

# Voice and Face Detection Robot

<sup>1</sup>Ms. Prerana Gangadhar Shewale, <sup>2</sup>Prof. Hon .Y .S

<sup>1</sup>ME(VLSI & Embedded System), <sup>2</sup>Asst Professor, <sup>1,2</sup>Dept of Electronics and Telecommunication, SNJB's Late  
Sau. Kantabai Bhavarlalji Jain College Of Engineering, Neminagar Chandwad, Maharashtra, India.

<sup>1</sup>prerana.shewale@gmail.com, <sup>2</sup>yogitarathod@gmail.com

**Abstract - Humans show universal consistency in recognizing emotions but also show a great deal of variability between individuals in their abilities. This paper introduces a human emotion according to face detection and voice . The interaction between human beings and computers will be more natural if computers are able to perceive and respond to human non-verbal communication such as emotions. Although several approaches have been proposed to recognize human emotions based on facial expressions or speech, relatively limited work has been done to fuse these two, and other, modalities to improve the accuracy and robustness of the emotion recognition system . Inter-personal human communication includes not only spoken language but also non-verbal cues such as hand gestures, facial recognition. We express our emotion according to face and voice. It improves the overall human computer interaction. The work presents an approach toward the face emotion. And speech recognition. In This paper emotion can be detected according to Robot, give the proper result related to voice and face recognition. Face detection is a very attractive research theme and it is also important both in theoretical research and in practical application. With the deepening of the research, face detection technology has achieved a lot of remarkable results. However, there are still many shortcomings, Feature extraction of image is necessary for face detection. However, these factors can tempestuously change the feature of a persons different Images.**

**Keyword - Detection of Emotion Information, voice Emotion Recognition, Face Emotion Recognition.**

## I. INTRODUCTION

Emotion recognition is the process of identifying human emotion, most typically from facial expressions. This is both something that humans do automatically but computational methodologies have also been developed. In the recent year, voice signal processing and face detection by using robot is the most challenging task it perform at intelligent work. Emotion recognition is used for a variety of reasons. In this paper output produce by the robot. To work on mechanical support by using GUI to calculate the Energy, Power, Pitch, MFCC by using emotion classification according to (Happy, Sad, Normal). The face detection work can be calculate by using LBP (Local Pattern) feature extraction and NMD (Normalized Minimum Distance ) to the classification to detect the face and visualize clear face. This work can do by the input of the software and output produce by the hardware.

In this age of modern technology , there is no field / sector which is free from the use of robotics. In 20<sup>th</sup> century , the concept of emotional intelligence is come across from questions arise in front of humans such as terrorism , cyber crime. So there is need of such a robot with emotional intelligence which will help humans handle human emotions.

## II. RELATED WORK

### 1. Emotion Recognition Using Face

Facial expressions give important clues about emotions. Therefore, several approaches have been proposed to classify human affective states. The features used are typically based on local spatial position or displacement of specific points and regions of the face, unlike the approaches based on audio, which use global statistics of the acoustic features. For a complete review of recent emotion recognition systems based on facial expression the readers are referred to .The calculations for recognizing by using the face to the LBP however this method can be describe using the overlapping region with multiple training image. This method can describe single training image per class Image can be converted into the pixel form to produce clarity of the image the process according image pre-process then to normalized face detection find the LBP to clear image. Pixel-based background subtraction technique.

### 2. Emotion Recognition Using Speech

Several approaches to recognize emotions from speech have been reported. A comprehensive review of these approaches can be found in and . Most researchers have used global supra segmental/prosodic features as their acoustic cues for emotion recognition, in which utterance-level statistics are calculated. For example, mean, standard deviation,

maximum, and minimum of pitch contour and energy in the utterances are widely used features in this regard human emotions by the use of pitch-related features. They implemented three different classifiers: Maximum Likelihood Bayes classifier (MLB), Kernel Regression (KR), and K-nearest Neighbors (KNN). The calculation for recognizing by using speech feature extraction using the Energy, Power, Pitch, MFCC by using emotion classification according to (Happy, Sad, Normal). To calculate the all values and finding the emotion.

System based on speech The most widely used speech cues for audio emotion recognition are global-level prosodic features such as the statistics of the pitch and the intensity. Therefore, the means, the standard deviations, the ranges, the maximum values, the minimum values and the medians of the pitch and the energy were computed using Praat speech processing software. In addition, the voiced/speech and unvoiced/speech ratio were also estimated. By the use of sequential backward features selection technique, a 11-dimensional feature vector for each utterance was used as input in the audio emotion recognition system.

### III. SYSTEM DESIGN

In this paper. Approach of this project face recognition by using the images like as Happy, Sad, and Normal. So the identification of this faces according to the way

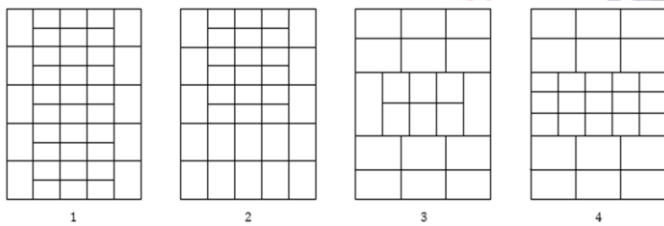


Fig.1 LBP can be assign in greed pattern

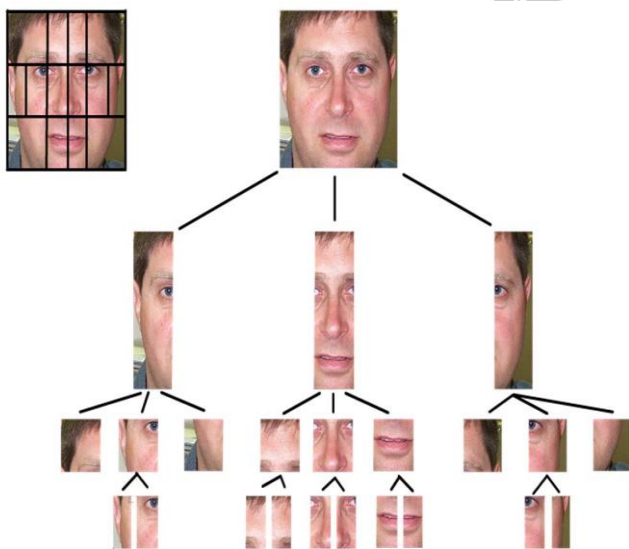


Fig.2. Image Division Pattern

In this figure it can divided to the part of subpart. And can be found the image classification. Speech recognition Speech data is an recorded data of an recorded data of an audio input device to finding the sample rate is 600.

#### a) Block Diagram

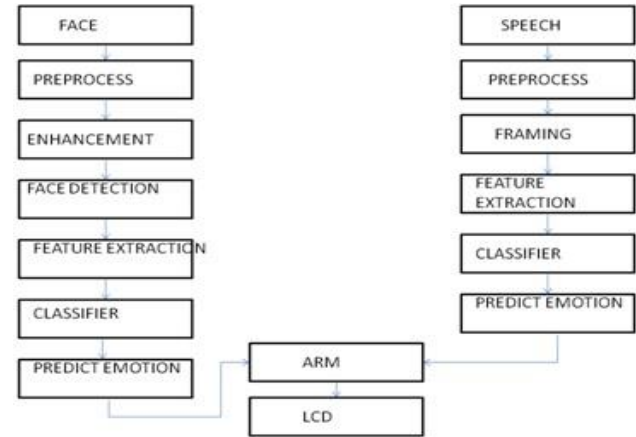


Fig. 3. Block Diagram

#### 1) Pre-process

Pre-process is the process is to improve the image data suppressed the undesired distortion as well as enhancement some image feature which is important to some process. And it including some colour images, in which taking the pixel value and dividing each pixel. The aim of **pre-processing** is an improvement of the **image** data that suppresses unwanted distortions or enhances some **image** features important for further **processing**. The pre-processing stage in speech recognition systems is used in order to increase the efficiency of subsequent feature extraction and classification stages and therefore to improve the overall recognition performance. Commonly the pre-processing includes the sampling step, a windowing and a de-noising step. At the end of the pre-processing the compressed and filtered speech frames are forwarded to the feature extraction stage. The general pre-processing pipeline is depicted in the following figure.

#### 2) Enhancement of Face

This process can be consist of direct observation of colour image, To improve the image feature according to colour. The pre-processing stage in speech recognition systems is used in order to increase the efficiency of subsequent feature extraction and classification stages and therefore to improve the overall recognition performance. Commonly the pre-processing includes the sampling step, a windowing and a de-noising step. At the end of the pre-processing the compressed and filtered speech frames are forwarded to the feature extraction stage. The general pre-processing pipeline is depicted in the following figure. By using Fourier transform, the resulting spectrum representation of face image is used as the feature of the face image for recognition.

#### 3) Feature Extraction

The face can be extracted by using sample in image processing the image can be classified using the Color, Shape, and Texture. The FFT can be used as the extract the spectral component, Feature extraction involves reducing

the amount of resources required to describe a large set of data. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power, also it may cause a classification algorithm to over fit to training samples and generalize poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. Feature plays a very important role in the area of image processing. Before getting features, various image preprocessing techniques like binarization, thresholding, resizing, normalization etc. are applied on the sampled image. After that, feature extraction techniques are applied to get features that will be useful in classifying and recognition of images. Feature extraction techniques are helpful in various image processing applications e.g. character recognition. As features define the behavior of an image, they show its place in terms of storage taken, efficiency in classification and obviously in time consumption also. Here in this paper, we are going to discuss various types of features, feature extraction techniques and explaining in what scenario, which features extraction technique, will be better. Hereby in this paper, we are going to refer features and feature extraction methods in case of character recognition application.

#### 4) Classifier

Classifier is determine the feature extraction of the given sample .Means depending on moment .And image can be classified into a small slot because this image can clearly visible to Robot and easy to detect the emotion.

As every time we face a complex system, we have to divide it into simpler sub-models. Here we have: Webcam image feed; Image classification; Text 2 Speech. We are going to program and test each of the components separately, and then join them. An essential step in understanding the processes underlying the general mechanism of perceptual categorization is to identify which portions of a physical stimulation modulate the behavior of our perceptual system. More specifically, in the context of speech comprehension,

#### 5) Experiment

In project we have to used the database by using the person to person and to calculate the performance of the system

LCD display :

The flight number is sent by the pilot at the time of sending request to land. This no. is display on LCD at ATC. This number is extracted from the message and then sent to the specified gate and then is displayed on the LCDs in the different gate where the passengers need to go to their

baggage. At last output can detect by using LCD display to detect the emotion by using face and speech by robot.

Application

In military application , In hospital In industry

#### 6) Result analysis

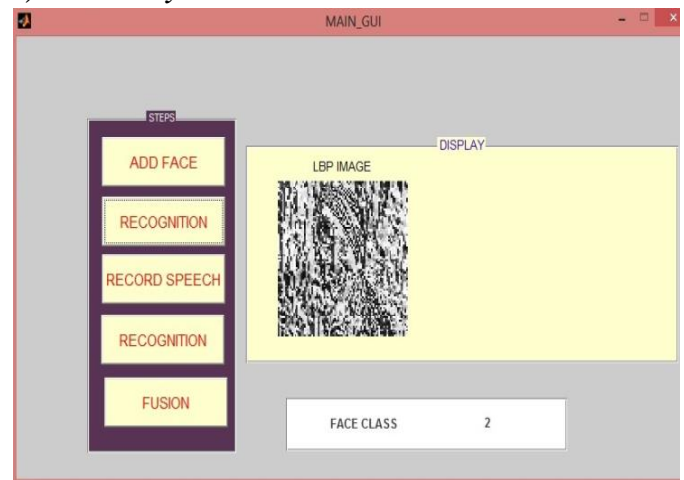


FIG 4. Main Graphics User Interface output display

Output display shows 5 blocks provided output step by step related to face image and speech,

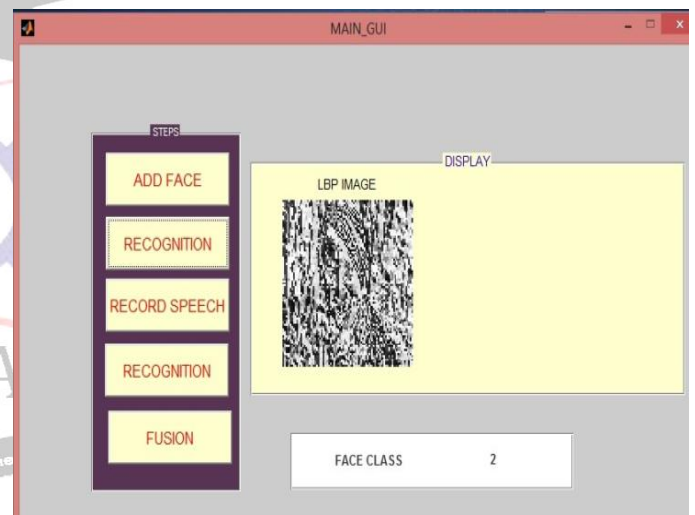


FIG 5. LBP image

In this window ,output display image converted to into LBP image, it produce clear output.

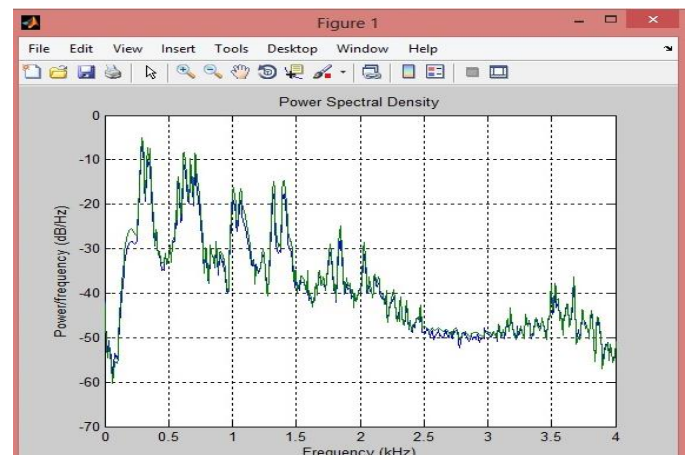


Fig 6. PSD of given signal

In this image produce power spectral density signal to recorded speech and to produce proper output.

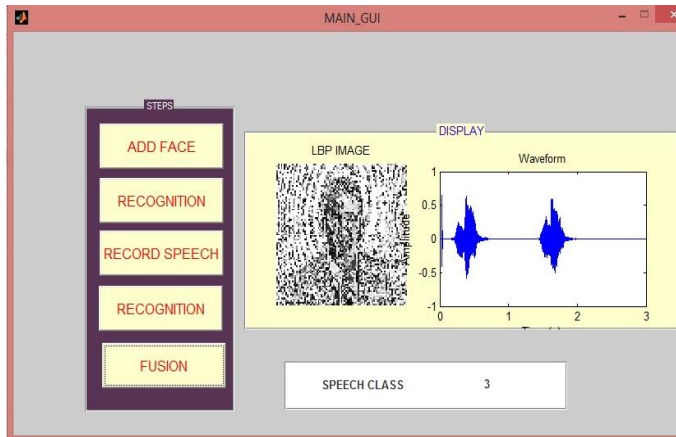


FIG 7. Speech Signal

In this window produce combine output of face and speech recognised and given to the proper result.

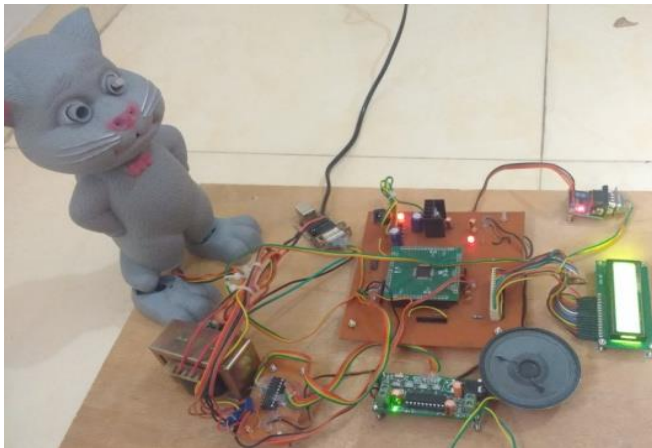


FIG 8. Hardware implementation of Robot relates to the voice and face detection.

#### IV. CONCLUSION

The results presented in this research show that it is feasible to recognize human affective states with high accuracy by the use of audio and visual modalities. Therefore, the next generation of human-computer interfaces might be able to perceive humans feedback, and respond appropriately and opportunely to changes 210 of users affective states, improving the performance and engagement of the current interfaces. In this paper we observe that to reduce the human efforts, we make emotional detection robot by using voice processing and face detection then to give the combine output produce display window. This window consist combine output and depend on output to produce proper result. This result is depend on the frequency of the voice and face image according to voice frequency it detect the emotion when it happy, sad or neutral this output given to the robot and to required proper output. To reduce human efforts like as Medicine : To help monitoring Monitoring : car driver – detect state and alert other car. Call- centre system – recognize voice , detect anger in speech. Music therapy – help patients deal with stress and depression and positive effective on Alzheimer.

#### REFERENCES

- [01] Alireza Zourmand, Hua-Nong Ting, and Seyed Mostafa Mirhassani, "Gender Classification in Children Based on Speech Characteristics: Using Fundamental and Formant Frequencies of Malay Vowels" Kuala Lumpur, Malaysia.
- [02] Hadi Harb, Liming Chen LIRIS CNRS FRE, "VOICE-BASED GENDER IDENTIFICATION IN MULTIMEDIA APPLICATIONS", 267 Dept. Mathématiques Informatique, Ecole Centrale de Lyon Ecully Cedex, France
- [03] Haizhou Li, Kong Aik Lee, Bin Ma, "Spoken Language Recognition: From Fundamentals to Practice", IEEE.
- [04] Md. Rashidul Hasan, Mustafa Jamil, Md. Golam Rabbani Md. Saifur Rahman, "SPEAKER IDENTIFICATION USING MEL FREQUENCY CEPSTRAL COEFFICIENTS" 3rd International Conference on Electrical & Computer Engineering ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh. ISBN 984-32-1804-4 565.
- [05] Tobias May, Steven van de Par, and Armin Kohlrausch, "Noise-Robust Speaker Recognition Combining Missing Data Techniques and Universal Background Modeling", IEEE TRANSACTIONS ON AUDIO, SPEECH, AND LANGUAGE PROCESSING, VOL. 20, NO. 1, JANUARY, 2012.
- [06] Santosh V. Chapaneri, "Spoken Digits Recognition using Weighted MFCC and Improved Features for Dynamic Time Warping", International Journal of Computer Applications (0975 – 8887) Volume 40– No.3, February 2012.
- [07] Pa Pa Win 1, Su Su Yi Mon, "Speech De-noising and Compression using Wavelet Transform", International Journal of Electronics and Computer Science Engineering V3-N2, pp 101-107.
- [08] T Subetha, Dr.S.Chitrakala, "Recognition of Human-Human Interaction using CWDTW", International Conference on Circuit, Power and Computing Technologies, 2016 .
- [09] Deepjoy Das, Alok Chakrabarty, "Emotion Recognition from Face Dataset Using Deep Neural Nets", Department of Computer Science & Engineering National Institute of Technology Meghalaya Meghalaya.
- [11] Supriya D. Kakade, "A Review Paper on Face Recognition Techniques", International Journal for Research in Engineering Application & Management (IJREAM), Vol-02, Issue 02, MAY 2016, pp 01-04.
- [12] Mamta Kumari, Israj Ali, "An Efficient Algorithm for Gender Detection using Voice Samples", 2015 International Conference on Communication, Control and Intelligent Systems (CCIS), KIIT University, Bhubaneswar, India.