Sign Language Recognition: Speech to Gesture

Panashe Manyonganise (h180024p@hit.ac.zw)
HIT
Philip Sumbureru
HIT

Research Article

Keywords: Sign Language, Speech Language, Text image processing

Posted Date: April 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1558869/v1

Abstract

Communication is an essential component of existence. Around 360 million people worldwide suffer from hearing loss, 32 million of them are children, and their lives are not as simple as they may be for human beings without boundaries. People who are hard of hearing or almost deaf find it difficult to use cell phones since they can't get to data anywhere due to a lack of administrators. Because of their hearing impairment, which is an invisible disability, they have difficulty reading and writing, as well as perusing and seeing all data on cell phones. The increase in engagement achieved by hard of hearing children over four years is equivalent to the addition of one year for hearing children. This group of people with disabilities does not have access to any visible text-based info. To provide benefits to those who are deaf or hard of hearing in order to promote their social integration and communication. This project provides a Python-based Sign Language Recognition system capable of identifying speech and converting it to the corresponding gesture.

1. Introduction

Gesture-based communication is a language that predominantly uses manual correspondence to convey meaning rather than acoustically transmitted sound samples. This can include simultaneously integrating hand states, hand direction and development, arms or body development, and outer appearance to express a speaker's concerns. Gesture-based communication translators are commonly used to facilitate correspondence between hearing impaired and hearing persons. Such activities require great effort from the translator, as gesture messages are unambiguous common languages with their own syntax, distinct from those expressed in language. Nonverbal communication is an important way for individuals to communicate with one another. Ordinary people can communicate their ideas and thoughts to others through dialogue. The use of gestures as a means of communication is the sole approach for specific strategy for the meeting impaired group. The consultation-impaired community has developed their own style of life and techniques for communicating among themselves and with the general public by employing sign gestures. Rather of transmitting their ideas and concerns audibly, they transmit them using ways for sign examples. Sign signals are a nonverbal visual language that is not the same as spoken language but serves a comparable purpose. It is sometimes difficult for the meeting weaker neighbourhood to communicate their ideas and innovation to the general public. This framework was inspired by an exceptional meeting of people who have problems communicating verbally. It is intended to be used without difficulty by those who are deaf or hard of hearing. The goal of this project is to create a system prototype that automatically recognizes the speaker's audio languages and translates them into sign language.

2. Related Work

This strategy is considered by Nasser H.D et al. [1], where the important features retrieved are SIFT (Scale Invariant Feature Transform) key-points. 4 They then created a language for recognizing dynamic gestures using a succession of hand positions.

[2] establishes a foundation for the use of Hidden Markov Models (HMM) by establishing a similar link between voice recognition and gesture recognition. HMMs may be used to represent time series data, and in this case, the hand's movement along the coordinate axis is monitored and each direction is considered a state. This work employs a vocabulary of forty gestures and obtains an accuracy of 85%. It also mentions the downside that as the vocabulary expands, so will the need to define the hand configuration as well as the hand trajectory, making the design of HMM more difficult and time intensive. We wanted a means to describe dynamic gestures in a more straightforward manner.

The system in [3] employs an intrinsic mobile camera for gesture detection and acquisition; gestures acquired are processed using Algorithms such as the HSV model (Skin Colour Detection), Large Blob Detection, Flood Fill, and Contour Extraction. The system can recognize one-handed sign representations of conventional alphabets (A-Z) and numeric values (0–9). This system's output is highly efficient, reliable, and has a high approximation of gesture processing and voice analysis.

The study [4] focuses on vision-based hand gesture recognition systems, offering a strategy based on a database-driven hand gesture recognition approach and the Thresholding technique, as well as an effective template matching using PCA. Initially, the hand region is split using the skin colour model in the YCbCr colour space. Thresholding is used in the following stage to differentiate foreground from background. Finally, for recognition, a template-based matching approach is constructed utilizing Principal Component Analysis (PCA).

[5] demonstrates this. Human computer interaction (HCI) and sign language recognition (SLR), which aim to create a virtual reality, 3D gaming environment, assist deaf-mute persons, and so on, make heavy use of hand gestures. The primary requirement of any hand gesture-based application system is the segmentation of the hand part from the other body parts and background; however, gesture recognition systems are frequently plagued by different segmentation problems, as well as problems such as coarticulation and recognition of similar gestures.

The fundamental goal of this study [6] is to develop and construct a low-cost wired interactive glove that can be interfaced with a computer running MATLAB or Octave and has a high level of accuracy for gesture detection. The glove uses bend sensors, Hall Effect sensors, and an accelerometer to register the orientation of the hand and fingers. As an error-controlling technique, the data is then delivered to the computer through automated repeat request.

Using skin colour segmentation, the algorithm proposed in [7] is capable of extracting indications from video sequences with less crowded and dynamic backgrounds. It differentiates between static and dynamic gestures and extracts the appropriate feature vector, which is categorized using Support Vector Machines (SVM). Speech recognition is based on the Sphinx standard module.

[8] Formalized paraphrase This work describes a Sign Language Recognition system that uses MATLAB to recognize 26 motions in Indian Sign Language (ISL). The suggested system is comprised of four modules: pre-processing and hand segmentation, feature extraction, sign recognition, and sign to text and

speech conversion. Image processing is used for segmentation. Different characteristics, such as Eigen values and Eigen vectors, are retrieved and employed in recognition. For gesture recognition, the Principal Component Analysis (PCA) technique was utilized, and the identified gesture was transformed into text and audio format.

This research [9] proposes a Hand Gesture Recognition system based on Dynamic Time Warping. The system is divided into three modules: real-time detection of the face region and two hand regions, tracking the hands trajectory in terms of direction between consecutive frames as well as distance from the center of the frame, and gesture recognition based on analysing variations in hand locations as well as the center of the face. The proposed technique overcomes not only the limitations of a glove-based approach, but also most of the vision-based approaches in terms of illumination condition, background complexity, and distance from camera, which can be up to two meters, by using Dynamic2Time Warping, which finds the optimal alignment between the stored database & query features. This results in improved recognition accuracy when compared to conventional methods.

A Wireless data glove, which is a conventional cotton driving glove coupled with flex sensors down the length of each finger and the thumb, is used in [10]. Mute persons can wear the gloves to make hand gestures, which will be transformed into speech so that regular people can comprehend what they're saying. A sign language often gives a sign for the entire word. It can also offer signs for letters to execute words for which there is no matching symbol in that sign language. The main function in this study is played by the Flex Sensor. Flex sensors are sensors whose resistance changes based on the degree of flexion. In this case, the equipment identifies sign language Alphabets and Numbers. It is now working on a prototype to bridge the communication gap between differentiable and regular persons. The software is written in embedded C. The Arduino software is used to monitor the operation of the program in the hardware circuitry, which is built with a microcontroller and sensors.

In [11], an overview of the basic investigation works based on the Sign Language acknowledgment framework is offered, and the created framework structured into the sign catching method and acknowledgment processes is discussed. The characteristics and flaws that contribute to the framework operating well or, in any event, will be highlighted by bringing up major difficulties linked to the established frameworks. Then, a unique technique for developing an SLR framework based on the integration of EMG sensors and an information glove is provided. For apportioning word limits for floods of words in persistent SLR, this technique relies on electromyography data obtained from hand muscles. The proposed framework was used to identify the words division issue, which will contribute to the uninterrupted sign acknowledgment framework's better acknowledgment capability.

[12] proposed a method in which the intended converter would act as a medium by sensing the marked pictures generated by the endorser and then converting those into text and so into dialogue. The flagged images are arranged to improve the algorithm's precision and efficacy.

They presented a system [13] in which they employed a flex sensor to collect data from deaf and stupid persons using sign language, a microprocessor AT89c51 to handle all activities, and an APR 9600 speech chip to store voice data. To communicate with the deaf and dumb, an LCD display and a speaker are employed as output devices. The software tools Keil and protos were used to compile software code and simulate the design.

Countless examination [14] works given out in the last two decades have been investigated. In those works, the several sub-parts and philosophies used for recognizing mostly hand signals have been illustrated. A brief correlation of the foundations, division tactics, and highlights used, as well as the acknowledgement strategies, has been completed and presented.

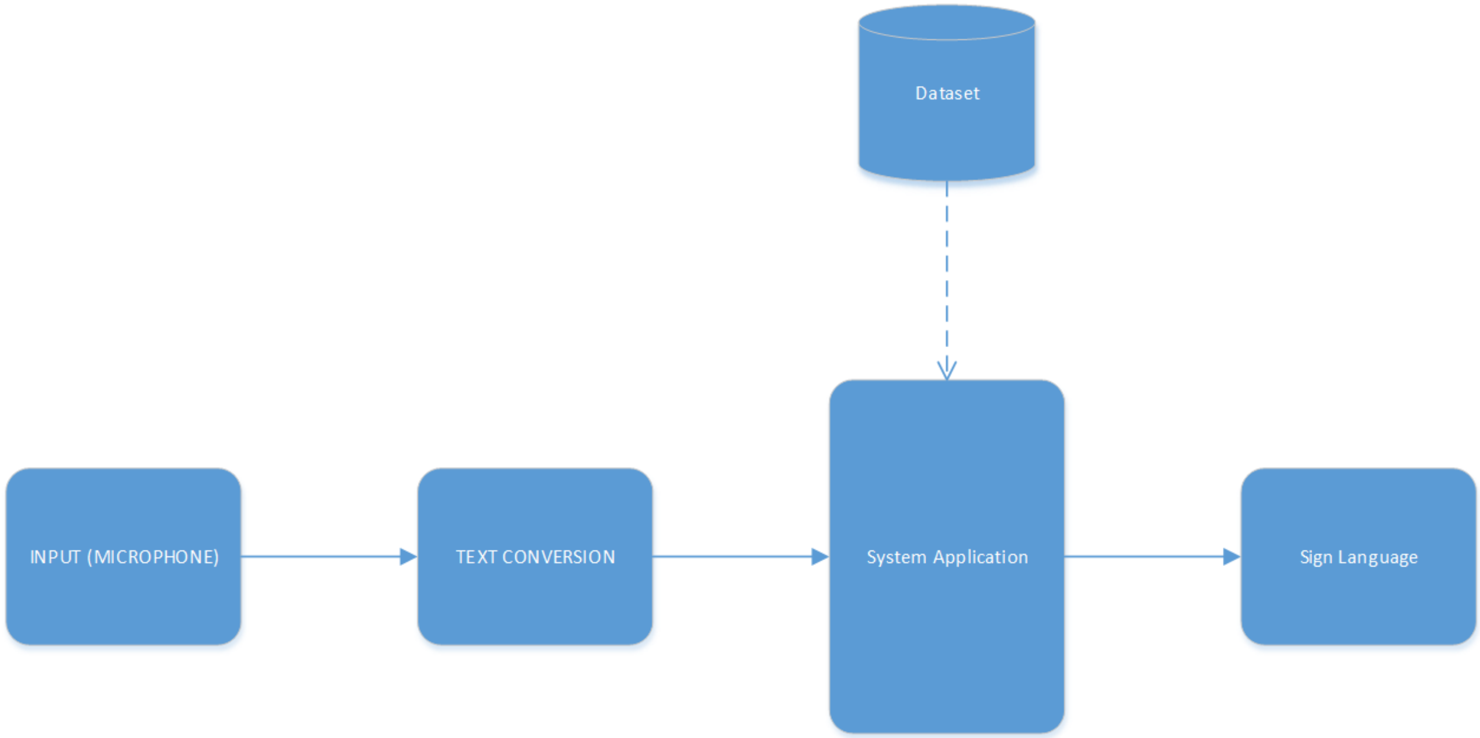
3. Objective

- Facilitate communication for the deaf population
- Provide accurate findings with appropriate data.
- Converting real-time covert speech to sign language.

4. Motivation

Although gesture-based communication is used all over the world to overcome correspondence barriers for hearing or speech impaired people who rely on gesture-based communication for most of their correspondence, there are no effective models that convert text to Indian communication by gestures. There is a scarcity of authentic and effective good assistance for correspondence. While significant progress has been achieved in computer recognition of communications by gestures of many nations, little effort has been done in ISL computerization.

5. System Architecture



6. Experimental Results And Discussion

The suggested framework is simple, and the subject is not need to wear any gloves or electromechanical devices. The technology hears the speech and converts it into alphabets and videos. As a result, the hand motion is visible to deaf persons. For example, ALEX is provided to the system as a speech. The system converts it to the alphabet, such as A, L, E, and X, and then matches the signs in the database to deliver output.

7. Conclusion

This sign language translator can accurately convert alphabets and words. All of the signs can be translated based on the supplied speech. The size of the database may be raised to improve recognition accuracy. We want to overcome the communication gap between persons who are deaf and others who are not.

References

1. N. H, Dardas and N. D. Georganas, "Real-time handGesture detection and recognition using bag-of-features and support vector machine techniques," *IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT*, 2011.
2. E. Starner, "Visual recognition of american sign language using hidden markov models," *Master's thesis, Massachusetts Institute of Technology*, 2015.
3. S. Patel, U. Dhar, S. Gangwani, R. Lad and P. Ahire, "Hand-Gesture Recognition for Automated Speech Generation," *IEEE International Conference On Recent Trends In Electronics Information*

Communication Technology, 2016.

4. M. K. Ahuja and A. Singh, "Static Vision Based Hand Gesture Recognition Using Principal Component Analysis," *3rd International IEEE Conference on MOOCs, Innovation and Technology in Education*, 2015.
5. S. Bhowmick, S. Kumar and A. Kumar, "Hand Gesture Recognition of English Alphabets using Artificial Neural Network," *IEEE 2nd International Conference on Recent Trends in Information Systems*, 2015.
6. T. Chouhan, A. Panse, A. K. Voona and S. M. Sameer, "Smart Glove With Gesture Recognition Ability For The Hearing AndSpeech Impaired," *IEEE Global Humanitarian Technology Conference*, 2014.
7. A. Kumar, K. Thankachan and M. M. Dominic, "Sign Language Recognition," 2016.
8. S. N. Sawant and M. S. Kumbhar, "Real Time Sign Language Recognition using PCA," *IEEE International Conference on Advanced Communication Control and Computing Technologies*, 2014.
9. W. Ahmed, K. Chanda and S. Mitra, "Vision Based Hand Gesture Recognition Using Dynamic Time Warping for Indian Sign Language," *International Conference on Information Science*, 2016.
10. H. J. R.B, P. A, R. A. M. S, S. K and Sumana, "Multiple Sign Language Translation into Voice Message," *International Journal of Engineering and Innovative Technology*, vol. 3, no. 10, 2014.
11. M. E. Al-Ahdal and N. M. Tahir, "Review in Sign Language Recognition Systems," *Review in Sign Language Recognition Systems*, 2012.
12. M. Gupta, M. Garg and P. Dhawan, "Sign Language to Speech Converter Using Neural Networks," *International Journal of Computer Science* & *Emerging Technologies*, vol. 1, no. 3, 2010.
13. J. Thilagavathy, D. murugan and S. Darwin, "Embedded Based Hand Talk Assisting System for Deaf and Dumb," *International Journal of Engineering Research* & *Technology*, vol. 3, no. 3, 2014.
14. A. R. Sarkar, G. Sanyal and S. Majumder, "Hand Gesture Recognition Systems: A Survey," *International Journal of Computer Applications*, vol. 71, no. 15, 2013.