Mood Based Music Player Using Real Time Facial Expression Extraction

[1]Prof.Sumeet Pate, [2]Miss.Shreya Zunjarrao, [3]Miss.Poonam Harane, [4]Mr.Akshay Choudhary
[1]Asst.Professor, [2, 3, 4] UG Student, [1,2,3,4] Department Of Computer Engg.ShivajiraoS.Jondhle College of Engineering & Technology, Asangaon, Maharashtra, India.
[1]*sumeetpate09@gmail.com*,[2]*zshreya24@gmail.com*,[3]*poonamharane1996@gmail.com*,
[4]*akshaychaudhari57@gmail.com*

Abstract-The face is a vital organ of an individual's body and it particularly plays a crucial role in extraction of an individual's behavior and emotion. With the presence of technology, it's helpful for a computer application to analyze human moods from facial video in real time [1]. The manual segregation of list of the songs and generation of an appropriate system based on an individual's emotional features is terribly tedious, time intensive and an upheld task[4]. The projected system supported facial features extracted will generate a system mechanical thereby reducing the hassle and time concerned in rendering the process manually. Facial expressions are captured by using a camera. The exactness of the emotion detection algorithm engaged in the system for real the time pictures is around 85-90%; whereas for static images it's around 98- 100 percent. Thus, it yields higher accuracy in terms of performance and procedure time.

Keywords- Emotion Recognition, Face Detection, Emotion Extraction Module.

I. INTRODUCTION

Facial expressions are of importance and provide meaningful insights about a person's emotions. The next generations of computer vision systems could be majorly based on current computer systems which deploy dynamic interactions with the user. Security, entertainment and human machine interaction (HMI) all use facial emotions as an important prospect. A human can express his/her emotions through their lips and eyes. The work describes a computer application in the form of an "Emotion based Music Player" which provides the users a added ease in accessing and playing their music playlist[1]. Having a large music playlist is a common thing in today's world. Music plays a very vital role in enhancing an individual's life as it is an important source of amusement for music aficionados and sometimes even imparts a therapeutic approach. In today's age, with ever increasing advancements within the field of multimedia content and consumption and, various intuitive music players are developed with various options, nevertheless the user has got to face the necessary work of manually browsing through the list of songs and select songs that compliment and suit his/her current mood. Current mood of individual user is extracted by the proposed system by scanning and interpreting the user's facial expressions. When the emotion is extracted, a playlist of songs matching to the mood will be provided to the user. It aims to provide better experience to the music lovers in listening their music. In the model, a few sample moods are included: Happy, Sad, Angry, And Neutral. The system incorporates image processing and facial detection processes. Images are still the major input , which are further processed to identify the emotion of the user. Images are captured at the launch of the application. Web-Cam is used to capture the images. The previously captured imagewill be saved and passed to the rendering phase.

II. AIM AND OBJECTIVE

a) Aim

1. The aim of the Mood Based Music Player Using Real Time Facial Expression Extractionto provide the user an enjoyable music listening experience by serving them with sensible music content based on his/her current mood.

2. Implementation and analysis of various viable techniques in extracting and classifying the emotions of the user and playing music accordingly.

b) Objective

The main objectives is that the proposed system will use face scanning and feature tracking to determine the user's mood and based on it gives a personalized play list.

III. LITERATURE SURVEY

An emotion based music player is system implemented by many researchers. The Proposed approaches have the centred solely on the number of the essential emotions.

Paper 1: The SMOODI provides sensible mood recommendation by bit and drag, and it's the three completely different views: Mood sq., cowl Flow and Mood Cloud. In the Mood sq., users will check the mood distribution of native clips and generate playlists[2].

Paper 2: In the paper [3], the author provides a real time face expression recognition approach with the help of Real Sense camera. Face detection and landmark detection is performed with the help of SDK of Real Sense camera. The different feature sets are tested for finding an optimal set and classification accuracy is noted for each feature set.

Paper 3: In the paper[5], the author describes making use of machine learning approach for detection of visual object that is capable of processing the images rapidly to obtain high detection rates. The work has been is divided into 3 key contributions which are as follows: First is an Integral image, second consists a learning algorithm that is based on AdaBoost and third is cascade.

IV. EXISTING SYSTEM

The previous music players consists of various features an. The features that are available in the previously developed Music players are as follows:

a) Manual selection of songs.
b) Party Shuffle.
c) Playlist

V. COMPARTIVE STUDY

Sr No.	Paper Title	Author's Name	Problem	Solution	Future work
1.	Geometrical Approaches for Facial Expression Recognition using Support Vector Machines.	Fernandes, J. de A., Matos, L. N., &Aragao, M. G. dos S.	Cannot be performed on different databases.	Perform experiments using different databases and verify the performance of the method in a wider scenario.	Others machine learning algorithms could be evaluated and adequacy of CNN (Convolutional Neural Network).
2.	EMOSIC- An Emotion Based Music Player For Android.	Nathan, K. S., Arun, M., & Kannan, M. S.	Limited number of songs due to less storage space and unable to identify complex and mixed emotions.	Making use of cloud storage.	• Export songs to a cloud database and allow users to download whichever song he wants. • Identify Complex and Mixed Emotions.
3.	Distinctive Image Features from Scale-Invariant Key points.	Lowe, D. G.	Features described use only a monochrome intensity image and not invariant illumination.	Distinctiveness can be derived by including the various illumination-invariant color descriptors.	•Deriving invariant and distinctive image features. •To individually learn the features for recognizing particular objects categories.
4.	A Media Player Which Operates Depending On Human Emotion.	Harshala Chaudhari, Amrapalli Waghmare, Reshma Ganjewar, Dr.Abhijit Bhanubakode.	Is unable to judge only one mood at a time.	Should work on the performance.	The future work is that it can be used in vehicles.

VI. PROBLEM STATEMENT

Problem being solved:

1) The current available systems require the user for doing the manual selection of the songs, but the proposed system will use facial scanning and face feature tracking to determine the user's mood based on it, will provide the user with a personalized playlist, thus making the process effortless for the user.

2) It will provide a better experience to the music connoisseurs and enthusiasts.

VII. PROPOSED SYSTEM

1) Facial Detection:

An efficient object detection method is to use the Haar feature based cascading classifier.It is an algorithm which follows a machine learning approach to increase its efficiency and precision. A varied degree of images are used for training the function. It is then acclimatized for detection of objects in different images. For this, Haartraits are used; Each feature is specially a single value that is obtained by deducting the sum of pixels falling under the

white rectangle from sum of pixels falling under black rectangle. [5]

2) Feature-Point Detection:

The feature points from the image are detected automatically. For face detection, firstly the conversion of RGB format image to binary format image is done. A black pixel is used as a replacement if average value of the pixel is obtained less than 110 orelse a white pixel is used.

3)Emotion Recognition System:

Face location, emotion classification and feature extraction are three parts of the process for mood determination

4) Prediction and Feature Extraction:

In order to predict the current emotion of the user, use of different feature extractorsto effectively locate important key points and distinguishing features from particular face regions.

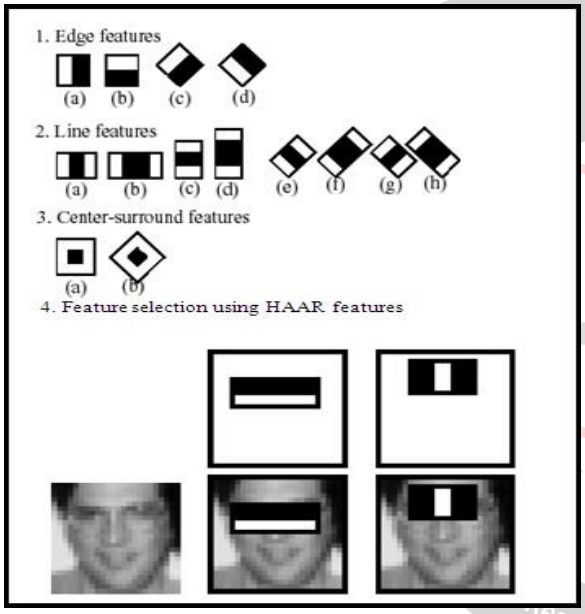

Fig.7.1:Haar features & Feature extraction from face.

VIII. ALGORITHM

The general idea of working of proposed system algorithm is given as follow:

Step.1: START.
Step.2: Capture image using webcam and save.
Step.3: Input image to the application.
Step.4: Face detection.
Step.5: Extract interest points on mouth and eye using HAAR Cascade method of face detection.
Step.6: Apply Bezier curve equation on the mouth and eye.
Step.7: Apply threshold.
Step.8: Device will recognize the emotions and will play music.
Step 9: If not detected then go to Step.2 else go to Step.10.
Step 10: According to emotions songs list will be open and music will be played.
Step 11: STOP

IX. MATHEMATICAL MODEL

Bézier curve is defined with parametric equations:

$\mathbf{f_x(t):=(1-t)^3p_{1x}+3t(1-t)^2p_{2x}+3t^2(1-t)p_{3x}+t^3p_{4x}}$ **and,**

$\mathbf{fy(t):=(1-t)^3p_{1y}+3t(1-t)^2p_{2y}+3t^2(1-t)p_{3y}+t^3p_{4y}}$

Formula for PCA (Principle Component Analysis):

Subtract each vector by Ψ and get *Φ*i, we want to find a projection vector *b* to minimize:

$E[|| bb^T\Phi_i - \Phi_i||^2] \quad E[|| (bb^T - I)\Phi_i||^2] \quad E[((bb^T - I)\Phi_i)^T(bb^T - I)\Phi_i]$

Formula for Eigen Vectors:

Σ □ $E[\Phi\Phi_{ii}^T]$□constant *□□^{T}(□is of *D*-by-*N*)

Fisher Linner Discriminant:

$$\text{inter-class: } |\tilde{m_1} - \tilde{m_2}| = |w^T(m_1 - m_2)| \qquad y: \tilde{m}_1, \tilde{m}_2 : \begin{bmatrix} \\ 1 \end{bmatrix} = \begin{bmatrix} \\ 1 \end{bmatrix} \{w^T x - w^T m_1\}$$

$$\text{intra-class: } \tilde{s}_i^2 = \sum_{y \in Y_i} (y - \tilde{m}_i)^2$$

$$\text{want to maximize: } J(w) = \frac{|\tilde{m}_1 - \tilde{m}_2|_2}{\tilde{s}_1^2 + \tilde{s}_2^2} \qquad x, w, m_1, m_2 : \begin{bmatrix} \\ D \end{bmatrix} \qquad S_B, S_w : \begin{bmatrix} \\ D \end{bmatrix}$$

$$|\tilde{s}^2 = \sum_{y \in Y_i} (w^T x - w^T m_i)(w^T x - w^T m_i)^T = \sum_{x \in X_i} w^T (x - m_i)(x - m_i)^T w = w^T S_i w$$

$$\tilde{s}_1^2 + \tilde{s}_2^2 = w^T S_1 w + w^T S_2 w = w^T S_w w$$

$$|\tilde{m}_1 - \tilde{m}_2|^2 = (w^T m_1 - w^T m_2)^2 = w^T (m_1 - m_2)(m_1 - m_2)^T w = w^T S_B w$$

$$\text{want to maximize: } J(w) = \frac{w^T S_B w}{w^T S_w w}$$

$$S_B w = \lambda S_w w$$

Fisher Multiple Discriminant:

$$S_B = \sum_{i=1}^{c} N_i (m_i - m)(m_i - m)^T \qquad \text{want to maximize: } J(W) = \frac{|W^T S_B W|}{|W^T S_w W|}$$

$$S_w = \sum_{i=1}^{c} \sum_{x \in D_i} (x - m_i)(x - m_i)^T \qquad \text{with } W = [w_1\ w_2\ \ldots\ldots w_m]$$

$$S_B w_i = \lambda_i S_w w_i$$

$$m \le c - 1$$

$$S_B, S_w : \begin{bmatrix} \\ D \end{bmatrix} \quad W : \begin{bmatrix} \\ m \end{bmatrix} \quad W_{PCA} : \begin{bmatrix} \\ N-c \end{bmatrix}$$

Problem: S_w is always singular

Fisherface solution:

$$W_{PCA} = \arg \max_W |W^T S_T W| \text{ where } S_T = \sum_x (x - m)(x - m)^T$$

$$W_{FLD} = \arg \max_W \frac{|W^T W_{PCA}^T S_B W_{PCA} W|}{|W^T W_{PCA}^T S_w W_{PCA} W|}$$

S_T is called the total scatter matrix

X. SYSTEM ARCHITECTURE

Image processing and use of facial detection are the two main corner stones of the proposed system. The input model for the system are still images of the user, which are used for further processing for determining the mood of user. The proposed system will capture an image of respective user at the initialization of the application. The

image scaptured by Web-Campreviously will be saved and further it will then be passed into the rendering phase. The techniques used for face detection are distinguished into 2 groups: holistic, wherever face is treated as a full unit and analytic, wherever co- incidence of characteristic facial components is studied.The expression is measured by a distance between positions of these points within initial image (neutral face) and peak image (affected face).The last part of the system is based on machine learning theory; exactly it's the classification task. The input given for the classifier could be a set of options that were retrieved from face region within the previous stage. Classification needs supervised coaching, so the training set should consist of labeled data.

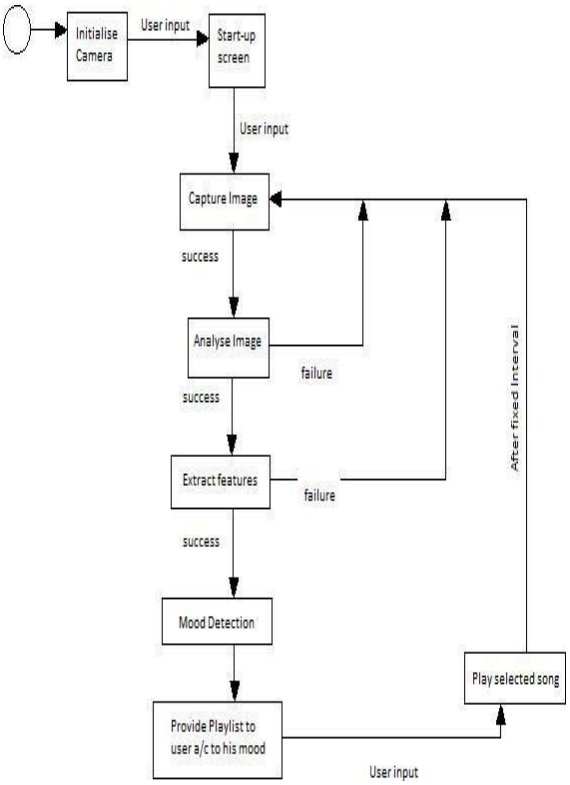

Fig.10.1: System Architecture

System for the emotion recognition is divided into 3 parts: face location determination, emotion classification and feature extraction. The project and system heavily focuses on the latter two parts, the feature extraction and classification. The extracted data is fed to a classifier which determines the emotion of the user and plays the song.

XI. ADVANATGES

1) Ease Of Use.

2) Mobile app can be expensive to build, maintain, and display.

3) Play's Songs as per user's mood.

4) No trouble of selecting the songs manually.

5) Free of cost.

XII .DESIGN DETAILS

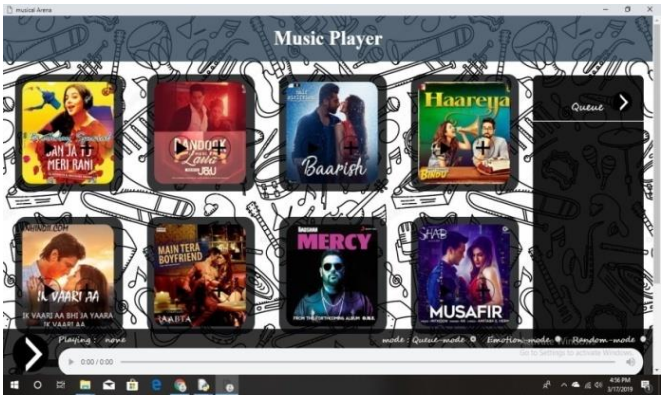

Fig.12.1: Initial Output

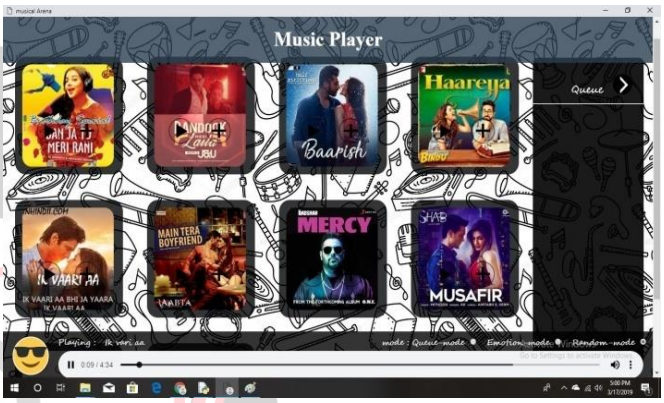

Fig.12.2: Result (user is happy)

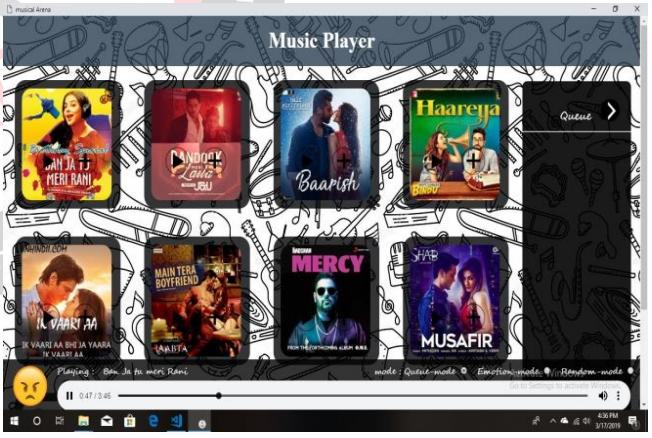

Fig.12.3: Result (user is angry)

XIII. CONCLUSION

We have tried to implement paper on Seangliet,Y.,Lee,B.S.,&Yeo,C.K. "Mood Prediction From Facial Video With Music Therapy On a Smartphone",2016 Wireless Telecommunications Symposium(WTS). The Proposed system is going to process image of facial expressions, recognize the actions with respect to basic emotions and afterwards play music based on emotions. Main strengths of the system are full automation independence. In future the application will be able to

export the songs to a particular cloud database and allow the users to download required songs and will also identify complex and mixed emotions.

REFERENCE

[1] Seangliet, Y., Lee, B.S., & Yeo, C.K. "Mood Prediction From Facial Video With Music Therapy On a Smartphone", 2016 Wireless Telecommunications Symposium(WTS).

[2] Seungjae Lee, Jung Hyun Kim, Sung Min Kim, & Won Young Yoo. (2011). Smoodi: Mood-based music recommendation player. 2011 IEEE International Conference on Multimedia and Expo.

[3] Patil, J. V., & Bailke, P. (2016). Real time facial expression recognition using RealSense camera and ANN. 2016 International Conference on Inventive Computation Technologies (ICICT).

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

[5] Viola,P., &Jones,M.(n.d.). Rapid object Detection using a boosted cascade of simple features. Proceedings of 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.