ISSN XXXX XXXX © 2018 IJESC

Research Article Volume 8 Issue No.3

Emotion Based Music Player

Vijay Chakole[1], Aniket Choudhary[2], Kalyani Trivedi[3], Kshitija Bhoyar[4], Ruchita Bodele[5], Sayali Karmore[6]
Professor[1], UG Student[2, 3, 4, 5, 6]
Department of Electronics Engineering
K.D.K. College of Engineering Nagpur, India

Abstract:
The human face is an important part of an individual's body and it especially plays an important role in knowing an individual's mood. Extracting the required input from the human face can now be done directly using a camera. This input can then be used in many ways. 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. Various algorithms have been developed and proposed for automating the playlist generation process. Facial Expression Based Music Player aims at scanning and interpreting the data and accordingly creating a playlist based the parameters provided. The scanning and interpreting includes audio feature extraction and classification to get a list of songs belonging to a similar genre or to get a list of similar sounding songs.

I. INTRODUCTION

The field of science is as big as the universe itself. Every passing day there are new developments; if not big or ground breaking, but constructive and leading towards a better tomorrow. Sound and graphics are two vast fields of Science and Engineering that not only intrigue but also attract learners to study them in detail to explore into their depths. Since then many such inventions have propelled us to this time where thinking of various ideas, which might not have been possible a few decades back, and more over implementing them is now possible.

Facial Emotion based Music Player is interactive, sophisticated and innovative desktop-based application to be used as a music player in a differentmanner. The application works in a different manner from the traditional software as it scans and classifies the audio files present on the device and according to the predefined parameters (Audio Features) present on the application in order to produce a set of mood-based playlists. The real time graphical input provided to the application is classified (Facial expression recognition) to produce a "mood" which will then be used to select the required playlist from the earlier set.

II. 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. Figure illustrates block diagram of proposed system.

Fig 1 Block Diagram of Proposed Algorithm

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, 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.

To increase the accuracy and an aim to obtain real time performance only features of eyes and mouth are appropriate enough to depict the emotions accurately. For extracting the features of mouth and eyes certain calculations and measurements are taken into consideration. Equations illustrate the bounding box calculations for extracting features of a mouth.

X (start pt of mouth) =X (mid pt of nose) – (X (end pt of nose) – (Xstartpt of nose)) .

X (end pt of mouth) =X (mid pt of nose) + ((Xendpt of nose) – (Xstartpt of nose)).

Y (start pt of mouth) =Y (mid pt of nose) +15 .

Y (end pt of mouth) =Y (start pt of mouth) +103.

Where (X (start pt of mouth), Y (start pt of mouth)) and (X (end pt of mouth), Y (end pt of mouth)) illustrates start and end points of the bounding box for mouth respectively, (X (mid pt of nose), Y (mid pt of nose)) illustrates midpoint of noise and ((Xendpt of nose), (Xstartpt of nose)) illustrates end and start point of noise. Classification is performed using Support Vector Machine (SVM) which classifies it into 4 classes of emotions.

III. 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 preprocessing 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 four type's viz. sad, happy, surprise, anger and others.
1. Songs that resemble cheerfulness, energetic and playfulness are classified under happy.
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 anger category.

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 is finally mapped and combined using an Emotion-Audio integration module. Illustrates mapping of Facial features and Audio features. For example, if an input facial image is categorized under happy, the system will display songs under happy, anger, surprise category.

Facial Features	Audio Features
Sad	sad
happy	happy
Anger	Anger
Surprise	Surprise

IV. RESULTS AND EXPERIMENTS:

Testing and implementation is performed using MATLAB R2013on Windows7/8, 32/64 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 for static and dynamic dataset experiment.

V.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 1.000 sec, which is very less thus helping in achieving a better real time performance and efficiency.

The system thus aims at providing the Windows operating system users with a cheaper, additional hardware free and accurate emotion based music system. 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. Also with its additional features mentioned above, it will be a complete system for music lovers and listeners.

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.

VI. REFERENCES:

[1]. Bill, David "Personalizing content based on emotions", US2010/0321519, San Francisco: Dec. 23, 2014.

[2]. Carlos A. Cervantes and Kai-Tai Song, ―Embedded Design of an Emotion-AwareMusic Player, IEEE International Conferenceon Systems, Man, and Cybernetics, pp 2528-2533 ,2013.

[3]. 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.

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

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

[6]. FatmaGuney, Emotion Recognition using FaceImages, Bogazici University, Istanbul, Turkey 34342

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

[8]. Chang, C. Hu, R. Feris, and M. Turk, Manifold based analysis of facial expression, Image Vision Comput, IEEETrans. Pattern Analogy. Mach. Intel. vol. 24, pp.05–614, June 2006.

[9]. Michael Lyon and Shigeru Akamatsu ―Coding Facial expression with Gabor wavelets, IEEE conf. on Automatic face and gesture recognition, March 2000.