IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

EMOTION BASED MUSIC RECOMMENDATION SYSTEM

M. Keerthana, M. Shruthi, S. Aravind Kumar

Department of Computer Science and Engineering, St. Peter's Engineering College, Hyderabad.

Abstract — The human face plays an important role in knowing an individual's mood. The required input are extracted from the human face directly using a camera. One of the applications of this input can be for extracting the information to deduce the mood of an individual. This data can then be used to get a list of songs that comply with the "mood" derived from the input provided earlier. This eliminates the time-consuming and tedious task of manually segregating or grouping songs into different lists and helps in generating an appropriate playlist based on an individual's emotional features. Facial expression-based music player aims at scanning and interpreting the data and accordingly creating a playlist based the parameters provided. Existing methods for automating the playlist generation process are computationally slow, less accurate and sometimes even require use of additional hardware like EEG or sensors. Thus, our proposed system focus on detecting human emotions for developing emotion-based music player, which are the approaches used by available music players to detect emotions, which approach our music player follows to detect human emotions and how it is better to use our system for emotion detection.

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. In today 's world, with ever increasing advancements in the field of multimedia and technology, various music players have been developed with features like fast forward, reverse, variable playback speed (seek & time compression),local playback, streaming playback with multicast streams. Although these features satisfy the user 's basic requirements, yet the user has to face the task of manually browsing through the playlist of songs and select songs based on his current mood and behaviour. The introduction of Audio Emotion Recognition (AER) and Music Information Retrieval (MIR) in the traditional music players provided automatically parsing the playlist based on various classes of emotions and moods.

AER is a technique which deals with classifying a received audio signal, by considering its various audio features into various classes of emotions and moods, whereas MIR is a field that extracts some critical information from an audio signal by exploring some audio features like pitch, energy, MFCC, flux etc. Though both AER and MIR included the capabilities of avoiding manual segregation of songs and generation of playlist, yet it is unable to incorporate fully a human emotion

controlled music player. Although human speech and gesture are a common way of expressing emotions, but facial expression is the most ancient

and natural way of expressing feelings, emotions and mood.

LITERATURE SURVEY:

Various techniques and approaches have been proposed and developed to classify human emotional state of behavior. The proposed approaches have focused only on the some of the basic emotions. For the purpose of feature recognition, facial features have been categorized into two major categories such as Appearance-based feature extraction and Geometric based feature extraction by zheng et. al [17]. Geometric based feature extraction technique considered only the shape or major prominent points of some important facial features such as mouth and eyes. In the system proposed by Changbo et. al [2], around a total of 58 major landmark points was considered in crafting an ASM.The appearance-based extraction feature like texture, have also been considered in different areas of work and development. An efficient method for coding and implementing extracted facial features together with multi-orientation and multi-resolution set of Gabor filters was proposed by Michael Lyons [10] et. al. An accurate and efficient statistical based approach for analyzing extracted facial expression features was proposed by Renuka R. Londhe et al. [13]. The paper was majorly focused on the study of the changes in curvatures on the face and intensities of corresponding pixels of images. Some of the drawbacks of the existing system are as follows

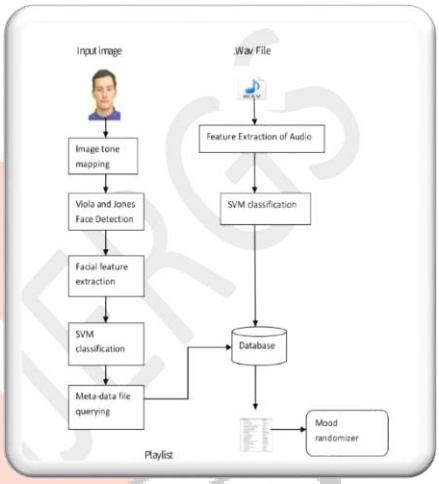
- Existing systems are very complex in terms of time and memory requirements for extracting facial features in real time.
- Based on the current emotional state and behavior of a user, existing systems possess a lesser accuracy in generation of a playlist
- Some existing systems tend to employ the use of human speech or sometimes even the use of additional hardware for generation of an automated playlist.

PROPOSED SYSTEM:

The proposed algorithm in our project involves an emotion music recommendation system that provides the generation of a customized playlists in accordance to the user ‘s emotional state.

The proposed system involves three major modules:

- ❖ Emotion extraction module,
- ❖ Audio feature extraction module
- ❖ An emotion-audio recognition module.



Emotion extraction module and audio feature extraction module are two separate modules and emotion-audio recognition module performs the mapping of modules by querying the audio meta-data file.

METHODOLOGY:

The proposed algorithm in this involves an emotion music recommendation system that provides the generation of a customized playlist in accordance to the user ‘s emotional state. The proposed system involves three major modules: Emotion extraction module, Audio feature extraction module and an Emotion-Audio recognition module. Emotion extraction module and Audio feature extraction module are two separate modules and Emotion-Audio recognition module performs the mapping of modules by querying the audio meta-data file.

A. EMOTION EXTRACTION MODULE:

Image of a user is captured using a webcam or it can be accessed from the stored image in the hard disk. This acquired image undergoes image enhancement in the form of tone mapping in order to restore the original contrast of the image. After image enhancement all images are converted into binary image format and the face is detected using Viola and Jones algorithm where the _Frontal Cart property 'of the algorithm is used that only detects upright and face forwarding features with a maximum threshold value set in the range of 16-20. The output of Viola and Jones Face detection block forms an input to the facial feature extraction block.

B. AUDIO FEATURE EXTRACTION MODULE:

In this module a list of songs forms the input. As songs are audio files, they require a certain amount of pre-processing Stereo signals obtained from the Internet are converted to 16-bit PCM mono signal around a variable sampling rate of 48.6 kHz. The conversion process is done using Audacity technique.

The pre-processed signal obtained undergoes an audio feature extraction, where features like rhythm toning is extracted using MIR 1.5 Toolbox, pitch is extracted using Chroma Toolbox and other features like centroid, spectral flux, spectral roll off, kurtosis,15 MFCC coefficients are extracted using Auditory Toolbox.
Audio signals are categorized into 8 types viz. sad, joy-anger, joy-surprise, joy-excitement, joy, anger, sad-anger and others.

1.Songs that resemble cheerfulness, energetic and playfulness are classified under joy.

2. Songs that resemble very depressing are classified under the sad.

3. Songs that reflect mere attitude, revenge are classified under anger.

4. Songs with anger in playful is classified under Joy-anger category.

5. Songs with very depress mode and anger mood are classified under Sad-Anger category.

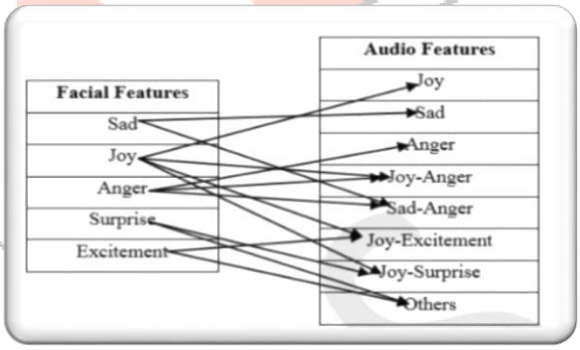
6. Songs which reflect excitement of joy is classified under Joy-Excitement category.

7. Songs which reflect surprise of joy is classified under Joy-surprise category.

8. All other songs falls under others category.

C. EMOTION-AUDIO INTEGRATION MODULE:

Emotions extracted for the songs are stored as a meta-data in the database. Mapping is performed by querying the meta-data database. The emotion extraction module and audio feature extraction module are finally mapped and combined using an Emotion-Audio integration module. Fig 2 illustrates mapping of Facial features and Audio features. For example, if an input facial image is categorized under joy, the system will display songs under joy, joy-anger, Joy-Excitement, Joy-Surprise Category.



RESULTS AND EXPERIMENTS:

Testing and implementation is performed using either MATLAB R2013a or latest version of MATLAB on Windows, 32-bit operating system and Intel i3 core processor. Facial expression extraction is done on both user independent and dependent dataset. A dataset consisting of facial image of 25 individuals was selected for user independent experiment and dataset of 10 individuals was selected for user dependent experimentation. An image of size 4000X3000 was used for static and dynamic dataset experiment.

A. EMOTION EXTRACTION:

Module	Time Taken(sec)
Emotion Extraction Module	0.9994
Emotion-Audio Integration Module	0.0006
Proposed System	1.0000

A user independent dataset of 30 images and user dependent dataset of 5 images is selected for extraction of emotions. Estimated time for various modules of Emotion Extraction Module is illustrated in Table 1.

Time Estimation of Various modules of Emotion Extraction Module

B. AUDIO FEATURE EXTRACTION:

A dataset of around 200 songs was considered for experimentation and testing of audio feature extraction module and the songs were collected from various Bollywood music sites like Djmaza.in, Songs.pk etc. Estimated accuracy for various categories of emotions is depicted in Table 2.

Emotions	Accuracy
Joy	89%
Sad	99%
Anger	99.8%
Surprise	88%
Excitement	95%
Joy-Excitement	96.4%
Joy-Surprise	100%

Estimated Accuracy for different categories of Audio Feature

C.EMOTION BASED MUSIC PLAYER:

The Proposed system is tested and experimented against an in-built camera, thus the total cost involved in implementation is almost negligible. Average estimated time for various modules of propped system is illustrated in Table 3.

Average Time Estimation of Various modules of Proposed System

Module	Time Taken(sec)
Face Detection	0.8126
Facial Feature Extraction	0.9216
Classification using SVM	0.1956
Emotion	0.9994

CONCLUSION AND FUTURE SCOPE:

Experimental results have shown that the time required for audio feature extraction is negligible (around 0.0006 sec) and songs are stored pre-handed the total estimation time of the proposed system is proportional to the time required for extraction of facial features (around 0.9994 sec).Also the various classes of emotion yield a better accuracy rate as compared to previous existing systems. The computational time taken is 1.000sec which is very less thus helping in achieving a better real time performance and efficiency.

The Emotion Based Music System will be of great advantage to users looking for music based on their mood and emotional behavior. It will help reduce the searching time for music thereby reducing the unnecessary computational time and thereby increasing the overall accuracy and efficiency of the system. The system will not only reduce physical stress but will also act as a boon for the music therapy systems and may also assist the music therapist to therapize a patient.

The future scope in the system would to design a mechanism that would be helpful in music therapy treatment and provide the music therapist the help needed to treat the patients suffering from disorders like mental stress, anxiety, acute depression and trauma. The proposed system also tends to avoid in future the unpredictable results produced in extreme bad light conditions and very poor camera resolution

REFERENCES:

[1]. Chang, C. Hu, R. Feris, and M. Turk, ―Manifold based analysis of facial expression,‖ Image Vision Compute, IEEE Trans. Pattern Anal. Mach. Intell. vol. 24, pp. 05–614, June 2006. International Journal of Engineering Research and General Science Volume 3, Issue 1, January-February, 2015 ISSN 2091-2730

[2]. A. Habibzai, ninavin, Mir kamalMirnia,‖ A new algorithm to classify face emotions through eye and lip feature by using particle swarm optimization.

[3]. Byeong-jun Han, Seungmin Rho, Roger B. Dannenberg and Eenjun Hwang, ―SMERS: music emotion recognition using support vector regression‖, 10thISMIR , 2009.

[4]. Alvin I. Goldmana, b.Chandra and SekharSripadab, ―Simulationist models of face-based emotion recognition‖.

[5]. Carlos A. Cervantes and Kai-Tai Song , ―Embedded Design of an Emotion-Aware Music Player‖, IEEE International Conference on Systems, Man, and Cybernetics, pp 2528-2533 ,2013.

[6]. Fatma Guney, ‖Emotion Recognition using Face Images‖, Bogazici University, Istanbul, Turkey 34342.

[7]. Jia-Jun Wong, Siu-Yeung Cho,‖ Facial emotion recognition by adaptive processing of tree structures‖.

[8]. K.Hevener,‖The affective Character of the major and minor modes in music‖, The American Journal of Psychology,Vol 47(1) pp 103-118,1935.

[9]. Samuel Strupp, Norbert Schmitz, and KarstenBerns, ―Visual-Based Emotion Detection for Natural Man-Machine Interaction‖.

[10]. Michael lyon and Shigeru Akamatsu, ―Coding Facial expression with Gabor wavelets.‖, IEEE conf. on Automatic face and gesture recognition, March [16]. Thayer ‖ The biopsychology of mood & arousal‖, Oxford University Press ,1989. [11]. Kuan-Chieh Huang, Yau-Hwang Kuo, Mong-Fong Horng ,‖Emotion Recognition by a novel triangular facial feature extraction method‖.

[12]. S. Dornbush, K. Fisher, K. McKay, A. Prikhodko and Z. Segall ‖Xpod- A Human Activity and Emotion Aware Mobile Music Player‖, UMBC Ebiquity, November 2005.

[13]. 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.

[14]. Russell, ―A circumplex model of affect‖, Journal of Personality and Social Psychology Vol-39(6), pp1161–1178 , 1980.

[15]. Sanghoon Jun, Seungmin Rho, Byeong-jun Han and Eenjun Hwang, ―A fuzzy inference-based music emotion recognition system‖,VIE,2008.

[16]. Thayer ‖ The biopsychology of mood & arousal‖, Oxford University Press ,1989.

[17]. Z. Zeng ―A survey of affect recognition methods: Audio, visual, and spontaneous expressions,‖IEEE. Transaction Pattern Analysis, vol 31, January 2009