IJMTST is UGC Approved Journal with Journal Id: 43137

International Journal for Modern Trends in Science and Technology, 8(S05):121-125, 2022

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST08S0520
Available online at: https://www.ijmtst.com/vol8si05.html

Conversion Techniques of Sign and Speech into Text using Raspberry Pi

N.M. Ramalingeswara Rao | N. Ramyasri | CH. Bhavya Manjusha | A. Sai Preetham | CH. Thirupathiswamy

Department of Electronics and Communication Engineering, Godavari Institute of Engineering and Technology(A), JNTUK, Kakinada.

To Cite this Article
N.M. Ramalingeswara Rao, N. Ramyasri, CH. Bhavya Manjusha, A. Sai Preetham and CH. Thirupathiswamy . Conversion Techniques of Sign and Speech into Text using Raspberry Pi. International Journal for Modern Trends in Science and Technology 2022, 8(S05), pp. 121-125. https://doi.org/10.46501/IJMTST08S0520

Article Info
Received: 26 April 2022; Accepted: 24 May 2022; Published: 30 May 2022.

ABSTRACT

Every day we see many people facing illnesses like deaf, dumb, etc. They face difficulty to interact with others. Previously developed techniques are all sensors based and they didn't give a general solution. This work explains a new technique of virtual talking without sensors. Web Camera is used to take the image of different gestures and that will be used as input to the OpenCV with Python. The software will recognize the image and identifies the text output which is displayed on the screen. This work explains two-way communications between the deaf, dumb and normal people which means the proposed system is capable of converting sign language and speech into text.

Keywords: Raspberry pi, Sign language, Web Camera.

1.INTRODUCTION

With the rapid development of computer technologies, devices/techniques have become indispensable in our daily lives. Human-computer interactive (HCI) devices such as personal computers, consumer electronics, mobile devices, etc., have also dramatically altered our lifestyle. The ease with which an HCI device can be understood and operated by users has become one of the major considerations when selecting such a device. Therefore, researchers must develop advanced and user-friendly HCI technologies which can effortlessly translate users' intentions into corresponding commands without requiring users to learn or accommodate the device. Many technologies have been developed to intuitively express users' intentions, such as handwriting, human body language, and gestures to naturally control HCI devices. These technologies have many applications in the fields of remote control, virtual reality, sign language, signature authentication, sports science, health care, and medical rehabilitation.

Deaf and Dumb people depend on sign language for communication. A real-time Sign Language Recognition system was designed and implemented to recognize 26 hand signs from the Indian Sign Language (ISL) by hand gesture/sign recognition system for text generation. The hand gestures are captured by using a webcam. These signs are processed for feature extraction using some color models. The extracted features are compared by using a pattern-matching algorithm. To calculate the sign recognition, the features are compared with the testing database. Finally, a recognized gesture is converted into text. This system provides an opportunity for deaf-dumb people to communicate with those who cannot understand sign language without the need for an interpreter.

In the existing systems, BSL uses a two-handed finger-spelling system, compared to the one-handed system used in ASL and FSL. Many people in America believe that one-handed finger-spelling makes for faster finger-spelling than two-handed systems. However, unreliable evidence has it that in a challenge between proficient BSL and ASL speakers, neither finger-spelling system proved to be faster; both finished reciting the alphabet at the same time. So that supposed "disadvantage" has been proven invalid.

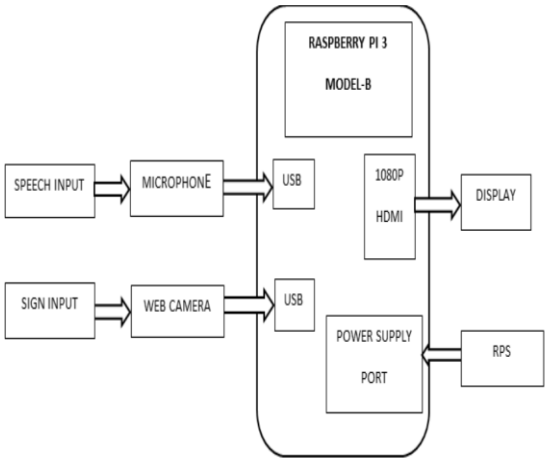
According to many Europeans, American signers tend to fingerspell "too much" compared to the rate of finger-spelling in many European sign languages, including BSL. This may be true because several examples of BSL signs for concepts that do not have a sign in ASL and are often finger-spelled for lack of formal signs in ASL. This is one of the advantages of BSL, but that is not essential to the language itself. On the other hand, many BSL signs are often derived from their initialized (English) base, while many ASL signs have been developed without initialization (including the influence of signed English systems), one might see that as a "disadvantage ".

Nowadays, people are not interested in speaking ASL when having a deaf relative or a friend. Hence, deaf people are often trapped and lonely. ASL requires the usage of a person's hands so if something happens to that person's wrist like a sprain and it disables that person from communicating. For example, there is a mother who strained her wrist from signing all of her life for her deaf daughter. The doctor also made her stop signing. This caused the communication with her deaf daughter to decrease since she had to understand by lip-reading from then on.

2. PROPOSED WORK

To enable communication between deaf, mute, and people who cannot use sign language, the sign language and speech conversion system proposed in this article is used to convert sign and speech into text, which is displayed on the monitor. To perform this conversion process, Raspberry pi 3 B is used as the core processing system for the execution. This system can also tackle the problem of miss-communication between sign language users as the meaning of hand signs changes from country to country. Since the output is in English, a language that is used globally.

This system has two inputs, one of them is a hand gesture which is taken in the form of an image and the other input is voice input which is given through a microphone. Both of these inputs are processed and the text output is displayed on the monitor. This process is clearly explained in the following block diagram.



The speech input is taken through a microphone which is connected to the raspberry pi through a USB port and the image input is taken through a web camera which is also connected using a USB port on the Raspberry pi. Both of these inputs are processed and the output is displayed on a laptop display connected through a 1080p HDMI port, we can also connect an external display if we wanted but in this proposed system we opted for a laptop/PC.

2.1 RASPBERRY PI 3 MODEL B

Raspberry Pi is a small single-board computer. By connecting peripheral devices like Keyboard, display, and mouse to the Raspberry Pi, it acts as a mini personal computer. Raspberry Pi is popularly used for real-time Image/Video Processing, IoT-based applications, and Robotics applications. Raspberry Pi is slower than a laptop or desktop but is still a computer that can provide all the expected features or abilities, at low power consumption.

The raspberry pi is powered through a Regulated Power Supply which provides 5v and 2.4A of current. The raspberry pi requires an operating system to work and, in this system, Ubuntu is used. Python is the programming language used to configure the operations in raspberry pi and in this proposed system OpenCV is also used to access specific libraries required for image

and speech processing like eSpeak synchronizer, Google synchronizer, etc.,

2.2 MICROPHONE

A microphone is a device that translates sound vibrations in the air into electronic signals and scribes them to a recording medium or through a microphone. Microphones enable many types of audio recording devices for communication, as well as music vocals, sound recording, and speech and it allows computer users to input audio into their computers.

2.3 WEB CAM

A webcam is a video camera that streams or feeds a video or an image in Real-time to or through a computer network, such as the Internet. Web-cams are a smaller version of a camera that sits on a desk, attach to a user's monitor it is also available as a built-in version where the webcam is built into the actual hardware itself. The webcam used in this system has Image resolution interpolated to 25 megapixels with 6 light sensors, 16 MP Image Resolution, USB Interface, Night Vision, Focus Range of 4 cm to infinity Image control color saturation, brightness, sharpness, and brightness are adjustable.

2.4 DISPLAY

In this system, the translated text outputs are displayed on the laptop display as the raspberry pi is connected to the desktop using an indirect connection by connecting both raspberry pi and the laptop/PC to the same network and accessing the IP address of the Raspberry pi.

2.5 POWER SUPPLY

To operate any hardware electrical/electronic component we need a certain amount of power supply so the device can get initialized. In this case, the raspberry pi uses 5 volts 2.4 amperes of power supply which is fed through a Regulated power supply/power charger which is connected by a micro-USB port.

2.6 SOFTWARE

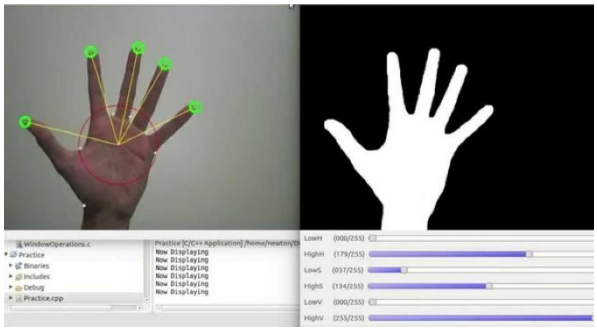
OpenCV is a Python open-source library, for computer vision in Machine Learning, Artificial Intelligence, image processing, etc. OpenCV supports a wide variety of programming languages like C++, Python, Java, etc. It can process videos and images to identify faces, objects, or even the handwriting of a human.

Python is a high-level, general-purpose, and high-level programming language. Python is currently the most widely used multi-purpose, very popular programming language. Python programming language (latest Python 3) is being used in Machine Learning applications, web development, along with all cutting-edge technology in Software Industry.

In this work, OpenCV with Python is used for the implementation of Sign language to text and text-to-speech conversions.

2.7SIGN TO TEXT CONVERSION

In this conversion method, we have the Sign/Hand gesture input which is in the form of an image and that is processed and converted into the equivalent text using DNN (Deep Nural Networks). The working of DNN goes as follows:

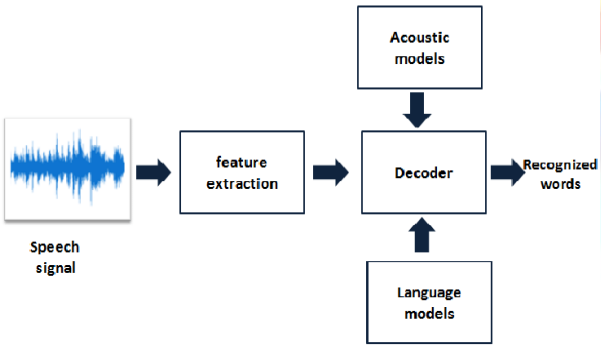


- A computer is given a piece of information, in this case, it's an image.
- The computer passes this image through its DNN, recognizing what it can be and sorting elements of it.

- When it reaches the end of this process it decides if the gesture is correct or not by comparing it with the predefined information.
- It's given feedback on this answer –yes or no– which the computer uses to strengthen its decision making.
- The process is repeated over and over with lots of different hand gestures until the computer can learn to recognize it instantly. Just like a brain would.

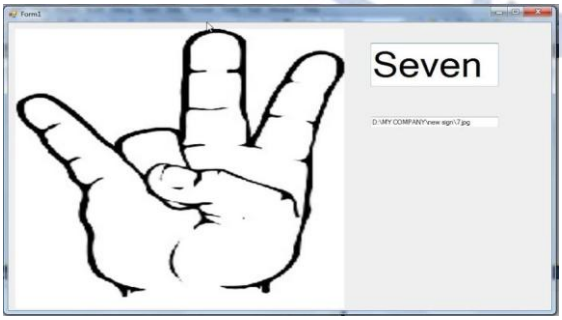
2.8 SPEECH TO TEXT CONVERSION

In speech-to-text conversion, the input is fed through a microphone which is connected to the raspberry pi through a USB connector. This input speech is processed using Google synchronizer where the speech is processed using a deep learning process called automatic speech recognition (ASR). Automatic Speech Recognition (ASR) is a technology that can allow human beings to use their voices to communicate with a computer interface in a way that, in its most sophisticated variations, resembles normal human conversation.



3. RESULTS

This system has two inputs taken in the form of image and voice using a webcam and microphone respectively. When these inputs are processed, the text output is generated and displayed on the screen.



4. CONCLUSION

This work associates the different needs of deaf and mute people using a Convolution neural network. With the use of this device deaf and mute people can vastly enhance their ability to communicate with others and improve their lives. The neural network model adjusts its weights and ensures that the predictions are accurate.

FUTURE SCOPE

Many works can be carried out as an extension of this work. This system predicts the need of the deaf and mute people but future systems may be developed that could communicate to their mobile devices, allowing the system to learn the needs of the user, thereby provisioning the development of re-commendatory systems as they have the relevant data related to the deaf and mute people that can easily be learned through neural network model.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Pooja Gupta, Dr. Ambuj Kumar Agrawal, Dr. Shahnaz Fatima, "Sign Language Problem and Solutions for Deaf and Dumb People", (2019)124-127.

[2] Sawant ramada, Deshpande Saylee, NalePranita, Nerkar Samiksha, Mrs. Archana, S. Vaidya "Intelligent Sign Language Recognition Using Image Processing" (2020)45-50.

[3] Mr. Kunal A. Wankhade1, Prof. Gauri N. Zade2, "Sign Language Recognition for Deaf and Dumb People Using ANFIS", (2014)1206-1209.

[4] Kiraly Patil, Gayatri Pendharkar, Prof. G. N. Gaikwad, "American Sign Language Detection", (2019)1-5.

[5] Shanmukha Swamy M N, Chethan M P, MahanteshGatwadi, Indian Sign Language Interpreter with Android Implementation (2020)36-40.

[6] K. Altun, B. Barshan, and O. Tunçel, "Comparative study on classifying human activities with miniature inertial and magnetic sensors," Pattern Recognit., vol. 43, no. 10, pp. 3605–3620, 2010.

[7] Card, Stuart K.; Thomas P. Moran; Allen Newell (July 1980). "The keystroke-level model for user performance time with interactive systems". Communications of the ACM. 23 (7): 396–410.

[8] G. Bailador, C. Sanchez-Avila, J. Guerra-Casanova, and A. de Santos Sierra, "Analysis of pattern recognition techniques for in-air signature biometrics," Pattern Recognit., vol. 44, nos. 10–11, pp. 2468–2478,2011.

[9] S. Kallio, J. Kela, P. Korpipää, and J. Mäntyjärvi, "User independent gesture interaction for small handheld devices," Int. J. Pattern RecognitArtif. Intell., vol. 20, no. 4, pp. 505–524, 2006.

[10] S. Katsura and K. Ohishi, "Acquisition and analysis of finger motions by skill preservation system," IEEE Trans. Ind. Electron., vol. 54, no. 6, pp. 3353–3361, Dec. 2007.

[11] http://www.raspberrypi.org/archives/3195
[12] https://developer.mbed.org/users/edodm85/notebook/HC-05-blue tooth/
[13] Sadeque Reza Khan and Farzana Sultana Dristy "Android Based Security and Home Automation System", International Journal of Ambient Systems and Applications (IJASA) Vol.3, No.1, March 2015".
[14] https://play.google.com/store/apps/details?id=com.app.control&hl =en
[15] Javale, Deepali, et al. "Home automation and security system using android ask." International journal of electronics communication and computer technology (ICT) 3.2 (2013).
[16] AnujaV.Nair, Bindu.V, "A Review on Indian Sign Language Recognition", International journal of computer applications, Vol. 73, pp: 22, (2013).
[17] Archana S. Ghotkar, Rucha Khatal, Sanjana Khupase, Surbhi Asati& Mithila Hadap, "Hand Gesture Recognition for Indian Sign Language", IEEE International Conference on Computer Communication and Informatics (CCI), pp: 1-4, (2012).
[18] Hu Peng, "Application Research on Face Detection Technology based on Open CV in Mobile Augmented Reality", International Journal of Signal Processing, Image Processing, and Pattern Recognition, Vol. 8, No. 2 (2015).