symposium on Computer Modelling and Simulation.

[11] Hemanth P1, Adarsh1, Aswani C.B1, Ajith P1, Veena A Kumar, "EMO PLAYER: Emotion Based Music Player", IRJET,Volume: 05 Issue: 04 | Apr-2018.

[12] Renuka R. Londhe, Dr. Vrushshen P. Pawar, ―Analysis of Facial Expression and Recognition Based on Statistical Approach‖, International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.

[13] . Prof. Jayshree Jha1, Akshay Mangaonkar2, Deep Mistry3, Nipun Jambaulikar4, Prathamesh Kolhatkar "Facial Expression Based Music Player",International Journal of Advanced Research in Computer and Communication EngineeringVol. 4, Issue 10, October 2015.

[14] Yang, yi-hsuan & Chen, Homer. (2012). Machine Recognition of Music Emotion: A Review. ACM Transactions on Intelligent Systems and Technology (TIST). 3. 10.1145/2168752.2168754.

[15] Blaiech, Hayfa & Neji, Mohamed & Wali, Ali & Alimi, Adel. (2013). Emotion recognition by analysis of EEG signals. 13th International Conference on Hybrid Intelligent Systems, HIS 2013. 312-318. 10.1109/HIS.2013.6920451.

ISSN: 2233-7857 IJFGCN