

## FACE DETECTION & EMOTION ANALYZATION

**Neha Dubey<sup>1</sup>, Suresh Jangam<sup>2</sup>, Shaik Saber Ali<sup>3</sup>,**

**Dr. B. Raveendranadh Singh<sup>4</sup>, Mahesh Akuthota<sup>5</sup>**

*<sup>1, 2, 3</sup> B.Tech (CSE), <sup>4</sup>Working as Professor & Principal, <sup>5</sup>Working as Asst. Professor*

*Department of (CSE) Visvesvaraya College of Engineering and Technology, M.P Patelguda,*

*Ibrahimpattanam (M), Ranga Reddy (D), Affiliated to JNTUH, (India)*

### ABSTRACT

*This venture proposes an Emotion based music player utilizing feeling investigation. That is, it extricates the client's demeanours and recognizes client's feeling. The music player will then play the melodies as per the classification of feeling identified. This framework is intended to give a superior satisfaction to music beaus in music tuning in. The working of our Proposed framework is: Scan client's expression for sound documents, Classify sound records utilizing sound element extraction module, Segregating sound documents as indicated by various state of mind, Capture continuous photograph of the client utilizing optional camera, Face identifying utilizing calculations like viola-Jones calculation[1], Expression acknowledgment utilizing calculations or existing API's, Listing out the playlist tunes of a recognized feeling, Playing an arbitrary sound from the playlist*

*The human face is a vital part of a human body and it particularly assumes an essential part to discover an individual's conduct and passionate state. Physically Sort Out the rundown of melodies and creating a proper playlist in light of an individual's passionate elements is an extremely dull, tedious, work serious and maintained assignment. Grouped calculations have been progress and produced for computerizing the playlist era process. However the Suggested existing calculations being used are computationally moderate, less impeccable and now and then even require utilization of extra equipment like EEG [2] or sensors. This proposed framework taking into account outward appearance established will produce a playlist naturally subsequently lessening the vitality and time included in rendering the procedure physically. Subsequently the proposed framework decreases the computational time included in acquiring the outcomes and the general expense of the outlined framework, in this way builds the exactness of the framework. Outward appearances are recorded utilizing an inbuilt camera. To get the precision of the feeling, location calculation utilized as a part of the framework for constant pictures is close-by 85-90%, while for static pictures it is around 98-100%. The proposed calculation on a normal computed estimation takes around 0.95-1.05 sec to deliver a feeling based music playlist. In this manner, it delivers better exactness as far as execution and computational time and lessens the planning cost, contrasted with the calculations utilized as a part of the writing study.*

**Key Word: Emotion Extraction Module, Music Information Retrieval, Artificial Neural Networks, Viola and Jones Face Detection, Audio Emotion Recognition [3], Audio Feature Extraction Module, Confusion Matrix [4]**

## I. INTRODUCTION

Music expect a basic part in updating an individual's life as it is an essential medium of incitement for music critical others and gathering of people individuals and on occasion even gives an accommodating philosophy. In today's world, with always extending degrees of progress in the field of sight and sound and development, distinctive music players have been delivered with parts, for example, brisk forward, inverse, variable playback speed (search for and time compression), local playback, spouting playback with multicast streams. Regardless of the way that these components satisfy the customer's crucial necessities, yet the customer needs to stand up to the errand of physically looking over the playlist of tunes and select tunes considering his present demeanour and behaviour. The presentation of Audio Emotion Recognition (AER)[5] and Music Information Retrieval (MIR)[6] in the traditional music players gave subsequently parsing the playlist in light of various classes of sentiments and demeanours.

AER is a strategy which oversees organizing a got sound sign, by considering its various sound segments into various classes of emotions and psyche sets, while MIR is a field that thinks some essential information from a sound sign by exploring some solid components like pitch, vitality, MFCC, flux and so on. Despite the fact that both AER and MIR fused the limits of avoiding manual confinement of tunes and period of playlist, yet it can't solidify totally a human feeling controlled music player. Yet human talk and movement are an ordinary strategy for imparting emotions, yet outward appearance is the most outdated and typical technique for conveying suppositions, sentiments and manner.

The guideline focus of this paper is to diagram a gainful and careful computation that would make a playlist considering current energetic state and lead of the customer. The count laid out requires less memory overheads, less computational and planning time, diminishing the cost of any additional gear like EEG or sensors. The outward appearance would arrange into 5 unmistakable sorts of outward appearances like inconvenience, rapture, stun, horrid, and intensity. A high exact sound extraction methodology is proposed that focuses huge, fundamental and huge information from a sound sign considering certain sound components in a much lesser time. An inclination model is recommended that requests a tune in light of any of the 7 classes of emotions viz unfortunate, joy offend, rapture stun, fulfilment vitality[7], joy, shock, and bleak disdain. The inclination extraction module and sound part extraction module is joined using an Emotion-Audio mix module. The proposed instrument fulfils an unrivalled capability and consistent execution than the present systems.

This paper is formed into: Section 2 gives the brief examination of composing survey. Territory 3 elucidates the methodology; Section 4 gives the test examination and results. Region 5 gives the completion of the paper and future work.

## II. LITERATURE SURVEY

Various frameworks and strategies have been proposed and made to portray human excited state of behaviour. The proposed approaches have focused quite recently on the part of the central emotions. With the final objective of highlight affirmation, facial segments have been arranged into two essential orders, for instance, Appearance-based component extraction and Geometric based component extraction by zheng. Geometric based component extraction technique considered only the shape or major discernible purposes of some basic facial

components, for instance, mouth and eyes. The structure proposed by Changbo. Around a total of 58 foremost noteworthy point centres was considered in showing up based extraction highlight like arrangement, have in like manner been considered in different scopes of work and headway. A gainful strategy for coding and executing removed facial segments together with multi-presentation and multi-determination set of Gabor channels was proposed by Michael Lyons.

Careful and beneficial truthful based procedure for looking at removed outward appearance components was proposed by Renuka R. Londhe. The paper was altogether revolved around the examination of the changes in rhythmic movements on the face and intensities of relating pixels of pictures. Fake Neural Networks (ANN)[8] was used as a part of the request isolated segments into 6 imperative comprehensive emotions like anger, detesting, fear, happy, disastrous, and stun. A Scaled Conjugate Gradient back-inducing count in association with two-layered sustenance forward neural framework was used and was productive as a part of getting a 92.2 % affirmation rate. In order to reduce the human effort and time required for manual segregation of tunes from a playlist, in association with different classes of sentiments and brain sets, diverse techniques have been proposed.

Thayer proposed an outstandingly accommodating 2-dimenesional (Stress v/s essentialness) model plotted on two tomahawks with sentiments outlined by a 2dimensional co-ordinate system, lying on either 2 tomahawks or the 4 quadrants moulded by the 2-dimensional plot. The music mind-set marks and A-V values from a total 20 subjects were attempted and dismembered in Jung Hyun Kim's work, and in light of the results got from the examination, the A-V plane was assembled into 8 regions(clusters), depicting demeanour by data mining capable k-infers bundling estimation.

Various procedures have been expected to think facial segments and sound components from a sound sign and not a lot of the systems formed have the ability to deliver an inclination based music playlist using human emotions and the present arrangements of the structures can make a robotized playlist using an additional hardware like Sensors or EEG systems in this way extending the cost of the diagram proposed. The rates of the disservices of the present system are according to the accompanying

Existing structures are particularly mind boggling similarly as time and memory necessities for isolating facial components persistently.

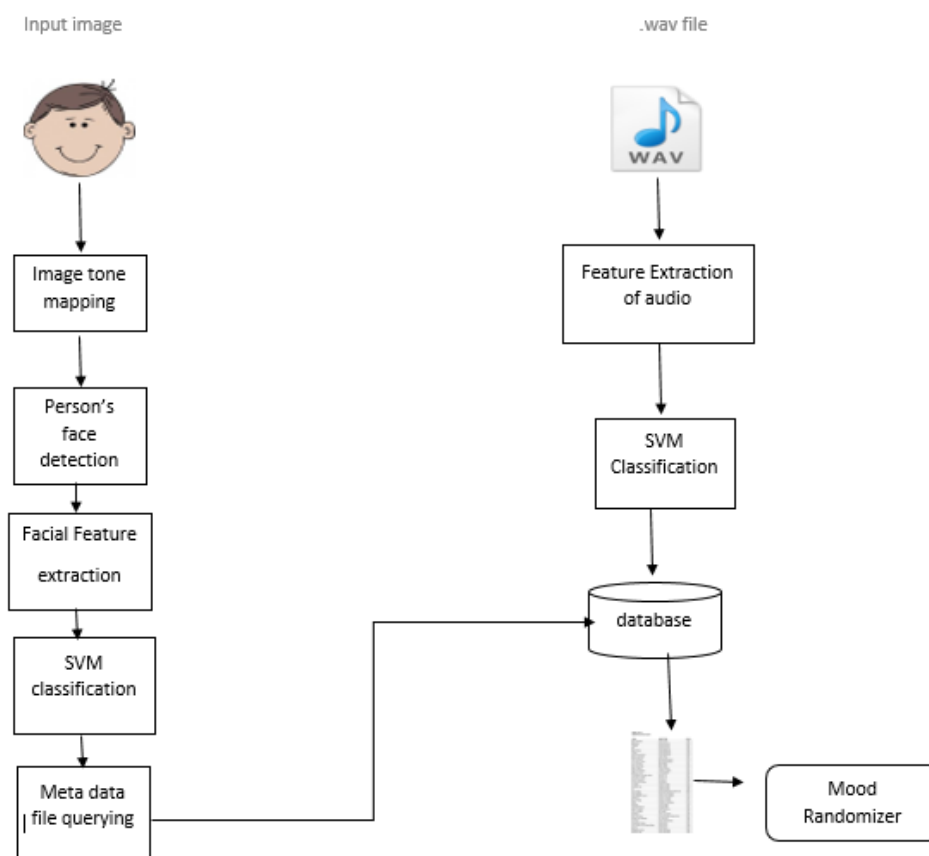
In perspective of the current energetic state and direct of a customer, existing systems have a lesser exactness in time of a playlist.

Some current structures have a tendency to use the usage of human talk or once in a while even the use of additional gear for time of an automated playlist, in like manner extending the total cost achieved.

This paper fundamentally focuses and spotlights on deciding the disservices incorporated into the present system by arranging a robotized feeling based music player for the time of revamp playlist in perspective of customer uprooted facial parts and thusly keeping up a vital separation from the work of any additional hardware. It also fuses a perspective randomized and goody work that moves the slant delivered playlist to another same level of randomized demeanour made playlist after some range.

### III. METHODOLOGY

The proposed algorithm includes an emotion music recommendation system that gives the generation of a customized playlist in understanding to the client's emotional state. The proposed system includes three noteworthy modules: Emotion extraction module, Audio highlight extraction module and an Emotion-Audio acknowledgment module. Feeling extraction module and Audio highlight extraction module are two separate modules and Emotion-Audio acknowledgment module performs the mapping of modules by questioning the sound meta-information record. Fig 1 shows square graph of proposed framework



#### 3.1 Emotion Extraction Module

Image of a user is captured using a webcam or it can be accessed from the stored image in the hard disk. This acquired image undergoes image enhancement in the form of tone mapping in order to restore the original contrast of the image. After image enhancement all images are converted into binary image format and the face is detected using Viola and Jones algorithm where the \_Frontal Cart property ' of the algorithm is used that only detects upright and face forwarding features with a maximum threshold value set in the range of 16-20. The output of Viola and Jones Face detection block forms an input to the facial feature extraction block.

To increase the accuracy and an aim to obtain real time performance only features of eyes and mouth are appropriate enough to depict the emotions accurately. For extracting the features of mouth and eyes certain calculations and measurements are taken into consideration. Equations (1), (2), (3) and (4) illustrate the bounding box calculations for extracting features of a mouth.

$$X(\text{start pt of mouth}) = X(\text{mid pt of nose}) - (X(\text{end pt of nose}) - (X_{\text{startpt of nose}})) \quad (1)$$

$$X(\text{end pt of mouth}) = X(\text{mid pt of nose}) + ((X_{\text{endpt of nose}}) - (X_{\text{startpt of nose}})) \quad (2)$$

$$Y(\text{start pt of mouth}) = Y(\text{mid pt of nose}) + 15 \quad (3)$$

$$Y(\text{end pt of mouth}) = Y(\text{start pt of mouth}) + 103 \quad (4)$$

Where  $(X(\text{start pt of mouth}), Y(\text{start pt of mouth}))$  and  $(X(\text{end pt of mouth}), Y(\text{end pt of mouth}))$  illustrates start and end points of the bounding box for mouth respectively,  $(X(\text{mid pt of nose}), Y(\text{mid pt of nose}))$  illustrates midpoint of nose and  $((X_{\text{endpt of nose}}), (X_{\text{startpt of nose}}))$  illustrates end and start point of nose. Classification is performed using Support Vector Machine (SVM) which classifies it into 7 classes of emotions.

### 3.2 Audio Feature Extraction Module

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 obtained from the Internet 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.

Audio signals are categorized into 8 types viz. sad, joy-anger, joy-surprise, joy-excitement, joy, anger, sad-anger and others.

1. Songs that resemble cheerfulness, energetic and playfulness are classified under joy.
2. Songs that resemble very depressing are classified under the sad.
3. Songs that reflect mere attitude, revenge are classified under anger.
4. Songs with anger in playful is classified under Joy-anger category.
5. Songs with very depress mode and anger mood are classified under Sad-Anger category.
6. Songs which reflect excitement of joy is classified under Joy-Excitement category.
7. Songs which reflect surprise of joy is classified under Joy-surprise category.
8. All other songs fall under ‘\_others’ category.

### 3.3 Emotion-Audio Integration Module

Emotions extracted for the songs are stored as a meta-data in the database. Mapping is performed by querying the meta-data database. The emotion extraction module and audio feature extraction module is finally mapped and combined using an Emotion-Audio integration module. Fig 2 illustrates mapping of Facial features and Audio features. For example, if an input facial image is categorized under joy, the system will display songs under joy, joy-anger, Joy-Excitement, Joy-surprise category.

## IV. CONCLUSION

Trial results have shown that the time required for sound segment extraction is inconsequential (around 0.0006 sec) and tunes are secured pre-given the total estimation time of the proposed structure is relating to the time required for extraction of facial components (around 0.9994 sec). Also the distinctive classes of feeling yield a

predominant precision rate when diverged from past existing systems. The computational time taken is 1.000sec which is less thusly helping in fulfilling an unrivalled consistent execution and profitability.

The system thusly goes for giving the Windows working structure customers with a less costly, additional hardware free and correct feeling based music system. The Emotion Based Music System will be of magnificent purpose of inclination to customers hunting down music in perspective of their demeanour and eager behaviour. It will diminish the filtering time for music along these lines lessening the pointless computational time and in this way extending the general accuracy and capability of the system. The system won't simply diminish physical extend yet will moreover go about as a safe house for the music treatment structures and may in like manner help the music guide to treatments a patient. In like manner with its additional parts determined above, it will be a completed structure for music sweethearts and group of on lookers individuals.

The future degree in the structure would to diagram a framework that would be helpful in music treatment and give the music pro the offer anticipated that would treat the patients some help with torment from disarranges like mental uneasiness, strain, exceptional despairing and damage. The proposed system furthermore tends to avoid in future the offbeat results conveyed in astonishing repulsive light conditions and uncommonly poor camera determination.

## REFERENCES

- [1] Chang, C. Hu, R. Feris, and M. Turk, —Manifold based analysis of facial expression,|| Image Vision Comput ,IEEE Trans. Pattern Anal. Mach. Intell. vol. 24, pp. 05–614, June 2006. International Journal of Engineering Research and General Science Volume 3, Issue 1, January-February, 2015 ISSN 2091-2730 756 [www.ijergs.org](http://www.ijergs.org)
- [2] A. habibzad, ninavin, Mir kamalMirnia,|| A new algorithm to classify face emotions through eye and lip feature by using particle swarm optimization.||
- [3] Byeong-jun Han, Seungmin Rho, Roger B. Dannenberg and Eenjun Hwang, —SMERS: music emotion recognition using support vector regression|| , 10thISMIR , 2009.
- [4] Alvin I. Goldmana, b.Chandra and SekharSripadab, —Simulationist models of face-based emotion recognition|| .
- [5] Carlos A. Cervantes and Kai-Tai Song , —Embedded Design of an Emotion-Aware Music Player|| , IEEE International Conference on Systems, Man, and Cybernetics, pp 2528-2533 ,2013.
- [6] FatmaGuney, || Emotion Recognition using Face Images|| , Bogazici University, Istanbul, Turkey 34342.
- [7] Jia-Jun Wong, Siu-Yeung Cho,|| Facial emotion recognition by adaptive processing of tree structures|| .
- [8] K.Hevener,|| The affective Character of the major and minor modes in music|| , The American Journal of Psychology, Vol 47(1) pp 103-118,1935.
- [9] Samuel Strupp, Norbert Schmitz, and KarstenBerns, —Visual-Based Emotion Detection for Natural Man-Machine Interaction|| .
- [10] Michael Lyon and Shigeru Akamatsu, —Coding Facial expression with Gabor wavelets.|| , IEEE conf. on Automatic face and gesture recognition, March 2000

## AUTHOR DETAILS

### NEHA DUBEY

Pursuing B.Tech (CSE) from Visvesvaraya College of Engineering and Technology, M.P Patelguda, Ibrahimpatnam (M), Ranga Reddy (D), Affiliated to JNTUH, India.

### SURESH JANGAM

Pursuing B.Tech (CSE) from Visvesvaraya College of Engineering and Technology, M.P Patelguda, Ibrahimpatnam (M), Ranga Reddy (D), Affiliated to JNTUH, India.

### SHAIK SABER ALI

Pursuing B.Tech (CSE) from Visvesvaraya College of Engineering and Technology, M.P Patelguda, Ibrahimpatnam (M), Ranga Reddy (D), Affiliated to JNTUH, India.

### ASST. PROF MAHESH AKUTHOTA

Working as Asst. Professor, Department of (CSE) from Visvesvaraya College of Engineering and Technology, M.P Patelguda, Ibrahimpatnam (M), Ranga Reddy (D), Affiliated to JNTUH, India.

### SRI. DR.BHALUDRARA VEENDRANADHSINGH

M.Tech, Ph.D. (CSE), MISTE, MIEEE (USA), MCSI

Professor & Principal. He obtained M.Tech, Ph.D(CSE)., is a young, decent, dynamic Renowned Educationist and Eminent Academician, has overall 23 years of teaching experience in different capacities. He is a life member of CSI, ISTE and also a member of IEEE (USA). For his credit he has more than 50 Research papers published in Inter National and National Journals.He has developed a passion towards building up of young Engineering Scholars and guided more than 300 Scholars at Under Graduate Level and Post Graduate Level.His meticulous planning and sound understanding of administrative issues made him a successful person.