SAHARA: Communication System for Deaf and Mute

Gargi Gaitonde, Kuruppathukattil Sreemol Bharathan, Aditee Arun Parkar
BTech in Electronics Engineering

***Abstract*—The Dynamics of the world changed after the invention of electricity and followed by the invention of the internet. These inventions drastically changed the standard of living. The technology prospered in all sectors with time. People started finding comfort in automation. Thus, we propose a smart communication system to enhance the standard of living for the deaf and mute. This system uses sign recognition based on dynamic hand gesture recognition techniques in real-time scenarios. Computer Vision-based techniques were used for image analysis. The captured video of the hand is converted to Hue, Saturation, Value (HSV) color space for processing and then segmentation was done based on skin pixels. Contour Extraction is used in parallel to get more accurate results. Then the classification of gestures was done by Support Vector Machine(SVM) Classifier. The respective message is displayed through a mobile application 'SAHARA'.**

***Index Terms*—Hand Gesture, Image processing , Android Studio, text to speech, accessive touch.**

I. Introduction

Any exchange of information between two entities is known as communication. One can communicate in four ways written, verbal, non-verbal and visual communication. The written form of communication is done by a written draft in anylanguage. The verbal form of communication is the most directand fastest form of communication. It is the way we speakwe use words, sentences, etc. In nonverbal communication, one uses facial expressions, posture, gesture physical touch,or even just eye contact. Sign language is a conventional form of nonverbal communication Any means of communication through bodily movements especially of the hands and armsis known as sign language. People who are deaf-mute usessign language to convey their message but most people are noteducated with this language. This becomes a communication barrier.

Communication is deeply tangled with human existence. It's an integral part of it. One cannot think about human life without communication. Are you able to imagine what would happen if you're not allowed to speak for an extended time? You would feel suffocated. That's, how a deaf-mute person feels. It's terribly tough for deaf-mute person to conveyany message to the world. Since the majority aren't trainedin obtainable linguistic communication, the communication becomes difficult. Here we tend to propose a wise speaking system that will facilitate any deaf-mute person in transfer their message to everyone through hand motions and gestures.

Image processing is a rapidly growing area that has diverse applications for instance multimedia computing, secured data communication, biomedical, biometrics, remote sensing, texture understanding, pattern recognition, content-based retrieval, compression, and many more. It is how a computer can sense pictorial data after processing an image. This technology can be a boon for the deaf-mute people who are not able to speak hence can't communicate. If the person has a different language than the receiver, it can also be used as a translator. It has been always considered a challenge for the development of a natural interaction interface that helped people interact with technology as they're wont to interact with the real world. An interface based on human sign language, where no devices are attached to the user, will assist the user from the real world to the virtual environment.

II. Literature Survey

There are multiple types of disabilities. There are people with vision loss, deaf or hard of hearing, mental health conditions, intellectual disability, acquired brain injury, autismspectrum disorder, and physical disability. [8]

According to World Health Organization (WHO), over 5 percent of the population around the world has disabilities in hearing loss. When a person is deaf, the person has little to no hearing and they often speak to others in sign language. The impact of hearing loss or sensation is losing the ability to have a proper conversation as they cannot hear the people talking to them nor can they hear themselves talking [5].The reason for deafness or hearing loss can be because of congenitalcauses or acquired causes. Congenital leads to deafness or hearing loss that is acquired at the time of birth while acquired leads to hearing loss at any age. Technology has kept on evolving throughout time. Assistive technology is one of the main reasons for it. So, Multiple ranges of adaptive, assistive, and rehabilitative devices are being covered by Assistive technologies have offered a wide variety to people subjected to special needs. [4]

For the past two decades, research and development continue within the sector of deafness and hearing loss via assistive technologies. They have been classified as both hardware and software-based including prosthetic implants. Assistive technology also allows people who has hard of hearing, complete deafness, or speech impaired to communicate using

a communication assistant with people having a standard telephone. [1]

Assistive technology has a great impact on modern society which has opened possibilities for people with disabilities. It assists them in day-to-day activities that are difficult or close to impossible and enables them to have an easier day. Deaf people often have a difficult time when it comes to commu- nication especially face to face. According to Gugenheimeret al. (2017) that Assistive Technology should not be just a tool for the blind but a collaborative tool that will lead the approach for using it efficiently.

When it comes to assistive technology, with the help of software and hardware capabilities must be counted depending on the disabilities of a target user. Multiple software can be used by the deaf or mute e.g. Text-to-Speech software, Speech-to-Text, and Haptic software.

Robots, humanoids, and other IoT (Internet of Things) as a means of HCI (Human Computer Interaction) use an important feature called Text-to-speech (TTS). It consists of two components; text processing, which refers to as segmentation of given text input into a sequence of syllabic speech units which later being converted to produce speech sounds that come from the other component called speech generation. It is the processing of text and translates the syllabic speech units into generating a human-like speech. [10]

Speech-to-Text recognition (STR) technology is a tool that synchronously transcribes text streams from the speech input onto a computer screen or visual aids. Haptics is the new technology that pertains to the human's sense of touch. Based on the article "Haptic Technology: A comprehensive review on its applications and future prospects" it is said to be the most recent technology to be implemented in the wide range of computer interface devices which promising to deliver experiential changes in the way that humans interact along with information and communication ideas. [7]

The vision-based method uses a web camera to capture images. After that, image segmentation has done. A feature like a palm, finger extracted from the input image. It uses different hand motion that is half-closed, fully closed, semi- closed was detected and also the angle between two fingers is noted down. Data is been saved in vector and with that vector, the recognition of alphabets is done [9] According to Paulo Trigueiros et al [11]used vision based technique with the help of Portuguese language. used the vision-based technique with the help of the Portuguese language. In their implementation, the hand gesture was captured in real-time which uses SVM algorithm which is used for classification purposes.

III. Methodology for Gesture Recognition

The figure 1 shows us the general block diagram of image processing steps used in hand gesture recognition.

- **Gesture Acquisition:** The input gesture made by theuser are captured in this stage.As we are using Vision- based technology to detect and recognize the gesture, the camera is used for motion tracking and capturing data. [2]In our case, we are using the camera of our mobile phone itself eliminating the need for a webcam or laptop to make it user friendly and portable.
- **Gesture Pre-processing:** It is mainly responsible for removing unwanted objects and making the captured gesture information into useful data for feature extraction.This block will enhance the useful information. Various image processing techniques such as skin color detection using HSV color scheme, Binarization of image, cropping, contour detection using Canny edge detection technique, cropping, noise filtering and morphological operations such as erosion and dilation.
- **Feature Extraction:** This block is responsible for collecting all the feature components of the input gesture and storing it in a vector. HOG(Histogram of Oriented gradients) technique is used for extracting features.
- **Classification using SVM:** In this block, the reference data stored in the database vector and the vector formed after feature extraction is classified to get a gesture of the closest match.
- **Gesture Recognition:** This block will recognize the complete gesture and give the corresponding output phrase/ text on the screen. The TTS button will convert text into their respective sound.

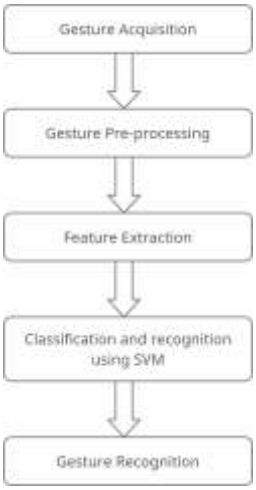

Fig. 1. Block Diagram

IV. System Features

In order to achieve communication between normal person and auditorily impaired/deaf people, we have proposed anapp which developed in Android Studio, converts the hand gesture made by the person into audio message.It comprises of prominent feature sign to speech, Text-to-Speech,Imageto Speech (Accessive Touch), Tutorials for American Sign Language basics, Profile details, Emergency notification. The Modules of Android App are briefly explained below.

1) Dashboard: Dashboard is the first page that appears after the successful sign in of the user. It consists of the Text to Speech Feature and Emergency Notification feature.The TTS

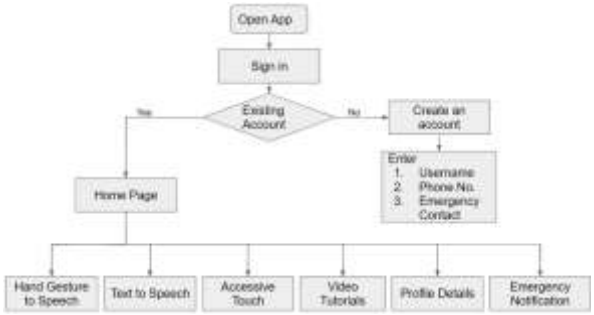

Fig. 2. Application Flowchart

Fig. 3. Dashboard

feature will convert the typed text into corresponding speech output. The emergency feature will send the notification to the respective emergency contact to alert them.

Fig. 4. Accessive Touch

2) Accessive Touch(Image to Speech): In this feature of our application, image buttons are created which when we click corresponding sound is played.

3) Video Tutorials for ASL: This feature contains all the video tutorials for learning American Sign Language for Beginners.

Fig. 5. Video Tutorial for ASL

4) Profile Details: This feature contains the name, phone number, emergency contact number of the user. It will be useful feature in case of any emergency situations.

Fig. 6. Sign-to-Speech feature

5) Hand Gesture to Speech: The gesture is captured by using the mobile camera. [6]In the pre-processing step the useful information is enhanced and unwanted objects are removed. For detecting the hand in the camera frame, skin color detection method is used to select the skin region.Itbasically identifies skin and non-skin pixels. So, HSV(Hue, Saturation, Value) color scheme technique is used.It is more efficient than RGB color model. The next step is to convertthe color image into a binary image and this process is called Binarization. Binary images are used because it becomes easy to separate the region of interest from the background.To precisely detect the hand from the background, edges of the hand are to be detected. So, Canny edge detection algorithmis used for this purpose.It is a multistage edge detection algorithm which applies a Gaussian filter to remove noise and smoothen the image. After the edges are detected, the next stepis to separate background noise and ROI, so we need to crop the ROI i.e.hand from the rest of the background. To improve the image quality, it is subjected to morphological operation such as erosion and dilation. Erosion is basically done toshrink the unwanted details or noise and dilation expands the image. [3]Feature extraction is done using the HOG feature descriptor.It is done by extracting gradient and orientation of edges. These values then are stored into an empty vector whichwill be compared with a reference database vector using SVM Classifier. If the nearest match is found the corresponding textis displayed on the screen. Using TTS we can convert the text into sound output.

Fig. 7. Hand Gesture Dataset

Fig. 8. Hand Gesture Recognition for Thankyou

Fig. 9. Hand Gesture recognition for Help Me

V. Result

The output for the gestures are displayed on the screen and using Text-to-Speech, it is converted into sound.the output of emergency notification is also obtained successfully. We have successfully implemented the Text-to-Speech feature and got the speech output.Also, the accessive touch was implemented,i.e. whenever we click on image button, corresponding speech output was obtained.

VI. Conclusion

Communication is greatly affected by disorders such as hearing or speech impairment. Utilizing a smartphone's capabilities and using it as an Assistive Technology, communication can be done without too much inconvenience. The focus of the research is to break the barrier between hearing and non-hearing or mute in terms of communication. The system was used to assist in the daily activities for the mute or deaf people and other people without disabilities whether it be usedinside of a household or outside. The purpose of it beingdeveloped as a mobile application is for people to be ableto use it anywhere at any time. The system was improvedas the researchers gather more data, conducted interviews in communities and with pathologists, and analyzed that data to further enhance the application for a better output

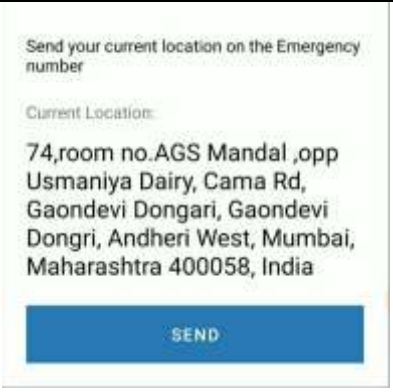

Fig. 10. Hand Gesture recognition for Help Me

Acknowledgement

It gives me immense pleasure to express my deep gratitude and sincere thanks to Dr.Mrs Shikha Nema, H.O.D., our Guide Prof. Bharat Patil all faculty members from Department of Electronics, Usha Mittal Institute of Technology for their valuable and useful support and comments for making this workshop a successful event. I'd not forget to mention that their approach kept my working environment alive and their encouragement promoted me to do my task rigorously.

References

[1] Emad E Abdallah and Ebaa Fayyoumi. Assistive technology for deaf people based on android platform. *Procedia Computer Science*, 94:295–301, 2016.

[2] Purva C Badhe and Vaishali Kulkarni. Indian sign language translator using gesture recognition algorithm. In *2015 IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS)*, pages 195–200. IEEE, 2015.

[3] Pham The Hai, Huynh Chau Thinh, Bui Van Phuc, and Ha Hoang Kha. Automatic feature extraction for vietnamese sign language recognition using support vector machine. In *2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom)*, pages 146–151. IEEE, 2018.

[4] Abhishek B Jani, Nishith A Kotak, and Anil K Roy. Sensor based hand gesture recognition system for english alphabets used in sign language of deaf-mute people. In *2018 IEEE SENSORS*, pages 1–4. IEEE, 2018.

[5] Esteban Alejandro Lopez, Orozimbo Alves Costa, and Deborah Viviane Ferrari. Development and technical validation of the mobile based assistive listening system: a smartphone-based remote microphone. *American journal of audiology*, 25(3S):288–294, 2016.

[6] JL Raheja, Anand Mishra, and Ankit Chaudhary. Indian sign language recognition using svm. *Pattern Recognition and Image Analysis*, 26(2):434–441, 2016.

[7] Izzad Ramli, Nursuriati Jamil, Noraini Seman, and Norizah Ardi. An improved syllabification for a better malay language text-to-speech synthesis (tts). *Procedia Computer Science*, 76:417–424, 2015.

[8] Mary Jane C Samonte, Renz A Gazmin, John Derrick S Soriano, and Martela Nicolai O Valencia. Bridgeapp: An assistive mobile communication application for the deaf and mute. In *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, pages 1310–1315. IEEE, 2019.

[9] RK Shangeetha, V Valliammai, and S Padmavathi. Computer vision based approach for indian sign language character recognition. In *2012 International Conference on Machine Vision and Image Processing (MVIP)*, pages 181–184. IEEE, 2012.

[10] M Sreelakshmi and TD Subash. Haptic technology: A comprehensive review on its applications and future prospects. *Materials Today: Proceedings*, 4(2):4182–4187, 2017.

[11] Paulo Trigueiros, Fernando Ribeiro, and Lu´ıs Paulo Reis. Vision-based portuguese sign language recognition system. In *New Perspectivesin Information Systems and Technologies, Volume 1*, pages 605–617. Springer, 2014.