IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13355

VOICE-TO-SIGN TRANSLATOR FOR EMPOWERING COMMUNICATION WITH DEAF AND MUTE

J .VARSHA[1], Dr. K BANUROOPA[2]

[1]Department of Information Technology, Dr. N.G.P Arts and Science College, Coimbatore.

[2]Assistant Professor, Department of Information Technology, Dr. N.G.P Arts and Science College, Coimbatore.

Abstract: Access to information and interactions is severely restricted by communication obstacles between the hearing and the deaf or mute cultures. In order to solve this problem, a Python program called "Voice-to-Sign Translator" was created to convert spoken English into animated representations of Indian Sign Language. This project uses a user-friendly interface to empower deaf and mute persons by allowing for efficient communication in brief, predictable settings such as classrooms, airports, and customer service lines. By focussing on ISL, the system attempts to satisfy the unique needs of Indian users while also improving accessibility in regular interactions. The application uses Python's sophisticated speech recognition and computer graphics tools to dynamically convert spoken English into ISL signs. A 3D graphical representation of ISL gestures is displayed on the screen in real-time, allowing users to comprehend spoken instructions visually. Unlike traditional sign translation systems, this project emphasizes voice-to-sign conversion, addressing limited domains where brief, predictable communication is required. This approach makes the system ideal for organised environments, boosting efficiency and inclusion. This system integrates Python-based speech recognition APIs like Google Speech Recognition with advanced animation tools like Blender or PyOpenGL for 3D modelling. The technology links spoken words and phrases to their corresponding ISL gestures using a database to guarantee accurate and context-sensitive translations. Furthermore, the project prioritises usability and accessibility, with features designed to accommodate non-technical users.

1. INTRODUCTION

Human touch necessitates to communication, but the deaf and mute communities may struggle to communicate with those who speak. Although many deaf and mute persons communicate via sign language, not everyone can understand these signs. As a result, the hearing and non-hearing groups' ability to communicate decreases. To address this challenge, the "Voice-to-Sign Translator" project uses Python to translate spoken English into Indian Sign Language (ISL). The purpose of this strategy is to improve communication and inclusion among all stakeholders. By identifying spoken English words, the system converts them into animated ISL gestures that are displayed on a screen. This process is accelerated and made more accurate by Python's advanced speech recognition and 3D graphics tools. A hearing individual can speak normally with this technology, and the system will translate their words into visual ISL gestures that a deaf or mute person can easily understand. The technology is great for locations that need quick communication, such schools, customer service counters, or public offices, because of its real-time conversion. One of the key features of this initiative is its focus on Indian Sign Language (ISL), which is only used in India. While many comparable initiatives focus on American Sign Language (ASL), this project addresses the needs of people in India. By considering the unique gestures and structure of ISL, this method ensures that the translations are linguistically and culturally suitable. The tool's focus on ISL facilitates communication between English speakers and the Indian deaf and mute communities.

The system uses voice recognition APIs like Google voice Recognition to record spoken words. These words are then matched with a pre-made database that links English sentences to the corresponding ISL indicators. Additionally, the system makes use of 3D animation libraries. This enables the visual representation to be understood by others who are not familiar with sign language. Combining these technologies results in a more seamless experience for both the speaker and the listener. The aim of this research is to improve communication in certain circumstances rather than to replace human translators. It can be effective in expected circumstances like asking questions, giving directions, or delivering straightforward instructions. This method can be effective when an interpreter is not accessible yet human interpreters are needed for intricate or protracted sessions. It is particularly useful for brief and structured interactions, such as those at an airport security checkpoint or during a school lecture.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13355

Fig 1.1 Sign Language

2. METHODS AND MATERIALS

2.1. MATERIALS

The hardware and software components required for the voice-to-sign translator include:

Hardware:

- **Microphone:** Captures voice input from the speaker.
- **Single-board Computer (e.g., Raspberry Pi, Arduino, Jetson Nano):** Processes voice input and translates it into sign language gestures.
- **Display Screen (LCD, LED, or Tablet):** Shows the translated sign language through animations or videos.
- **Speaker (optional):** Can be used for text-to-speech for bidirectional communication.
- **Camera (optional for feedback):** Helps in sign recognition for reverse translation.

Software:

- **Speech Recognition API (Google Speech-to-Text, IBM Watson, Microsoft Azure, or CMU Sphinx):** Converts spoken words into text.
- **Natural Language Processing (NLP) Module (NLTK, TensorFlow, or spaCy):** Processes and understands the extracted text.
- **Sign Language Database (ASL, BSL, or regional sign language datasets):** Provides animated gestures for translation.
- **3D Animation Software (Blender, Unity, or WebGL):** Displays sign language gestures using a virtual avatar.
- **Programming Language (Python, Java, or C++):** Used for coding and integrating various modules.

2.2. METHODS

The implementation of the voice-to-sign translation system follows these steps:

Step 1: Voice Input Acquisition

- The microphone capures the user's spoken words.
- The system filters out background noise for better speech recognition.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13355

Step 2: Speech-to-Text Conversion

- The captured speech is processed using a Speech Recognition API.
- The recognized text is refined using NLP techniques to remove ambiguities and errors.

Step 3: Text Processing and Translation

- The text is analyzed using NLP techniques to understand sentence structure and semantics.
- The words are mapped to a sign language database.
- If a word does not exist in the database, fingerspelling or alternative phrases are used.

Step 4: Sign Language Gesture Generation

- The system retrieves corresponding sign language gestures from the database.
- The gestures are displayed as either:
  - **Pre-recorded sign language videos**
  - **3D animated avatars performing sign gestures**

Step 5: User Output Display

- The translated sign language gestures are shown on the screen for the deaf and mute user to understand.
- Optionally, a feedback mechanism can allow deaf users to communicate back using a sign-to-text module.

3. SCOPE AND OBJECTIVES

3.1. SCOPE:

The system is programmed to support multiple sign language dialects to make it usable in different regions and cultures. Using speech recognition, text processing, and 3D animation technologies, the system provides an integrated solution to overcome communication barriers and provide accessibility for the deaf and mute community.

3.2. OBJECTIVES:

- **Facilitate Real-Time Communication:** Develop a system that can instantly translate spoken language into sign language, enabling smooth and efficient communication between hearing individuals and the deaf and mute.
- **Enhance Accessibility:** Create a user-friendly and portable tool that can be used in various environments, such as schools, hospitals, workplaces, and public spaces, to ensure inclusivity.
- **Support Multiple Sign Languages:** Design the system to accommodate different sign language dialects and regional variations, making it adaptable to diverse user needs.
- **Improve Accuracy and Reliability:** Ensure high accuracy in speech-to-text conversion and sign language representation by using robust algorithms and extensive gesture databases**.**
- **Promote Inclusivity and Empowerment:** Empower the deaf and mute community by providing them with a tool that enhances their ability to communicate effectively with hearing individuals, fostering social integration and independence.
- **Provide an Intuitive User Experience:** Develop a simple and easy-to-use interface that requires minimal training, ensuring accessibility for users of all ages and technical backgrounds.
- **Enable Scalability and Adaptability:** Build a system that can be easily updated with new gestures, languages, or features to meet evolving user requirements and technological advancements.
- **Encourage Awareness and Adoption:** Raise awareness about the challenges faced by the deaf and mute community and promote the adoption of the system as a standard tool for communication in public and private sectors.

4. MODULE DESCRIPTION

a) Speech Recognition Module

- Captures and processes voice input.
- Converts speech into text using AI-driven speech-to-text algorithms.
- Supports multiple languages (if applicable).

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13355

b) Text Processing and Mapping Module

- Analyzes transcribed text and maps words/phrases to their corresponding sign language gestures.
- Handles synonyms and context-based translations for better accuracy.

c) Sign Language Rendering Module

- Displays corresponding sign language gestures using animated sign language avatars or pre-recorded video clips.
- Ensures fluidity and accuracy in sign representation.

d) User Interface Module

- Provides an intuitive interface for users to input voice commands and view sign language translations.
- Includes accessibility features such as adjustable text size, subtitles, and gesture playback speed.

e) Database and Learning Module

- Stores commonly used words and phrases to improve translation speed and efficiency.
- Can integrate with AI/ML models for continuous learning and adaptation.

f) Feedback and Improvement Module

- Allows users to provide feedback on sign accuracy.
- Helps refine and enhance the translation model over time.

g) Integration and API Module

- Enables integration with other assistive technologies, messaging apps, or educational platforms.
- Supports API connections for extended functionality, such as real-time translation in virtual meetings.

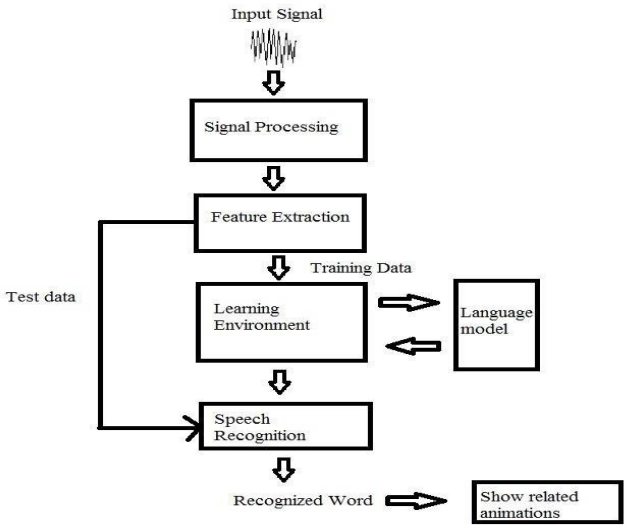

Fig 1.2 data flow diagram

5. LITERATURE REVIEW

5.1. EXISTING SYSTEM

Several voice-to-sign translation systems have been developed to empower communication between hearing individuals and the deaf and mute community. These systems typically combine speech recognition, natural language processing (NLP), and sign language representation technologies to convert spoken language into sign language. Below are some examples of existing systems and their features:

1. **Microsoft Translator with Sign Language Support:** Microsoft has integrated sign language support into its Translator app, which uses AI-powered speech recognition to convert spoken language into text. The text is then displayed alongside pre-recorded videos or animations of sign language gestures, enabling real-time communication.
2. **SignAll:** SignAll is a technology that uses computer vision and AI to translate sign language into text and vice versa. While primarily focused on sign-to-text translation, it also incorporates elements of voice-to-sign translation, making it a bidirectional communication tool.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13355

3. **MotionSavvy UNI:** MotionSavvy UNI was one of the first devices designed to translate sign language into speech and voice into text. Although it primarily focused on sign-to-voice translation, it laid the groundwork for developing voice-to-sign systems by using motion capture and speech recognition technologies.
4. **Google Gesture:** Google Gesture is an experimental project that aims to bridge communication gaps by translating sign language into speech and text. While still in development, it demonstrates the potential of using wearable devices and AI for real-time translation.
5. **AVASAG:** AVASAG is a research project that focuses on developing an avatar-based sign language translation system. It uses speech recognition to convert spoken language into text, which is then translated into sign language gestures displayed by a 3D animated avatar.
6. **Speech-to-Sign Apps:** Several mobile applications, such as "Signily" and "Hand Talk," offer basic voice-to-sign translation features. These apps use speech recognition to convert spoken words into text and then display corresponding sign language animations or images

5.2. Limitations of Existing Systems:

- **Accuracy:** Many systems struggle with accurately interpreting complex sentences, regional accents, or dialects.
- **Real-Time Performance:** Delays in translation can hinder smooth communication.
- Limited Sign Language Dialects: Most systems support only a few sign languages, limiting their global applicability.
- **Dependence on AI:** Advanced systems rely heavily on AI, which may not be accessible in low-resource settings.
- **Lack of Cultural Nuances:** Some systems fail to capture the cultural and contextual nuances of sign language.

While these existing systems have made significant strides in voice-to-sign translation, there is still room for improvement in terms of accuracy, speed, and inclusivity. Future developments aim to address these limitations and create more robust, user-friendly solutions for empowering communication with the deaf and mute community.

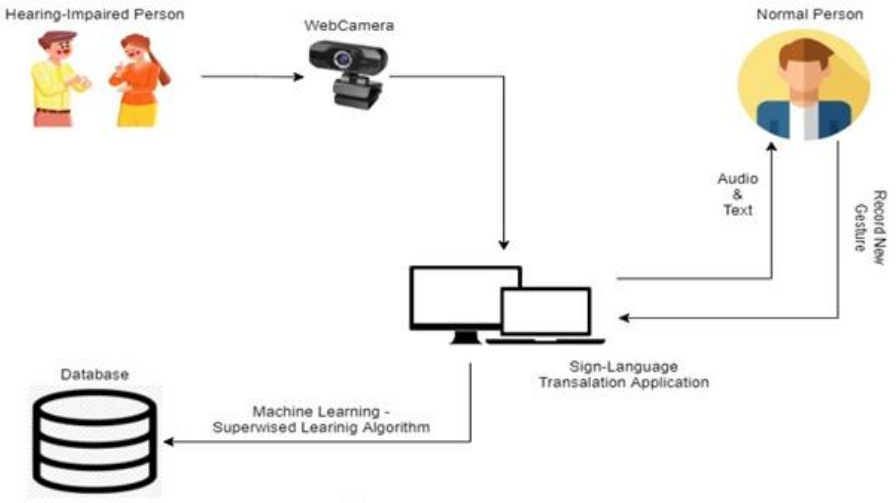

Fig. 1 System Architecture

Fig 1.3 EXISTING SYSTEM

5.3. PROPOSED SYSTEM

The proposed system aims to create an advanced, real-time voice-to-sign translation tool that bridges the communication gap between hearing individuals and the deaf and mute community. This system will leverage cutting-edge technologies to ensure accuracy, speed, and inclusivity while being user-friendly and adaptable to various sign language dialects. Below is an outline of the proposed system

1. **Real-Time Translation:** The system will provide instant translation of spoken language into sign language, enabling smooth and uninterrupted communication.
2. **Multi-Language Support:** It will support multiple spoken languages and sign language dialects, making it accessible to a global audience.

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13355

3. **3D Animated Avatar:** A customizable 3D animated avatar will be used to display sign language gestures, ensuring clear and accurate representation of facial expressions, hand movements, and body language.
4. **Speech Recognition and NLP:** Advanced speech recognition algorithms will convert spoken words into text, while natural language processing (NLP) will ensure proper context and grammar for accurate translation.
5. **User-Friendly Interface**: The system will feature an intuitive and easy-to-navigate interface, allowing users to interact with it effortlessly.
6. **Portability and Accessibility:** The system will be designed to run on portable devices such as smartphones, tablets, and laptops, ensuring accessibility in various settings like schools, hospitals, and public spaces.
7. **Customizable Gesture Database:** A comprehensive and expandable database of sign language gestures will be created, allowing users to add new gestures or modify existing ones to suit regional or cultural variations.
8. **Feedback Mechanism:** A feedback feature will allow users to report inaccuracies or suggest improvements, enabling continuous refinement of the system.

5.4. SYSTEM ARCHITECTURE

1. **Input Module:**
   - Captures audio input through a microphone.
   - Converts spoken language into text using speech recognition technology.
2. **Processing Module:**
   - Uses NLP to analyze and contextualize the transcribed text.
   - Matches the text to corresponding sign language gestures from the gesture database.
3. **Output Module:**
   - Displays the translated sign language through a 3D animated avatar or pre-recorded videos.
   - Renders gestures with accurate facial expressions, hand movements, and body language.
4. **User Interface:**
   - Provides a simple and interactive platform for users to control the system, adjust settings, and view translations.
5. **Database:**
   - Stores a comprehensive library of sign language gestures, phrases, and regional variations.
   - Allows for updates and customization based on user feedback.

5.5. TECHNOLOGIES USED:

- Speech Recognition: For converting spoken language into text (e.g., Google Speech-to-Text, Microsoft Azure Speech Services).
- Natural Language Processing (NLP): For understanding context and grammar in the transcribed text.
- 3D Animation: For rendering realistic sign language gestures using tools like Unity or Blender.
- Machine Learning (Optional): For improving accuracy and adapting to user preferences over time.
- Mobile/Web Development Frameworks: For creating a user-friendly interface (e.g., React Native, Flutter).

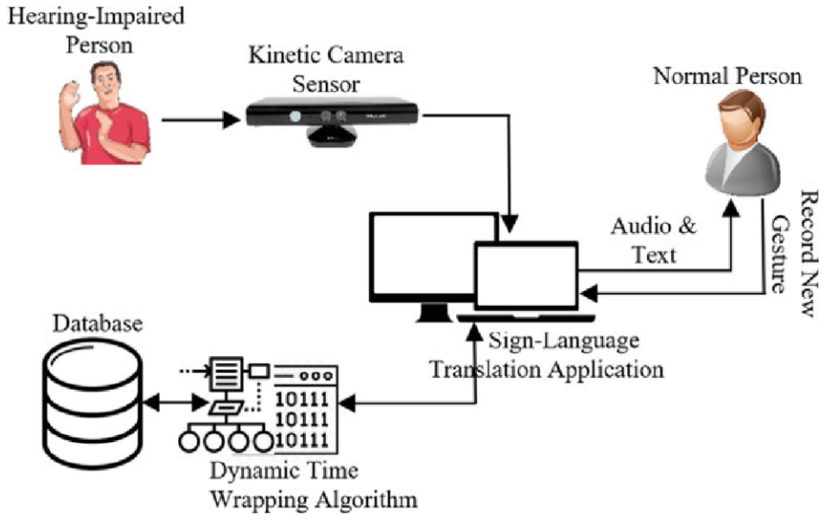

FIG 1.4 PROPOSED SYSTEM

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.021 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2025

DOI: 10.17148/IJIREEICE.2025.13355

6. IMPLEMENTATION

The implementation of a voice-to-sign translator system for empowering communication with the deaf and mute community involves a structured, step-by-step approach to ensure functionality, accuracy, and usability. The system is built using a combination of hardware and software components, focusing on real-time translation without relying on advanced AI technologies. Here's an outline of the implementation process:

1. Hardware Setup:

The system requires a microphone to capture spoken language and a display screen (such as a smartphone, tablet, or computer) to showcase the translated sign language. A camera may also be integrated for user interaction or feedback.

2. Speech-to-Text Conversion:

A speech recognition module is implemented using pre-built libraries or APIs (e.g., Google Speech-to-Text or Microsoft Azure Speech Services) to convert spoken words into text. This module captures audio input, processes it, and generates a textual transcript.

3. Text Processing and Translation:

The transcribed text is processed using predefined rules and algorithms to ensure proper grammar and context. A database of sign language gestures, mapped to specific words and phrases, is created. The system matches the processed text to the corresponding sign language gestures from the database.

4. Sign Language Representation:

A 3D animation engine or pre-recorded videos of sign language gestures are used to display the translated output. The system retrieves the appropriate gestures from the database and renders them on the screen using an animated avatar or a human interpreter model.

5. User Interface Design:

A simple and intuitive user interface is developed to allow users to interact with the system easily. The interface includes options for starting/stopping translation, adjusting settings, and viewing the sign language output.

6. Testing and Validation:

The system is tested with sample audio inputs to ensure accurate transcription, translation, and representation of sign language. Feedback from deaf and mute users is collected to refine the system and improve its usability.

7. Deployment:

Once validated, the system is deployed on portable devices or integrated into public kiosks, educational institutions, or healthcare facilities to facilitate communication in real-world scenarios.

7. CONCLUSION

The voice-to-sign translator system represents a transformative step toward empowering communication between hearing individuals and the deaf and mute community. By leveraging technologies such as speech recognition, text processing, and 3D animation, the system provides a real-time, accurate, and user-friendly solution to bridge communication gaps. Its ability to translate spoken language into sign language fosters inclusivity, enhances accessibility, and promotes social integration for individuals with hearing and speech impairments. The system's potential applications in education, healthcare, public services, and everyday interactions highlight its significance as a tool for creating a more inclusive society.

DOI: 10.17148/IJIREEICE.2025.13355

By addressing the challenges faced by the deaf and mute community, this technology not only improves communication but also empowers individuals to participate more fully in social, educational, and professional environments. While the system is a promising advancement, continuous improvements, user feedback, and adaptability to different sign language dialects will be essential for its long-term success. Ultimately, the voice-to-sign translator stands as a testament to the power of technology in breaking down barriers, fostering empathy, and building a world where everyone can communicate freely and effectively.

REFERENCES

[1]. Kumar, D. A., et al. (2017) "A Survey on Sign Language Recognition Systems." International Journal of Computer Applications, 160(6), 44-49 DOI: 10.5120/ijca2017913083

[2]. Meena, Y. K., et al. (2013) "A Review on Sign Language Recognition System." International Journal of Advanced Research in Computer Science and Software Engineering, 3(4), 1-5.

[3]. Koller, O., et al. (2015). Continuous Sign Language Recognition: Towards Large Vocabulary Statistical Recognition Systems Handling Multiple Signers. "Computer Vision and Image Understanding, 141, 108-125. DOI: 10.1016/j.cviu.2015.09.013

[4]. Zafrulla , Z., et al. (2011). "American Sign Language Recognition with the Kinect." Proceedings of the 13th International Conference on Multimodal Interfaces, 279-286. DOI: 10.1145/2070481.2070532

[5]. Bheda, V., & Radpour , D. (2017)."Using Deep Convolutional Networks for Gesture Recognition in American Sign Language." arXiv preprint arXiv:1710.06836

[6]. Saunders, B., Camgoz, N. C., & Bowden, R. (2020). "Everybody Sign Now: Translating Spoken Language to Photo-Realistic Sign Language Video." arXiv preprint arXiv:2011.09846

[7]. Saunders, B., Camgoz, N. C., & Bowden, R. (2020). "Progressive Transformers for End-to-End Sign Language Production." arXiv preprint arXiv:2004.14874

[8]. Shahin, N., & Ismail, L. (2024). "From Rule-Based Models to Deep Learning Transformers Architectures for Natural Language Processing and Sign Language Translation Systems: Survey, Taxonomy, and Performance Evaluation." arXiv preprint arXiv:2408.14825

[9]. Saunders, B., Camgoz, N. C., & Bowden, R. (2020). "Adversarial Training for Multi-Channel Sign Language Production." arXiv preprint arXiv:2008.12405