

Mood Therapist: Emotion Based Music Player

¹Isha Shyam Repal, ²Sathvika Selvandurai, ³Prof. Prita K Patil

¹Student, ²Student, ³Assistant Professor

¹Computer Department,

¹Vidyalankar Institute of Technology, Mumbai, India

Abstract : In this fast paced and stressful world, music plays a very important role in a lot of people's lives. It could be for entertainment purposes and sometimes even for therapeutic reasons. To satisfy basic user's requirements there are a lot of music players available with basic features and a large library.

One's tastes in music vary mostly according to their current mood and have to manually select the songs to soothe them. Hence to be even more personal and efficient understanding human emotions is desirable for better human-machine interaction. Humans express emotions in different ways like facial expression, gestures, speech and even by written text. In this paper we propose a system which uses facial features and speech as inputs to detect the mood and give the corresponding playlist. It also takes the user's feedback into consideration while doing so.

IndexTerms - Face Recognition, Machine Learning, Music Analysis

I. INTRODUCTION

Music plays a very important role in enhancing an individual's life as it is an important medium of entertainment for music lovers and listeners and sometimes even imparts a therapeutic approach [1]. In today's world, various music players have features like fast forward, variable speed, genre classification. Although these features satisfy the user's basic requirements, yet the users have to manually make playlists for any current moods and suit their tastes. That is the requirements of an individual, a user sporadically suffered through the need and desire of browsing through his playlist, according to his mood and emotions. Using the old music players, a user manually browsed through his playlist and selected songs to comfort one's self. This process is very draining for a user who needs their soothing playlists at that very instant and has to search for songs instead.

Emotions are synonymous with the aftermath of interplay between an individual's cognitive gauging of an event and the corresponding physical response towards it. In the past decade, considerable amounts of research have been done on emotion recognition from voice, visual behavior. Very good progresses have been achieved in this field, and several commercial products have been developed, such as smile detection in camera. A facial expression and human speech is a discernible manifestation of the emotive state, motive, and psychopathology of a person. Hence this paper proposes a methodology that aims at minimizing the drawbacks and shortcomings of the existing technology. The objective of this paper is to provide a music playlist in response to the user's mood captured. It categorizes six emotions: angry, sad, happy, neutral, fear, surprise.

The emotion recognition stage is based on image processing, semantic analysis and machine learning. The music analysis is done by reading the MP3 metadata of a music file.

The main objective of this paper is to detect emotions from facial features, speech input and to give a mood dependant music playlist as an output. The image processing step requires turning the image to grayscale and resizing it. This is followed by extracting multiple features using different techniques and adapting different classifiers to determine the mood of the user. The sentences spoken by the user are converted to text to get the exact feelings of a person in a concise manner. Emotions from text are detected based on the combination of detected keyword, learned patterns, and other supplementary information. Thus, an analysis is made to determine the best solution for the emotion recognition problem based this project.

Using the bit stream from mp3 files, we extract metadata to determine the required information for each particular song. Using the determined emotion, create a playlist of songs for the user.

II. LITERATURE REVIEW

A. Mood Therapist:

Mood therapist is a music player which plays songs according to the mood of the user. The mood of the user is extracted by reading the facial features and text converted from speech.

According to the user feedback like number of clicks, a playlist called 'Frequently Played' is created which helps to create a more user-specific playlist according to the detected mood.

B. Emotion Detection from Face:

There are different algorithms used for feature extractions and the computation time and accuracy is different for different algorithm. The basic facial expressions are sadness, surprise,, anger, neutral, disgust and happiness. We have to design playlist according to the emotional state of the user.[1]

Several methods have been proposed to detect and recognize emotions from facial features with certain algorithms and very few which can make a music playlist based on these emotions.

Jung Hyun Kim.[10] created probability based music mood model and implementation of music recommendation system using the mood model.

Object Detection using Haar feature-based cascade classifiers is an effective object detection method. It is a machine learning based approach. In this approach a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. First of all, lot of positive images (images of faces) and negative images (images without faces) are used

to train the classifier. Then features are extracted from it. Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle. [4]

C. Emotion Detection from Text:

Detecting emotional state of a person by analyzing a text document written by them maybe challenging but very essential due to the fact that often text are often direct using emotional words, but also sometimes use many slang words which have many hidden concepts. Recognizing the emotion of the text plays a key role in the human-computer interaction [12].

III. PROPOSED WORK

The work flow of the system goes as follows:

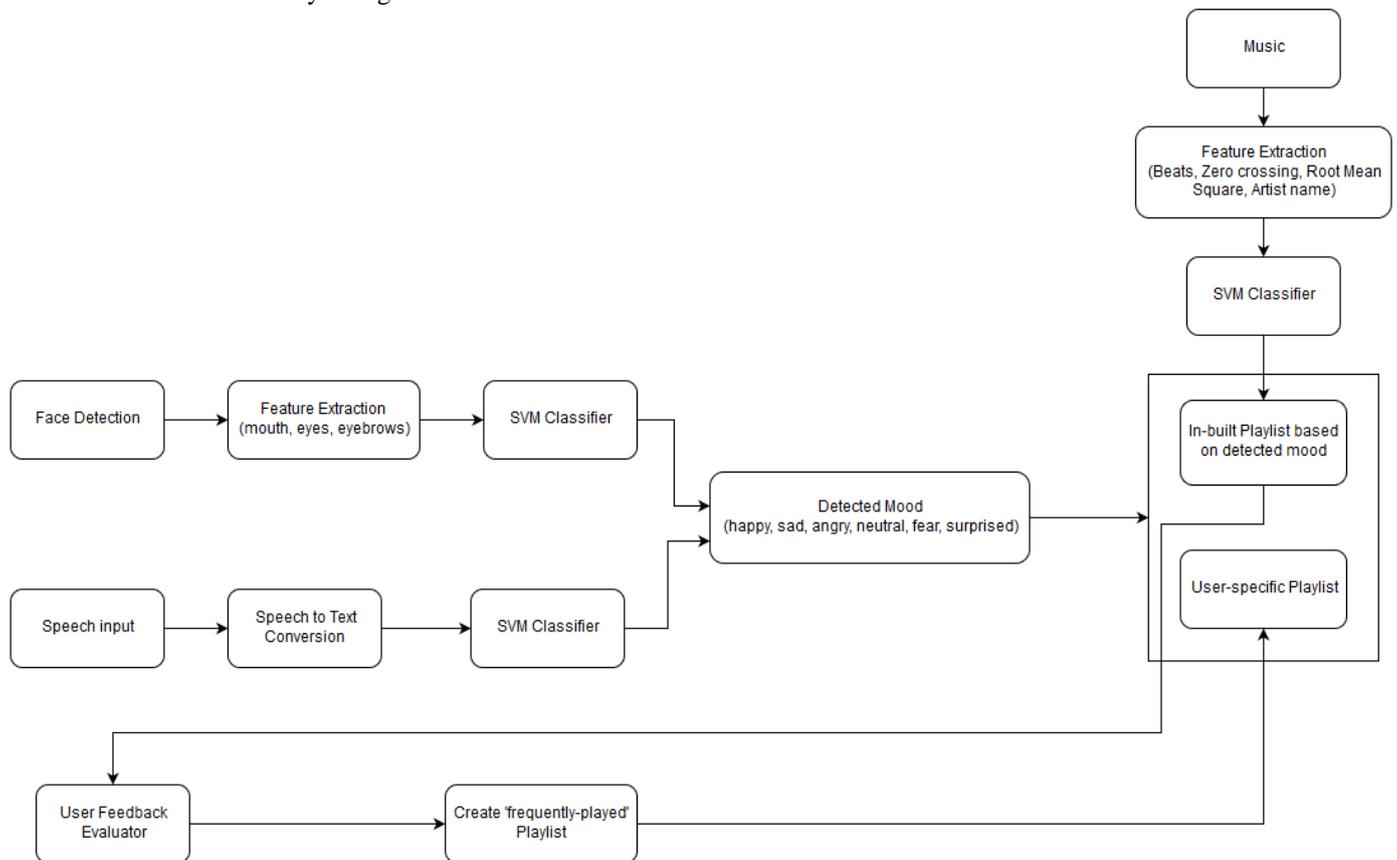


Figure 1: Workflow of Mood Therapist

The different stages of our proposed system are as follows:

- Face Recognition
- Emotion detection from face
- Speech to Text conversion
- Emotion detection from text
- Feature extraction from songs
- Playlists of songs according to the moods
- Music playlist based on User Feedback

Each stage is explained briefly below.

A) Face Recognition:

Among the different algorithms we are using Haar Cascade which uses lot of positive and negative images for training. The platform we have opted for is OpenCV. Open CV (Open Source Computer Vision Library) is a library of programming functions mainly focused on real-time computer vision. The library is cross-platform. It mainly aimed at real-time image processing. It provides multiple numbers of functions for face recognition and facial detection. Open CV comes with a trainer as well as detector.[4]

B) Emotion Detection from Face:

A person's face is the most expressive part of a person's body. A change in head position or light setting would affect the mood expressed. A photo extracts a user's data like characteristic positions on the face such as the corners of the mouth, nose, cheekbones, eyebrows, etc. The algorithm composed of three stage: image processing stage, facial feature extraction stage, and emotion detection stage. In image processing stage, we use the proposed image processing algorithm developed in previous study To extract more effective feature, we proposed the new feature extraction method in facial feature extraction stage. The proposed feature extraction method consists of three features regions: eye region, mouth region, and axillary region. There are only vague patterns are given as the input of system in emotion recognition problem. To overcome this difficulty, the fuzzy classifier is adopted in emotion detection stage. When the extracted features are given, the fuzzy classifier returns the recognized emotion.

C) Speech to text conversion:

In our project, we are using the Google Speech-to-Text API. This API uses Synchronous Recognition. In this the audio data is sent to the API which performs recognition on that data and returns the results only after all the audio has been processed. Synchronous recognition requests are limited to audio data of 1 minute or less in duration. The other two methods are Asynchronous Recognition and Streaming Recognition.

D) Emotion Detection from Text:

In computational linguistics, the detection of human emotions in text is very important from an applicative point of view. Emotion is expressed as joy, sadness, anger, surprise, hate, fear and so on. The keyword pattern matching problem can be described as the problem of finding occurrences of keywords from a given set as substrings in a given string [14]. This problem has been studied and algorithmic solutions have been given thought to. In the context of emotion detection this method is based on certain predefined keywords. These words are classified into categories such as disgusted, sad, happy, angry, fearful, surprised etc.

E) Feature extraction from songs:

The following were considered as candidate features for the classification process:

- Tempo: the speed or pace of the piece, measured in beats per minute, the “rhythm” of a song
- Energy: obtained by integrating power over density
- Mode: indicates if a piece is played in major or minor key
- Key: identifies which of the keys the song has been played (A, B, C, D, E, F, G)
- Harmony: relative weighting between notes, characterized as chords or modes.

F) Playlists of songs according to the moods:

Faster tempos are associated with high-energy songs, and slower tempos with lower energy, sadder songs. Loudness of a song can be connected with anger, while softer songs would suggest tenderness, sadness, or fear. Higher overall pitch can be an indicator of happiness, carefree and light moods within a song, while lower pitch implies a darker, sad, and serious tone.

G) Music playlist based on User Feedback:

The system keeps a track of the songs the user ‘liked’ and also songs which were selected multiple times and makes a special playlist ‘Frequently played songs’. Then these songs are classified according to the moods they were selected in. with the help of this playlist a new user-specific playlist is created for each emotion along with the system-created one. Next time when the same emotion is detected priority is given to the user specific playlist but the user can choose the system-created playlist too if needed.

IV. IMPLEMENTATION AND ANALYSIS

The proposed system is implemented using incremental software process model using technologies OpenCV, Spotify, Google cloud speech to text and python programming language.

An implementation of proposed system aims to satisfy Mood Based Music System which will be of great advantage to the users looking for music based on their mood and emotional behavior. It will help to reduce the searching time for music thereby reducing the unnecessary computational time and thereby increasing the overall accuracy and efficiency of the system. Also with its additional features like user specific playlist as per user feedback & emotion capture from voice to text will be a complete system for music lovers and listeners.

From our analysis, proposed system can be compared with existing systems as :

Table 1: Comparison between existing emotion based music player systems and proposed system

Parameter	Music play system using facial expression as input	Music play system using speech as input	Proposed System
Facial Detection	✓	×	✓
Speech to text	×	✓	✓
Analysis of text for emotions	×	✓	✓
Feature extraction from music	✓	✓	✓
Dividing music into playlists into respective moods	✓	✓	✓
User Feedback	×	×	✓

Frequently played playlist	×	×	✓
User-specific playlist	×	×	✓

V. CONCLUSION

In this paper, a suitable emotion based music player is designed. The principle idea is emotions being detected from facial expressions and speech/text inputs helping the system get a better idea of the mood of the user. And with the user feedback system the playlists become tailor-made for the user.

REFERENCES

- [1] 2017, IEEE International Symposium on Signal Processing and Information Technology (ISSPIT). Karthik Subramanian Nathan, Manasi Aruny and Megala S Kannan, College of Engineering, Guindy, India. EMOSIC - An Emotion Based Music Player For Android
- [2] 2016, ECE Seminar. Nisamudeen P. EMO player - the emotion based music player
- [3] The University Of Manchester, The School of Computer Science. Umer Iftikhar, Emotion Based Music Player.
- [4] 2014, Nikhil Zaware, Tejas Rajgure, Amey Bhadang, D. D. Sapkal, Department of Computer Engineering, India. Emotion Based Music Player.
- [5] 2016, Sushmita G. Kamble, Asso. Prof. A. H. Kulkarni, Department Of Computer Science and Engineering. Facial Expression Based Music Player.
- [6] 2017, Sneha Lukose, Savitha S. Upadhya, Electronics and Telecommunication Fr.C.R.I.T, Mumbai, India. Music Player Based on Emotion Recognition of voice signals.
- [7] 2017, Shlok Gilda, Husain Zafar, Chintan Soni and Kshitija Waghurdekar. Smart Music Player Integrating Facial Emotion Recognition and Music Mood Recommendation.
- [8] 2017, Fabio Calefato, Filippo Lanubile, Nicole Novielli. University of Bari "Aldo Moro". EmoTxt: A Toolkit for Emotion Recognition from Text.
- [9] 2015, Jacek Grekow, Emotion Detection Using Feature Extraction Tools.
- [10] JungHyun Kim, Seungjae Lee and WonYoung Yoo, "Implementation and analysis of mood-based music recommendation system," 2013 15th International Conference on Advanced Communications Technology (ICACT), PyeongChang, 2013, pp. 740-743., IEEE.
- [11] Galen Chuang, EmotiSphere: From Emotion to Music, 2015 TEI15 Proceedings of the Ninth International Conference on Tangible, Embedded Embodied Interaction, Pages 599-602., ACM.
- [12] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, "Emotion recognition in human-computer interaction," in IEEE Signal Processing Magazine, vol. 18(1), Jan. 2001.
- [13] Nicu Sebea, Ira Cohenb, Theo Geversa, and Thomas S. Huangc "Multimodal Approaches for Emotion Recognition: A Survey", USA.
- [14] Chun-Chieh Liu, Ting-Hao Yang, Chang-Tai Hsieh, Von-Wun Soo, "Towards Text-based Emotion Detection: A Survey and Possible Improvements", in International Conference on Information Management and Engineering, 2009.