Journal for Research | Volume 03| Issue 01 | March 2017
ISSN: 2395-7549

Real Time Intelligent Emotional Music Player using Android

Maruthi Raja S K
UG Student
Department of Computer Science & Engineering
Saranathan College of Engineering. Trichy, India

Kumaran V
UG Student
Department of Computer Science & Engineering
Saranathan College of Engineering. Trichy, India

Keerthi Vasan A
UG Student
Department of Computer Science & Engineering
Saranathan College of Engineering. Trichy, India

Kavitha N
Assistant Professor
Department of Computer Science & Engineering
Saranathan College of Engineering. Trichy, India

Abstract

Expressing and recognizing emotional behavior of human play an important role in communication systems. Facial expression analysis is the most expressive way to display human emotion. Three types of facial emotion are recognized and classified: happy, sad, anger. Depending on the emotion, the music player will play the song accordingly which eliminates the time-consuming and tedious task of manually segregating or grouping songs into different lists and help in generating an appropriate playlist based on individual's emotional features. This paper implements an efficient extraction of facial points using Beizer Curves which is more suitable for use in mobile devices. To enhance the audibility level of audio and memory shortage group play concept is also introduced. More than one android mobiles communicate with each other through Peer to Peer connection using Wi-Fi direct. Along with music player editing features like audio trimming and precise voice recording. Later the customized audio file can be set as alarm, ringtone, notification tones. Since all the emotional recognition is done for the real time images, it outperforms well than the any other existing face recognition algorithms or music play applications.

Keywords: Bezier Curves, Face Detection, Facial Expression Recognition, Music Player, Peer to Peer

I. INTRODUCTION

Human Beings have the ability to express and recognize emotions by temperament, disposition, personality and mood. Computer seeks to emulate the human emotions by digital image analysis. The computer vision applications face the issue due to the fact that the real world is three dimension but the computer works in two dimensions [1]. In our day to day life we interact with each other directly either by face to face or indirectly by phone calls. In some professional interaction like call centers, interaction with people is important. With great advancement in technology in terms of different techniques, people moving along with each other should be aware of their present emotions of the person with whom, he or she is interacting. It is widely accepted from psychological theory that human emotions can be classified into six archetypal emotions: surprise, far, disgust, anger, happiness and sadness. Facial motion and the tone of the speech plays a major role in expressing these emotions.

The physique and the tone of the face influences the energy in the utterance of the speech, which can be intentionally modified to communicate different feelings [2-4]. Human beings can easily recognize the changes in these signals along with the information felt by other sensory organs. Based on psychological studies, which show that visual information modifies the perception of speech, is it possible to assume that human emotion perception follows a similar trend [5-7]. This paper analyses the use of digital image to recognize three different human emotions: sadness, happiness and anger. The utilization of the camera present in smart phones has brought about a wide range of unique business applications. Due to the development of smartphone applications and technology, mobile camera has experienced considerable changes in the user interface.

II. OVERVIEW

The main user of our app will be of all age group. The user does not require any technical knowledge in the computer field. The application is provided with a simple user interface for easier access to the system. We have developed an effective application running on smart phones, for playing music based on real-time facial recognition. The proposed framework is so straight forward and proficient that it runs easily on the cell phones. The application is developed using Java language which can be adjusted and maintained easily.

III. EXISTING SYSTEM

The existing system has only the boring traditional type music players in android, in which the users can only simply listen to the songs that are already stored in the playlists and even if the user is the any sad mood or a happy mood all the songs in the playlist will be played by the player in the regular order, which the user may not mostly like to listen to those songs and in that case, he/she has to manually change the song or skip the song.

IV. PROPOSED SYSTEM

This paper is about detecting facial emotion of the user and playing corresponding playlist of song according to the mood of the user. The face recognition algorithm consists of 2 functions: Face Detection and Recognition. The input image is processed by skin color segmentation and it detects the human skin color. After that it detects human face, it separates eyes and lips from the face. Then Bezier curve is drawn for lips and eyes. Then it is compared with the Bezier curve of the eyes and lips that are stored in the database for each mood and finds the nearest Bezier curve from the data store. Based on the result, the respective mood is selected and the songs are played accordingly.

A. Algorithm

The algorithm for facial recognition used is Bezier Curve algorithm. It consists of 2 major functions. The flow of the proposed app is shown in Fig 1.

- Face Detection
- Emotion Recognition

The given image is processed for locating the eyes and lips in the face by using color segmentation process. The color segmentation is done by initialized spatial filtering in the pixels of the image. The segmentation is also affected by lighting of the image. Using feature map the algorithm detects the face location and also detect the location of eyes and lips from the detected face. The Bezier curve is applied on the points obtained from feature map. Then the facial emotion of the user is identified by calculating the variance in the curve value by comparing with the curve value of the in the predefined images that are stored in the database.

B. Skin Color Detection using YCBCR

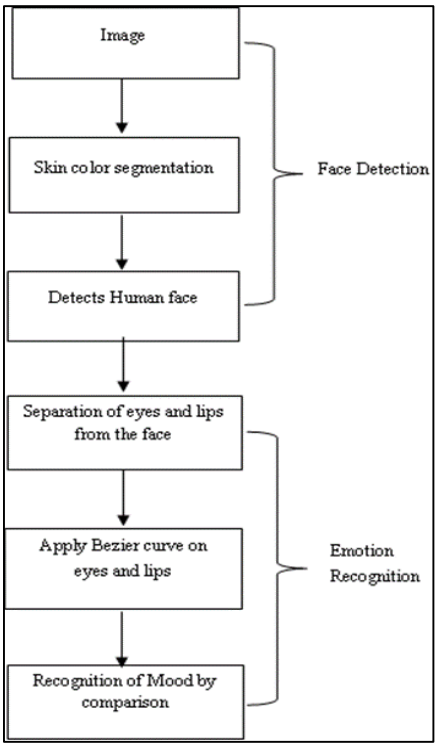

Fig. 1: Architecture of the proposed method

Initially color compensation and color space transformation is done. Using equation (1), luminance information is represented by a single component Y and color information is stored as two color difference component Cb and Cr.

$$\begin{aligned} Y &= 0 + (0.299 * R) + (0.587 * G) + (0.114 * B) \\ Cb &= 128 - (0.168736 * R) + (0.331264 * G) + (0.5 * B) \\ Cr &= 128 + (0.5 * R) - (0.418688 * G) - (0.081312 * B) \end{aligned} \quad (1)$$

After conversion of color model, for accurate face detection illumination calibration is done in pre-processing. Illumination condition equalize the intensity value in an image as per equation (2).

$$Y1 = \left(\frac{y-min_1}{max_1-min_1}\right)(max_2 - min_2) + min_2, \text{ if } (y \leq K_1 \text{ or } K_h \leq y) \quad (2)$$

Where min_1 and max_1 are minimum and maximum value of Y component on input image, min_2 and max_2 are the value of the transformed space, K_1 = 30 and K_h = 220. The values are estimated from training data sets of skin patches in sample database. Histogram equalization is done to enhance the brightness in one direction.

C. Eye and Mouth Detection

The illumination impact is reduced to the brightness component, we are trying to extract the eye and mouth region from an image using Eyemap and Mouthmap [8]. The region of eyes can be found by its symmetry intrinsic feature. Both eyes should be present inside the image to detect the skin region. Eyes are characterized by low red component and high blue one in the CbCr planes using equation (3). Eyemap feature is calculated by

$$\text{Eyemap} = {}^{1}/_{3}\left((\alpha.(C_b)^2 + \beta.(\hat{C_r}) + \left(\frac{C_b}{C_r}\right)\right) \quad (3)$$

$(C_b)^2$, $(\hat{C_r})$ and $\left(\frac{C_b}{C_r}\right)$ are normalized between the range 0 to 255.

$\hat{C_r}$ is the negative of C_r. α is greater than 1 and β is lesser than 1. It effects on variation in red and blue color. It generally has RGB pattern. Mouth is characterized by a high red color component because of lips and little blue, So mouth region has high response in the $(C_b)^2$ because of high red color and low response in the $\left(\frac{C_b}{C_r}\right)$ because of low red color. Mouthmap is constructed by equation (4)

$$\text{Mouthmap} = C_r^2 * \left(C_r - \frac{\alpha.C_r}{255.C_b}\right) \quad (4)$$

D. Bezier Curve on Eyes and Mouth

The Bezier curve generates control points and the curve passing through the first and last control points. These points are then blended to form P(t) shown in equation (5). It describes the path of Bezier polynomial function between P_0 and P_L.

$$P(t) = \sum_{k=0}^{L} P_k \, BEZ_{k,L}(t) \quad (5)$$

Bernstein polynomial $BEZ_{k,L}(t)$ is defined in equation (6) as

$$BEZ_{k,L}(t) = \left(\frac{L}{k}\right) t^k (1-t)^{L-k} \quad (6)$$

The individual Bezier curve coordinates are represented in eqution (7) by following parameters

$$x(t) = \sum_{k=0}^{L} x_k \, BEZ_{k,L}(t) \qquad y(t) = \sum_{k=0}^{L} x_k \, BEZ_{k,L}(t) \quad (7)$$

To apply Bezier curve, the control points of each interest regions are to be extracted. The control points are located in the area of left eye, right eye and mouth. Then, big connect is applied for finding the highest connected area within each interest regions from the eyemap and mouthmap. From this, we find the horizontal and vertical directional pixels to fix the start and end point. Thus, four boundary points of each region is found. Then the Bezier curve for left eye, right eye and mouth are obtained by drawing tangents to a curve passing through the four control points.

E. Recognition of Mood

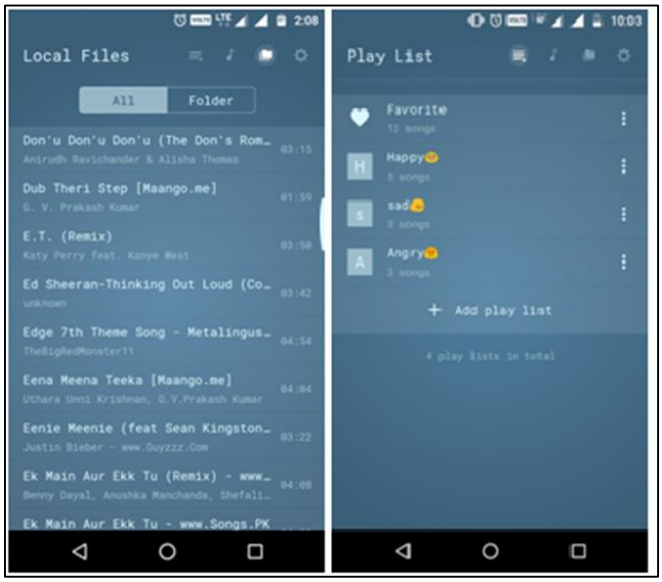

Fig. 2(a): Playlist View Fig. 2(b): Categorizing the songs

Fig. 2: The app view for the playlist

The Bezier curves obtained from the above equation and its coordinates are compared with the curves which are previously sored in the database. After matching the coordinates with the database value, the most likely value of the coordinate is matched. And the corresponding mood for the particular emotion is detected.

After detection of the mood the corresponding playlist of song is played. The playlist should be predefined with the songs of the respective mood.

V. RESULTS

The results of the proposed application are shown here. The Fig 2 shows the playlist of the mobile and the categorization of the songs is done as a preprocessing step. Then a photo is given as input through to the app through the camera device available in the mobile or any image from the gallery can also be selected. The Fig 3 show the happy, sad and angry mood images along with the matching values of the emotions stored in the database. The maximum matched value is taken as the mood of the person in the photograph and the result is displayed accordingly. Along, with the emotional behavior identification, the respective music is also played. The Fig 4 show the song played and user2 trying to connect to the current music playing device through Wi-Fi Direct and gets connected and user2 also listens to the same song as user1. This displays the "group play" concept.

Table - 1
Accuracy Table

Emotional Behaviour	*Accuracy (%)*
Happy	*97.6*
Sad	*92.5*
Angry	*89.9*

The Table 1 shows the accuracy percentage of the various emotions for the proposed app. The accuracy for the happy and sad emotions are better when compared to anger. The anger emotion is sometimes misclassified as happiness due to the exageration of the eyes and lips of the face.

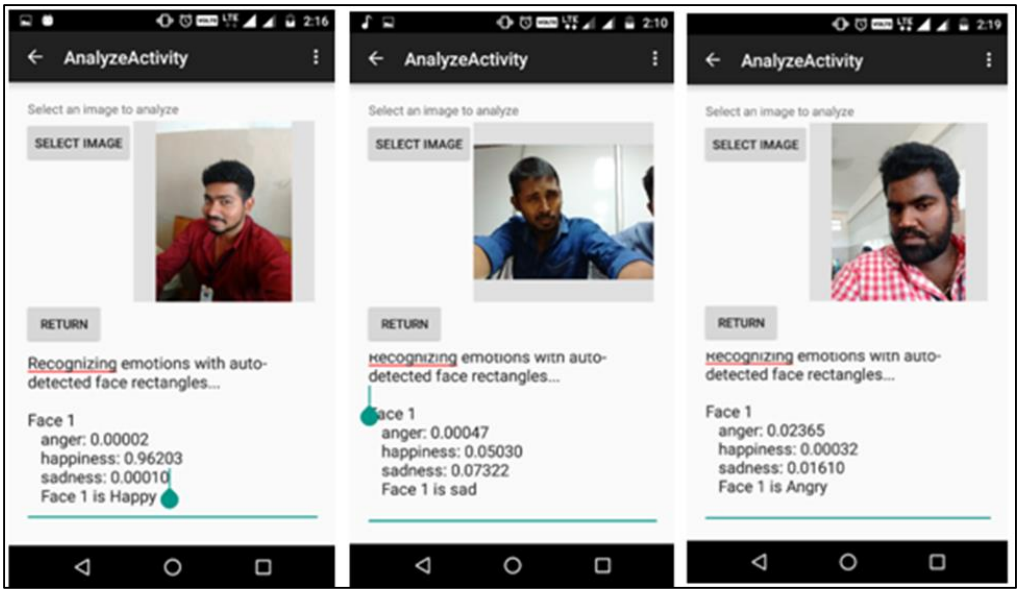

Fig. 3(a): Input for Happy face Fig. 3(b): Input for Sad face Fig. 2(c): Input for Anger face

Fig 3: Visualized results of the happy, sad, angry mood

VI. CONCLUSION

In this article a real time intelligent emotional music player is proposed which establishes a relationship between the playlist and the mood of the person listening to the music. It has been structured around two core researches, each of which investigates the connection between music classification and user's emotion through facial expressions. The music classification is done manually and the input for the app is given through the mobile camera or from the gallery. Since it uses YCbCr color space, a better classification of the skin color is done to segregate the face from the image. Each emotion's Beizer Curve value is stored in the database and the emotional behavior of the given image is analyzed and compared with the same to find the respective behavior of the person in the given image. The result presents the value of the emotion along with the song of the respective mood. Also a group play concept is introduce to rectify the memory shortage of any mobile device. Our result shows that it is nearly accurate to provide songs based on user's emotion with a little refinement in the expression acquired from the listener in the process of evaluating and analyzing the same. These findings, as a whole, suggest that emotional functions of music can well be explored from real life. It is evident that a relationship between human perception of music and emotion with a music features can be drawn to provide a better music system.

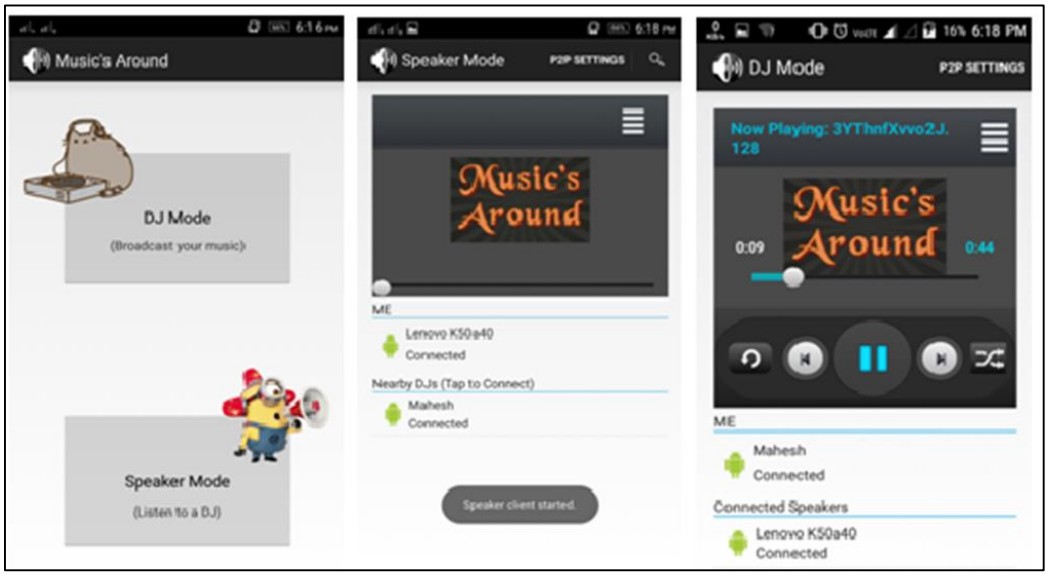

Fig. 4(a): User2 trying to connect Fig. 4(b) User2 connected through Wi-Fi Direct Fig. 4(c): User2 also listening to the same song as User1
Fig. 4: Group Play Concept

REFERENCES

[1] R. Adolphs. Recognizing Emotion From Facial Expressions: Psychological and Neurological Mechanisms. Behavioral and Cognitive Neuroscience Reviews, 2002.
[2] I. Cohen, A. Garg, and T. S. Huang. Emotion Recognition from Facial Expressions using Multilevel HMM.In Proc. of the 2000 NIPS Workshop on Affective Computing, Denver, Colorado, USA, December 2000.
[3] P. Ekman and W. Friesen. Facial Action Coding System (FACS): Investigator's Guide. Conculting Psychologists Press, 1978.
[4] R. C. Gonzalea, R. E.Woods, and S. L. Eddins. Digital Image Processing using Matlab. Prentice Hall, 2004.
[5] G. H. and P. M. Automatic Temporal Segment Detection and Affect Recognition from Face and Body Display. IEEE Transactions on Systems, Man, and Cybernetics - Part B, 39(1):64–84, 2009.
[6] R.-L. Hsu, M. Abdel-Mattaleb, and A. K. Jain. Face Detection in Color Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(5):696–701, 2002.
[7] S. V. Ioannou, A. T. Raouzaiou, V. A. Tzouvaras, T. P. Mailis, K. C. Karpouzis, and S. D. Kollias. Emotion Recognition through Facial Expression Analysis based on a Neurofuzzy Network. Neural Networks, 18(4):423–435, 2005.
[8] Bezier curve. Wikipedia. http://en.wikipedia.org/wiki/Bezier_curve.