I J C T A, 10(8), 2017, pp. 435-441 ISSN: 0974-5572

Emotion Based Music Player with Sleep Alert System

Kiran Yesugade[1], Anjali Gupta[2], Shriya C. Patankar[3], Varsha Desai[4] and Shweta M. Kumbhar[5]

ABSTRACT

In this paper we propose a smart emotion based music player with sleep alert system. The objective of this paper is to introduce needs and applications of facial expression recognition. The proposed system uses Intel's real sense camera to capture the user's facial expression. By analyzing the expressions, the system will get idea about user's emotion. After getting idea about emotion , the music player will start playing music as per the user's emotion. Simultaneously based on frequency and timing of eyes blinking and yawning the sleepy state of the user will be detected and accordingly user will be alerted.

Keywords: Feature extraction, Euclidean distance, Intel Real Sense camera, Emotion recognition.

1. INTRODUCTION

The field of science is as big as the galaxy itself. Every passing day there are new developments; if not big or shocking, but constructive and leading towards a better future. Sound and Graphics are two huge fields of Science and Engineering that not only interesting but also attract learners to study them in detail to explore into their knowledge. Since then many inventions have pull-back us to this time where thinking of various ideas which might have been difficult a few decades back and more over implementing them is now possible.

The face is an expressive organ of an human's body which makes him or her more descent and one's emotional state can be judged easily. Extraction of individual's behavior and the emotional state is perceived through senses. Music has a great impact on changing person's mood and enjoy every moment by enhancing one's life. In today's world , with increasing advancements in the field of multimedia and technology, multiple music players have been developed with features like fast-forwarding, reverse, variable playback speed ,locally playback, streaming playback with multi-streams.

Although these enhanced and user-friendly features are integrated in the current music players but still user needs to select the songs as per his requirement because these music players are not emotion specific.

The main purpose of this paper is to develop the system with more efficient features overcoming the drawback of existing music players. This system play the playlist as per current emotions and expressions of the user which are captured by using camera and then performing analysis of the viewer's expressions, the system is able to determine his/her emotion and play the music from the playlist.

[1] Computer Engineering, Bharati Vidyapeeth College of Engineering for Women, Katraj, Pune, India, *Email: kiran.yesugade@gmail.com*

[2] Computer Engineering, Bharati Vidyapeeth College of Engineering for Women, Katraj, Pune, India, *Email: guptaanjali8877@gmail.com*

[3] Computer Engineering, Bharati Vidyapeeth College of Engineering for Women, Katraj, Pune, India, *Email: shreeyapatankar@gmail.com*

[4] Computer Engineering, Bharati Vidyapeeth College of Engineering for Women, Katraj, Pune, India, *Email: desaivarsha123@gmail.com*

[5] Computer Engineering, Bharati Vidyapeeth College of Engineering for Women, Katraj, Pune, India, *Email: shwetakumbhar95@gmail.com*

2. LITERATURE SURVEY

Human's emotions can be recognized by using two different categories: Facial expression and voice recognition. Human's face plays major role in expressing any feelings. So, the accuracy rate of emotion recognition is quite good in facial expression recognition. There are many works done in this field following are some of them.

Yaun-Chin Yu et al.[5] , presented magic mirror table for smart home by using BPNN(Back Propagation Neural Network) algorithm he detected the user's emotion. They achieved an F-score of 60% when tested on six emotions. It also provides silent mode adaption if user is out of the operation range and applicable only for home members.

Mu-Chun Su et al.[4] Proposed SOM-based Facial Expression Recognition System by using Viola and Jones used for face detection and SOM is used for emotion recognition. Self-organizing feature maps(SOM) technique is simple to use and efficient for detection of face from image. The principal aim of SOMs is to transition of patterns from arbitrary dimensionality into the responses of 1 or 2 dimensional lists of neurons and also transform adaptively in a topological fashion. Monika Dubey et al.[1], presented a review on Automatic Emotion Recognition using Expression on face. This paper includes introduction of facial emotion recognition system, Application, comparative study of popular face expression recognition techniques & phases of automatic facial expression recognition system.

M. Pantic et al.[8] ,presented Dynamics of facial expression recognition of facial actions and their temporal segments using facial expression dynamics analysis , particle filtering, rule-based reasoning ,spatial reasoning ,temporal reasoning from face profile image sequences. An overall of 87% of recognition rate is achieved. M. Chung et al.[7], proposed Cognitive face analysis system for future interactive TV in which Ada-LDA learning algorithm is used. This paper, introduces a novel architecture of the future interactive TV and propose a real-time face analysis system that can detect and recognize human faces and even their expressions, and therefore understand their internal emotional states. M. Panticet et al.[9], presented a analysis of facial expressions which uses Facial Action Coding System (FACS).A system shows that in real time would it is possible to form a big step in achieving a human-like interaction between man and machine.

P. Zhou et al.[6], presents Music therapy on basis of the heart rate variability. As the music treatment goes on, subjective shows sign of significant relaxation. It shows that experiments were performed upon 16 undergraduates and 4 middle-aged women and as the music treatment goes on, subjective show significant relaxation. Guojun Yang et al.[2],presents a system for Eye Tracking using Monocular Camera and detail study of gaze estimation applications which gives valuable information in terms of human behavior. It requires the system to extract the information of the human eye location then estimate the direction of the gaze. Ismail Shaykha et al.[3], shows Non-Intrusive Facial Expression and Emotional Recognition for driver's proper monitoring using FEER. The closure rate and frequency of the eyes is then combined with the rate and frequency of yawning in a weighted combination.

3. METHODOLOGIES

3.1. Intel Real sense 3D camera

The Intel® Real Sense Camera SR300 is a subassembly product that implements a short range, coded light 3D imaging system.The small size of the SR300 subassembly provides system integrators flexibility to design into a wide range of products.The broad range of 3D mode configurations and synchronization capabilities of the SR300 enable the product to be an optimal solution for 3D imaging applications.

3.1.1. Features Of the Camera

- Onboard Imaging ASIC
- Depth Capture from 0.2 to 1.5m

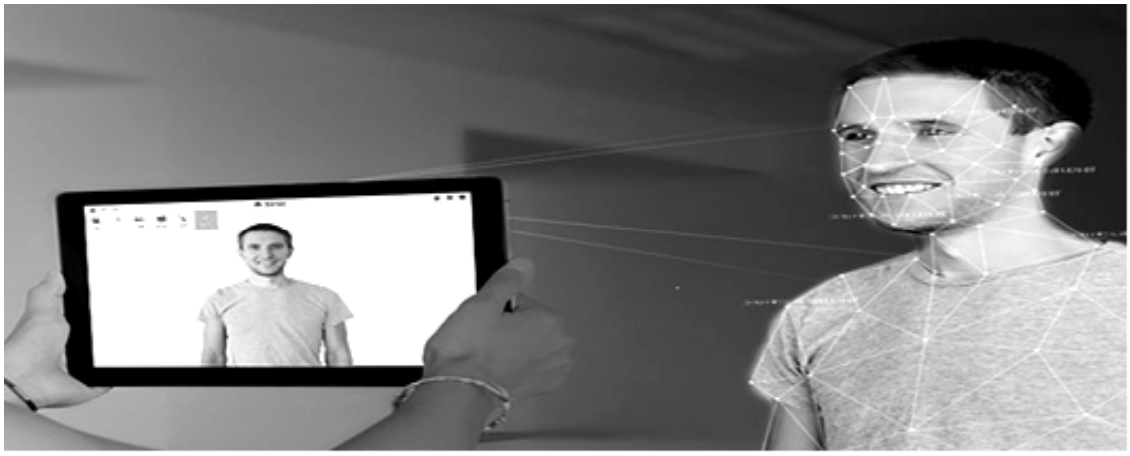

Figure 1: Intel Real Sense camera with user

- Infrared (IR) Laser Projector System
- Synchronized Depth, Colour, Infrared Video
- Texture Mapping of Depth to Colour
- Depth Unprojection to World Coordinates
- Up to 60FPS Depth at 640 × 480 (VGA)
- Up to 30FPS Colour at 1920 × 1080 (FHD)
- Up to 200FPS Infrared at 640 × 480 (VGA)
- 110mm Width × 12.6mm Height
- 4.1mm Maximum Thickness (3.8–4.1mm)
- Green Activity LED
- Class 1 Laser Compliant
- Skype* 2.0

3.1.2. Applications

- Face Analytics and Tracking
- Scanning and Mapping
- Scene Segmentation
- Hand and Finger Tracking
- Augmented Reality

3.2. Visual studio

MVS is an Microsoft's integrated development environment (IDE) for Visual Studio. VS code is available on a cross platform ,it aims to provide developer tools and services for any platform and any language. IDE, Develops, code editor and more. Also there is built-in syntax highlighting and matching code in the IDE and easy code navigation .It also provides friendly prompts on language features but also smart Intellisense that is local context in custom code. In-built languages contain C, C++ and C++/CLI , VB.NET (via Visual Basic .NET) also F# (as of Visual Studio 2010) and C# (via Visual C#).

3.3. C#

The language is strongly-typed, object-oriented, and component oriented and utilizes a unified type system. Unlike C or C++, C# handles memory and resource management for developer, lending to the concept of managed code. C# uses a garbage collection mechanism to release memory and resources that are no longer referenced in the application code, helping to prevent memory leak issues.

- Strongly-typed - the languages enforces type checking on objects in the code meaning it is type-safe.
- Object-oriented - C# offers the developer all the tenets of OOP such as encapsulation, inheritance, and polymorphism.
- Component-oriented - C# permits the creation of software components for self-contained, self-describing packages of functionality.
- Unified type system - all C# types, from primitive to reference types, inherit from a single root known as Object.

3.4. Serialize database

In the reference of data storage, serializing data means storing the data inside TEXT/BLOB columns. Serialization is executed by CLR (Common Language Runtime) to save an object‘s current state information to a temporary such as ASP.NET cache or permanent storage like file, database, etc. so as to be used later to update an object with this same information. For example, in a file or memory buffer, or transferred across a network connection link.

3.5. Euclidean distance

The Euclidean distance or Euclid an metric is the “ordinary” distance between two points in Euclidean space. Using this distance, Euclidean space becomes a metric space. In general, for an n-dimensional space, the distance is

$$d(p,q) = \sqrt{(q1-p1)^2 + (q2-p2)^2 + \ldots + (qn-pn)^2}$$

$$= \sqrt{\sum_{i=1}^{n} (qi-pi)^2}$$

Euclidean distance is the only metric that is the same in all direction, that is, rotation invariant. This fits very nicely with the general qualities of our universe, which is also rotation invariant. All other metrics are dependent on how the coordinate system is rotated to be meaningful.

4. PROPOSED SYSTEM

The proposed architecture in this paper basically involves an music player system which plays the music according to detected emotional state of user.

The proposed system architecture involves following steps:-

- The first step will be user registration with login and password. In that user can manage his own playlist according to his choice.
- The playlist will be categorized according to user’s emotion like happy, sad, joy, anger, etc.
- In this proposed system, the real sense camera captures image of user. From the captured image the face will be detected. The detected face provides the landmark points like eyes, eyebrows, mouth , lips, nose , etc.

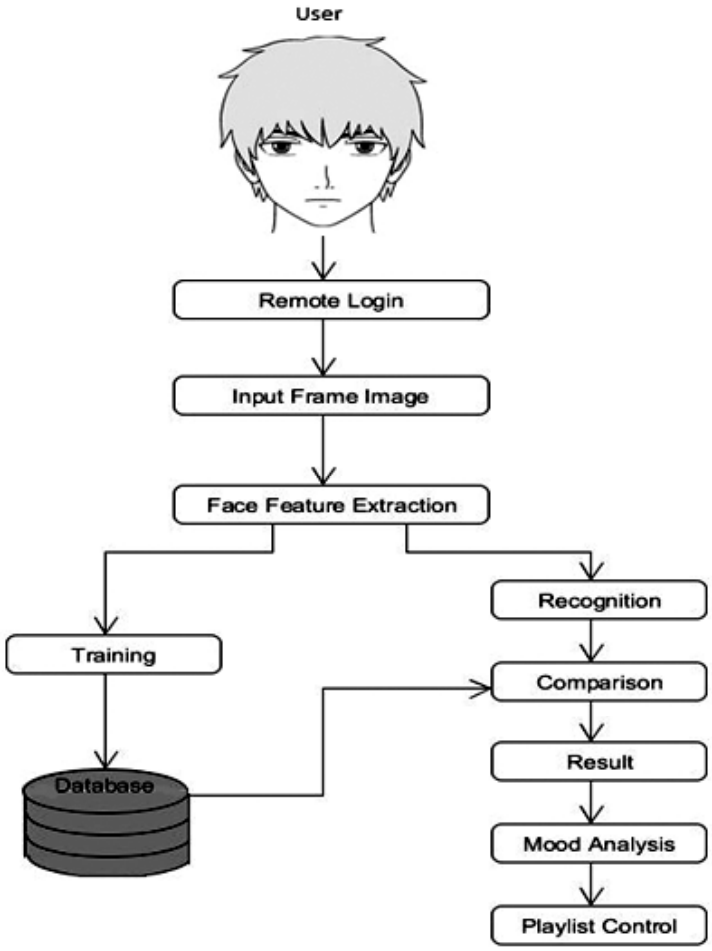

Figure 2: Architecture of proposed system

- In the trained database we have set of landmark points stored for various emotions such as happy, sad, joy as a default set of vectors.
- By using Euclidean distance, the distance between the two points is calculated. Then this distance is compared with the database and accordingly emotion will be recognized.
- After recognition the resultant mood will be compared with the song database and song will be played from the playlist.
- Also in the same system there is provision of eye feature extraction where sleepy eyes are detected.
- Similar to the mood analysis sleep analysis is done using the same system.

The system consists of following modules:

4.1. Image Acquisition

Dynamic image or image sequences are used for facial expression recognition.3-D facial image is most popular for facial image recognition. For image acquisition Intel Real Sense Camera is used.

4.2. Pre-processing

Pre-Processing plays the main role in overall process. Pre-Processing involves default set storage of noiseless and smoothed image's vector points as trained database. Later these vector points are used for comparison purpose.

4.3. Feature Extraction

Image sequence consists of input images that has been detected using Intel Real Sense Camera where the facial expression information is extracted. Geometrical and appearance features extraction are the two method mainly used in facial feature extraction. Mouth , eyes , eyebrows, nose and cheeks are the intransient facial features of different shapes and sizes that are extracted in geometrical way. In the same way the system extracts geometrical features from the real time input.

4.4. Recognition

Recognition stage follows the output of feature extraction stage. Recognition stage identifies the facial image and grouped them according to certain classes and help in their proficient recognition. Recognition is a complex process because it may get affected by many factors.

4.5. Comparison and mood analysis

Comparison is done on basis of the landmark points from the current input image and the stored dataset . The displacement of each landmark point is calculated by the difference between its real time input and stored image's landmark point. After that specific mood can be analyzed.

5. MATHEMATICAL MODEL

Let S = {I, FP, FE, Fu, So, PL}

I = {I1, I2, ……, In}
where 'I' is set of input images, these images are captured through the Real sense 3D camera.

FP = { FP1, FP2, ……..., FPn}
where 'FP' is Feature points of face. This is the finite set with total 76 feature points.

FE = {FE1, FE2,…….FEn}
where 'FE' is finite set of facial expressions like happy, sad, angry, sleepy etc.
Our system is going to recognize facial expressions based on the comparison of feature points.

Fu={login(), featureExtraction(), recognition(), compare(), managePlaylist()}

'Fu' is the set of functions given below-

login()- this function takes user ID and password as an input, if these entries are correct then user is able to login successfully in the system.

featureExtraction()- this function takes detected face image as an input and extract all total 76 feature vector points and produces it as an output.

recognition()- it takes feature points as an input and recognizes all these points.

compare()- this function compares the recognized feature points with the predefined dataset to determine facial expression.

managePlaylist()- this function plays the songs according to user's detected mood.

So = { So1, So2, ……, Son}
where 'So' is the set of songs of various categories like happy songs, sad songs, alert ringtones etc.

PL = playlist
Playlist plays the song from the categorized songs.

6. CONCLUSION AND FUTURE SCOPE

The Music Player has changed in various different ways since it was early introduced. Today different users like to get more out of different softwares and applications, so the designing phase of applications and the mind thought behind it has changed. The users prefer more interactive & user friendly yet simple to use applications.

Human face is the most expressive organ , From face one's feelings are understandable to other one. However machine is not so smart that it will understand these emotions and so to make human machine interaction more efficient this system is proposed in which music player is emotion based. such a system requires to recognize Facial Expression based on images captured or based on input video. Later human face is detected and the facial extraction of features is done. Based on these features the system computes the comparison result from input image landmark points and trained database points. Mood analysis is done in order to play the desired music based on user's mood. Finally based on these statistics the system concludes the person's emotional state and same goes for the sleep detection and this avoids the risk of prone accidents.

The future scope regarding system would to develop a mechanism where the songs are segregated into different playlists based on the feature extraction process. Therefore lists of similar sounding songs or songs belonging to similar genres are generated. Using such a system, great advantage can be provided to users who are expecting music to be played based on their mood. It will help in reducing the time required for segregating and searching the suitable song as per mood, in the playlist thereby increasing the efficiency of the music player.

REFERENCES

[1] Dubey, Prof. Lokesh Singh, "Automatic Emotion Recognition Using Facial Expression: A Review", International Research Journal of Engineering and Technology, vol. 03, 2016.

[2] Guojun Yang and Jafar Saniie ,IEEE , "Eye Tracking using Monocular Camera for Gaze Estimation Applications", 2016.

[3] Ismail Shaykha, Ahmad Menkara, Michel Nahas, Milad Ghantous, "FEER: Non-Intrusive Facial Expression and Emotional Recognition for driver's vigilance Monitoring", Lebanese International University, Beirut, Lebanon, September 2015.

[4] Mu-Chun Su, Chun-Kai Yang, Shih-Chieh Lin,De-Yuan Huang, Yi-Zeng Hsieh, andPa-Chun Wang, "An SOM-based Automatic Facial Expression Recognition System", International Journal on Soft Computing, Artificial Intelligence and Applications (IJSCAI), Vol. 2, No. 4, August 2013.

[5] Yaun-ChinYu,Member,IEEE,Shingchern D. You,and Dwen-Ren Tsai, "Magic mirror table for Social emotion alleviation in the smart home", IEEE transactions on Consumer Electronic,vol. 58, No. 1, FEB 2012.

[6] P. Zhou, F.-F. Sui, A. Zhang, F. Wang, and G.-H. Li, "Music therapy on heart rate variability," 3rd International Conference on Biomedical Engineering and Informatics (BMEI),vol. 3, pp. 965–968, 2010.

[7] H.A. Kwang, and M. Chung, "Cognitive face analysis system for future interactive TV," IEEE Transactions on Consumer Electronics vol. 55, no. 4, pp. 2271–2279, 2009.