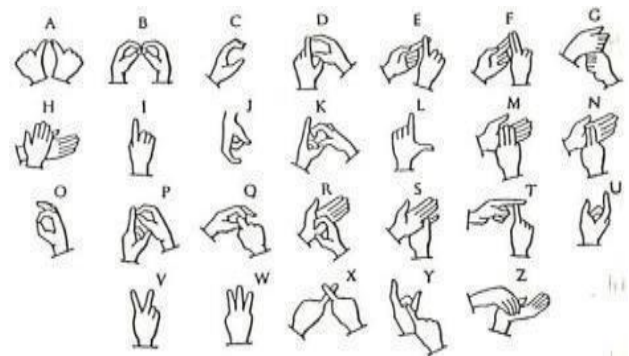# Real Time Speaking System for Speech and Hearingimpaired People - Literature Survey

Uma N M[1], Syeda Rabiya Hussainy[2], Syeda Hafsa Ameen[3], Subahini.A[4], Prof.Sundari Tribhuvanam[5]

Uma N M, Electronics and Communication, Atria Institute of Technology, Bangalore, Karnataka
Syeda Rabiya Hussainy, Electronics and Communication, Atria Institute of Technology, Bangalore, Karnataka
Syeda Hafsa Ameen, Electronics and Communication, Atria Institute of Technology, Bangalore, Karnataka
Subahini.A, Electronics and Communication, Atria Institute of Technology, Bangalore, Karnataka
Professor. Sundari Tribhuvanam, Electronics and Communication, Atria Institute Of Technology, Bangalore, Karnataka

---***---

Abstract— *Sign language plays a vital role in communication between audio-vocal challenged and normal people. In this paper, we propose a new approach that recognizes the hand gesture based on Indian Sign Language and convert them into text and speech output. This system uses the vision-based technique where the hand gestures and facial expressions are captured using web-camera and the various technologies. The captured images are processed with image processing, and classified with neural network, Open CV to recognize the hand gesture and facial expressions and convert it into text and speech using microcontroller based hardware (raspberry pi).*

Keywords— INDIAN SIGN LANGUAGE, HAND GESTURE, FACIAL

EXPRESSIONS, IMAGE PROCESSING, RASPBERRY

I. INTRODUCTION

Sign Language also called as the gesture-based language used by hearing impaired people to communicate with others. The term Deaf-mute is used to refer people who have hearing and speech impairment [32]. Needs of deaf or mute community have long been ignored and the problems have been documented by various organization [33].

Indian Sign Language has evolved in India over last 100 years [34].According to 2011 Indian Censes ,it was found that there were 1.3 million people suffering with hearing impairment [35]. There is a huge communication gap between deaf, dumb and normal people. Scientific groups are now pacing up to abridge this gap and make communication easy.

Figure(1) :Indian Sign Language

Gestures are non-verbal communication which can help people form clear thoughts and use more declarative language.

Gestures include movement of hand, face and other parts of body. Hand gestures play important role in communication particularly in sign language

The hand gesture can be recognized using three approaches:

1. Vision based approach: This approach uses a camera(s) which are used to capture hand gesture and process the image. This image considers various factors like skin colour, velocity, recognition time and variation in light.

2. Instrument-Glove based: In this approach, sensors are used to find the position of hand and its movements and recognises the gesture.

3. Colored-Marker based approach: In this approach, hand is represented using 3 colors glove for tracking different parts of the hand like palm and fingers.[1]

In this paper, we propose a technology that recognizes the facial expressions and hand gestures of Indian Sign Language and convert those to text and further to speech output using a vision-based technique. In this work, we present a system that uses the following technologies.

(1)   Image processing: It is a technique that processes that image to extract useful information from it. The hand gestures that are captured by webcam are processed using image processing technique.

(2)   Background Subtraction: It is a method in which an image's foreground is separated from the background for further processing .The steps involved in it are: pre processing, background subtraction, foreground detection and data validation.[37] (3) Haar Cascade: It is a machine learning based on approach in which the features: Edge Feature, Line features, Centersurround features are used to detect the image (face in our case)[36]

(3)   Open-source computer vision (Open CV): Opensource computer vision library is used for real time computer vision applications. It is mainly used for image processing, video capture and analyses. (4) Raspberry pi: It is a inexpensive, easy available, credit card sized computer. The recognized image is converted to text and speech output using Raspberry pi

II.   Literature Survey

Shilpa L Khandade and S T Khot [1] proposed a system that recognises the hand gesture using MATLAB. This paper also mentions the different input approaches in hand gesture recognition and its methods

Sanish Manandhar, Sushana Bajracharya, Sanjeev Karki,Ashish Kumar Jha [2] in their work they have built a wearable glove controller that converts the given sign into text, audio and pictorial form using flex sensors and random forest algorithm and obtained 92.14% accuracy.

The work of Muthu Kumar,S Poorani [3] used a vision-based hand gesture recognition for Indian Sign Language using local binary pattern and SVM classifier and obtained 92.14% accuracy.

Mandeep Kaur Ahuja and Amardeep Singh [4] have used a vision-based technique where the hand gestures are recognised upon skin colour model and thresholding approach along with template matching using PCA. It is implemented using MATLAB and obtained an accuracy of 91.25%. Wne-pinn fang [5] proposed a paper which implements an intelligent hand gesture recognition extraction and recognition for home care application by using two parts of algorithm (a)still image recognition and (b)interframe recognition P.Subha Rajam and G Balakrishnan [6] have proposed a technology in which fingertip position are identified and converted into text

using image processing and pattern recognition. The developed 32 signs can be used to recognise the letters of Tamil language to achieve an accuracy of 92.125%. B. Laxmi, Rasheed Ahamed, Harshali Rani and Ravi Kishore Kodali [7] in their work, different gestures based signs are recognised and converted to speech in real time using text to speech module, has used CNN for designing converter and implemented on raspberry pi controller, the accuracy is 99%. Anchal Sood and Anju Mishra [8] they have proposed a paper where the input is given in the form of text as well as audio, using image matching, feature matching and recognition after searching from the database and processing, get an output in the form of text and audio.

Muhammed Yaqoob Javed and Muhammed Majid

Gulzar, Syeda Tahir Hussain Rizvi, Mjunadi Azif, Zaineb Iqbql [9] have proposed a device that recognized hand gesture by computer vision program, where the algorithm is based on ANN and open CV is used to write the program. The system is implemented using Arduino microcontroller, laptop, LCD, arduino voice box and web camera. Hence the gesture is recognised and arduino receives the data serially and displays the text on LCD and speech is produced in the voice box. Shreyashi Narayan Sawant and M.S.Kumbhar [10] have proposed a system that performs hand gesture recognition which is based on MATLAB using PCA technique. This system has four modules: pre-processing and hand segmentation, feature extraction, sign recognition and sign to text and voice conversion 26 combinations of Indian signs are developed and stored in database. The final result obtained is converted to the corresponding text and voice form. Yeresime Suresh and J. Vaishnavi, M Vindhya,

Mohammed Sadiq Afreed Meeran, Supritha Vemala

[11], their work proposed a system that interprets the concepts of Machine learning using neural network technologies. Open CV libraries are used for computer vision, Machine learning and image processing. Here canny edge detection is used, and prediction is done using Convolution Neural Network (CNN) algorithm.

Geethu G. Nath and Anu V.S [12] have proposed the concept of Arm Cortex processor board where the captured hand gestures are implemented. Beagle bone black is the Linux operating system supported with open CV library. A USB camera is used to capture image, the pre-processing is done by the ARM Cortex A8 processor board. Rashmi R Koli and Tanveer I.Bagban [13] have proposed a method that uses CNN algorithm that recoginzes the hand gesture and convert it into text. Frame feature extraction is done using python programming language.

Areesha Gul and Batool Zetra, Sadia shah, Nazih Javeed,Muhammed ,Imran Saleem[14] has proposed a twoway smart communication system using Leap Motion

Controller (LMC) to track the hand movement in real time. They make use of Hidden Markov Model (HMM) for hand gesture recognition. The raspberry pi reads input from LMC and speech is converted to text using google API, and displayed on raspberry pi screen.

In this work of Surendra Kumar Keshari, Shruti Tyagi, Niketa Tomas Smiti Goel [15] proposed a system that reduces the communication gap by connecting the real time gesture based signs to text and finally to speech. This method which was used to create this system is divided into two parts 1) Gesture based calculator which performs the arithmetic operations using gesture and pre-defined operations 2) say out loud which converts words into speech. It uses open CV for gesture recognition. Tanzila Ferdous Ayshes, Sadia Afrin Raka, Quazi Ridwan Harib,Md, Horrain, Rashedeu M Rahman[16] proposed a system that uses fuzzy rule based hand gesture recognition system for the Bengali characters. It is implemented using Harris algorithm, using MATLAB.

Yellapa Madhuri and Anitha G Anburajuan M [17] propose an interative application program developed using LABVIEW software and incorporated into mobile phone vision analysis function are performed in the operating system and thereby provides speech output. The hardware requirements are minimized. Their translator was able to translate alphabet(A-Z) and numbers (0-9).

Purva Chaitanya Bandhe and Vaishali Kulkarni [18] there work use ANN (Artificial Neural Network) for classification. Gestures here Indian Sign Language and translated to English, the help of hand crafted feature extraction technique and ANN. The ANN here is employed in TensorFlow using Keras at backend out of 500 videos,300 video is used for training the model. The accuracy achieved was 98%. Amiya Kumar Tripathy and Dipti Jadhav,Steffi A Barreto,Daphne Rasquinha,Sonia S Mathew [19] acquire real time images and convert them to speech and text. In this paper they have considered only fingering spelling of alphabet signs. The processing of signs is done using microsoft visual studio as an

IDE and open CV modules

Diponshes someshwar and Dharmik Bhanushali [20] proposed system that uses CNN for training and recognizing the Hand gesture with the help of Tensor flow and open CV. The python speech to text library is used to obtain the output in the form of text.

Shweta S Shinde and Rajesh M Autu [21] have developed a real time system for hand gesture The acquired image is processed using MATLAB. Then feature extraction is done using angle and peak calculation. Lastly the hand gestures are converted to speech.

K.H.Wanjale, Amit Bhoomkar, Ajay Kulkarni, Somath Gosavi[36] has proposed a system to track a face in real time, in which they have used Adaboost algorithm and Haar features algorithm to detect facial expressions.PCA(Principal Component Analysis) is implemented for face detection. To track the face and error correcting graph matching method is implemented.

Rucha D. Pathari and Sachin M. Bojewar [37] have proposed a system where in moving objects are detected in videos, for this background subtraction is done which involves four major steps: pre processing, background modeling, foreground detection and data validation. Kalman or enhanced Kalman filters are used to remove noise from the output file.

M. Satiyan and R. Nagarajan [38] have proposed a system that uses Haar-like feature extraction method to recognise facial expressions. The subjects face is fixed with a set of luminance stickers, and they perform facial expressions which are recorded, then a set of 2D coordinate values are obtained using tracking software. With these values Haar-like feature extraction method is employed to capture the facial expressions.

Shivam Gupta [39] has implemented a system where using both real time and static images emotions can be detected. For this Haar filter from open CV is used to detect faces. For static images data organization, extracting faces, training and classification is done. In real time testing the landmark detector, extracting features from the faces is done using webcam.

Soukaina Chraa Mesbahi, Mohamed Adnane Mahraz, Jamal Riffi and Hamid Tairi [40] have presented a method to recognise hand gesture using convexity defect and background subtraction. In background subtraction using difference between current frame and reference frame the moving object is detected. Using one, two, three, four and five fingers one by one five hand gestures are tested

III.   Methodology

There are various means of communication, sign language being one of them which involves movements of hands and arms. Indian sign language is used in the deaf and mute community all over India.
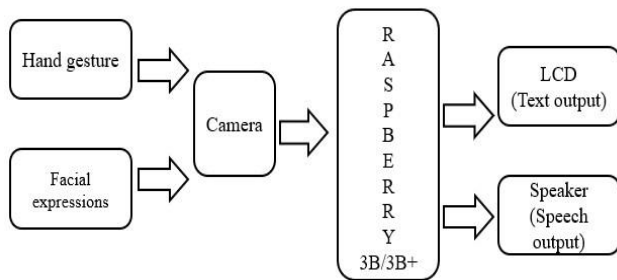
Figure(2) : Proposed Block Diagram

Various gestures used in the Indian sign language are given as an input to a webcam. The image is captured in real time and are analysed using image processing (shown in figure.2). Image Acquisition: Some of the devices used for data acquisition are images, gloves and markers [21]. In our system, vision-based technique is used where the real time image is captured using webcam.

Image Processing: Initially, the captured image is processed by converting it into grey scale, then the Facial Expressions are identified using Haar algorithm .The detected face is further processed for identification of eye, eyebrow, nose, corner of face, mouth regions.[39].The hand region is detected by using background subtraction where the foreground is separated from the background followed by thresholding, where we select the areas of interest of an image. After thresholding, contour detection is done. Finally, this recognised image is sent to raspberry pi. And this Raspberry pi processes the image and identifies the hand gestures, facial expressions and converts them into text and speech [7]

 i. Software Description

Open-source computer vision (Open CV): Open-source computer vision library is a free library for computer vision used in image processing for real time applications. In open CV once the image is captured it is first converted into gray channel. RGB channel and grey scale channel are the 2 channels in which images are considered. Even when there is a slightest motion of the hand gesture in real time, opencv will continuously suggest its predictions. In opencv there are some features that support object detection. Here the programming language used is python with the required libraries from open cv. [20]

Haar Cascade: Haar cascade classifier is an algorithm used in detecting the human faces and objects. The Haar classifier cascades are first trained in order to detect the human facial features such as mouth, eyes and nose. Ada boost algorithm and Haar feature algorithms are implemented to train the classifiers. Therefore a person's face is detected in the video with the help of this algorithm system [36].

Convolution Neural Network (CNN):

CNN is a class of deep feed artificial neural networks.

It is being used in image, pattern and speech recognition. Twodimensional image information is identified using CNN algorithm [13]. After preprocessing and thresholding, CNN model receives the image. Then the hand gestures in the image are correctly classified using this trained model.[7]

 ii. Hardware Description

Raspberry pi is a small PC board which is used in real time projects. We use raspberry pi to implement our model. The raspberry pi is embedded with an image processing algorithm, which monitors the hand fingers

[7].Microcontroller processes the image, and LCD displays the text output. [30] Controller recognizes the text and gives audio output using speaker, for this text-to-speech is used by the controller which recognizes the text and converts it and passes it on to the speaker.[7]

IV. CONCLUSION

The main purpose of our project is to minimize the communication gap between the normal world and the deafmute people, to make their everyday life a little simple. We have implemented a system that is capable of recognising human hand gesture and simultaneously produce corresponding speech and text output. This system recognises the facial expressions and hand gestures without any glove based data acquisition devices. The recognition is done in real time for obtaining text and speech output. Therefore, our system aids the SpeechHearing impaired community by reducing the communication gap.

REFERENCES

1. S.L. Khandade and S.T.Khot, "MATLAB Based Gesture Recognition", International Conference on Inventive Computation Technologies (ICICT),jan, Sanish Manandhar1

2. S.Bajracharya, S.Karki and A.Kumar Jha, "Hand Gesture Vocalizer for Dumb and Deaf People", SCITECH Nepal, Vol. 14, No. 1,Sept .2019.

3. K.Muthukumar, S.Poorani and S.Gobhinath, "Advances in Natural and Applied Sciences", Published by Anesi Publication, 2017, Special 11(6): pp 314-321

4. M.K.Ahuja, A.Singh, "Static Vision Based Hand Gesture Recognition Using Principal Component Analysis", IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE), 2015.

5. W.Fang, "An Intelligent Hand Gesture Extraction and Recognition System for Home Care Application", Sixth International Conference on Genetic and Evolutionary Computing, Kitakushu, Japan -2012.

6. P.S.Rajam, G.Balakrishnan, "Real Time Indian Sign Language Recognition System to aid Deaf-dumb People", (IEEE) 13th International Conference on Communication Technology, Jinan, China, 2011.

7. B.Lakshmi, R.Ahamed, H.Rane, and R.K.Kodal, "Assistive SIGN LANGUAGE Converter for DEAF AND DUMB", International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), July.2019.

8. A.Sood, A.Mishra, "AAWAAZ: A Communication System for Deaf and Dumb", International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2016.

9. M.Y.Javed, M.M.Gulzar, S.T.H.Rizvi, M.Asif and Z.Iqbal, "Implementation of Image Processing Based Digital Dactylology Converser for Deaf-Mute Persons", 2016 International Conference on Intelligent Systems Engineering (ICISE),Pakistan, Jan. 2016.

10. S.Sawant, M. S.Kumbhar, "Real Time Sign Language Recognition using PCA", IEEE International Conference on Advanced Communications, Control and Computing Technologies, Ramanathapuram, India, 2014.

11. Y.Suresh, J Vaishnavi, M Vindhya, Md.Meeran and S.Vemala, "MUDRAKSHARA - A Voice for Deaf/Dumb People", 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), July 2020.

12. Geethu G Nath, Anu V S, "Embedded Sign Language Interpreter System For Deaf and Dumb People", International Conference on Innovations in information Embedded and Communication Systems (ICIIECS), Coimbatore, India, March 2017.

13. R.Koli, T.I.Bagban, "Human Action Recognition Using Deep Neural Network", IEEE Sensors Journal, Vol 20, Issue: 6, March15, 15 2020.

14. A.Gul, S.Shah, B.Zehra, N.Javed and Md.Saleem, "Two-way Smart Communication System for Deaf & Dumb and Normal People", 2020 International Conference on Information Science and Communication Technology (ICISCT), KARACHI, Pakistan,2020.

15. S.Keshari, S.Tyagi, N.Tomar and S.Goel, "Aphonic's VOoice: A Hand Gesture Based Approach to Convert Sign Language to Speech ", International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT), India, Sept. 2019.

16. T. Ayshee, S.A. Raka, Q. Hasib, Md. Hossain and R.M Rahman, "Fuzzy Rule-Based Hand Gesture Recognition for Bengali Characters", IEEE International Advance Computing Conference (IACC), published by IEEE, Feb. 2014.

17. Y.Madhuri, Anitha.G and Anburajan.M, "Vision Based Sign Language Translation Device": International Conference on Information Communication and Embedded Systems (ICICES), Chennai, India, Feb. 2013.

18. P.C. Badhe and V.Kulkarni, "Artificial Neural Network based Indian Sign Language Recognition using hand crafted features",11th International Conference on Computing,Communication and Networking Technologies (ICCCNT), July 2020.

19. A.K.Tripathy, D.Jadhav, S.A.Barreto, D.Rasquinha and S. Mathew, "Voice for the Mute", 2015 International Conference on Technologies for Sustainable Development (ICTSD-2015), Feb 2015, Mumbai, India.

20. D.Someshwar, D.Bhanushali and V. Chaudhari, "Implementation of Virtual Assistant with Sign Language using Deep Learning and TensorFlow", Proceedings of the Second International Conference on Inventive Research in Computing Applications (ICIRCA-2020), July 2020.

21. S.Shinde, Rajesh M. Autee, V.K. Bhosale, "Real Time Two Way Communication Approach for Hearing Impaired and Dumb Person Based on Image Processing", :IEEE International Conference on Computational Intelligence and Computing Research (ICCIC),Chennai, India, Dec. 2016.

22. Suganya R and T. Meeradevi, "Design of a Communication Aid for Physically Challenge(ICECS 2015),Feb. 2015.

23. N.S.Soni, M.Nagmode and R.D.Komati, "Online Hand Gesture Recognition & Classification for Deaf & Dumb", in: International Conference on Inventive Computation Technologies (ICICT),Coimbatore, India,2016.

24. S.Rathi and U.Gawande, "Development of Full Duplex Communication System for Deaf and Dumb People", 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, Jan. 2017.

25. S. Chattoraj and K. Vishwakarma, "Assistive System for Physically Disabled People using Gesture Recognition", 2017 IEEE 2nd International Conference on Signal and Image Processing, Singapore, Aug. 2017.

26. P.Sridevi, T.Islam, U.Debnath, N.A Nazia, R.Chakraborty and C.Shahnaz, "Sign Language Recognition for Speech and Hearing Impaired by Image Processing in MATLAB",: IEEE Region 10 Humanitarian Technology Conference (R10HTC),Malambe, Sri Lanka, 2018.

27. P.K.Datta, A. Biswas, A.Ghosh and N.Chaudhury, "Creation of Image Segmentation Classifiers for Sign Language Processing for Deaf and Dumb", 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)Amity University, Noida, India. June 4-5, 2020

28. S.Upendran and Thamizharasi. A, "American Sign Algorithms", International Journal of Scientific and Technology research vol 9, issue 1,Jan 2020, pp1597- 1579.

29. https://en.wikipedia.org/wiki/Deaf-mute

30. https://m.economictimes.com/industry/services/educ ation/ national-level-research-institute-onsignlanguage-to-be-setup/articleshow/49058725.cms

31. https://www.bbc.com/news/world-asiaindia39101899

32. https://www.npr.org/sections/goatsandsoda/2018/0 1/14/57 5921716/a-mom-fights-to get-an-educationfor-her-deaf daughters

33. K H wanjale ,Amit Bhoomikar, Ajay Kulkarni,Somnath Gossavi,"Use Of Haar Cascade Classifier For Face Tracking System in Real Time Video", International Journal Of Engineering Research & Technology (IJERT),April 2013

34. Rucha D .Pathari,Sachin M Bojewar, "Background Subtraction Algorithm with Post Processing in Video Surveillance", International Journal Of Engineering Research & Technology (IJERT),January 2014.

35. M. Satiyan and R. Nagarajan , "Recognition of Facial Expressions using Haar-like Feature Extraction Method",International Conference on Intelligent and Instrumentation, Communication and Computational Technologies (ICCICCT),Kanyakumari, India,2014.

36. Shivam Gupta "Facial Emotion Recognition in Real Time Language Interpreter System for Deaf and Dumb Advanced Systems,Manila,Phhelliphines,February,2011 Individuals", :International Conference on Control

37. Fahad Ullah, "American Sign Language Recognition System for Hearing Impaired People Using Cartesian Genetic Programming ", Proceedings of the 5th International Conference on Automation, Robotics and Applications, Dec 68, 2011, Wellington, New Zealand.

38. V P Fantin, G MohanBabu,J M Jasmine, P MurugaBoopathi and S K Nagarani, "Smart Communication Haptic Glove For Deaf and Dumb people", International Journal of Scientific Research and Review,Volume 8, Issue 3,pp-591594, 2019.

39. Zulfia Afreen. M. A, Ramachandran. A, "Study Of Image Processing, Recognition And Computer Vision and Static Images",International Conference on Inovative System and Control (ICISC),Coimbatoor, India, DOI : 10.1109/ICIAS.2010.5716228, January 2018

40. Soukaina Chraa Mesbahi, Mohamed Adnane Mahraz, Jamal Riffi, Hamid Tairi, "Recognition based on Convexity approach and background subtraction", International Conference on Intelligent Systems and

41. Computer Vision (ISCV), Fez, Morocco, DOI :10.1109/ISACV.2018.8354074, April 2018

42. K Chetan Kumar, Neha N, Kamatham Tejaswini, Namratha V, Sundari Tribhuvanam, "Emotion Recognition Based on Facial Expressions and Hand Gesture Using Image Processing-A Survey", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, April 2020.