International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 01 | Jan 2020 www.irjet.net p-ISSN: 2395-0072

Musical Therapy using Facial Expressions

Pranjul Agrahari[1], Ayush Singh Tanwar[2], Biju Das[3], Prof. Pankaj Kunekar[4]

[1,2,3]Student, Department of Information and Technology, Atharva College of Engineering, Mumbai, Maharashtra, India.
[4]Professor, Department of Information and Technology, Atharva College of Engineering, Mumbai, Maharashtra, India.

Abstract – *Most people agree that music is expressive and its expressivity can be easily related to people's emotions, almost all types of human emotions directly related to the respective music genre. Human's emotions like happiness, anger, sadness, fear, depression and tenderness can typically be determined through their facial expressions. Music could change human's emotions, could have an influence on their mood, and finally could affect their health. Musical Therapy is one among the oldest strategies used for treating some of the psychological diseases. The combination of Musical therapy and facial emotion detection results in an intelligent system that sorts a music collection based on the genre conveyed by each music and then recommends a well-suited music playlist to the psychiatrist for the patients based on their current mood. The image is subjected to facial detection and emotion recognizing the techniques of the patients. The music that best matches this emotion is then recommended playlist that gives the best results to relax and calm patients.*

Key Words: **Image Processing, Facial Expression Recognition, Music Genre Classification, Music Therapy, Fisherface Machine Learning Algorithm.**

1. INTRODUCTION

Human beings have the natural intelligence to look at someone's face and guess their mood. This intelligence if learnt by an electronic device – laptop, robot automaton or a mobile device –will have valuable application within the real world. Music is considered to be a good medium of emotional communication. Emotional expression is taken into account a vital therapeutic think kind of treatment. In each style and scientific literature, it's connected to psychological wellbeing, whereas inhibition of feeling appears to play task within the development of assorted diseases, as well as physical illnesses.

Using traditional musical therapy, in which the therapist had to analyze the patient's mood manually by using different observing techniques, that doesn't include any computer machine or any other specific intelligent system. After finding out the mood of the patient, the therapist selects songs that would soothe his/her mood and emotional experience. This task was effortful, more time consuming and an individual often faced the dilemma of landing at an appropriate list of songs. Even if they acknowledge the mood of the user then their choice of songs for creating a playlist is such it'll simply decide songs reflective the present mood of the user and can't attempt to enhance his mood in any way. So, if the user is gloomy, he are supplied with an inventory of songs with unhappy feeling which might degrade his mood additional and can cause depression. So, the system planned during this paper can notice the feeling of the user from his facial expressions. It'll then give the user with a play list of songs, taking note of that the user can feel higher.

Problem Statement

Compared with the traditional music therapy, our proposed Facial expression based music therapy does not need a therapist to do all excessive observations, and the therapy is an adaptive one it can be personalized to each individual patient. There is a choice of different therapies targeting pain management, depression, stroke rehabilitation, stress management, etc. Healthcare providers can provide better service by using additional information about patients' emotional state during treatment.

2. LITERATURE REVIEW

Olga Sourina and other co-authors proposed Real-time EEG-based emotion recognition for music therapy in which they used two main algorithms, the first is Information acquisition algorithm and second is Music therapy algorithm [1]. Here authors distinguished between different type of therapy, the first one deals with pain related issues: the music pieces targeting a happy emotion are played to the patient to destruct his/her attention from the pain. The second strategy of therapy deals with depression: here the song targeting the sequence of sad, pleasant and satisfied emotions are played to user. The third type of therapy is used to decrease anxiety level of the patients. Electroencephalography (EEG) signals are used to detect the human emotions. Fractal dimension (FD) values of EEG could reveal geometric complexity of the signals. It has been proved that FD could be applied in real-time EEG signal processing to identify different brain states.

Ludwika Konieczna–Nowak worked on the correlation between Music Features with Emotional Experiences, also worked on Implications for Music Therapy. Emotions in music may well be elicited in three other ways. First area unit episodic associations, that link sure musical items with specific life circumstances and things coming back from the autobiographies of listeners. Second, painting associations return from the musical likeness to nonmusical phenomena. The ultimate choice, structural expectations, suggests that emotions area unit derived from structural options of the

composition. Expressing Emotions in Gestalt Therapy Gestalt psychotherapists focus on emotional experiences [2]. Gestalt therapy techniques are based on existential dialogue that a therapist carries out with a patient.

Ramya Ramanathan and other co-authors proposed an intelligent music player system that is based on human emotional expression and then recommend music to the user. Here they first work on facial recognition techniques, along with learning techniques for correlating the facial expression with an emotion. The learning can be achieved by using ANNs and HMMs [3] for clustering and classification of the emotions, which can be trained datasets of faces such as the Cohn Kanade dataset. After the next phase is music classification, using the Arousal-Valence model of emotion proposed by Thayer. Haar classifier is mostly used for face detection where the classifier is trained with declared face data which enables it to different faces accurately. The music feature extraction can be performed in python by using MIR packages such as LibROSA and PyAudio Analysis. They have used MATLAB for plotting Thayer's Graph using arousal value and Audio Feature Extraction. Music is subject to k-means clustering, to obtain clusters of music based on their emotional expression which is reflecting on their face.

Aurobind V. Iyer and other co-authors proposed mood enhancing music recommendation system which is based on Canny Edge Detection and Viola Jones algorithm. In this method there are four major steps are involves Integral Image, Feature Discussion, AdaBoost learning algorithms, Cascading Classifier respectively. The above steps are to be used for detecting image from an image after that it has to recognize emotion carried out by that face. For emotion recognition Eigenface and Fisherfaces are methods used for dimensionality reduction. Eigenface method uses Principal Component Analysis (PCA) [4]. It is a statistical technique that is used for data dimension reduction, data compression and also uses eigenvector properties for determining object orientation. Fisherfaces method uses fisher's Linear Discriminant (FLD) provides higher between-class scatter as compared to PCA.

F. ADAT and other co-authors proposed that Anthropometric model for facial feature points localization and Main axis localization [6]. To realize the anthropometric model following stages:

1) Main axis localization:

a) Eyes axis localization using gradient projection;

b) Mouth axis localization using color information;

c) Symmetry axis localization using gray image;

2) Facial feature points localization using the proportional position.

3) Facial feature points detection using the Shi & Thomasi method for more accuracy.

The confusion matrix of different emotions with Kohn-Kanade and FEEDTUM database respectively. This matrix shows the effectiveness of classification method with the SVM.

Kyogu Lee and Minsu Cho proposed a system for Mood Classification from Musical Audio Using User Group-Deendent Model. They used adjectives of USPOP's mood label data and Last. fm's tags as classification keywords. Among those adjectives, they selected 19 terms related to music moods, and found the optimal number of clusters, which was three. They then performed PCA on the mood adjectives using tracks as observations and verified that the top adjectives in each of the three clusters are close to each other [7].The important part is preprocessing, acoustic feature vectors extracted from the raw audio are high-dimensional which likely causes a problem in estimating the model parameters, known as the curse of dimensionality. In order to reduce the dimension of the feature vectors and maximize the separability between classes, they perform Linear Discriminant Analysis before estimating the model parameters.

Shlok Gilda and other co-authors study the usage of convolutional neural networks (CNNs) in the context of emotion recognition CNNs are known to simulate the human brain when analyzing visuals; however, given the computational requirements and complexity of a CNN, optimizing a network for efficient computation is necessary[11]. The Emotional module has two different sections Dataset Description and Model Description. A multi-layered CNN programmed to evaluate the feature of images. The layers are Input layer, Convolutional layer, Dense layer and output layer. The second module is music classification module which consists of Preprocessing, feature Description of music dataset. The module is recommendation module is responsible for generating a playlist of relevant songs for the user using mapping techniques which results in playlist generation.

Renuka s. Deshmukh and other co-authors proposed facial emotion detection using machine learning approach, the machine learning algorithmic program are going to be used for the aim of classification of the emotions. The emotional expression API are going to be further developed which will provide us an output within the kind of emotion that are classified supported the given input to the system. The paper includes only the partial results of implementation; the more implementation is carried on. The performance measures are going to be used counting on the further implementation that's in method [12]. A system for human expression recognition by analyzing key facial regions using PCA and neural networks. Emotion recognition during this paper relies on observation of contours, particularly of facial expression displayed in still footage. Facial expressions used are obtained by edge detection and specializing in specific facial regions of eyes and the mouth. Therefore, classification and feeling recognition is performed through those two facial regions. The choice of these two regions, as the basis

for feeling recognition is intuitive since the most visual indication of emotions is visible in those regions.

3. DESIGN METHODOLOGY

For implementing this project the design is divided into three major part, the first part is extracting facial expressions from image after performing some important image processing technique and machine learning approach, the second part is classification of music genre and last part is choosing/playing proper music for therapy which give the best result for that respective patient. Multiple images are captured from a web camera. To predict the emotion accurately, we might want to have more than one facial image. Blurred images can be an error source (especially in low light conditions) and hence, the multiple images are averaged to get an image devoid of any blur. Histogram equalization is an image processing technique used to enhance the contrast of the image by normalizing the image throughout its range. This image is then cropped and converted to greyscale so that only the foreground of the image remains, thereby reducing any ambiguity.

For music emotion recognition, the music dataset needs to first be obtained. Initially, a small dataset of songs from a local collection is used, after which a subset of the thousands of Song Dataset can be used for further testing and training. The music feature extraction can be performed in python by using MIR packages such as LibROSA and PyAudio Analysis. These are python packages that aid the analysis and retrieval of musical features such as tempo, rhythm, timbre, etc. which, along with the lyrics, are further used for clustering by unsupervised learning, using the k-means algorithm. For the efficient execution of this technique, the centroids are manually set to select the locations of the clusters. Once this has been accomplished, each song must be labeled with appropriate descriptors, to make the search process more efficient once the facial emotion recognition has been completed. The music database can then be queried using appropriate descriptors, to obtain the desired music playlist automatically.

A) Facial Features Extraction:

A picture is taken throughout the runtime of the appliance, which once preprocessing, is foreseen to belong to one of the emotion categories by the fisher face classifier. The model additionally permits the user to customize the model so as to cut back the variance among the categories any, initio or sporadically, such that solely variance would be that of feeling modification.

Fisherfaces method:

Discriminant Analysis is best approach to scale back dimensionality especially in classification domain. Linear discriminant analysis is one among the foremost popular Discriminant Analysis. It's a way wont to find a linear combination of features which may be wont to separate classes of various objects. It can thus be used for dimensionality reduction as well as a linear classifier.

Fisherfaces method uses fisher's Linear Discriminant (FLD) provides higher between-class scatter as compared to PCA. It focuses to maximize the ratio of between-class to within class scatter. As a result of which the tightly spaced clusters are formed.

As the data set that's used for learning labeled, this information is used to create a more reliable model to scale back dimensionality. It starts by creating a picture matrix wherein each column may be a vector that represents pixel intensities of the image.

There are many factors contribute in conveying emotions of a patient. Pose, speech, facial expressions, behavior and actions are a number of them. From these above mentioned factors, facial expressions have a better importance since they're easily perceptible.

Parts of the Face	Emotion
lip corner pulled, open eyes, open mouth, cheeks raised	Happy
Eyebrow pulled down, open eyes, lip tightened	Anger
Outer eyebrow down, inner eyebrows raised, eyes closed, lip corner down	Sad
Outer eyebrow down, inner eyebrow up, mouth open	Fear
Eyebrow up, open eyes, jaw dropped	Surprise
Lip corner depressor, lower lip depressor, eyebrows down, nose wrinkled	Disgust

Table 1: Emotion and Facial part

Above mentioned table is giving brief idea of different facial parts relation with different type of Emotion/Expression carried by Humans.

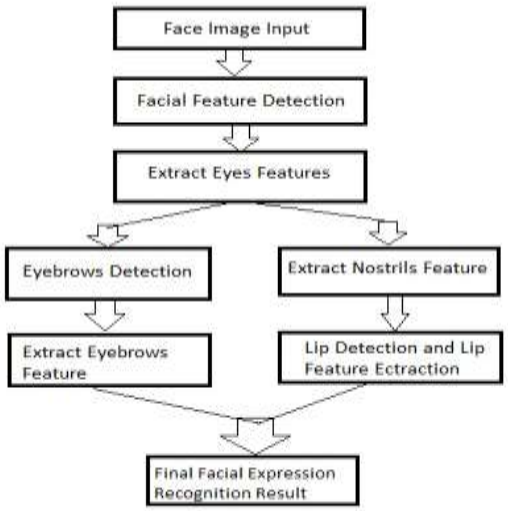

Fig 1: Facial Emotion Extraction

The machine learning algorithms are going to be used for the aim of classification of the emotions. The emotion API are going to be further developed which will give us an output within the sort of emotion that are classified supported the given input to the system.

B) Music Genre Classification:

Categorizing music files consistent with their genre may be a challenging task within the area of music information retrieval (MIR).With the expansion of online music databases and easy access to music content, therapist find it increasing hard to manage the songs that improves patients health, a method to categorize and organize songs is based on the genre, which is identified by some characteristics of the music like rhythmic structure, harmonic content and instrumentation.

GTZAN which was firstly proposed by G. Tzanetakis in is one among the foremost popular dataset used for music signal processing. It contains 1,000 music with 30-second, 22050 Hz frequency and 16 bits. Genres within the GTZAN are blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae and rock and every one of those genres have 100 music. In this section, to classify music consistent with their genre some machine learning algorithms are discussed. The classification algorithms utilized in this part are Naive Bayes, Random Forest, K Nearest Neighbors, Decision Tree and Support Vector Machine.

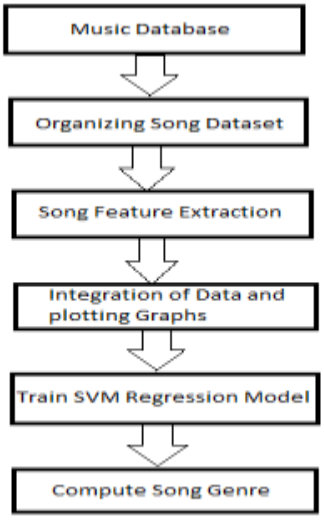

Fig 2: Music Genre Classification

Support Vector Machine (SVM):

In machine learning, support-vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. Given a set of training examples, each marked as belonging to at least one or the opposite of two categories, an SVM training algorithm builds a model that assigns new examples to one category or the opposite , making it a non-probabilistic binary linear classifier (although methods like Platt scaling exist to use SVM during a probabilistic classification setting). An SVM model may be a representation of the examples as points in space, mapped in order that the samples of the separate categories are divided by a transparent gap that is as wide as possible. New examples are then mapped into that very same space and predicted to belong to a category supported the side of the gap on which they fall.

Applications-

1. Classification of images can also be performed using SVMs
2. Hand-written characters can be recognized using SVM
3. The SVM algorithm has been widely applied in the biological and other science field.
4. SVMs are helpful in text and hypertext categorization.

A number of emotions is collected and put up in the list. Each emotion category has a number of songs listed in it. When the user's expression is classified with the help of SVM algorithm, songs belonging to that category.

C) Recommendation Therapy Module:

This module is responsible for generating a playlist of relevant songs for the patient. It allows the user to modify the playlist based on therapist preferences and modify the class labels of the songs as well. Classified songs are mapped to the user's mood. After the mapping procedure is complete, a playlist of relevant songs is generated. Similar songs are grouped together while generating the playlist. Similarity between songs was calculated by comparing songs over 50 ms intervals, centered on each 10 ms time window. After empirical observations, we found that the duration of those intervals is on the order of magnitude of a typical song note. Feature values like an audio file were compared to the values (for an equivalent features) corresponding to audio files belonging to an equivalent class label. The recommendation engine features a twofold mechanism.it recommends songs based on: Patient's perceived mood and Therapist's preference also.

4. RESULT OF IMPLEMENTATION

The implementation is carried out in Python3.6 or above. Here, for the facial expression recognition purpose testing was carried out on dynamic images to achieve real time performance. The images were taken through the in-built camera for various individuals. The detected faces are then accustomed train the face, which works on reduction of variance between classes. Fisher face recognition method proves to be efficient because it works better with additional features. Here noise and blurriness is off from the image

following which the image is converted from RGB to Grayscale then resized into 300*300 pixels. The results of the proposed application are shown here.

Recognizing Different Facial Features:

The primary objective is to detect face from a captured image then detecting different facial features to extract patient's emotion in real time environment.

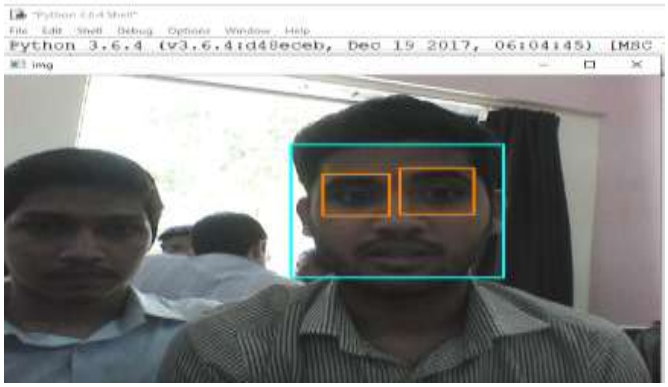

Fig 3: Real-time Detection of Face and Eyes

Fig 4: Real-time Detection of Mouth

Facial Expression Extraction:

Different type of expression of Patients can be extracted like Happy, Sad, Neutral, Disgust, Anger, Fear, and Surprise.

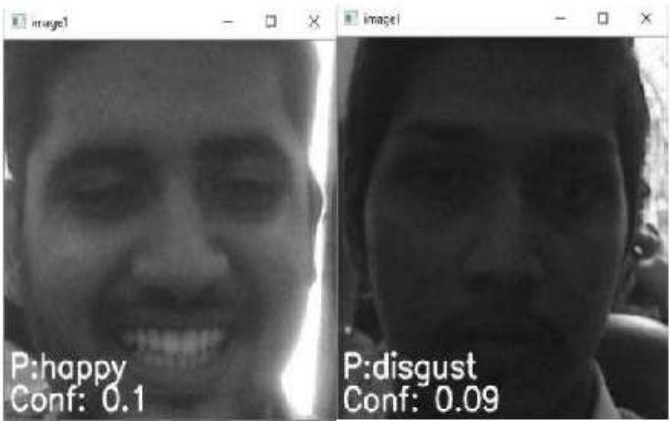

Fig 5: Happy and Disgust Emotions

The GUI for Music Therapy:

Fig 6:Graphical User Interface

5. CONCLUSION

The intelligent system is created keeping in mind the simplicity of use by the therapist and to enable them the power of machine learning for their day to day work. This musical treatment system was made keeping in mind the ease of use and intuitive interface of the music player for listening music of the required mood. Once the user opens the application and runs the program, the program automatically detects the emotions of the person in front of camera and plays the music accordingly. The online music player is built by using HTML, CSS and java script, making it very robust and sturdy. The main highlight of the music therapy application is it automatically plays music and has an option for further improving the trained model by choosing the training of model again.

ACKNOWLEDGEMENT

We express our gratitude to **Prof. Pankaj Kunekar**, our project guide who gave us all the guidance and motivation and at any time available to us for doubts solving, grateful for giving us the required help, point by point recommendations and furthermore support to do the venture. We are excited and happy to offer our thanks to the Head of the Information Technology Department **Prof.**

Deepali Matse, for her endorsement of this undertaking. We also want to profoundly offer our genuine thanks to our regarded head **Prof. Dr. Shrikant Kallurkar** and the administration of Atharva College of Engineering for giving such a perfect environment to develop this undertaking with well-furnished library with all the most extreme important reference materials and forward-thinking IT Laboratories We are amazingly appreciative to all staff and the administration of the school for giving every one of us the offices and assets required.

REFERENCES

[1] Olga Sourina ,Yisi Liu, Minh Khoa Nguyen "Real-Time EEG-Based Emotion Recognition For Music Therapy" Published by journal on Multimodal User Interfaces – Springer International Publishing on 2012 5:27-35.

[2] Anna Bukowska, Krzysztof Stachyra, Amelia Oldfield "Musical Therapy and Emotional Expression: A kaleidoscope of Persspective" Edited by Ludwika Konieczna–Nowak Published by Katowise ISBN: 978-83-942090-3-2 on 2016.

[3] Ramya Ramanathan, Radha Kumaran, Ramrohan, Rajat Gupta, Vishalakshi Prabhu "An Intelligent Music Player Base On Emotion Recognition" Published By IEEE International Conference On Computational Systems And Information Technology For Sustainable Solutions 2017 Page no 299-303.

[4] Aurobind V. Iyer, Viral Pasad, Karan Prajapati "Emotion Based Mood Enhancing Music Recommendation" Published by IEEE International conference on RTEICT, May 2017,India. Page No: 1573-1577.

[5] Pratik Gala, Raj Shah, Vinee Shah, Yash Shah, Mrs. Sarika Rane "MoodyPlayer: AMusic Player Based On Facial Expression Recognition" Published by IRJET Volume: 05 Issue: 04 on April-2018 ISSN: 2395-0056.

[6] F. ADAT,C. MAAOUI and A. PRUSKI "Human –Computer Interaction Using Emotion Recognition From Facial Expression" Published by IEEE UKSim 5th European Symposium on Computer Modeling And Simulation on 2011.

[7] Kyogu Lee and Minsu Cho "Mood Classification from Musical Audio Using User Group-Deendent Model" Published by IEEE 10th International Conference on Machine Learning and Application on 2011 (978-0-7695-4607-0/11).

[8] JungHyun Kim, Seungjae Lee, WonYoung Yoo "Implementation And Analysis Of Mood-Based Music Recommendation System" Published by ICACT on 2013 ISBN 978-89-968650-1-8.

[9] Sushmita G. kamble and Asso. Prof. A.H. Kulkarni "Facial Expression Based Music Player" Published by IEEE International Conference on Advance Computing, ICACCI on Sep, 2016.

[10] Karthik Subramanian Nathan, Manasi Arun and Megala S kannan "EMOSIC-An Emotion Based Music Player for Android" Published by IEEE, ISSPIT on 2017.

[11] Shlok Gilda, Hussain Zafar, Chintan Soni and Kshitija Waghrdekar "Smart Music Player Integration Facial Emotion Recognition And Music Mood Recommendation" Published by IEEE WISPNET on 2017.

[12] Renuka S. Deshmukh, Vandana Jagtap and Shilpa Paygude "Facial Emotion Recognition system Through Machine Learning" Published by ICICCS on 2017.

[13] Swathi Swaminathan and E. Glenn Schellenberg "Current Emotion Research In Music Psychology" Published Emotion Review Vol. 7 ,No.2 on April 2015 (189-197), ISSN 1754-0739.

[14] Maruthi Raja S K, Kumaran V, Keerthi Vasan Aand Kavitha N "Real Time Intelligent Emotional Music Player Using Android" Published by Journal for Research Volume 03, Issue 01 on March 2017.

[15] Prof. Vijaykumar R. Ghule, Abhijeet B. Benke , Shubham S. jadhav and Swapnil A. Joshi "Emotion Based Music Player Using Facial Recognition" Published by IJIRCCF Volume 05, Issue 02 on Feb 2017.