

SPEAKING SYSTEM FOR MUTE PEOPLE USING HAND GESTURE RECOGNITION USING DL

Akash Shikare¹, Pratiksha Patil², Manisha Gaikwad³

¹Akash Shashikant Shikare, Student at PES Modern COE in Department of ECE Engineering, Pune, Maharashtra, India

²Pratiksha Bapurao Patil, Student at PES Modern COE In department of ECE Engineering, Pune, Maharashtra, India

³Manisha Sanjay Gaikwad, Student at PES Modern COE in Department of ECE Engineering, Pune, Maharashtra, India

⁴Seema Bhalgaonkar, Professor at PES Modern COE in department of ECE Engineering, Pune, Maharashtra, India

Abstract - It's very difficult for mute people to convey their message to normal people in their day-to-day life. Since normal people are not trained on sign language, the communication becomes very difficult. In an emergency, when a mute person travelling or amongst new peoples communication with nearby people or conveying a message becomes very difficult. Here we proposed a smart speaking system that helps mute people in communicating to normal people using hand motions and gestures. The system makes use of a camera to capture hand movement. Text to speech convertor use to convert text into voice. The system is powered by a battery powered circuitry to run it. A raspberry pi 4 is used for processing the data and operating the system. The system consists of 10 stored messages like "help", or "where is the washroom" and so on that help mute people convey basic messages. The system captures hand images and hand motion in different varitaions. It also has a trigger in order to indicate that the person wishes to activate the system and speak something. This ensures the system does not speak when the person is just involuntarily making geatures.

Key Words: Raspberry pi 3, Ultrasonic sensor HC SR04, Feature extraction technique.

1. Introducton:

Overview:

The need of this system is to give output in day-to-day life for "Image Processing Based Sign to Speech Converter for Mute People" using ML and CNN. It explains the aim and declaration for the evaluation of system. It will also explain system constraints, interface and interactions with other external applications. The aim of the proposed project is to overcome the challenge of skin colour detection for natural interface between user and machine. This project is developed for the impaired people and would be beneficial as they can communicate with everyone. In our system a camera is placed in front of the physically impaired person. The physically impaired person will place a hand with particular action in front of the camera .When he makes the gestures,

the camera will capture the exact positions of the fingers and hand and perform image processing and then feature extraction would be done. 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 by some text-to-speech algorithms so that it would be audible to everyone through the speakers. By using this system the physically impaired people and mostly mute people would be benefited as they can communicate with anyone freely which indeed would be great achievement for the mankind. The system can be defined using a flowchart which has three main steps, they are as follows: Learning, Detection and Recognition.

Learning: It involves two aspects such as

– Training dataset: This is the dataset that consists of images of different types of hand gestures that are used for training the system based on which the system performs the actions.

– Feature Extraction: In feature extractions main features of images are stored and use to train the model. Features can be edges or contours etc.

Detection:

– Captures scene: Captures the images through a camera, which is used as an input to the system.

– Image Pre-processing: Images that are captured through the camera are compared with the images from dataset to recognize the valid hand movements that are needed to perform the required actions.

– Hand Gesture Detection: The requirements for hand detection involve the input image from the camera.. Distance should also be maintained between the hand and the camera. Approximate distance that should be between hand and the camera is around 30 to 100 cm.

Recognition:

– Gesture Recognition: The number of fingers present in the hand gesture is determined by making use of defect points present in the gesture and then used to detect what gesture would it be.

– Performing action: The recognized gesture is used as an input to the raspberry pi. It detects the message and then text to speech converter convert into voice.

1.1 Objective

The great challenge lies in developing an economically feasible and making hardware independent system so that physically impaired people can communicate easily.

- Hand gestures are a form of non-verbal communication that is been be used in several fields.
- The setup consists of a single camera, to capture the gesture formed by the user, connected with raspberry pi and captures the hand image as an input to the proposed algorithm and produce voice messages as output through the speakers.
- The sample algorithm divided into four main steps, which includes segmentation, feature extraction and classification.
- Finally Output display through speaker.

1.2 Advantages

- Cheap
- Compact and small systems
- Flexible and easy to use for anyone.

1.3 Application

- Communication made easy for the disables.
- Can be used in special schools and NGO.

2. Earlier Works:

Anbarasi Rajamohan, Hemavathy R. Dhanalakshmi proposed a Deaf-Mute Communication converter. Mute people has always found it difficult to communicate with normal people. The low accuracy of this paper was overcome in our project. As the technology we used is ML and CNN there was growth in accuracy.

Prem Jacob and Puneeth proposed a Hand gesture Recognition for disables using CNN. In this paper there was no hardware implemented to make this device or project portable. The approach was pretty clear using CNN to convert hand gestures into speech. But we overcome the disadvantage and make the project portable to use with two different and effective approach.

3. Proposed Methodology:

For implementation of this system we require different hand gestures and text assigned to these gestures and camera is required for capturing the gestures. The person would be doing gestures in front of the camera. While implementing the system

there are different modules. The modules involved are as follows:

3.1 Machine Learning Approach:

Generation of Database : First of all while processing the images, it is necessary to prepare a proper database of 10 images of each symbol total of 50 images are captured in order to increase the accuracy in detection of images.

Image pre-processing and segmentation: Image Pre-processing consists of image acquisition, segmentation and filtering methods. Grayscale of image is done and Segmentation of hand is carried out to separate the hand from the background. Principal Component analysis (PCA) algorithm is used for segmentation purpose and some edge detection techniques are used to gather some edge features. The segmented hand image are represented certain features. Filtering methods like median filters, Gaussian filters etc are used to remove noises from images so that we can get smooth contours.

Feature extraction: Feature Extraction Technique is used in this approach. There are many algorithms available which extracts features from the images. Some of these are HOG(Histogram of gradient), PCA(Principal Component analysis), Local Binary pattern etc. These algorithms selects the main features of the image or a pattern which is unique. These patterns or features are then used for training purpose.

Sign Recognition: It uses principle component algorithm analysis to Extract features from the image from the database. The PCA algorithm has 2 phases:

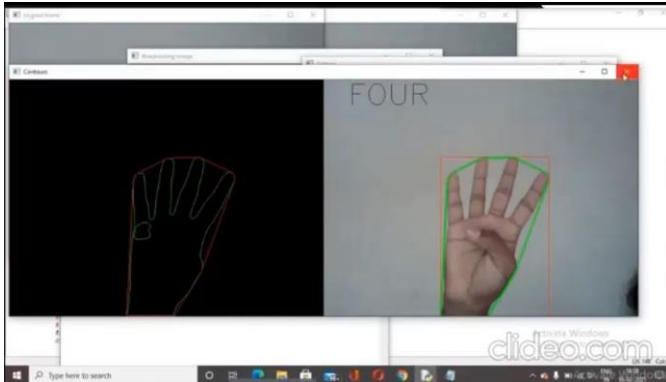
- Training phase
- Recognition phase

In training phase, each gesture is represented as a column vector. PCA finds the eigen vectors of the covariance matrix of the gestures and then they are multiplied by each of their gesture matrix to obtain their gesture space projection.(Jain.A, ET.AL)

In recognition phase, a subjected gesture is projected onto gesture space and the euclidean distance is computed between this projection and all known projections.

Gestures to Voice: The identified image is converted into voice or speech signal using Text-to-speech algorithm. GTTS algorithm (Google Text-to-

speech, by Google) is used to convert any text to a .mp3 file storing the voice of the text message. The stored mp3 file runs and the output voice can be heard through speakers.



3.2 Deep Learning Approach:

Hand Databases: There are many databases available on internet. As for this approach we need huge amount of image. We cannot make such huge databases so we used a hand gestures recognition database from kaggle.com (GTI Leap motion, kaggle.com). It had 20,000 PNG hand images. We use these images to train the CNN model.

Image pre-processing: The input image had to be resized and grayscaled as the images from the database were grayscaled images. We used noise removal filters to make the input image match the quality of the images from database.

CNN model implementation: We used a CNN model which can detect hand gestures with a prediction accuracy of 99.9%. The captured and processed image from the above step is now the input to the CNN model. The model detects the gestures using Convolution Methodology and Neural Networks. This approach does not need any feature extractions because of neural networks.

4. Results:

This paper proposes an electronic system that can be used for communication between mute people and normal people.

- The design is compatible and faster.
- Response time of 2 to 3 seconds.
- Compact and portable.
- Efficient communication between impaired (mute in this context) and normal people.
- sign language involves different gestures and hand movement, improves small motor skills of differently abled people.

- A mobile app can be built for the efficient use of the design and to make it user-friendly.

5. Conclusion:

- Sign language is a tool for communication between the mute community and normal people. As it is difficult for the mute community to communicate with people who don't understand the hand sign language.
- This system prototype was designed to automatically recognize sign language to help normal people to communicate more effectively with mute people.
- This system recognizes the hand gestures using web-cam and these recognized hand gestures are converted into speech so that normal people can easily understand.
- The project aims to low the communication gap between the world and the mute community.

6. Acknowledgement:

This work is done, supervised and supported by the students and faculty members of the Department Electronics And Telecommunication Engineering, PES Modern College Of Engineering, Pune.

7. References:

1. Ibrahim Patel¹, Dr. Y. Srinivas Rao² proposes a "Automated speech synthesizer and converter in cue symbol generation for physically impairs". International Journal of Recent Trends in Engg, Vol 2, No. 7, Nov. 2009
2. Anbarasi Rajamohan, Hemavathy R., Dhanalakshmi proposes a "Deaf-Mute Communication Interpreter". Proceedings of the International MultiConference of Engineers and Computer Scientists 2009 Vol I IMECS 2009, March 18 - 20, 2009, Hong Kong.
3. Jain, A., Jain, L., Sharma, I., & Chauhan, A. (n.d.). Image processing based speaking system for mute people using hand gestures. international journal of engineering sciences & research technology, 7(4), 368-374..
4. "Vehicle tracking and fuel level identifying using android" by A.vijayabhasker reddy, IJCRT, Dec 2017, Volume 6, Issue 1.

5. "Sign language recognition using sensor gloves" by Mehdi, S.A.; Khan, Y. N. Neural Information Processing, 2002.ICONIP '02. Proceedings of the 9th International Conference, Volume: 5 Publication Year: 2002, IEEE Conference Publications.
6. Kunal Kadam, Rucha Ganu, Ankita Bhosekar, Prof. S. D. Joshi, "American Sign Language Interpreter", Proceedings of the 2012 IEEE Fourth International Conference on Technology for Education .
7. "Deaf-Mute Communication Interpreter" by Anbarasi Rajamohan, Hemavathy R., Dhanalakshmi© 2018 JETIR May 2018, Volume 5, Issue 5 www.jetir.org (ISSN-2349-5162) JETIR1805715 Journal of Emerging Technologies and Innovative Research (JETIR) www.jetir.org 685 M. International Journal of Scientific Engineering and Technology (ISSN : 2277-1581) Volume 2 Issue 5, pp : 336-341, 1 May 2013
8. "Microcontroller and Sensors Based Gesture Vocalizer" by Salman Afghani, Muhammad Akmal, Raheel Yousaf . Proceedings of the 7th WSEAS International Conference on signal processing, robotics and automation.
9. Yongjing Liu, Yixing Yang, Lu Wang, Jiapeng Xu, "Image Processing and Recognition of Multiple Static Hand Gestures for Human-Computer Interaction", IEEE 2013 Seventh International Conference on Image and Graphics.
10. P.Subha Rajam, Dr. G. Balakrishnan, "Real Time Indian Sign Language Recognition System to aid Deaf-dumb People", 2011 IEEE.
11. <https://www.kaggle.com/gti-upm/leapgestrecog>
12. <https://www.geeksforgeeks.org/convert-text-speech-python/>