ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

Facial Expression Based Music Player

Prof. Jayshree Jha[1], Akshay Mangaonkar[2], Deep Mistry[3], Nipun Jambaulikar[4], Prathamesh Kolhatkar[5]

Lecturer, Information Technology, Atharva College of Engineering, Mumbai, India [1]

Student, Information Technology, Atharva College of Engineering, Mumbai, India [2,3,4,5]

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.

Keywords: Music player, Face Detection, Facial Expression Recognition, Audio Feature Recognition, Viola Jones Algorithm.

I. INTRODUCTION

The field of science is as big as the universe itself. Every passing day there are new developments; if not big or groundbreaking, 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.

Now in the present time, where clicking a photo and listening to music 'on the go' is just a part of anyone's daily life, providing any improvements in the working of such technologies that in turn make the user experience better are always appreciated. With the improvements in technology the level of sophistication in software has also increased. Also with the idea of 'keeping it simple', developing sophisticated applications is a challenge.

Facial Expression based Music Player is interactive, sophisticated and innovative mobile (Android) based application to be used as a music player in a different manner. 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. LITERATURE SURVEY

The potential abilities of humans to be able to provide inputs to any system in various ways has caught the attentions of various learners, scientists, engineers, etc from all over the world. The paper by Hafeez Kabini et al [1] addressed the problem of the existing methods typically handle only deliberately displayed and exaggerated expressions of prototypical emotions despite the fact that deliberate behavior differs in visual appearance, audio profile, and timing from spontaneously occurring behavior, by taking efforts to develop algorithms that can process naturally occurring human affective behavior have recently emerged. They introduced and surveyed these recent advances and discussed human emotion perception from a psychological perspective. They examined available approaches to solving the problem of machine understanding of human affective behavior, and discuss important issues like the collection and availability of training and test data.

The 'mind' is a term that has always attracted scientists towards understanding it in a wholesome manner. The most natural way to express emotions is using facial expressions. We humans, often use nonverbal cues such as hand gestures, facial expressions, and tone of the voice to express feelings in interpersonal communications. Nikhil Zaware et al [2] stated that it is very time consuming and difficult to create and manage large playlists and to select songs from these playlists.

Thus, it would be very helpful if the music player itself selects a song according to the current mood of the user using an application to minimize the efforts of managing playlists. In their paper, they stated a way to automatically detect the mood of the user and generate playlist of songs which is suitable for the current mood. The image is captured using webcam and that image is passed under different stages to detect the mood or emotion of the user. 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.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

The application also includes the facility of sorting songs based on mp3 file properties so that they can be added into appropriate playlists according to the mood.

To detect facial expression as indicator to cast a music playlist is one task carried out by Setiawardhana et al [3] in their technical paper. They work by doing facial expression detection system input performed offline by taking photograph of a subject with nearest position from the camera where facial position should not be tilted. The image is identified as a combination of color and feature extraction is performed based on location of eyebrow, eye, and mouth. They use Artificial Neural Network Back propagation method for facial expression detection. The output data is an index, which automatically select and play the music. In this way, the music is modified according to the changes of facial expression. They designed a system to detect three facial expressions: normal, angry, and happy expression. The similarity between features values from each expression influence the ability to differentiate each expression. Offline system evaluation is performed with back propagation neural network method, for learning process, it reaches convergent value with lowest error value when using 10 unit neuron on hidden layer, learning rate value is 0.0625325 and mean square error value is 0.0135.

The paper by Henal Shah et al [4] conveys our proposed intelligent music player using sentimental or emotion analysis. The Emotions are a basic part of human nature. They play a vital role throughout the life. Human emotions are meant for mutual understanding and sharing feelings and intentions. The emotions are represented in verbal and through facial expressions. One can also express his emotions through written text. This paper mainly focus on the methodologies available for detecting human emotions for developing emotion based music player, the approaches used by available music players to detect emotions, the approach a music player follows to detect human emotions and how it is better to use the proposed system for emotion detection. It also gives brief idea about our system's working, playlist generation and emotion classification.

Anukriti Dureha [5] suggested manual segregation of a playlist and annotation of songs, in accordance with the current emotional state of a user, as a labour intensive and time consuming job. Numerous algorithms have been proposed to automate this process. However the existing algorithms are slow, increase the overall cost of the system by using additional hardware (e.g. EEG systems and sensors) and have less accuracy. This paper presents an algorithm that automates the process of generating an audio playlist, based on the facial expressions of a user, for rendering salvage of time and labour, invested in performing the process manually. The algorithm proposed in this paper aims at reducing the overall computational time and the cost of the designed system. It also aims at increasing the accuracy of the designed system. The facial expression recognition module of the proposed algorithm is validated by testing the system against user dependent and user independent dataset. Experimental results of the process indicate that the user dependent results give 100% accuracy, while user independent results for joy and surprise are 100 %, but for sad, anger and fear are 84.3 %, 80 % and 66% respectively. The overall accuracy of the emotion recognition algorithm, for user independent dataset is 86%. In audio, 100 % recognition rates are obtained for sad, sad-anger and joy-anger but for joy and anger, recognition rates obtained are 95.4% and 90 % respectively. The overall accuracy of the audio emotion recognition algorithm is 98%. Implementation and testing of the proposed algorithm is carried out using an inbuilt camera. Hence, the proposed algorithm reduces the overall cost of the system successfully. Also, on average, the proposed algorithm takes 1.10 sec to generate a playlist based on facial expression.

Thus, it yields better performance, in terms of computational time, as compared to the algorithms in the existing literature.

III. PROPOSED SYSTEM

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 pre-defined order to get the desired output. The working can be stated as follows:

1. The proposed System works by first providing a simple enough interface which prompts the user to scan the memory for audio files when the application is opened.
2. Then after the files are detected, they are scanned for audio features and these features are extracted.
3. Then the extracted feature values are subjected to classification according to the parameters provided.
4. These parameters include a limited set of genre types based on which the audio feature values will be processed.
5. After this, the songs are segregated into different playlists based on the feature extraction process. Hence lists of similar sounding songs or songs belonging to similar genres are generated.
6. In the next step, the user camera is invoked with proper permissions and a real time graphical input (image) is provided to the system.
7. The system first checks for the presence of a face in the input using the face detection process , then classifies the input and generates an output which is an emotion (mood) based on the expression extracted from the real time graphical input.
8. After this the classified expression acts as an input and is used to select an appropriate playlist from the initially generated playlists and the songs from the playlists are played.

THE BLOCK DIAGRAM FOR THIS SYSTEM IS GIVEN AS FOLLOWS AND IS SELF-EXPLANATORY.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

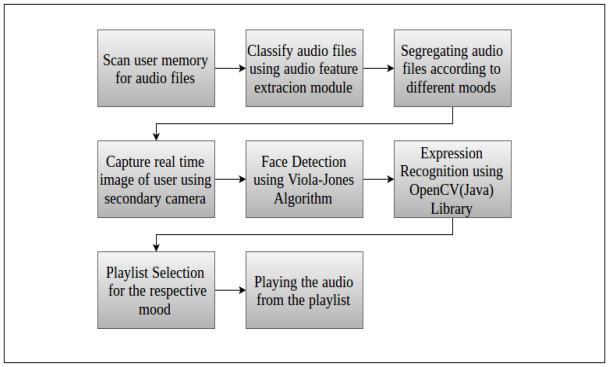

FIG 1: PROPOSED SYSTEM

Working of the system

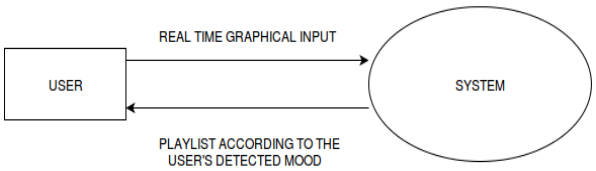

FIG 2: DFD LEVEL 0 DIAGRAM OF THE PROPOSED SYSTEM

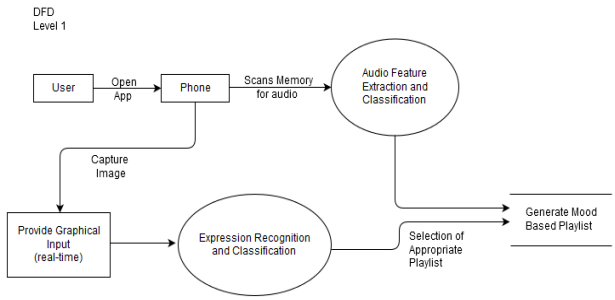

Fig 3: DFD Level 1 diagram of the proposed system

The technologies and algorithms to be used in achieving this goal are explained in the next section.

IV. METHODOLOGIES

A. OpenCV

Open CV is a library of programming functions mainly aim at real time computer vision. It is a C++ implementation library. There is a javacv library which is derived from OpenCV using this we will implement viola and Jones face detection algorithm. Face detection is important as it will classify only if face is present. Expression recognition is also done using graphical based classification method. Audio files will be scanned and features will be extracted from them and according to the mood we get the playlist.

B. Android Studio

Android Studio is the official IDE for Android application development, based on IntelliJ IDEA.

C. Viola-Jones Algorithm

The Viola–Jones object detection framework is the first object detection framework to provide competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones. Although it can be trained to detect a variety of object classes, it was motivated primarily by the problem of face detection. This algorithm is implemented in OpenCV as `cvHaarDetectObjects()`.

D. Facial Expression Recognition

The output of Viola and Jones Face detection block forms an input to the facial feature extraction block. The required features are extracted using the JavaCv library that is imported in the Android Studio java folder. This helps in classifying that the input real time image has 'x' as its expression i.e. the user's expression.

E. Audio Feature Extraction

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 that are obtained from the Internet and 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.

V. CONCLUSION

Music Player has changed in many different ways since it was first introduced. Now-a-days people like to get more out of different applications, so the designing of applications and the thought process behind it has changed. The users prefer more interactive & sophisticated yet simple to use applications.

The proposed system (Facial Expression based Music Player) presents a music player capable of generating a playlist from the songs' audio features and thereby providing the user with an easy way to get the playlist. The proposed model makes use of the Viola-Jones algorithm implemented using java and OpenCV (javacv) to carry out one phase of its functioning and svm classifier is used to carry out the audio feature extraction and classification.

The references used provide us with vital information about the different techniques and strategies followed to carry out their respective individual systems. Based on the knowledge obtained the above content not only provides an in-depth knowledge of the proposed Software development project system but also tries to incorporate the information from various sources in order to execute the tasks using open-source resources. The various aspects of the project have been presented in the above pages in an adequate manner.

The details of the various aspects of the projects have been provided with a well defined set of diagrams.

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 10, October 2015

ACKNOWLEDGMENT

It gives us great pleasure in presenting this project report titled:"Facial Expression Based Music Player".

On this momentous occasion, we wish to express our immense gratitude to the range of people who provided invaluable support in the completion of this project. Their guidance and encouragement has helped in making this project a great success.

We express our gratitude to our project guide **Prof Jayshree Jha**, who provided us with all the guidance and encouragement and making the lab available to us at any time. We also would like to deeply express our sincere gratitude to Project coordinators.

We are eager and glad to express our gratitude to the Head of the Information Technology Dept. **Prof Neelima Pathak**, for her approval of this project. We are also thankful to her for providing us the needed assistance, detailed suggestions and also encouragement to do the project.

We would like to deeply express our sincere gratitude to our respected principal **Prof. Dr. Shrikant Kallurkar** and the management of Atharva College of Engineering for providing such an ideal atmosphere to build up this project with well equipped library with all the utmost necessary reference materials and up to date IT Laboratories

We are extremely thankful to all staff and the management of the college for providing us all the facilities and resources required.

REFERENCES

[1] Hafeez Kabini, Sharik Khan, Omar Khan, Shabana Tadvi "EMOTION BASED MUSIC PLAYER" International Journal of Engineering Research and General Science, Volume 3, Issue 1, 2015.

[2] Nikhil Zaware, Tejas Rajgure, Amey Bhadang, D.D. Sakpal "EMOTION BASED MUSIC PLAYER" International Journal of Innovative Research & Development, Volume 3, Issue 3, 2014.

[3] Setiawardhana, Nana Ramadijanti, Peni Rahayu "FACIAL EXPRESSIONS RECOGNITION USING BACKPROPAGATION NEURAL NETWORK FOR MUSIC PLAYLIST ELECTIONS" Jurnal Ilmiah Kursor, Volume 6, Issue 3, 2012.

[4] Henal Shah, Tejas Magar, Purav Shah and Kailas Devadkar "AN INTELLIGENT MUSIC PLAYER USING SENTIMENTAL ANALYSIS" International Journal of Innovative and Emerging Research in Engineering, Volume 2, Issue 4, 2015.

[5] Anukriti Dureha "AN ACCURATE ALGORITHM FOR GENERATING A MUSIC PLAYLIST BASED ON FACIAL EXPRESSIONS" International Journal of Computer Applications, Volume 100-No.9, 2014.

[6] Byeong-jun Han, Seungmin Rho, Roger B. Dannenberg, Eenjun Hwang "SMERS: MUSIC EMOTION RECOGNITION USING SUPPORT VECTOR REGRESSION" 10th International Society for Music Information Retrieval Conference (ISMIR 2009).

[7] Benoit Mathieu, Slim Essid, Thomas Fillon, Jacques Prado, Ga¨el Richard "YAAFE, AN EASY TO USE AND EFFICIENT AUDIO FEATURE EXTRACTION SOFTWARE" Institut Telecom, Telecom ParisTech, CNRS/LTCI.