An Intelligent Emotion Based Music Player

Aman Ghosh
Computer Science and Engineering
Chandigarh university
Mohali,India
cookie.senpai74@gmail.com

Asad Amanullah
Computer Science and Engineering
Chandigarh university
Mohali,India
asadamanullahahsan@gmail.com

Debdulal Das
Computer Science and Engineering
Chandigarh university
Mohali,India
way2deb007@gmail.com

Harsh Kumar Gupta
Computer Science and Engineering
Chandigarh university
Mohali,India
hkgupta1310@gmail.com

Himanshu Affaria
Computer Science and Engineering
Chandigarh university
Mohali,India
himanshuaffaria7@gmail.com

Gaurav Kumar Bharti
Computer Science and Engineering
Chandigarh university
Mohali,India
gauravkumarbharti7@gmail.com

Abstract: - This paper suggests an emotion-based music player that produces a playlist based on user captured images. A manual playlist and song selection to meet current emotions are more time consuming and uninteresting. Many people will pick songs at random from the playlist. As a result, some songs chosen by the user do not correspond to their present mode. Furthermore, no widely used music player can play songs based on the mood, which includes the emotion like neutral, sad, angry and happy. This system makes extensive use of images processing and facial recognition technology. The suggested model requires .jpeg format as input. Model performance is evaluated by loading 64 still photos (16 for each emotion category) into it. To see how accurate, it is at detecting emotions, the proposed model has an 80% recognition rate based on the testing results. The system is validated by comparing it to the user in user dependent and user independent. It is aimed to provide a better enjoyment to music lovers in music listening.

Keywords: - Facial Recognition, Emotion Detection, Emotion Detection module, Camera, Music, Open CV

I. INTRODUCTION

Emotions are bodily sensations linked to a person's mood, temperament, personality, or character. The classifications of basic emotions developed by Paul Ekman are anger, happiness, sadness, and neutral **[1].**

The "Emotion Based Music Player" is a device that detects an individual's feelings and plays music lists based on that emotion. It requires the individual to represent his emotion through facial expressions. After that, the device will detect the facial expressions of the individual and analyze them then accordingly interpret the emotion displayed at the time by the person. After determining the emotion of the individual, the music player will play the songs that best suit the current emotion of the individual. The device will only focus on the facial expression and analyze it and not on the movement of the head or face or blinking of eyes or sneezing; such movements will not be detected under facial expression to select the songs. The world is changing very fast and with this ever-changing world the technology and human mind are developing side by side which leads to further development in the technology of music player, decades after decades new updates are added to enhance its functioning and perfect it further, for instance, features like fast forward, reverse, variable playback speeds, multicast streams, etc. were

included and are still worked on. Although features satisfy the user's basic requirements at that time but users must do the task of manually browsing through the playlist of songs and selecting the preferred song based on their current mood and behavior. Music plays an important role in enhancing the life of an individual many problems of a person's life are specified by music and it transforms an individual's mood and feelings in no time and works as an important medium of entertainment. Facial expressions give important clues about a person's emotions. felt.

The major goal of this project is to create an "Emotion Based Music Player" for music enthusiasts that will act as a platform for individuals to automatically play and listen to songs based on their emotions. Its goal is to give music fans with a better kind of entertainment. The specific goals are to present a facial expression detection model, reliably recognize the four main emotions (normal, happy, sad, and surprise), and incorporate the music player into the proposed model.

II. LITERATURE REVIEW

Barbara Raskauskas stated in the year 2009 that "Music can fill the silence while also masking the loudness.
Whether or not the song has words, music is enjoyable and communicates to us. I've never met somebody who didn't enjoy music in some way. Even my deaf acquaintance stated she enjoyed music because she could feel the vibrations it produced. Music provides global happiness."[2]
"People enjoy music for much the same reason they love sex, drugs, gambling, and excellent cuisine, according to new research," Emily Sohn said. When people listen to harmony, their brains release dopamine, a type of neurotransmitter produced by the body that is involved in addiction and motivation.[3]

Mary Duenwald produced an article in 2005 that discussed facial expressions.
Around the world, there are generally seven categories:

I. **Sadness**: The eyes droop and the brows' inner corners lift.
II. **Surprise**: The top eyelids and brows are both surprised. The jaw drops open as the shoulders rise.
III. **Anger**: The lower and upper eyelids both swell in anger. Draw together by squeezing in. The jaw clenches, forward, lips crushed against each other.
IV. **Contempt**: On one's face is a look of contempt half of the upper lip on one side of the face upwards tightens
V. **Disgust**: The person's nose wrinkles in disgust. While the top lip rises, the lower lip falls protrudes.
VI. **Fear**: Fear causes the top lids to lift and the eyes to expand. The brows furrow and the lips purse horizontally extend.
VII. **Happiness:** The corners of the lips, lifted and shaped a smile. The cheeks rise up and outside corner of the brows pull down.[4]

Face detection, according to Joseph C. Hager, is utilised as one of the inputs to other image processing activities such as face and emotion recognition.[5]

W.K. Teo, Liyanage C De Silva, and Prahlad Vadakkepat suggested a method of merging feature detection and extraction with facial expression recognition into an integrated system in 2004. Small and precise aspects of the face, such as the eyebrows and lips, can be detected and located using the integration projection.[6]

Jagdish Lal Raheja and Umesh Kumar had introduced Back Propagation Neural Network approach. In order to recognise human face expressions from collected images, the edge, the tasks of detection, thinning, and token detection are completed as the image is being processed.[7]

The exploration of feature-based facial emotion detection within a two-layer perceptron architecture was described by Zhengyou Zhang (1998). The results reveal that face expression identification is primarily a low-frequency operation that may be

carried out with a spatial resolution of 64 pixels X 64 pixels.[8]

Eva Cerezo and her colleagues developed an automated real-time system for emotion recognition that tracked facial features and used a simple emotional classification algorithm. The developed technology was integrated into the Maxine system, which is a 3D virtual scene and character management engine that enhances user engagement.[9]

Facial expression identification using local binary patterns was reported by Caifeng Shan, Shaogang Gong, Peter W. McOwan. Local Binary Patterns features perform reliably and robustly over a relevant range of low-resolution facial picture resolutions, according to the authors' studies. This means that a low-resolution image can be processed and the emotion properly determined.[10]

Daniel T. Bishop, Costas I. Karageorghis, and Georgios Loizou presented their findings on the use of music to manipulate the emotional condition of young tennis players. The study included fourteen young tennis players in all. Participants select to listen to music in order to provoke distinct emotional states, according to the study.[11]"Increasing the pace and intensity of a musical excerpt may raise the amplitude of an effect," Frijda wrote in one of his articles. [12]

Automatic music emotion annotation, according to Wai Ling Cheung and Guojun Lu (2008), is crucial. Hybrid sampling, data-driven detection thresholds, and synonymous correlations between emotional states make computerised music emotion annotation viable and practical.[13]

Proposed System

The proposed system captures the user's facial expressions, extracts face features based on those expressions, and then separates them into unique user moods through fisher face algorithm. Songs that fit the user's feelings are displayed and played after the emotions have been separated.

III. METHODOLOGIES

A. Research Methodologies

Numerous research efforts on the project's history and related works have been made. Available materials that are personally accessible from the Resource Centre for Information (IRC). Sources Books, papers, journals, and thesis reports are all involved. and other internet sources the information gathered includes current image processing progress and applications of face recognition and image processing.

B. Development Methodology

The entire progress of the project has been done in a sequential manner. Where each and all steps are inter-dependent to each other. The entire work is divided into 4 phases.

First Phase involved planning and brainstorming on research undertakings, the steps that would be involved in the project depending upon the need in various phases, and the expected result that need to be delivered. In addition to this, Literature reviews were analyzed and along with-it similar technologies were also studied extensively.

Second Phase started with extensive research analysis. Tasks like Gathering of Data, Data Analysis and Model Analysis were executed. Then various techniques which were instrumental in developing the Face detection were learnt.

Third Phase is the design and development phase. The framework and interface of the entire system was developed. Facial expression or emotions are detected and analyzed and then it was integrated to the music player in order to get an output with respect to the input which is the analyzed facial emotion.

Fourth Phase involved Testing and evaluation of the designed/developed system. Bugs and errors were found and then they were resolved and fixed accurately. In this phase, accuracy of the system was increased by extensively testing the functionality of the system and its process of face detection, emotion analysis and then the
corresponding music playing step.

Phase 1	Planning, discussion, brainstorming on steps to be taken, research and tasks to be performed and result to be delivered. Literature reviews and similar technology were analyzed in a detailed manner.
Phase 2	Research analysis, where tasks like Data Gathering, Data Analysis, and Model Analysis executed. In addition to it, various techniques were analyzed and learnt to detect facial expressions.
Phase 3	This is the design and development phase. Facial Emotion is detected and analyzed. Interface and framework of the system are designed and then it was integrated to the music player.
Phase 4	Testing and evaluation of the developed Music Player system. Bugs and errors are fixed. Accuracy is increased by testing its functionality in face detection for the various emotions.

IV. Results

With the help of this paper, we can detect the emotion of a person through. Their facial Expression and accordingly our paper will play the song according to your mood.

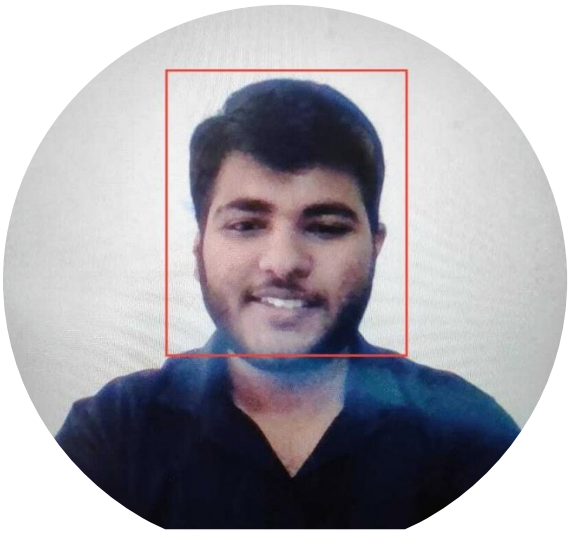


Fig. 3 Face detection using Haar Cascade through web cam

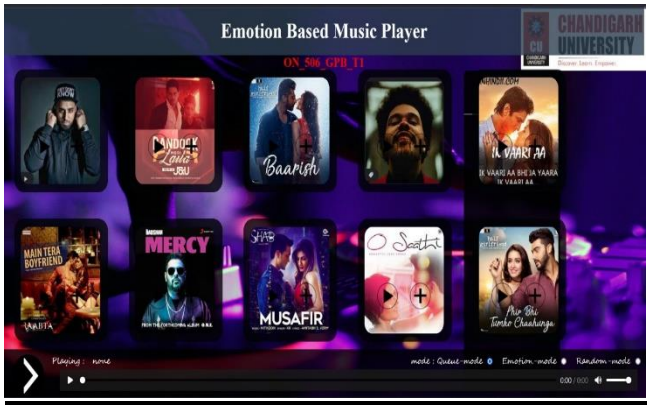


Fig. 4 Proposed Interface of the music player.

Fig. 5 Emotion detected and song has played according to the emotion.

A. Proposed Model

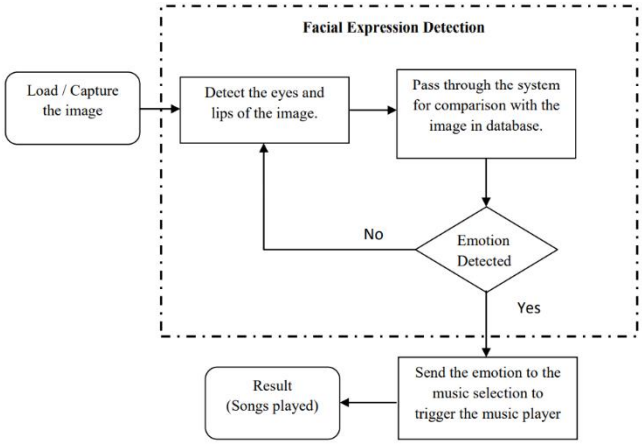


Fig 5: Flow Chart of the Proposed Model

From the developer's perspective, the suggested model for this FYP will focus on two key functions: expression detection and a list of songs performed for each emotion category. The system is capable of detecting expressions. Developed primarily to recognize four major expressions: Happy, Sad, Angry, and Neutral. Normal and unexpected on the other hand, music will be available in each location/category. The device will play the eight songs after detecting the individual's emotion through the music player.

B. System Function

Table 1: Function of the buttons.

1.	"Play Song" button	Enable the user to play the song from display tray.
2.	"Add to Queue" button	Enable the user to add songs to the queue from the display tray.
3.	"Next" button	Enable the user to play next song from the queue
4.	"Play list" Button	Enable the user to play and pause the currently running song.
5.	Volume button	Mutes and unmutes the song playing in the player.
6.	Volume slider	Enables the user to increase and reduce the volume of the song
7.	Track slider	Enables the user to navigate through the song freely
8.	Download button	Allows the user to download the currently playing song to their system.
9.	Playback speed button	Enables the user to increase or decrease the speed of the song
10	Queue-mode	Plays the song in user selected playlist
11	Emotion-mode	Plays the songs according to users' mood by detecting user's expressions
12	Random mode	Plays songs randomly from the directory.

The proposed model is designed to recognise the user's emotion using the methods, First and foremost, when an image is loaded, a conversion is performed in which the colourful image is converted to a grayscale image. The model will then be given a cropped face height image, which will be separated into four pieces. The eyes are separated into two halves (left and right). Following the extraction of the features, the thinning process will be carried out until the eye and lip boundaries are obtained. Finally, the boundary is turned to a number and compared to the data in the dataset to identify the data's mood.

C. System Testing Results

Once the system development is complete, functional testing is performed. The goal is to see if the model is functional and use-able. The findings are presented in the table below.

Table 2: System functional testing results

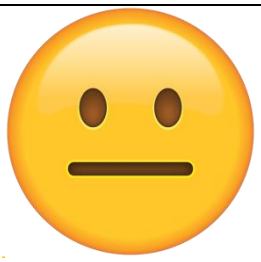
Component	Expected Function	Testing Result	
		Positive	Negative
Play Song	Play the song from tray.	✓	
Add to Queue	Play next song from the queue.	✓	
Next	Play next song from the queue.	✓	
Play list	Play and pause the running song.	✓	
Volume	Mutes and unmute the song playing	✓	
Volume slider	Increase and decrease the volume	✓	
Track slider	Navigate through the song	✓	
Download	Download the currently playing song	✓	

Playback speed	Increase or decrease the speed of the song.	✓	
Queue-mode	Plays the song in user selected playlist.	✓	
Emotion-mode	Plays the songs according to users' mood	✓	
Random-mode	Plays songs randomly from the directory.	✓	

D. Emotion Accuracy Testing Results

The model saves a set of photos for each category (normal, sad, surprise, and joyful) for comparison purposes. In order to recognise emotion, the model will compare the loaded image to the dataset. The saved dataset is shown in the table below.

Table 3: The dataset of images in proposed model

Images	Emotion
	Angry
	Happy
	Sad
	Neutral
	Error

The suggested model is put to the test with a series of similar emotion photos to see how accurate it is at detecting the emotion. Ten photographs were examined for each emotion category, and the results are presented in the tables below.

Table 4: Testing result for "Angry" Expression

Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative

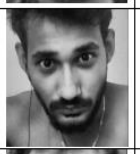
	✓			✓	
	✓			✓	
	✓			✓	
		✓		✓	
	✓				✓

Table5: Testing result for "Happy" Expression

Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
		✓			✓
	✓			✓	
	✓			✓	
		✓		✓	
	✓				✓

Table 6: Testing result for "Sad" Expression

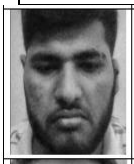
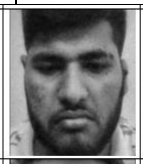
Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative
	✓			✓	
	✓			✓	
	✓			✓	
	✓			✓	
		✓		✓	

Table 7: Testing result for "Neutral" Expression

Sample	Testing Result		Sample	Testing Result	
	Positive	Negative		Positive	Negative

		✓		✓	
	✓			✓	
	✓			✓	
	✓			✓	
	✓				✓

Summary of Result are shown in the table below:

Table 8: Summary of the results tested

Emotion	No. of Samples	No. of Recognized Sample	RR
Angry	10	8	80%
Happy	10	7	70%
Sad	10	9	90%
Neutral	10	8	80%
Total	40	32	80%

The following formula is used to determine the RR (Recognition Rate):

$$RR = \frac{Classified\ Character}{Total\ Number\ of\ Character} \times 100\%$$

RR= (32/40) ×100

RR= 80%

According to the given results, the suggested model has an 80 percent recognition rate (RR).

Based on observations made during the prototype evaluation period, there are various restrictions that prohibit the suggested model from performing accurately. Posture of the object in the image. If the image's object is upright and the individual's face is clearly visible, the proposed model will work correctly. A slight slant in an individual's head is acceptable and visible.

Figure 4: Images which are accurately detected

The photos below are examples of images that the suggested model fails to detect.

Figure 5: Image which cannot be detected accurately

V. Conclusion

A. Conclusion

The emotion recognition of the photographs loaded into the suggested model is the most important aspect of this study. The main focus is on the emotion detecting feature. The proposed approach aims to improve an individual's entertainment through the integration of emotion recognition technology with a music player. The proposal can recognize the four emotions of the photos placed into it: normal, happy, angry and sad. After the proposed model has compared and detected the user's emotion, the music player will play the appropriate song(s).
In terms of usefulness and accuracy, the suggested model has undergone both system and emotion accuracy testing, with positive results. The proposed model correctly identified 32 of the 40 photos that were imported. The Recognition Rate was 80%. Furthermore, the proposed model is a computer programmer that may run on a variety of operating systems and machines.

Users may now choose music in a more engaging and straightforward way with this Emotion Based Music Player. Music fans will no longer have to sift through a huge list of tracks to choose the right song to play, but rather will be able to match the mood in the song selection.

B. Future Work

Every application or system, including the Emotion Based Music Player, is subject to changes and improvements. In terms of performance and features, the proposed model can yet be improved. First and foremost, the limitation in detection can be reduced. As previously stated, emotion detection has various drawbacks. When extracting facial features for comparison, increasing the number of facial features can help enhance accuracy.
For contrast, the proposal removes only the lips and eyes at the moment. In the future, other face traits like brows and cheek movement could be incorporated in the comparison.
Furthermore, the future model can be improved by removing or reducing noise in loaded or taken photos. In the future, noise reduction software can be included in the model to eliminate noise from either still or recorded images, increasing the accuracy of Emotion identification.
Apart from the foregoing, the proposed model can be improved by incorporating automatic image resolution, brightness, and contrast adjustments.
The quality of the photographs loaded has a significant impact on the accuracy of emotion detection in the current application.
As a result of the auto adjustment, the user can load any image quality or capture images with any webcam. The future model will have the ability to modify the image quality that can be identified and processed.
Furthermore, for improved interactivity between Real-time emotion detection technique can be applied to the model depending on the user and application. Once the app is activated, the future model will identify and extract facial features, allowing the emotion to be identified in real time.

REFERENCES

[1]. Ekman, P. & Friesen, W. V. (1969): The repertoire of nonverbal behavior: Categories, origin, usage, and coding: Semiotica, Vol. 1: 49-98.

[2]. Barbara Raskauskas (2009). Why Do People Listen to Music? Retrieved on October 9 2012 from http://voices.yahoo.com/whypeople-listen-music2608185.html

[3]. Emily Sohn (2011) Why Music Makes You Happy? Retrieved on October 9 2012 from http://news.discovery.com/human/musicdopamine-happiness-brain-110110.html

[4]. Mary Duenwald (2005). The Physiology of Facial Expressions. Retrieved on October 9 2012 from

http://discovermagazine.com/2005/jan/physiology-of-facial-expressions

[5]. Joseph C. Hager (2003). *et.al:* Introduction to the data face site: Facial expressions, emotion expressions, nonverbal communication, physiognomy. Retrieved 10 October 2012, from http://face-and-emotion.com/index.html.

[6]. W.K Teo, Liyanage C De Silva and Prahlad Vedakkepat. (2004). *et.al*: Facial expression detection and recognition system: Journal of The Institution of Engineers, Singapore. Vol 44.

[7]. Jagdish Lal Raheja, Umesh Kumar. *et.al* (2010): Human facial expression detection from detected in captured image using back propagation neural network: International Journal of Computer Science & Information Technology. (IJCSIT). Vol.2(1).

[8]. Zhengyou Zhang. (1998). *et.al:* Feature-based facial expression recognition: Sensitivity analysis and experiments with a multi-layer perceptron: Journal of Pattern Recognition and Artificial Intelligence. Vol. 13(6):893-991.

[9]. Eva Cerezo1, Isabelle Hupont, Critina Manresa, Javier Varona, Sandra Baldassarri, Francisco J. Perales, and Francisco J. Seron. (2007): Real-Time Facial Expression Recognition for Natural Interaction: In J. Martí *et al.* (Eds.): IbPRIA 2007, Part II, LNCS 4478, (pp. 40-47). Springer-Verlag Berlin Heidelberg.

[10]. C. Shan, S. Gong, and P.W. McOwan. *et.al," Facial* expression recognition based on local binary patterns: A comprehensive study, "Image vis. Compute., vol. 27, no. 6, pp. 803-816, May 2009.

[11]. Frijda, N.H. (1986). *et.al*: The emotions. New York: Cambridge University Press.

[12]. Daniel T. Bishop, Costas I. Karageorghis, and Georgios Loizou. (2007): A Grounded Theory of Young Tennis Player' Use of Music to Manupulate Emotional State: Journal of Sport & Exercise Psychology: 584-607.

[13]. Wai Ling Cheung and Guojun Lu. (2008): Music Emotion Annotation by Machine Learning.doi: 10.1109/MMSP.2008.4665144.