Open access Journal International Journal of Emerging Trends in Science and Technology
Impact Factor: 2.838 DOI: http://dx.doi.org/10.18535/ijetst/v3i05.12

Emotion Based Music System

Authors
Miss. Disha Rabashette[1], Miss. Rhutuja Tale[2], Miss. Ashwini Hinge[3], Miss. Karishma Padale[4], Miss. Rutuja Chavan[5], Prof. Nutan Deshmukh[6]
[1,2,3,4,5]Students B.E, MKSSS Cummins College of Engg. For Women, Pune 411052
Email: disharabashette@gmail.com, Contact No:-8975733594.
[6]Associate Professor, Dept of Computer Engineering, MKSSS Cummins College of Engg For Women, Pune 411052
Email: *nutuna.deshmukh@cumminscollege.in, Contact No:-9130008813*

Abstract
The human face is an important organ of an individual's body and it especially plays an important role in extraction of an individual's behavior and emotional state. Manually segregating the list of songs and generating an appropriate system based on an individual's emotional features is a very tedious, time consuming, labor intensive and upheld task. Various algorithms have been proposed and developed for automating the system generation process. However the proposed existing algorithms in use are computationally slow, less accurate and sometimes even require use of additional hardware like sensors. This proposed system based on facial expression extracted will generate a system automatically thereby reducing the effort and time involved in rendering the process manually. Thus the proposed system tends to reduce the computational time involved in obtaining the results and the overall cost of the designed system, thereby increasing the overall accuracy of the system. Testing of the system is done on both user dependent (dynamic) and user independent (static) dataset. Facial expressions[1] are captured using an camera. The accuracy of the emotion detection algorithm used in the system for real time images is around 85-90%, while for static images it is around 98- 100%.The proposed algorithm on an average calculated estimation takes around 0.95-1.05 sec to generate an emotion based music system. Thus, it yields better accuracy in terms of performance and computational time and reduces the designing cost, compared to the algorithms used in the literature survey.

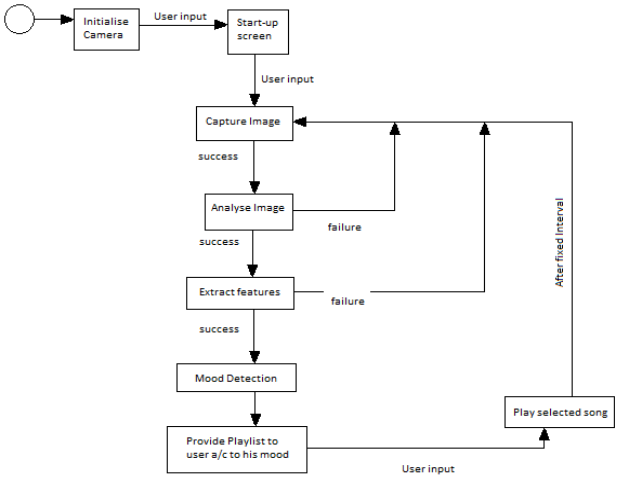
Keywords—*Emotion Recognition, Emotion Extraction Module, Audio Feature Extraction Module, Artificial Neural Networks, Face Detection.*

Introduction

Facial expressions give important clues about emotions. Computer systems based on affective interaction could play an important role in the next generation of computer vision systems. Face emotion can be used in areas of security, entertainment and human machine interface (HMI) [11]. A human can express his/her emotion through lip and eye. The work describes the development of Emotion Based Music Player, which is a computer application meant for users to minimize their efforts in managing large playlists. Generally people have a large number of songs in their database or playlists. Thus to avoid trouble of selecting a song, most people will just randomly select a song from their playlist and some of the songs may not be appropriate for the current mood of the user and it may disappoint the user. As a result, some of the songs are not matching to the user's current emotion. Moreover, there is no commonly used

application which is able to play songs based on the current emotions of the user. The proposed model will extract user's facial expressions [1] and features to determine the current mood of the user. Once the emotion Is detected, playlist of songs suitable to the mood of the user will be presented to him. It aims to provide better enjoyment to the music lovers in music listening. In the model, following moods are included: Happy, Sad, Thinking [3]. The system involves the image processing and facial detection processes. The input to the model is still images of user which are further processed to determine the mood of user.

The system will capture the image of the user at the start of the application. The images are captured using webcam. The image captured previously will be saved and passed to the rendering phase. The mood of the user may not be same after some time; it may or may not change. Thus the image is captured after every decided interval of time. And then that image will be forwarded to the next phase [5]. For the purpose of analyzing the image first of all, the images are converted from RGB format to binary format [7]. For that the average value of RGB for each pixel is calculated and if the average value is greater than 110, it is replaced by white pixel and otherwise it is replaced by black pixel. Then efforts are done to find the forehead from the binary image.



Current Music SYSTEMS

The features available in the existing Music players present in computer systems are as follows:

- Manual selection of Songs
- Party Shuffle
- Playlists
- Music squares where user has to classify the songs manually according to particular emotions for only basic emotions .Those are happy, thinking, sad and angry.

Limitations of existing system:

- It requires the user to manually select the songs.
- Randomly played songs may not match to the mood of the user.
- User has to classify the songs into various emotions and then for playing the songs user has to manually select a particular emotion.

Literature Survey

1. Face detection methods:

The techniques for face detection can be distinguished into two groups: holistic, where face is treated as a whole unit and analytic, where co-occurrence of characteristic facial elements is studied. Pantic and Rothkrantz proposed system which process images of frontal and profile face view. Vertical and horizontal histogram analysis is used to find face boundaries. Then, face contour is obtained by thresh holding the image with HSV color space values. Kobayashi and Hara used image captured in monochrome mode to find face brightness distribution. Position of face is estimated by iris localization.

2. Feature extraction methods:

Pantic and Rothkrantz selected a set of facial points from frontal and profile face images. The expression is measured by a distance between position of those points in the initial image (neutral face) and peak image (affected face) [9].

Cohn et al. developed geometric feature based system in which the optical flow algorithm is performed only in 13x13 pixel regions surrounding facial landmarks.

Shan et al. investigated the Local Binary Pattern method for texture encoding in facial expression description. Two methods of feature extraction were proposed. In the first one, features are extracted

from fixed set of patches and in the second method from most probable patches found by boosting.

3. Expression Recognition:

The last part of the FER system is based on machine learning theory; precisely it is the classification task. The input to the classifier is a set of features which were retrieved from face region in the previous stage. Classification requires supervised training, so the training set should consist of labeled data. There are a lot of different machine learning techniques for classification task, namely: K-Nearest Neighbors, Artificial Neural Networks, Support Vector Machines, Hidden Markov Models, Expert Systems with rule based classifier, Bayesian Networks or Boosting Techniques [4].

Major problems encountered in the above stages are different scales and orientations of face. They are usually caused by subject movements or changes in distance from camera. Significant body movements can also cause drastic changes in position of face in consecutive frames what makes tracking harder.

Complexity of background and variety of lightning conditions can be also quite confusing in tracking. For instance, when there is more than one face in the image, system should be able to distinguish which one is being tracked.

Proposed System

1. Facial Detection:

Object Detection using Haar feature-based cascade classifiers is an effective object detection method. It is a machine learning based approach. In this approach a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. First of all, lot of positive images (images of faces) and negative images (images without faces) are used to train the classifier [2]. Then features are extracted from it. For this, haar features shown in below image are used Each feature is a single value obtained by subtracting sum of pixels under white rectangle from sum of pixels under black rectangle. Two rectangular features detect contrast between two vertically or horizontally adjacent regions. Three rectangular features detect contrasted region placed between two similar regions and four rectangular features detect similar regions placed diagonally (Fig. 1). Rectangle features can be computed very rapidly using an intermediate representation for the image which we call the integral image Figure 1:

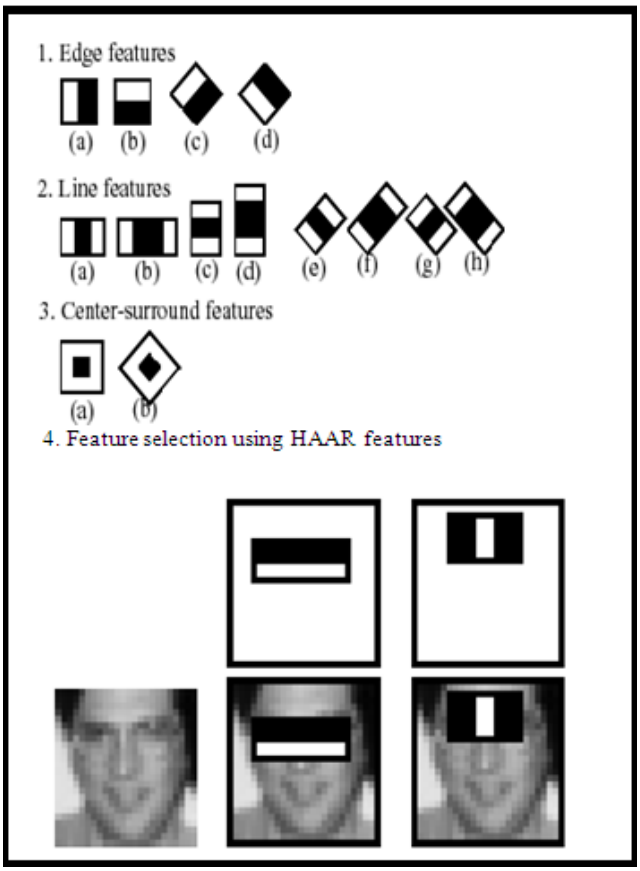


Fig:1 Common Haar features &Feature extraction from face

The feature capitalizes on the observation that the eye region is often darker than the cheeks. The second feature compares the intensities in the eye regions to the intensity across the bridge of the nose. The face detection procedure includes some steps which are consecutively performed on the input image. Firstly, the classifier trained for face detection searches for a face in the image .In case when face is not found in the image, further processing is omitted and system returns appropriate error message.

2. Feature-Point Detection:

The feature points are detected automatically. For face detection, first we convert binary format image

from RGB format image [8]. For converting binary image, we calculate the average value of RGB for each pixel and if the average value is below than 110, we replace it by black pixel and otherwise we replace it by white pixel. By this method, we get a binary image from RGB image [12].

Then next stage is to find the forehead from the binary image. System will start scan from the middle of the image, after that it will look for continuous white pixels after a continuous black pixel. In this we want to find the maximum width of the white pixel by searching vertical both left and right site. Then, if the new width is smaller half of the previous maximum width, then we break the scan because in that case we will reach to the eyebrow. Then we cut the face from the starting position of the forehead and its height will be 1.5 multiple of its width [5]. In this processed image we will only have eyes, nose and lip. Then we will cut the RGB image according to the binary image. This stage can also be achieved using Haar Cascades Technology [10]. Haar Cascade Handle this with much more optimized and efficient way. Lower face is extracted from face.

Now we will have an image which contains lip portion of the face. Now the next step is to extract the expression features from lip .To extract the feature we just have to measure the distance between upper lip & lower lip. Also the system will consider the position of contour points of lip. By lip feature we can significantly determine two features happy & surprise mood. The lip from lower face is detected.

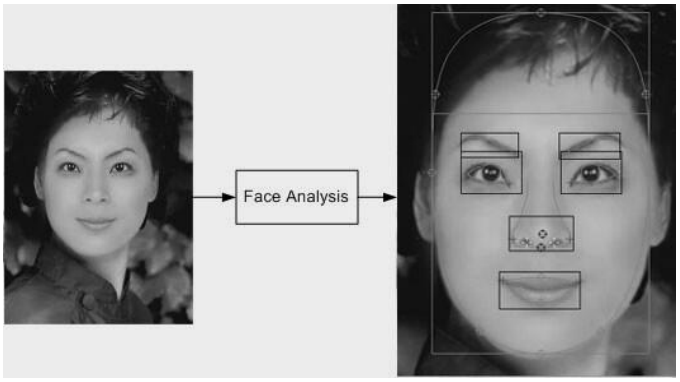


Fig 2 . Feature point detection of eyes, upper lips, lower lips, nose.

Conclusions and Future Work

The aim of this paper was to explore the area of automatic facial expression recognition for implementation of an emotion based music player. Beginning with the psychological motivation for facial behavior analysis, this field of science has been extensively studied in terms of application and automation. Manual face analysis used by psychologists was quickly replaced by suitable computer software. A wide variety of image processing techniques was developed to meet the facial expression recognition system requirements. Apart from theoretical background, this work provides ways to design and implement Emotion based music player. Proposed system will be able to process the image of facial behavior, recognize displayed actions in terms of basic emotions and then play music based on these emotions. Major strengths of the system are full automation as well as user and environment independence. Even though the system cannot handle occlusions and significant head rotations, the head shifts are allowed. In the future work, we would like to focus on improving the recognition rate of our system. Also, we would like to develop a mood-enhancing music player in the future which starts with the user's current emotion (which may be sad) and then plays music of positive emotions thereby eventually giving a joyful feeling to the user. Finally, we would like to improve the time efficiency of our system in order to make it appropriate to use in different applications.

Acknowledgement

We would like to express our gratitude to our project guide Prof. Nutan Deshamukh and also our Head Of Department Prof. Kelkar for their valuable advice and support.

References

1. Chang, C. Hu, R. Feris, and M. Turk, ―Manifold based analysis of facial expression, Image Vision Comput ,IEEE Trans. Pattern Anal. Mach. Intel. vol. 24, pp. 05–614, June 2006.

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. Fatma Guney, 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. Michael Lyon and Shigeru Akamatsu, ―Coding Facial expression with Gabor wavelets, IEEE conf. on Automatic face and gesture recognition, March 2000.
9. Renuka R. Londhe, Dr. Vrushshen P. Pawar, ―Analysis of Facial Expression and Recognition Based On Statistical Approach, International Journal of Soft Computing and Engineering (IJSCE) Volume-2, May 2012.
10. Thayer The biopsychology of mood & arousal, Oxford University Press, 1989.
11. Z. Zeng ―A survey of affect recognition methods: Audio, visual, and spontaneous expressions, IEEE. Transaction Pattern Analysis, vol 31, January 2009
12. Bill, David "Personalizing content based on emotions", US2010/0321519, San Francisco: Dec. 23, 2014.