Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

Identification of Rhythmic Sounds Patterns by Emotion Recognition Using Landmark and Euclidean Distance Techniques

Muhammad Saood Sarwar[1*], Zain-ul-Abideen[2]

[1,2]Faculty of Computer Science and Engineering, Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan.*

Email: [2]zain-ul-abideen@giki.edu.pk
Corresponding Email: [1]saood.sarwar@giki.edu.pk*

Received: 03 February 2023 **Accepted:** 19 April 2023 **Published:** 30 May 2023

***Abstract:** The human face is a prominent part of the human body; one can easily guess the emotional state of a person using visual expressions like facial expressions. A camera can be used to extract the required features from the human face. These features can be used from different perspectives and this information can be used to roughly estimate the emotional state. This information sounds patterns that match the mood collected from the previously provided information. Manually searching for sound patterns from playlists is difficult and time-consuming. This eliminates the tedious and repetitive task of physically separating sound recordings or collecting them in separate recordings, allowing you to create the right playlists for your current emotion. Therefore, the proposed system automatically generates playlists based on the emotions and expressions of the human face, as the human face plays an important role in detecting and extracting emotions from the person. This paper focuses on the strategies available for recognizing human emotions to develop an emotion-based rhythmic sounds patterns player, and the methodologies used by existing rhythmic sound patterns players to distinguish emotions. It additionally gives brief thought regarding our system working, emotion classification, and song playlist generation. The methodology that is proposed in this research paper contains less computation with an accuracy between 90%-95%.*

***Keywords:** Emotion-Based Rhythmic Sounds Patterns Player, Emotion-Based Music Playlist, Emotion Recognition, Emotion Classification, Human Emotions.*

1. INTRODUCTION

Most rhythmic sounds patterns-cherishing clients wind up in an odd circumstance when they do not discover tunes to suit their mindset in the circumstance. Regardless of the way that these highlights meet the necessities of the client, the client faces the undertaking of physically choosing the melodies through the playlist of tunes dependent on their present state of mind and conduct because manual searching of rhythmic sound patterns from the playlist is troublesome and tedious.

Thus, our proposed system will produce a playlist dependent on the feelings and appearances of a human face because the human face assumes a significant part in the acknowledgment and extraction of feelings of a distinctive individual. Our proposed system will utilize Image Processing to perceive looks. The face is caught utilizing an in-constructed camera. Here we are utilizing the Cohn Kanade data set to order the feeling. Considering the recognized feeling the playlist is created consequently. Emotions are classified into different categories happy, sad, disgusted, normal, and angry. Various strategies and approaches have been suggested and created to recognize current human emotions through facial expressions. Existing systems are exceptionally highly complex as far as time and memory prerequisites for extricating facial expressions. Because of the current enthusiastic state of a user, existing systems have a lesser precision. Rahul Rirve and S. Jagdale Proposed an Automatic Facial Expression recognition system in the research paper EmoPlayer: An Emotion-Based Music Player.

In this system, there are three steps Face detection, Feature Extraction, and Classification. The first step Face Detection is done by Voila Jones Algorithm. The second step is Features Extraction by Landmark algorithm which extracts 68 landmark points on the face. The third step Classification is done by SVM. This paper is based on this research paper with some amendments like Image Enhancement by Histogram Equalization, calculating Euclidean distance from each point to all other points after getting landmarks points from the Landmark algorithm and with rhythmic sounds patterns Recommendation Algorithm. After these amendments, this increases the accuracy of the system. This paper aims to plan a legitimate and precise calculation that will produce a playlist dependent on the current mindset of a client. For face identification, we have utilized the Haar Cascade classifier. After extricating your face from the image, the following stage is to recognize the feeling. The landmark algorithm is used for feature extraction and SVM is used for the classification of emotions into different categories. If the quality of a camera is not good, then there might be a problem for the system to understand the emotions of the user. If the user will use this system at night.

So, the system cannot recognize human emotions. The approach our rhythmic sound patterns player takes to identify human emotions, and how it is more efficient to use our system for emotion detection. In case the system fails to extract features due to any reason, then the system tries to capture the frame again until finds the frame with the correct detection and orientation of the face.

2. RESEARCH METHODOLOGY

A webcam is used to capture the system's input image. This image is subjected to image enhancement, which involves using an algorithm for photographs with poor contrast to

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

restore good image contrast. As a result, SVM is used for training and classification, which successfully supports multi-class classification.

Voila Jones Algorithm

Haar-Cascade is a machine learning object recognition system that uses characteristics provided by Paul Viola and Michael Jones to recognize things in an image or video. The use of the algorithm in EBMP is to detect a face in frames so that further process of feature extraction can be executed. Effective object detection method. In this approach, each kernel is utilized to compute many characteristics. We must find the total of pixels beneath white and black rectangles for each feature computation. Every feature is a single value calculated by subtracting the sum of pixels in the white rectangle from the total of pixels in the black rectangle. If a window isn't a face region, it's not a face area. If it isn't, toss it out in one shot. With extra calculations, need not process it again. Figure 1 shows the detection of faces by the Voila Jones algorithm.

Fig.1 Detection of the face by Voila Jones algorithm

Fig.2 Face Detection by Voila Jones algorithm

Histogram Equalization

After the detection of the face, that image will be enhanced by Histogram Equalization this increases the global contrast, particularly when the usable information of the picture is addressed by close differentiation value.

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

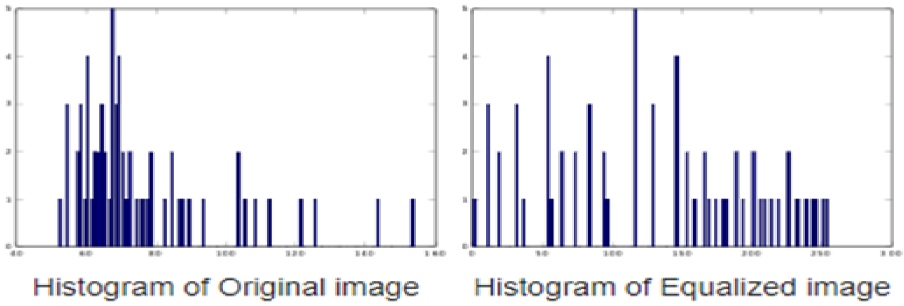

Fig.3 Image enhancement using Histogram Equalization


Fig.4 Image Enhancement using Histogram Equalization

Landmark Algorithm

We can study other feelings by reading their facial expressions. Landmark Algorithm can be utilized to determine the location of facial muscles that helps in recognizing feelings. These focuses should be properly addressed to accomplish an effective feeling of acknowledgment. By using the Landmark Algorithm, the extraction of 68 facial landmarks including landmark points on eyes, nose, eyebrows, lips, and jaw, and the classification of 5 basic emotions are presented in this algorithm. The landmark model is used to determine 68 (x, y)-coordinate locations that map on the facial image. These 68 coordinates can be visualized on the image.


Fig.5 Facial landmarks using Landmark Algorithm

Euclidean Distance and Rhythmic sounds patterns Recommendation Algorithm

Distance varies from each landmark point to other points for different facial expressions of humans. After extracting 68 facial points on the face, calculate the Euclidian distance from

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

each point to all other points. This increases the number of features. It is a self-developed algorithm to recommend rhythmic sound patterns according to the detected mood of the user. The recommended rhythmic sounds patterns list is generated by a formula, according to the formula the list will show ten songs consisting of three pre-lists i.e.

- Three songs from the pre-list of high-rated songs of detected mood type from the user rating list.
- Three songs from the pre-list of high-rated songs of detected mood type from the public rating list.
- Four random songs from the pre-list of the detected mood-type songs list.

Rules for the algorithm

- In case the user rating list or public rating list of the detected mood type contains less than three songs then the no. of random songs of the detected mood type will increase.
- In case of similarity of songs in the pre-lists, the algorithm will drop that song from any of the pre-list and will take the song next to that in that pre-list.

Top-rated songs from history = x → maximum 3
Overall high rated songs = y → maximum 3
Random songs = z => 10-x-y

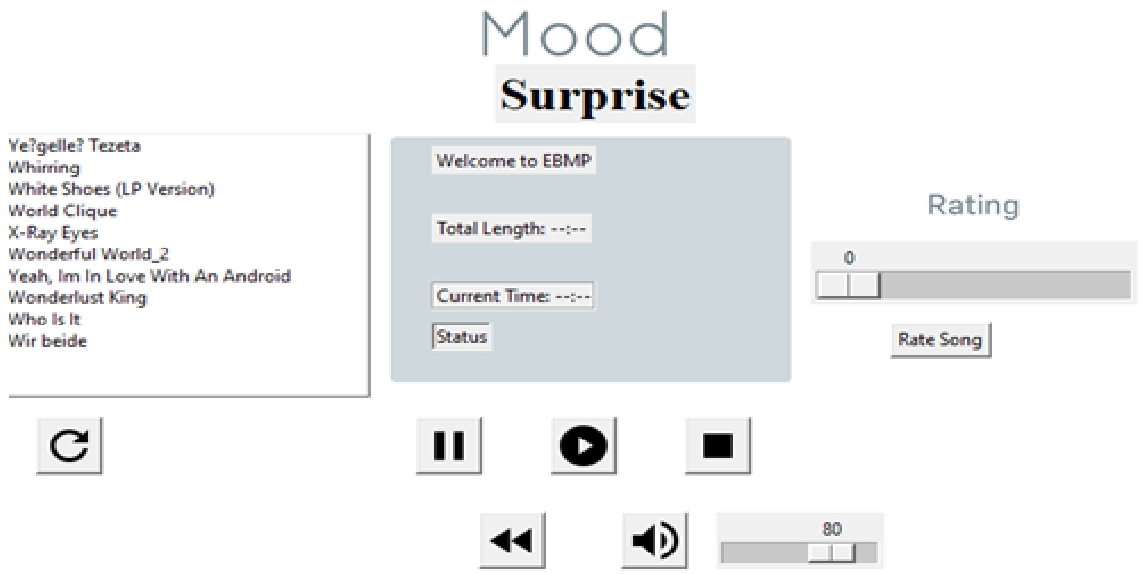

Fig.6 Rhythmic sound patterns Recommendation

SVM Classifier and analysis

Support Vector Machine is used for supervised learning techniques for classification problems. We have used a Multi-class Support Vector Machine (SVM) classifier for the classification of different kinds of emotions. These emotions are classified into five emotion

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

classes which are HAPPY, SAD, DISGUST, ANGRY, and SURPRISE. Existing solutions contain high computations and calculations which slow down the system. Thus, this system provides a cheaper and less computation system for emotion detection by Landmark Algorithm which identifies 68(x, y) landmark points on the human face and then calculates the distance of each point to all other points. After the classification by SVM Classifier system will be able to fetch rhythmic sound patterns from the database according to that emotion. Euclidean distance is calculated by using the Euclidean formula (1).

$$\text{dist}((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2} \quad (1)$$

3. RESULTS

The results obtained from our proposed system demonstrate its effectiveness in generating playlists based on the user's current mood. One notable advantage of our system is its comparatively faster response time compared to other algorithms. This can be attributed to the optimized computational processes implemented in our system, resulting in quicker playlist generation. Moreover, the accuracy of our system is commendable, ranging between 90-95%. This high level of accuracy is achieved through the incorporation of various enhancements and amendments to the existing EmoPlayer system. For instance, the inclusion of Image Enhancement by Histogram Equalization helps to improve the quality of the captured images, thus enhancing the accuracy of emotion recognition.

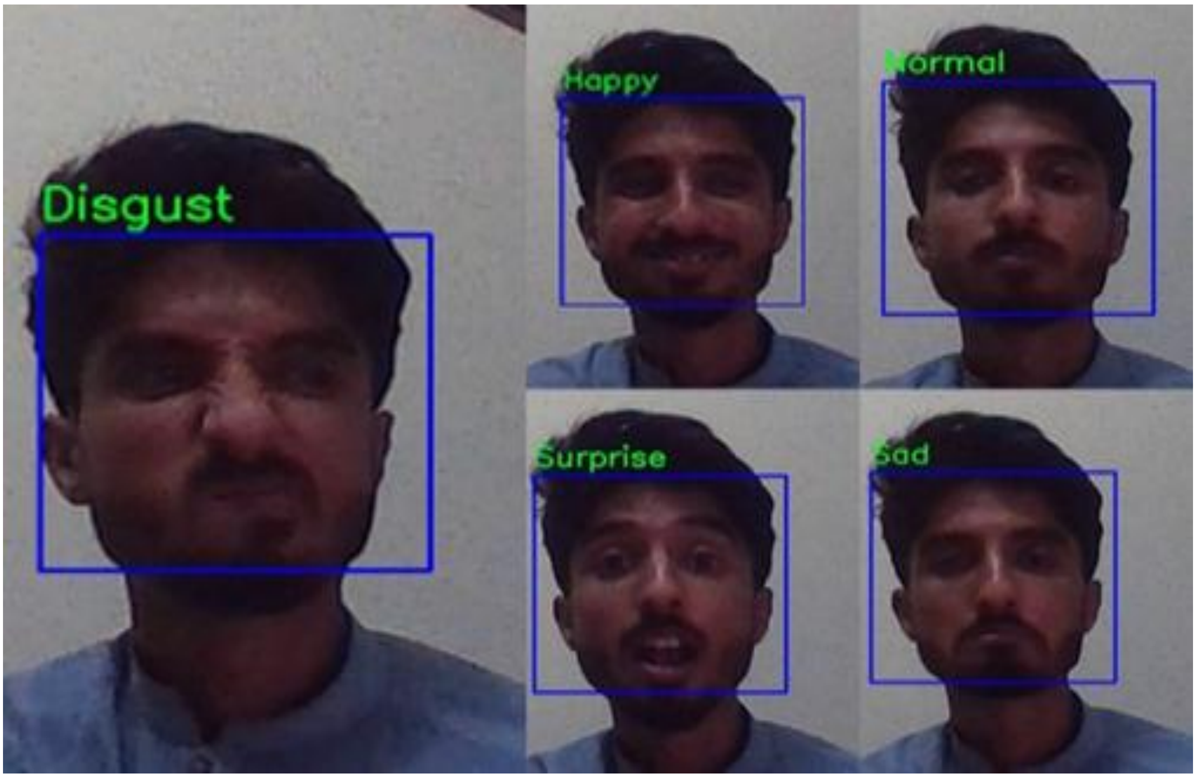

Fig.7 Different Emotions Classification.

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

Another significant improvement is the calculation of the Euclidean distance from each landmark point to all other points after obtaining the landmark points from the Landmark algorithm. This additional step enables a more precise analysis of facial expressions and contributes to the overall accuracy of the emotion classification process. Furthermore, the introduction of the rhythmic sound patterns Recommendation Algorithm adds a personalized touch to the playlist generation process. By considering the emotions and expressions extracted from the human face, our system can recommend songs that are more likely to resonate with the user's current mood, thereby enhancing the overall listening experience. It is important to note that the success of our system is dependent on the quality of the camera used for face identification. In cases where the camera's quality is poor, there may be challenges in accurately recognizing and extracting facial expressions, leading to a decrease in the system's accuracy. Additionally, our system may encounter difficulties in recognizing human emotions when used in low-light conditions or during nighttime. To mitigate potential issues, our system incorporates a mechanism to capture multiple frames until a frame with correct face detection and orientation is obtained. This ensures that the system can extract accurate features and perform precise emotion classification. Overall, the results obtained from our research indicate that our proposed system effectively generates playlists based on the user's current mood, offering a valuable and accurate algorithm for music recommendation.

Table 1: Confusion Matrix for Each Emotion

Emotions	**Disgust**	**Happy**	**Normal**	**Sad**	**Surprise**
Disgust	47	0	1	0	0
Happy	0	56	0	0	0
Normal	3	1	39	5	1
Sad	0	0	1	22	0
Surprise	0	0	2	0	83

System Flow

Below are two diagrams describing the activity and data flow diagrams of a system. The activity diagram of system visually represents the sequence of actions and behaviors within the system, providing an overview of how different components interact to accomplish specific tasks. A level 0 data flow diagram of system provides a high-level overview of the system, showing the major processes or subsystems, external entities, and the flow of data between them, highlighting the interactions and boundaries of the system.

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

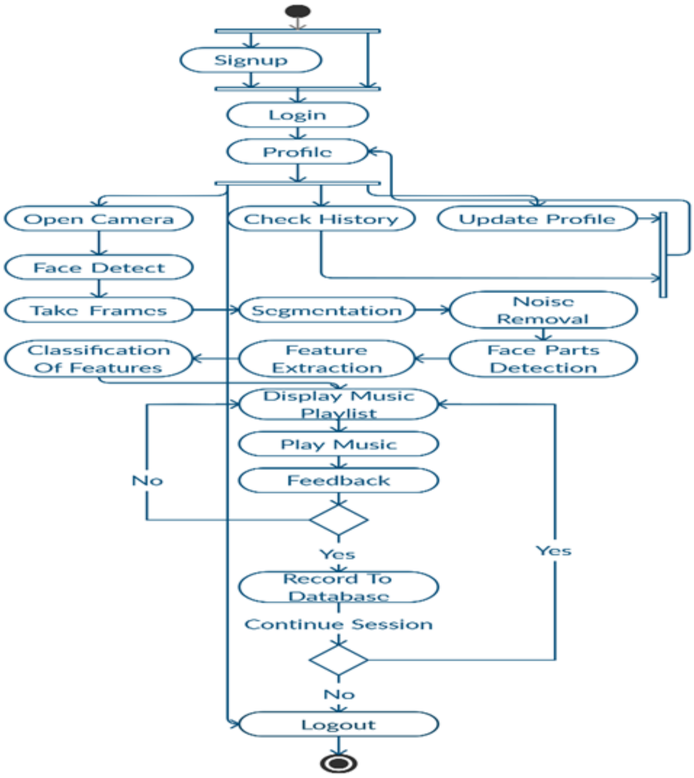

Fig.8 Activity Flow Diagram of complete System.

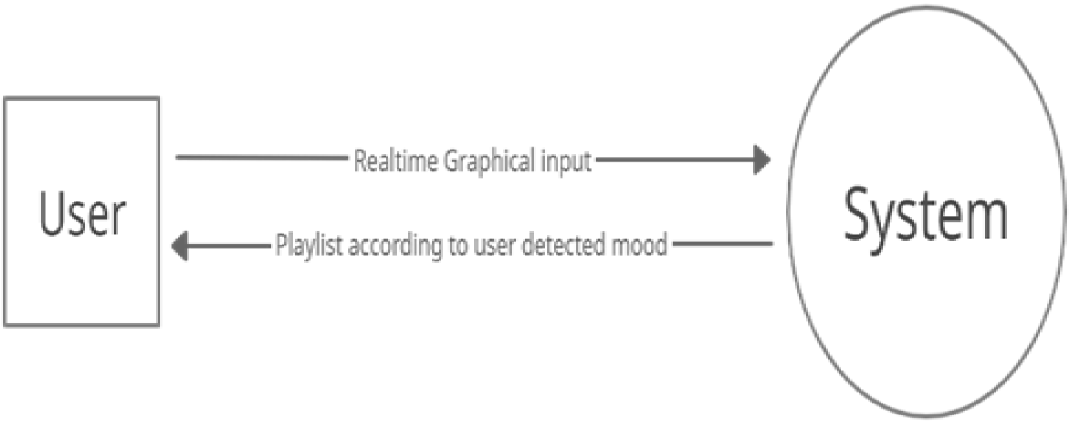

Fig.9 Data Flow Diagram (DFD) Level 0 of System

4. CONCLUSION

The system is expected to give a less expensive, extra equipment-free, and precise emotion prediction to the Windows operating system. Emotion Based rhythmic sounds patterns player will be of extraordinary benefit to the clients searching for rhythmic sounds patterns dependent on their disposition and enthusiastic conduct. The system will assist with

Journal of Image Processing and Intelligent Remote Sensing
ISSN 2815-0953
Vol: 03, No. 03, April-May 2023
http://journal.hmjournals.com/index.php/JIPIRS
DOI: https://doi.org/10.55529/jipirs.33.27.35

diminishing an opportunity to look through the rhythmic sound patterns as indicated by the mindset of the client. Diminishing the superfluous chance to process, this expands the general precision and proficiency of the system.

Data Availability

Face expression recognition (FER) dataset (Oheix, 2019). CK+ dataset for facial expression recognition (Shawon, 2018). Evaluation of Musical Features for Emotion Classification

5. REFERENCE

1. https://www.kaggle.com/jonathanoheix/face-expression-recognition-dataset
2. //www.kaggle.com/shawon10/ckplus
3. https://code.soundsoftware.ac.uk/projects/emotion-recognition/repository/show/4.%20dataset%20(audio)
4. http://www2.projects.science.uu.nl/memotion/emotifydata
5. "Emotion Based Music Player Using Facial Recognition," by Prof. Vijaykumar R. Ghule, Abhijeet B. Benke, Shubham S. Jadhav, Swapnil A. Joshi ., in International Journal of Innovative Research in Computer and Communication Engineering
6. "EMO-MUSIC (Emotion based Music player)," by Sarvesh Pal, Ankit Mishra, Hridaypratap Mourya and Supriya Dicholkar.
7. "EMO PLAYER: Emotion Based Music Player" Hemanth P,Adarsh1,Aswani C.B,Ajith P, Veena A Kumar in International Research Journal of Engineering and Technology (IRJET)
8. "Emotion based Music Player" by Shital Rewanwar , Siddhika Pawar , Swati Bade , Sujata Oak in International Journal of Advanced Research in Computer and Communication Engineering

