A

Dissertation Report on

Gesture Recognition for Dumb People

Submitted by

Ms. Manvita Asnodkar

(Exam Seat Number 2010MEN009)

Under The Guidance of

Dr. Nilesh Bhaskarrao Bahadure

For The Award of The Degree of

Master of Technology

in

Electronics Engineering

DEPARTMENT OF ELECTRONICS ENGINEERING

SANJAY GHODAWAT UNIVERSITY, KOLHAPUR

This Project is Sponsored by

The Institution of Engineers (India)

Kolkata

February 2020

CERTIFICATE

The report titled

Gesture Recognition for Dumb People

duly completed by

Ms. Manvita Asnodkar
(Exam Seat Number 2010MEN009)

as a partial fulfillment of the requirements for the degree of

Master of Technology
in
Electronics Engineering

from
SANJAY GHODAWAT UNIVERSITY, KOLHAPUR

Dr. Nilesh Bhaskarrao Bahadure
Project Guide

Dr. Nilesh Bhaskarrao Bahadure
Head of Department

Prof. Dr. M. T. Telsang
Academic Dean

External Examiner

Date: March 7, 2020
Place:

DEPARTMENT OF ELECTRONICS ENGINEERING,
SANJAY GHODAWAT UNIVERSITY, KOLHAPUR

Declaration

I undersigned have submitted the synopsis for the proposed project work entitled ”Gesture Recognition for Dumb People” declare that we have submitted the synopsis after through study & is not copied from some source.

Name **Sign**

Ms. Manvita Asnodkar

Acknowledgements

I am pleased to present this dissertation report entitled to my college as part of academic activity.

I would like to express my deep sense of gratitude to my guide Guide Name for his valuable guidance, encouragement and kind co-operation throughout the dissertation work. I feel proud to present my dissertation under his guidance. I am thankful to HOD Name and Expert if any for their encouragement and support.I am also thankful to all the teaching staff and non-teaching staff for their co-operation to complete my dissertation work. Last but not the least I am very thankful to all my friends, parents and those who helped me directly or indirectly throughout this dissertation work.

Ms. Manvita Asnodkar

SGU, Kolhapur

Abstract

In our country around 2.78% of peoples are not able to speak (dumb). Their communications with others are only using the motion of their hands and expressions. We proposed a new technique called artificial speaking mouth for dumb people. It will be very helpful to them for conveying their thoughts to others. Some peoples are easily able to get the information from their motions. The remaining is not able to understand their way of conveying the message. In order to overcome the complexity the artificial mouth is introduced for the dumb peoples. This system is based on the motion sensor. According to dumb people, for every motion they have a meaning. That message is kept in a database. Likewise all templates are kept in the database. In the real time the template database is fed into a Microcontroller and the motion sensor is fixed in their hand. For every action the motion sensors get accelerated and give the signal to the Microcontroller. The Microcontroller matches the motion with the database and produces the speech signal. The output of the system is using the speaker. By properly updating the database the dumb will speak like a normal person using the artificial mouth. The system also includes a text to speech conversion (TTS) block that interprets the matched gestures.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Deaf-mute people need to communicate with normal people for their daily routine. The deaf-mute people throughout the world use sign language to communicate with other people. However, it is possible only for those who have undergone special training to understand the language. Sign language uses hand gestures and other means of non-verbal behaviors to convey their intended meaning. It involves combining hand shapes, orientation and hand movements, arms or body movement, and facial expressions simultaneously, to fluidly express speaker's thoughts. The idea is to create a sign language to speech conversion system, using which the information gestured by a deaf-mute person can be effectively conveyed to a normal person. The main aim of this work is to design and implement a system to translate finger spelling (sign) to speech, using recognition and synthesis techniques. The modules to be present in the proposed system are,

1. Finger spelling (gesture) recognition module.

2. Text-to-Speech synthesis module.

Other applications of hand-gesture recognition systems include character-recognition, gesture recognition to remotely control a television set, home au-

tomation, robotic arm controller and gesture recognition for wheel chair control, games. In a vision-based gesture recognition system, a camera is used for capturing the image/video of the gesture. A person should know the sign language to understand the language; this becomes complicated when a person who has inability to speak or hear wants to convey something to a person or group of persons, since most of them are not familiar with the sign language. The development of the most popular devices for hand movement acquisition, glove-based systems started about 30 years ago and continues to engage a growing number of researchers. Communication involves the exchange of information, and this can only occur effectively if all participants use a common language. Sign language is the language used by deaf and mute people that uses gestures instead of sound to convey or to express fluidly a speaker's thoughts. A gesture in a sign language is a particular movement of the hands with a specific shape made out of them. The main aim of this paper is to present a system that can efficiently translate Sign Language gestures to both auditory voice and text. Several languages are being spoken all around the world. So this system aims to give the voice output in various regional languages [**?**, **?**].

Each typical person sees, tunes in and after that responds to the circumstances by talking himself out. Individuals, fundamentally the hard of hearing and the imbecilic, depend on some kind of gesture based communication for imparting their emotions to others. For the most part idiotic individuals utilize gesture based communication for correspondence, yet they discover trouble in speaking with other people who don't comprehend communication via gestures. Thus, there is a boundary in correspondence between these two groups. This venture intends to lower this obstruction in correspondence. It turns into the issue for two people who knows two diverse dialects, so it turns into an issue to chat with each other thus they require an interpreter physi-

cally which may not be constantly helpful to orchestrate and this same sort of issue happens in the middle of the Normal Person and the Deaf individual or the Normal Person and the Dumb individual. The fundamental point of the venture is to build up a financially savvy framework which can offer voice to voiceless. With the proposed work is signaled are changed over into discourse. It implies that correspondence boundary between two unique groups can be made productive [**?**].

Communication is the best media used by the people to communicate with each other. The problem arises when normal people and deaf-dumb people want to communicate with each other. Sign Language is a language which is used for communication by the deaf and dumb people. This project is used to reduce the communication barrier between the deaf-dumb people and the normal people. The Sign language interpreter developed uses a hand glove fitted with flex sensors that can interpret the English letters, numbers and some words in American Sign Language (ASL) & some one-handed letters in Indian sign language (ISL).

The proposed system is sign language recognition system for the vocally disabled people who use sign language for communication. The use of specially designed sensor gloves connected to system while a disabled person wearing the gloves makes the signs and gestures. System will analyze these gestures and synthesize the sound for the corresponding word or letter for normal people to understand. Proposed system is to utilize single handed gestures for American Sign Language (ASL) and some one handed gestures for Indian sign language (ISL) for implementation of Sign Language Interpreter [**?**].

The need of this system is to give output in day to day life for ”Image Processing Based Sign to Speech Converter for Dumb People” using PCA algorithm. It will explain the aim and whole declaration for the evaluation of system. It will also explain system constraints, interface and interactions with other

external applications. An attempt has also been made to explore about the need and motivation for interpreting ISL, which will provide opportunities for hearing impaired people in industry. The aim of the proposed project is to overcome the challenge of skin color detection for natural interface between user and machine. This project is developed for the physically impaired people and would be beneficial as they can communicate with everyone. In our system a web-cam is placed in front of the physically impaired person. The physically impaired person will place a finger with particular action in front of the camera .When he makes the gestures, the webcam will capture the exact positions of the fingers and perform image processing using principle component analysis algorithm. The co-ordinates captured will be mapped with the one previously stored and accordingly exact image from the database will be identified. Continuing in this way physically impaired person will be able to go through the entire sentence that he wants to communicate. Later on this sentence will be translated into speech so that it would be audible to everyone. By using this system the physically impaired people would be benefited as they can communicate with everyone freely which indeed would be great achievement for the mankind [**?**].

Sign language is the language used by deaf and dumb people and it is a communication skill that uses gestures by hand movements instead of sound to convey, shapes, orientations and movement of the hands, arms or facial and body expressions to express continuously thoughts a speaker's. Signs are used to communicate words and sentences to audience. A gesture sign language at particular movement of hands with a specific shape made out of them. A sign language usually provides sign for entire words. Thus for all letters It provide sign to perform words that donâĂŹt have corresponding sign in that sign language. In this project Flex Sensor and accelerometer sensor plays the major role, Flex sensors are sensors that depends on the amount of bends when it

varies in resistance. In this project data glove is implemented to capture the hand gestures of a user [**?**]

Figure 1.1 Kinect sensor

Specifications of kinect sesnor is listed below:

1. Brand: Microsoft
2. Model: Kinect Sensor for Xbox 360 (Free Kinect Sports)
3. Color: Black
4. Category: motion controller
5. Plateform: Xbox
6. Interface: Wired
7. Sales Package: Kinect Sensor for Xbox 360, Kinect Sports Game

The Kinect contains three vital pieces that work together to detect your motion and create your physical image on the screen: an RGB color VGA video camera, a depth sensor, and a multi-array microphone. The camera detects the red, green, and blue color components as well as body-type and facial features. This little black box contains multiple features that allow gamers to

be the controller and feel as if they are actually in the game. When we heard about this sensor giving players' the ability to transform their entire bodies into a game controller, we were convinced that we just had to investigate what really goes on inside of this revolutionary device.

1.1 Hardware Description of Kinect Sensor

The Kinect contains three vital pieces that work together to detect your motion and create your physical image on the screen: an RGB color VGA video camera, a depth sensor, and a multi-array microphone.

The camera detects the red, green, and blue color components as well as body-type and facial features. It has a pixel resolution of 640x480 and a frame rate of 30 fps. This helps in facial recognition and body recognition.

The depth sensor contains a monochrome CMOS sensor and infrared projector that help create the 3D imagery throughout the room. It also measures the distance of each point of the player's body by transmitting invisible near-infrared light and measuring its "time of flight" after it reflects off the objects.

The microphone is actually an array of four microphones that can isolate the voices of the player from other background noises allowing players to use their voices as an added control feature.

These components come together to detect and track 48 different points on each player's body and repeats 30 times every second.

Putting both hardware and software together give the Kinect the ability to generate 3D images and recognize human beings within its field of vision. It can analyze the person in front of it and go through multiple "filters" to try and determine which type of body structure matches with the correct type programmed in its system. Data is constantly being transferred back and forth between the Kinect and the objects in its field of vision while you simply

enjoy the fun of being a character in a game, without holding anything in your hands.

1.2 Working of Kinect Sensor

Microsoft Kinect is poised to shake up the video game console experience. Announced and demonstrated as Project Natal in June 2009, Kinect seems almost magical the way it can ”see” every movement of your body and reproduce it within the video game you’re playing. Plus, it recognizes your face and voice so it can pick you out in the room and know who you are, even if you’re playing with a group of friends. November 2010 marks Kinect’s historic and anticipated release as a new addition to Microsoft’s Xbox 360 product line.

As it turns out, Kinect isn’t magical at all. It’s a highly innovative combination of cameras, microphones and software that turns your body into the video game controller. The name Kinect is inspired by the words ”kinetic,” which means to be in motion, and ”connect,” which means it ”connects you to the friends and entertainment you love” . It’s not just the games that get you moving, either: Kinect turns your Xbox 360 into a voice-activated console with video capturing and facial recognition, applicable for everything from selecting a TV show to creating digital artwork.

Microsoft has also designed Kinect as an enhancement to the Xbox Live experience. Xbox Live Video Chat makes use of Kinect’s cameras and microphones for a webcam-like live chat with multiple friends at once . Plus, Microsoft teamed with ESPN to create a live interactive sports experience for Xbox Live Gold users. Not only can you watch sporting events in HD from ESPN, but you can also ”join” other fans in rooting for your favorite team and answering sports trivia questions.

This article recalls the earliest buzz around Project Natal, describes the hardware, software and development behind the project, and explores how Microsoft Kinect could change the video game console market compared to the Nintendo Wii and the PlayStation Move. The working of Kinect sensor is shown in Figure 1.2.

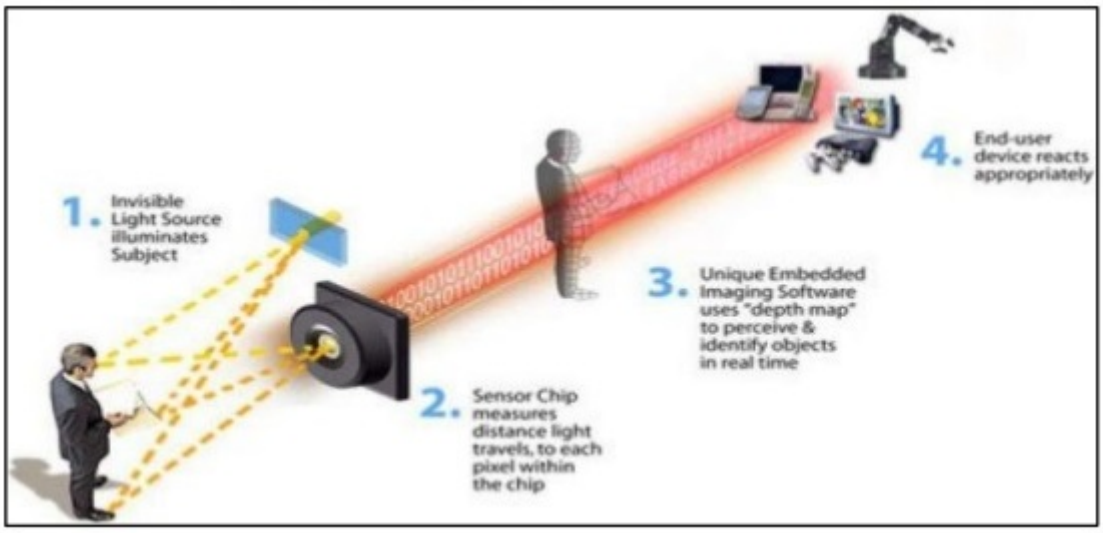

Figure 1.2 Working of Kinect sensor

Chapter 2

Literature Review

The detailed literature review has been studied and some of the most sought study is presented here. Jadhav at el.[**?**] Presented ”Hand Gesture Recognition to Speech Conversion in Regional Language”.This methodology provides a map for developing a Digital wireless glove which is fitted with Flex sensors and accelerometer.This system includes a voice playback IC to give the real time speech output. according to Author, to increase Capability they must have to increase No. of flux sensors and Accelerometer. Channaiah at el.[**?**] developed ”Hand Gestures Recognition System for Deaf and Dumb and Blind People” They uses Web Camera and Microcontroller. In this They Design a code to detect colors and to send the corresponding id’s. According to them, Their Project gives speech as output and must be analyzed and converted to textual display on the screen of ALCD. Sunita at el.[**?**] developed ”Electronic Support System for Deaf and Dumb to Interpret Sign Language of Communication” In this they used flex sensor and ARM7TDMI. According to Author, their Project will only recognize Alphabets and numbers.

V.Padmanabhan, M.Sornalatha [**?**] Presented ”Hand gesture recognition and voice conversion system for dumb people”. In this Paper they used a Flex sensor for sensing the action its output is given to microcontroller for matching

input with database. According to them this Project Overcomes the necessary Time Difficulties of Dumb People and Improves their manner. Prajapati at el.[**?**] Investigated ”Hand Gesture Recognition & voice conversion For Deaf and Dumb People”. In this Paper they are used Web Camera and PCA Algorithm. According to Author Implementation of this system gives up to 90% accuracy and works successfully in most of the test cases. Jagadish at el.[**?**] Presented ”Hand Gesture Recognition System for Deaf and Dumb Persons”. In this project they used The methodology which provides a map for developing a digital wireless glove which is fitted with Flex sensors. This system includes a voice playback IC to give the real time speech output in regional language. According to Author, This paper has the potential of minimizing this communication barrier by working as an automated translator and converting sign language directly into vocal and textual format.

Kunal at el.[**?**] investigated ”Sign Language Recognition For Deaf And Dumb People Using ANFIS” In this they used International Neural Networks, Hidden Markov Model (HMM) According to author,Their system gives maximum efficiency, having low cost, and is an optimal mixture of methods, giving results against complex backgrounds as well, should be preferred. Jain at el.[**?**] proposed ”Image Processing based speaking system for mute People using hand gesture” In this system a webcam will capture the hand gesture and perform image processing using principle component analysis algorithm (PCA). Author says, Their design is more compatible and faster responsive when compared to existing design using PCA. Pradeepa and Tharanyaa[**?**] presented ”an artificial speech system for dumb and blind people” here they used a accelerometer sensor, flex sensor & ultrasonic sensor. According to author this project will mostly helpful for deaf and dumb people. Dawanel at el.[**?**] developed ”A Review on Hand Gesture Recognition for Deaf and Dumb People Using GSM Module” In this They used flux sensor along with gloves and GSM

modem for Displaying messages. According to author, their Project is require more number of flux sensor for high accuracy.

2.1 Problem Statement

1. Communication barrier between the deaf and dumb community and the normal people.

2. The mute/deaf individuals have a communication problem dealing with other people.

3. It is hard for such individuals to express what they want to say since sign language is not understandable by everyone.

2.2 Objective

1. A prototype for checking the feasibility of recognizing gestures using image processing.

2. To Design a prototype this makes possible to convert hand gestures into speech which can be understood easily by normal people.

3. To develop a useful tool that uses gesture recognition for reducing the communication barrier between the deaf and dumb community and the normal people.

Chapter 3

Implementation of Project

The overall implementation of the project is explained in this section.

3.1 Methodology

Image processing is a method to perform some operations on an image, in order to get an enhanced image or to extract some useful information from it. It is a type of signal processing in which input is an image and output may be image or characteristics/features associated with that image. Nowadays, image processing is among rapidly growing technologies. It forms core research area within engineering and computer science disciplines too. There are two types of methods used for image processing namely, analogue and digital image processing. Analogue image processing can be used for the hard copies like printouts and photographs. Image analysts use various fundamentals of interpretation while using these visual techniques. Digital image processing techniques help in manipulation of the digital images by using computers. The three general phases that all types of data have to undergo while using digital technique are pre-processing, enhancement, and display, information extraction.

3.1.1 Image Matching Process

Image matching is a process in which the captured image is compared with the images stored in database.

Feature Extraction

Principal component analysis (PCA) is one of the statistical techniques frequently used in signal processing to the data dimension reduction or to the data de-correlation. This Project deals with two distinct applications of PCA in image processing. The first application consists in the image color reduction while the three color components are reduced into one containing a major part of information. The second use of PCA takes advantage of eigen vectors properties for determination of selected object orientation. Various methods can be used for previous object detection. Quality of image segmentation implies to results of the following process of object orientation evaluation based on PCA as well. This Project briefly introduces the PCA theory and Results are documented for the selected real pictures. Principal component analysis (PCA) belongs to linear transforms based on the statistical techniques. This method provides a powerful tool for data analysis and pattern recognition which is often used in signal and image processing as a technique for data compression, data dimension reduction or their de-correlation as well. There are various algorithms based on multivariate analysis or neural network that can perform PCA on a given data set. This project introduces PCA as a possible tool in image enhancement and analysis. ASL Fisher is a feature extraction algorithm which is an in-built function in MATLAB. It basically used to extract the Euclidean distance between two images.

Classification Method

K-Nearest Neighbor is a Learning algorithm that Defer in the decision to generalize beyond the training examples till a new query is encountered. Whenever we have a new point to classify, we find its K nearest neighbors from the training data. The distance is calculated using n Euclidean Distance. ”Support Vector Machine” (SVM) is a supervised machine learning algorithm which can be used for both classification and regression challenges. However, it is mostly used in classification problems. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate. Then, we perform classification by finding the hyper-plane that differentiate the two classes very well. Support Vectors are simply the co-ordinates of individual observation. Support Vector Machine is a frontier which best segregates the two classes.

3.1.2 System Architecture

The figure below represents the System Architecture of our system that basically show each component of the system, how the system works, and the flow of the system and so on. Images that are taken from the web camera or Kinect sensor goes under pre-processing stages to enhance the feature of an image. Then there is a removal of object and background from the images which later convert into binary form. Feature extraction and reorganization helps to match the images that is stored in database and we get the desired output in the form of text and converts that text to speech. The system architecture flow is shown in Figure 3.1.

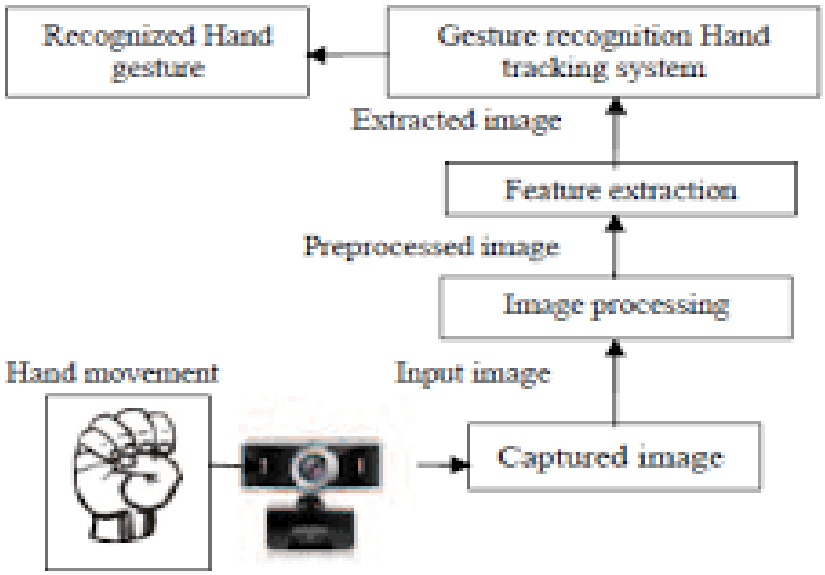

Figure 3.1 Block diagram of system architecture

3.1.3 Implementation of the System

Here we are going to discuss about how we are implementing our system and is represented in a flowchart manner in Figure 1.

Training of System

User have to enter number of samples to store in the database. The number of samples should be more than 5 in order to get better accuracy. User have to select the folder where the images will get saved. Click on start video to open the web camera or Kinect sensor in order to start the process of database creation. Click capture image to store the number of images in the training folder as per the no. of sample specified. When the number of images will be equal to number of captured image then done storing will get displayed that means database creation is done successfully.

Image Pre-Processing

The captured images go under pre-processing stage in order to enhance the property of an image. Pre-Processing is basically done to remove the object and background of an image and focus on the hand gestures only. The pre-processed image is then represented in the form of black and white pixels which basically means binarized image.

Feature Extraction and Recognition

PCA algorithm is used in order to extract the feature of an image. PCA algorithm is applied on the captured images in order to extract the best featured image from the database. PCA converts the images into some independent linear set of variables which refers to the information in the original data which is referred as principal components. Following are the steps that is used to extract the feature of an image using PCA:
Step 1: Convert all images into the column matrix.
Step 2: Evaluate the Mean column matrix of column matrix.
Step 3: Calculate the difference for each vector set.
Step 4: Calculate a Covariance matrix.
Step 5: Calculate Eigen value and mean Eigen value for Covariance matrix.
Step 6: Sort the Eigen value.
Step 7: Calculate mapping eigenvectors and for matching, project the data.
After applying all the above steps, we get the reduced features of PCA and then calculate the result. After feature Extraction, Recognition of an image is done by KNN and SVM algorithm. Recognized hand gesture is converted into text which converts into speech.

The overall working of the system is shown in Figure 3.2

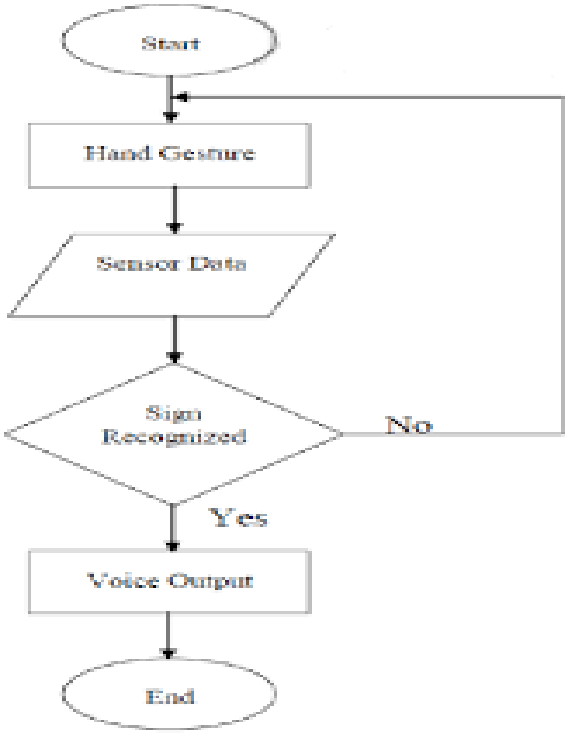

Figure 3.2 Flowchart

3.2 Advantages of the System

1. Easily Understandable
2. Low cost
3. Reduce the Barrier between the normal person and disable People.

3.3 Disadvantages of the System

1. Disturbances add Fastly.
2. Require Calm Environment.
3. Not useful in crowdy Places.

3.4 Expected Outcome/Result

Sign language may be a helpful gizmo to ease the communication between the deaf or mute community and additionally the standard people. This project aims to lower the communication gap between the mute community and additionally the standard world. The projected methodology interprets language into speech. The system overcomes the necessary time difficulties of dumb people and improves their manner. Compared with existing system the projected arrangement is compact and is possible to carry to any places. This system converts the language in associate passing voice that's well explicable by blind and ancient people. The language interprets into some text kind displayed on the digital display screen, to facilitate the deaf people likewise. In world applications, this system is helpful for deaf and dumb of us those cannot communicate with ancient person. The foremost characteristic of this project is that the gesture recognizer may be a standalone system, that's applied in common-place of living. It's in addition useful for speech impaired and paralyzed patient means those do not speak properly and in addition used for Intelligent Home Applications and industrial applications.

Chapter 4

Conclusion

Hand Gesture recognition and voice conversion for dumb and deaf person will successfully executing using image processing. The method takes image as input and gives text and speech as an output. Implementation of this system gives up to 90% accuracy and works successfully in most of the test cases [**?**].

References

[1] B.D.Jadhav, Nipun Munot, Madhura Hambarde, and Jueli Ashtikar. Hand gesture recognition to speech conversion in regional language. *International Journal of Computer Science and Network*, 04:161–166, 2015.

[2] Channaiah Chandana, K. Nikhita, P. Nikiitha, N. K. Bhavani, and J. Sudeep. Hand gestures recognition system for deaf and dumb and blind people. *International Journal of Innovative Research in Computer and Communication Engineering*, 05:10058–10062, 2017.

[3] Shital P. Dawane and Hajiali G.Sayyed. A review on hand gesture recognition for deaf and dumb people using gsm module. *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 05:8499–8503, 2016.

[4] Abhishek Jain, Lakshita Jain, Ishaan Sharma, and Abhishek Chauhan. Image processing based speaking system for mute people using hand gestures. *International Journal of Engineering Sciences & Research Technology*, 03:368–374, 2018.

[5] V. Padmanabhan and M. Sornalatha. Hand gesture recognition and voice conversion system for dumb people. *International Journal of Scientific Engineering Research*, 05:427–431, 2014.

[6] U. Pradeepa and J. P. Tharanyaa. An artificial speech system for dumb and blind people. *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, 05:476–479, 2016.

[7] Rupesh Prajapati, Vedant Pandey, Nupur Jamindar, Neeraj Yadav, and Neelam Phadnis. Hand gesture recognition and voice conversion for deaf and dumb. *International Research Journal of Engineering and Technology (IRJET)*, 05:1373–1376, 2018.

[8] R.Jagadish, R.Gayathri, R. Mohanapriya, R. Kalaivani, and S. Keerthana. Hand gesture recognition system for deaf and dumb persons. *Indo-Iranian Journal of Scientific Research (IIJSR)*, 02:139–146, 2018.

[9] Dr. M.R.Dixit Sunita V. Matiwade1. Electronic support system for deaf and dumb to interpret sign language of communication. *International Journal of Innovative Research in Science and Engineering and Technology*, 05:8683–8689, 2016.

[10] Mr. Kunal A. Wankhade and Prof. Gauri N. Zade. Sign language recognition for deaf and dumb people using anfis. *International Journal of Science, Engineering and Technology Research (IJSETR)*, 03:1206–1210, 2014.

Publication