MoodSound: A Emotion based Music Player

Mayank vyas

ABSTRACT: Music could be a important component of life. folks take it's facilitate to evoke their emotions and prefer being attentive to songs consistent with their mood. It takes heaps of efforts to search out applicable music from the list for the actual spirit. Music players in today's world don't seem to be giving priority to the spirit of someone. The aim of this paper is to develop music system that considers human emotions under consideration. The spirit is taken from facial expressions through the digital web camera. we've used the CNN classifier to create a neural network model. This model is trained and subjected to observe mood from facial expressions exploitation

OpenCV. Songs happiness to specific sentiments square measure classified on the premise of tempo feature in beats per minute. A system generates music listing supported that detected mood. MoodSound player automatically play that generated music listing to improves user's mood.

1. Introduction

MoodSound system working on principle of human based emotion which uses the local integrated camera and process the image and check captured photo and analysis the emotion of that image and enhancing emotional state.

Emotion of any human is easily capture thorough facial expression and detect mood by help of many researcher's article which say that "wide open eye" it means that observe human maybe happy or sad in nature, whereas the "wrinkled nose" defines anger or disgust emotion. Which we used in our system that detect the mood by help of image processing and machine learning techniques. In world of music term "tempo" known as speed of pace of given piece, which is usually detect in "BPM" which is Beats Per Minute. Songs are having similar or same emotions which are detected using "tempo". Information provided by "tempo" than we provide those song's which is categories by moods.

MoodSound system aims to suggest and play the suitable music, supported user's current emotion with the assistance of image process techniques through user's facial expressions.

2. Related work:

Sentiment-Based Music Play System, has implemented RASPBERRY-PI based module for mood detection and generate music playlist. To capture the voices and conversations in the room, a microphone is connected to Raspberry Pi. They have converted that speeches into text for speech-to-text conversion. They have used the Python library for Speech Recognition 3.1.3. The generated text is passed through naïve Bayesian classifier for the sentimental analysis of the text. They have used the tempo of the song as the attribute to be related to the sentiment in the room. For example, BPM values for the mood filled with happiness, anger, and sorrow will be 125, 130, 85 BPM respectively. The system will play a song on the basis of sentimental analysis of text generated from speeches in the room.

There is ample work has been done in the field of face based emotion detection. Proposed mood based music player, it scans memory for audio files, and classifies audio files using audio extraction module. After dividing audio files into mood based segregation, it captures image from device camera. Feature detection is done with the help of Viola and jones algorithm . With help of OpenCV libraries, it recognizes the emotional state and device plays music accordingly the mood.

In[4], they have proposed a mood detection system. They have used modern CNN for building framework. Their architecture is fully-convolutional neural network that contains 4 residual separable convolutions where each convolution is followed by a batch normalization operation and ReLU activation. Architecture has 60000 parameters. which is corresponds to reduction of 10x of Naive and 80X of original CNN. They have achieved 66% of accuracy in mood detection.

They have proposed a bimodal emotion recognition system with the combination of facial expressions and speech signals. The models obtained from a bimodal corpus with six acted emotions and ten subjects were trained and tested with different classifiers, such as Support Vector Machine, Naive Bayes and K-Nearest Neighbor. It's result reveals that facial expression gives better performance compared to speech and combination of both also gives better performance with SVM classifier.

3. Methodology:

The MoodSound is web based program that mainly focus on captured image using different logical stages like capture picture, detection of face, face feature extraction, mood detect by help of "tempo", playlist generated, play song's. As shown in fig.1.

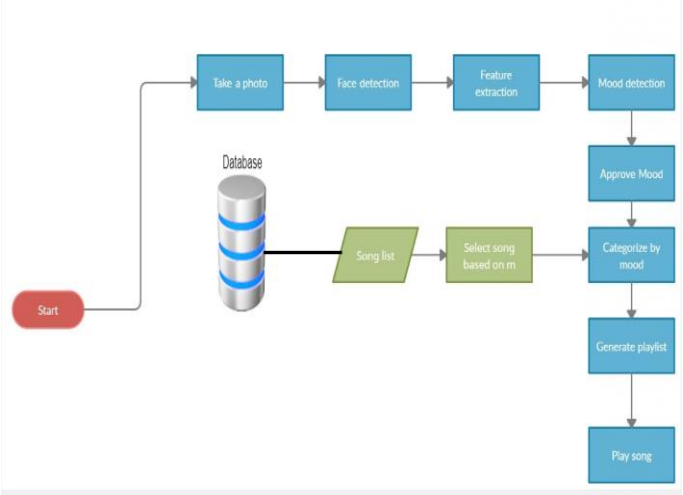

Fig.1: Workflow diagram

3.1Face Detection:

System can take real time input from internet camera. We've used OpenCV libraries to implement face detection through Haar cascade classifier. Haar cascade classifier enforced piecemeal by application of options sorted into criteria such it'll discard the unwanted half from more analysis. Face detection algorithmic rule detects face and more it'll extract needed options from it. It work on the concept of dimensionality reduction. Which reduce the execution time among classification. Principal Component Analysis (PCA) is a one kind of conversion from correlated variables to uncorrelated in the form of mathematical values. It is mostly used for the observing data and from that by some probabilistic calculation generate models. The flow of Fisherface is like it takes classified images then it will reduce the dimension of the data and by calculating its statistical value according the given categories it stores numeric values in .xml file. While prediction it also calculate the same for given image and compare the value with the computed dataset values and give according result with confidence value. Resizing image we have chosen for dataset it mostly related to the size which can give an precise output. The size is chosen such like the model can able to easily distinguish face from image by haarcascade model. And the size what we get from real time scan is not always same as data (very less difference) so, We resize it to the exact model data size. In our case we have chosen 350*350.

3.2Mood Detect:

Facial expression detection in Fisherface works with the help of trained models. Reason behind this is to allow user to take dataset according to their use. Suppose if we take a huge amount of dataset of around 25-30k it will give nice accuracy no doubt but if the situation is like that the user of the devices are a few people. Now in such condition if we take some precise dataset with around 400-450 images as input releted to the user then it will also give good accuracy with the benefit of less amount of dataset and less storage on memory to operate. As well as small memory of data give output fast which result in quick response time. Here we first tried with Cohn-Kanade dataset then we made some classification in the as our need make it to train our model. For training, we have used Fisherface method of cv2 library. For training data model we have make a python code which grab all the classified images from folders and map it with it's emotion. These data we at an instance stored in dictionary and then use. train method to train model. In training module dataset is divided into 3 classes for 3 different mood which is define as +1,+2,+3 which assigned for happy, neutral and sad which same label for this modal.

3.3 Generate playlist:

By using Echonest API, we can calculate the tempo of song i.e. BPM (beats per minute). According to the tempo values, we have categorized the songs according to the sentimental categories. BPM values for the mood filled with happiness, anger, and sorrow will be 125, 130, 85 BPM respectively. Finalized mood is imported to this playlist generation module, so that it can recommend appropriate playlist on basis of finalized mood.

3.4 Song play:

We have implemented the linking of python with javascript through eel library. Which provide us the privilege to access python methods from javascript as well as vice versa. Here the striating flow will be in python code as the library is implemented in python then it transfers the control to html, JS. And according to the result we show emotions.

According to which we can classify emotion directory for playing song we have chosen this 4 emotions.

3.5 Classification: Classification may be a general method associated with categorization. It's the action or method of clump one thing. MoodSound is design in that way to recognised the emotion of user and recognize mood and generate song on that bases. For recognition of image, MoodSound is using Support Vector Machine (SVM). There area unit 2 phases in getting ready AN economical classifier.

3.5.1Coaching phase: The coaching dataset is split into 3 categories for 3 different moods i.e. happy, neutral, and sad, for every of those corresponding mood we added the values +1 ,+2, and +3 is allotted within the coaching file because the correct label for this model to coach. Here giving feature purpose along side its correct label is at the most necessary because the learning of the model is supervised learning. The feature purpose area unit the position of the landmark on the face that were calculated within the previous step of feature extraction these feature purpose area unit regenerate in a very single dimension from a 2 dimension X,Y arrangement.

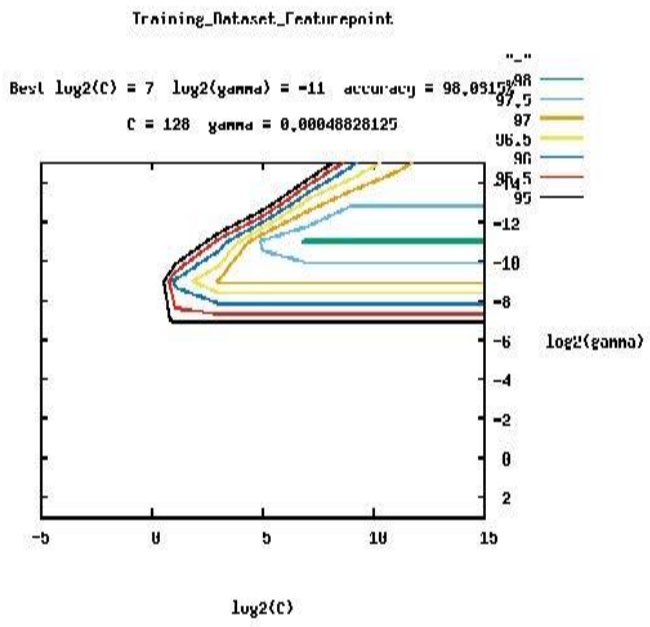

Fig.2: Optimized parameter for C-SVM with RBF kernel

For more accuracy in our emotion based music system we use Dlib and LIBSVM libraries were used.As shown in figure 3 defining accuracy plot of C-SVM for Indian and standard datasets and in figure 4 defining the accuracy plot of Nu-SVM for Indian and standard datasets.

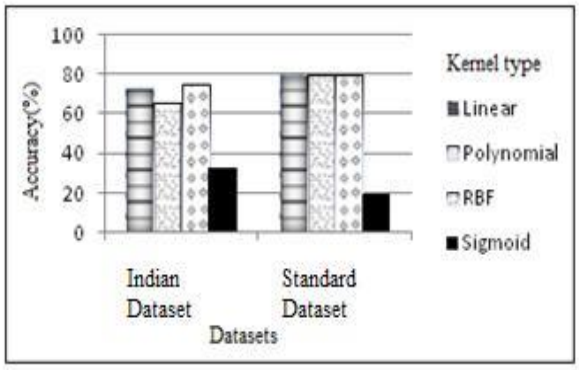

Fig.3: Accuracy plot of C-SVM for Indian and Standard datasets

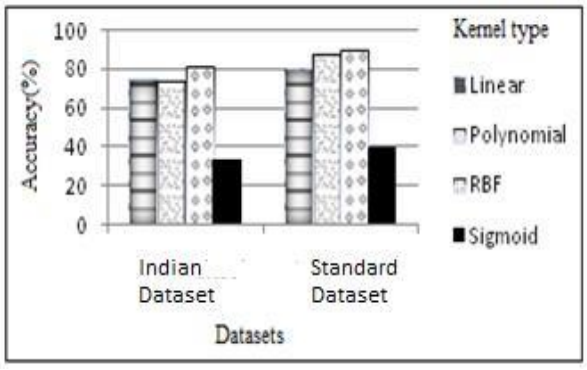

Fig.4: Accuracy plot of Nu-SVM for Indian and Standard datasets

Dlib library is basically used for face and facial features detections. These detected points were classified using SVM classifier which was created with the help of LIBSVM library.

4. RESULTS:

In our model we have not stick on one image for testing, While the code will run it will take around 10 images in a short time(1-2 sec) and for all those images it will compute result and according to the average value of that it will give result. Apart from that we have make two codes one work on single face at a time while another work with multiple faces in the images. As shown in figures 5 MoodSound capture image normally and converted Into grayscale focus on face and generated mood which is happy and after generating music automatically played in system as shown in figures 6.

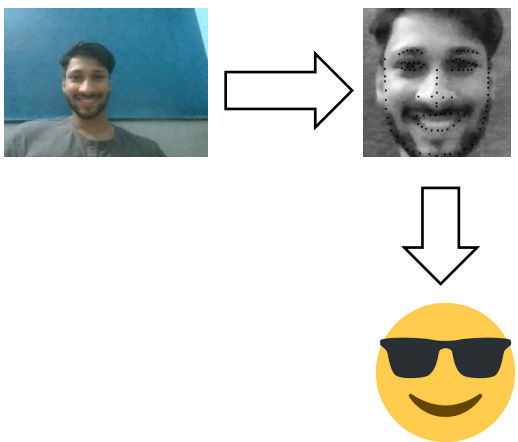

Fig.5: Mood detection phases of MoodSound(user is happy)

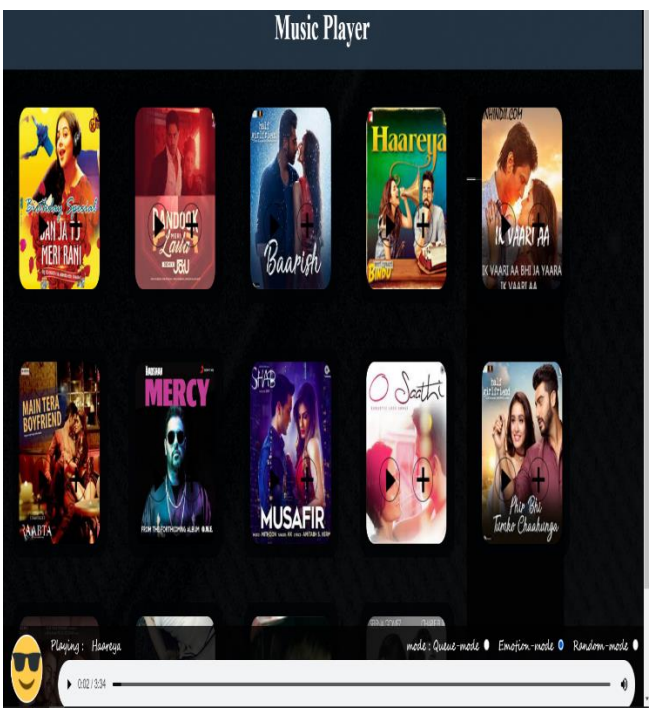

Fig.6: GUI of MoodSound player(user happy)

5. CONCLUSION and Future Scope:

MoodSound: A emotion based music player we have implemented real-time facial emotion recognition using CNN algorithm in OpenCV. Recognized emotion is passed to music player which has sorted music list according to mood. This music list is sorted by using the tempo of songs. By passing this mood, the system plays appropriate songs in that list to improve user's mood.

This system minimizes the efforts of user to select music according to user's current emotional state. It may reduce physical stress and also act as a boon for the music therapy system. MoodSound will help user in enjoying their music much more depending on their current mood and give new experience of listening music's. In future, this system can be enhanced with capability of detecting mood using voice, facial expressions and body postures to get more accuracy.

REFERENCES

[1] V. Patchava, P. Jain,R. Lomte, P. Shakthi, H. B. Kandala, "Sentiment Based Music Play System "

[2] P. Pathak, Prof. S. Aranjo, "Mood based Music Player Karan Mistry".

[3] Viola, P., & Jones, M. J., "Robust Real-Time Face Detection", International Journal of Computer Vision, 57(2), 137-154 (2004).

[4] O. Arriaga, G. Paul, M. Valdenegro, "Real-time Convolutional Neural Networks for Emotion and Gender Classification"

[5] Face detection in OpenCV using Haar cascade

[6] The Echo Nest Developer Center.

[7] Kaggle FER2013 open source dataset for face

[8] A. R. Patel, A. Vollal, P. B. Kadam, S. Yadav, R. M. Samant,"MoodyPlayer: A Mood based Music Player".

[9] C. Jenefer, S. Leena, M. Nirmala, Dr. J. SelvaKumar, "A Survey: Expression Based Music Play