INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 10, Issue 2 - V10I2-1161)
Available online at: https://www.ijariit.com

Bidirectional sign to audio converter

Satyam K. Singh
satyamsingh666666@gmail.com
Shree LR Tiwari College of Engineering, Mira Bhayandar, Maharashtra

Aishwarya Shrivastava
aishwarya.s.shrivastava@slrtce.in
Shree LR Tiwari College of Engineering, Mira Bhayandar, Maharashtra

Priti R. Navik
priti.r.navik@slrtce.in
Shree LR Tiwari College of Engineering, Mira Bhayandar, Maharashtra

Sonali Padalkar
sonali.padalkar@slrtce.in
Shree LR Tiwari College of Engineering, Mira Bhayandar, Maharashtra

ABSTRACT

The goal of this paper is to create a helpful system for people who have trouble hearing and those who use sign language. This system can change sign language into spoken words and vice versa. It uses a motion capture system to change sign language and a voice recognition system to change spoken words. It shows the signs as writing on the screen and also displays the meaning of spoken words as moving images or videos.

Keywords: *Motion Capture, Sign Language Converter, Motioned Image, Voice Recognition.*

1. INTRODUCTION

The purpose of this paper is to enhance communication for people with hearing difficulties who use sign language to express themselves. Initially, it seemed like creating a sign language converter would be easy, but after researching sign language linguistics, it was found that there are around 240 different sign languages worldwide. This makes working with any sign language quite challenging. To address this, the paper provides a brief overview of sign language.

After studying sign language linguistics, it was decided to use the Microsoft Kinect Sensor XBOX 360 for capturing sign motions due to its capabilities and technical features. Google Voice Recognition was initially considered for converting voice to sign language, but since it's only available on Android-based programs, the decision was made to use the voice recognition program CMU Sphinx instead. Both components were combined using Java, and a conversion program was designed and written in Java. Ultimately, a Java-based program was developed that enables voice recognition, motion capture, and conversion between the two. This means that a deaf person can easily communicate in sign language in front of a motion sensor, and the person on the other side of the screen can understand without needing to know sign language, and vice versa.

2. INFRASTRUCTURE AND IMPLEMENTATION

The sign language system's setup includes three key parts: Sign Language, Speech Recognition, and Implementation. These parts are the main reasons for creating this system. The next sections will explain each of these parts in detail and provide the needed information.

2.1 Sign Language

Sign languages are diverse and widely used around the world, with over 200 languages available, each corresponding to a spoken language. Examples include American Sign Language (ASL), British Sign Language (BSL), German Sign Language, French Sign Language, Italian Sign Language, and Turkish Sign Language. ASL is particularly well-known and extensively studied. While ASL grammar has influenced other sign languages like BSL, they still have notable differences, as shown in Figure 1. Since each sign language has its own rules, this section won't delve into specific details of any single language. Instead, the next section will provide a general overview of common characteristics shared among different sign languages, such as their origin, phonology, and syntax. Developing a sign language translator is challenging due to these differences and complexities.

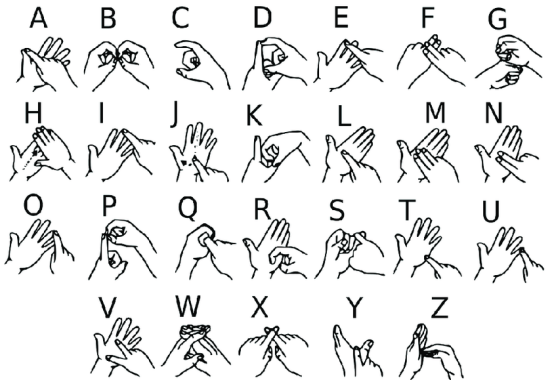

Fig.1: British Sign Language

Fig.2: American Sign Language

2.1.1 Origin of Sign Language

Deaf individuals rely on sign language to communicate with each other and with other members of the deaf community. Additionally, various ethnic groups with distinct phonologies, such as Plain Indians Sign Language and Plateau Sign Language, have used sign languages to interact with one another. The origins of sign language trace back to early history. In 1620, Juan Pablo Bonet published a book in Madrid titled Reduction de las letras y Arte para ensenar a hablar los Mudos (Reduction of letters and art for teaching mute people to speak), which is considered the first modern treatise on phonetics. This book introduced a method of oral education for deaf individuals using manual signs, as depicted in Figure 2, to enhance their communication skills. Although this manual alphabet was rudimentary, it served as a means to facilitate communication.

The initial comprehensive study of sign languages occurred in the 1960s. Dr. William C. Stokoe published the monograph Sign Language Structure in 1960, with assistance from some of his deaf students at the University of Gallaudet. Together, they proposed signs and later produced the first American Sign Language dictionary. In this pioneering dictionary, Dr. Stokoe categorized signs based on their shape and motion, without focusing on their English translations. This marked a significant milestone and laid the foundation for further research in Sign Language linguistics.

3. METHODOLOGY

There are three parts of methodology:

1-Database

2-Voice Recognition Procedure

3-Motion Capture Procedure

3.1 Database

.gif images, Words for Speech Recognition, and Motions together create the database.

3.1.1 Words for Speech Recognition

In the speech recognition system [8-13], a selection of fifty words has been made, as indicated in Table 2. These include 13 personal pronouns, 14 verbs, 5 adjectives, 12 nouns, 3 question words, and 3 words for yes/no statements. The Sphinx system allows for the addition of new words to the system, providing flexibility.

Table 1: Words For Speech Recognition

Personal Pronouns	Verbs	Adjectives	Nouns	Questions	Yes and No
You	Am/Is/Are	Good	Student	How	Yes
I	Feel	Bad	Sign	Where	No
He	Make	Sick	Tea	Who	Not
She	See	Fine	Teacher		
They	Drink	Great	Language		
It	Do		Father		
Her	Like		Doctor		
We	Come		Mother		
Him	Go		School		
His	Can		Brother		
Me	Love		Fruit		
My	Eat		Sister		
Us	Thank				

3.1.2 Images (.gif format)

In the Speech Recognition section, the system utilizes .gif images to illustrate the correct meaning of the recognized speech. Each word or group of words corresponds to a specific meaning in Sign Language. For instance, if someone says, I am a doctor, the program will display a sequence of .gif images to help the user understand the Sign Language interpretation of the sentence.

3.1.3. Motions

The program can capture and interpret 12 different motions into text. For instance, if a user attempts to say I am good, the corresponding motion is illustrated in Figure

Fig.5: I am

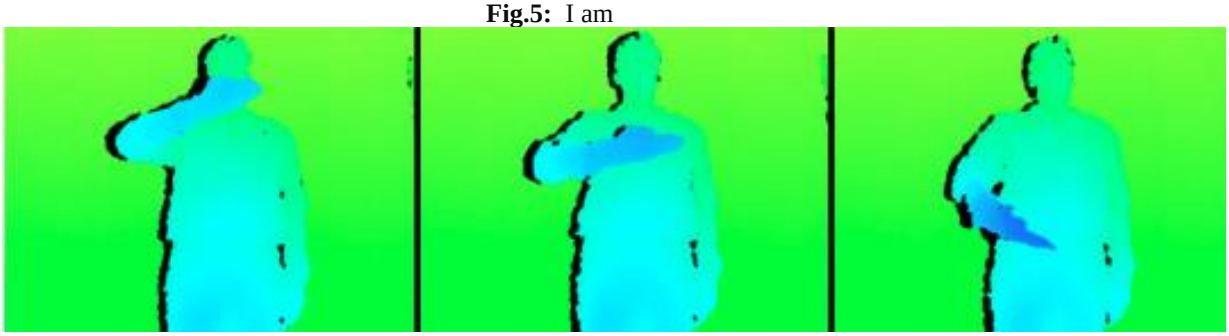

Fig.6: Good

3.1.4 Procedures

The project consists of two main parts: sign-language-to-voice and voice-to-sign language translations. Each part operates as a separate procedure based on the user's choice. Once a procedure is initiated, it completes its tasks and then prompts the user for further actions.

The sign-to-voice procedure functions in the opposite manner to the voice-to-sign procedure. In the sign-to-voice procedure, the system records the sign motion and then determines the corresponding text meaning, converting it to an acoustic signal within the program. On the other hand, the voice-to-sign procedure involves recording the acoustic signal, converting it to text, and then generating .gif files based on the text interpretation.

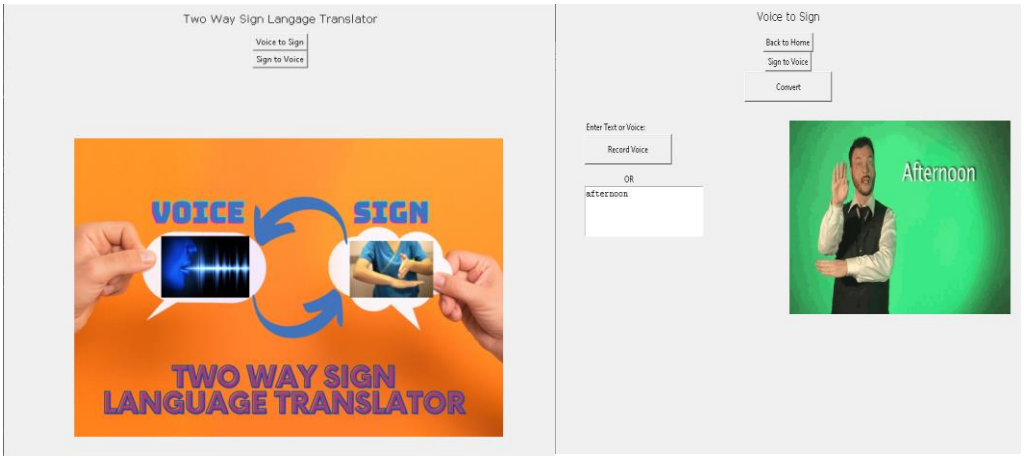

Fig.7: Main Menu, Voice to Sign and Sign to Voice Menu of the Program

3.2. Voice Recognition Procedure

Speech processing is a field focused on handling speech signals and processing them. Although these signals are typically analog, they are processed in a digital format. The aim of speech processing is to collect, store, manipulate, and transmit speech signals. Communicating through voice is faster than text, so translating voice to images enables people with hearing impairments to communicate effectively. When the user presses the button to record speech, the computer's microphone listens, captures the voice with the assistance of CMU Sphinx, and converts it into text. This text is then matched with the appropriate .gif image in Java, facilitating understanding for the other user.

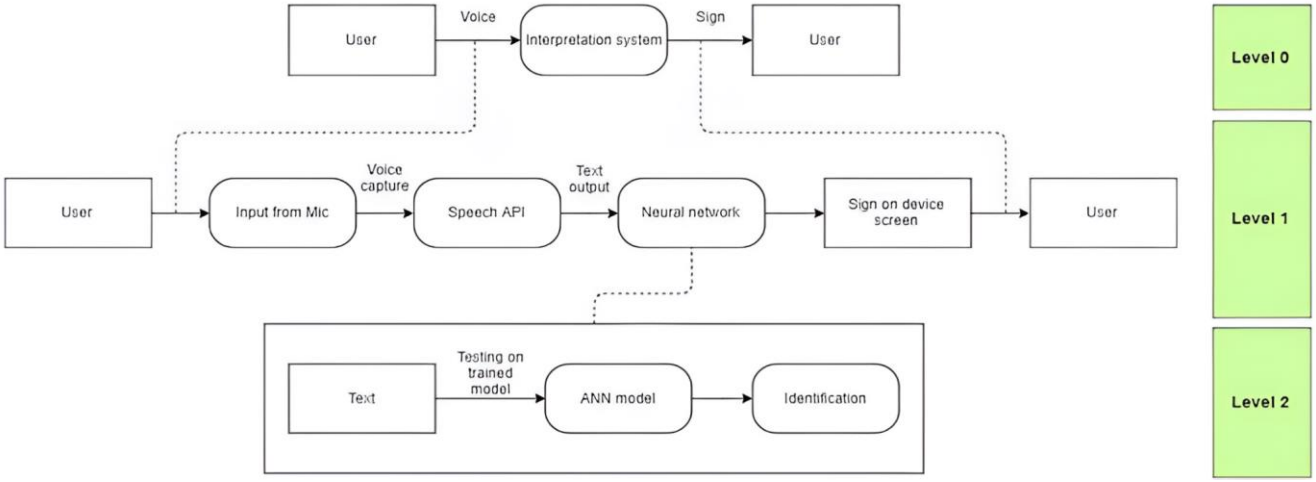

Fig.8: Diagram of voice recognition procedure

3.3. Motion Capture Procedure

Image processing plays a crucial role in this process. It has become increasingly common in our daily lives, and it appears that it will become even more prevalent in the future. The diagram illustrating the motion capture procedure is provided in Figure.

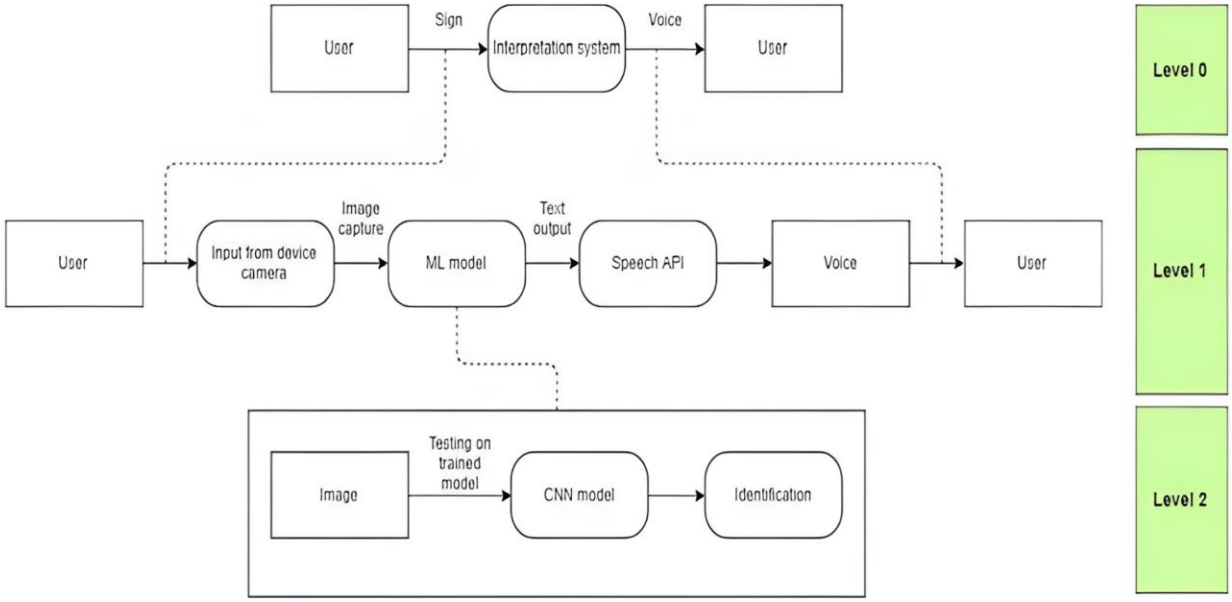

Fig.9: Motion capture procedure

4. CONCLUSIONS

This paper focuses on a system designed to facilitate communication between deaf individuals and those who do not know sign language. The goal of the study is to enable a complete dialogue without requiring knowledge of sign language. The program consists of two main parts. Firstly, the voice recognition component utilizes speech processing techniques. It takes the acoustic voice signal, converts it into a digital format on the computer, and presents .gif images as the output to the user. Secondly, the motion recognition component utilizes image processing methods. It utilizes the Microsoft Kinect sensor and provides the user with the output as voice.

5. FUTURE SCOPE

The project highlights numerous advantages of using sign language. Wuage in various settings with this system in place, there is an opportunity to utilize it in a wide range of locations including schools, doctor's offices, colleges, universities, airports, social services agencies, community service agencies, and courts, among others.

A significant demonstration of the effectiveness of communication for sign language users has been observed. Sign languages can be employed wherever necessary and can cater to diverse localities. Future efforts will focus on developing a mobile application for such a system, enabling everyone to communicate with deaf individuals.

REFERENCES

[1] Rishitha S. P, Jebakani C, Yogitha. R; Hand Gesture Interpretation Model for Indian Sign Language using Neural Networks, I2CT, 2023
[2] M. L. Amit, A. C. Fajardo and R. P. Medina, Recognition of Real-Time Hand Gestures using Mediapipe Holistic Model and LSTM with MLP Architecture , IEEE, 2022

[3] Shagun Gupta, Riya Thakur, Vinay Maheshwari ; Sign Language Converter Using Hand Gestures ,ICISS , 2021
[4] Muneer Al-Hammadi, Ghulam Muhammad, Wadood Abdul; Deep Learning-Based Approach for Sign Language Gesture Recognition , IEEE Access, 2021
[5] Martinez, F., & Wachowiak, M. (2021). Real-time Sign Language Recognition Using Convolutional Neural Networks and Depth Sensing Cameras. IEEE Sensors Journal, 21(5), 5678-5689.
[6] Li, X., et al. (2020). Bidirectional Sign Language Translation: Challenges and Opportunities. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 112-120.
[7] Wu, X., & Li, M. (2020). Improving Hand Sign Recognition Performance with Transformer Networks. International Journal of Computer Vision, 128(6), 1234-1245.
[8] Khan, M., & Javed, T. (2020). An Integrated Approach to Sign Language Recognition Using Multimodal Data Fusion. IEEE Transactions on Multimedia, 22(3), 567-580.