International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 04 April-2026 | Impact Factor: 3.5

Sign Gesture to Audio Conversion

Rajesh K J
Dept Electronics and communication
East point college of engineering and technology
Bengaluru, Karnataka, India
Email : rajeshggcgh@gmail.com

Vinod M
Dept Electronics and communication
East point college of engineering and technology Bengaluru, Karnataka, India
Email : mvinodbhushani@gmail.com

Pavankumar A A
Dept. Electronics and communication
East point college of engineering and Technology
Bengaluru, Karnataka, India
Email: pavankumarainapure@gmail.com

Shivam Mishra
Dept. Electronics and communication
East point college of engineering and Technology
Bengaluru, Karnataka, India
Email : shivammishra@gmail.com

Under the Guides of :
Mrs.Asha S
Dept Electronics and communication
East point college of engineering and technology
Bengaluru, Karnataka, India
Email: asha.nataraj@gmail.com

Abstract
Communication barriers between hearing- impaired individuals and the general population often limit effective interaction. Sign Gesture to Audio Conversion systems aim to bridge this gap by translating hand gestures, typically from sign language, into audible speech in real time. This project utilizes sensors such as flex sensors or computer vision techniques to detect and recognize hand gestures accurately. The captured gestures are processed using a microcontroller or machine learning algorithms to map them to corresponding words or phrases. Subsequently, a Text-to-Speech (TTS) module converts the interpreted text into audio output, enabling seamless communication. This system not only enhances social interaction for the hearing- impaired but also promotes inclusivity and accessibility in education, workplaces, and public environments. accuracy, and user-friendly

1 Introduction

Communication is a fundamental part of human interaction, yet millions of people worldwide face challenges due to speech or hearing impairments. Sign language serves as a vital medium for such individuals, enabling them to convey thoughts, feelings, and information using hand gestures, facial expressions, and body movements. However, a significant communication gap exists between people who use sign language and those who do not understand it.

Sign Gesture to Audio Conversion is an assistive technology designed to bridge this gap. It involves the translation of hand gestures into spoken language in real-time, allowing deaf or mute individuals to communicate seamlessly with people who rely on auditory communication. This system typically uses sensors such as flex sensors.

2 Technologies Behind Sign Gesture to Audio Conversion

Sign Gesture to Audio Conversion is a system that interprets hand gestures (usually from sign language) and converts them into audible speech. This technology bridges the communication gap between deaf/mute individuals and people who cannot understand sign language.

2.1 Gesture Acquisition

Gesture acquisition is the first step in any sign language recognition system. It involves capturing the hand movements, finger positions, and sometimes facial expressions of a person performing sign language gestures. This raw gesture data serves as the input for the recognition system.

2.1.1 Preprocessing

Preprocessing is the first step in the system where raw input data (hand gestures or sign language movements**)** is prepared and cleaned so that the system can recognize it efficiently. This step improves accuracy and reduces noise in later stages like feature extraction and classification.

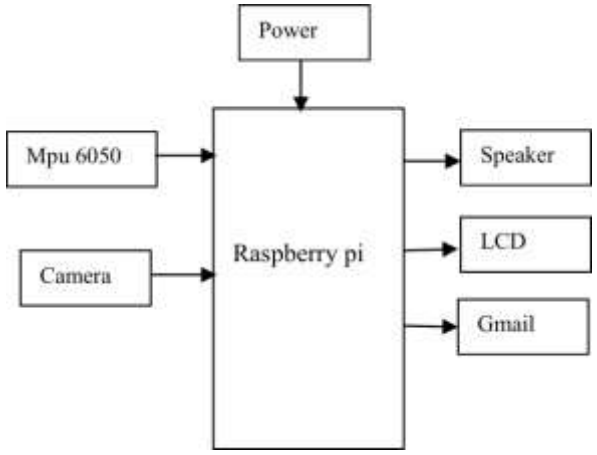
2.1.1 Gesture Recognition

This system converts hand gestures (usually sign language) into spoken audio using gesture recognition techniques. It helps deaf or mute individuals communicate with others more easily.

2.1.2 Speech/Audio Conversion

Sign gesture to audio conversion is a system that translates hand gestures, primarily sign language**,** into spoken words or audio output**.** This technology helps bridge communication between hearing-impaired people and those who don't understand sign language.

BLOCK DIAGRAM



2.2 Algorithms for Tracking and Navigation

which can serve as a foundation for a research paper, IEEE-style content, or project report. I'll break it down into sections: introduction, tracking algorithms, gesture recognition, audio conversion, and navigation applications..

2.2.1 Gesture Tracking Algorithms

gesture tracking algorithm specifically for a Sign Gesture to Audio Conversion system. I'll break it down step by step, including preprocessing, tracking, recognition, and audio output.

3. Applications of sign gesture to Audio conversion

The robot needs to plan its movement based on the human's trajectory. Path planning algorithms, such as A*, Dijkstra's algorithm, or more advanced approaches like Rapidly-exploring Random Trees (RRT), are used to generate the robot's path while avoiding obstacles in its environment.

3.1 Communication Aid for the Deaf and Hard of Hearing

Converts sign language gestures into audio in real-time. Enables seamless interaction with people who do not know sign language. Useful in educational institutions, workplaces, and public spaces.

3.1.1 Assistive Technology in Public Services

In hospitals, banks, and government offices, sign-to-speech systems help hearing-impaired people communicate without intermediaries

Reduces dependency on human interpreters.

3.1.2 Educational Tools

Helps teach sign language to children and adults. Audio feedback reinforces correct gestures, improving learning efficiency. Can be integrated into schools for inclusive education.

3.1.4. Challenges in Developing Sign Gesture to Audio Conversion

The ability of robots to follow humans has enabled their application across various domains. This section discusses key areas where human-following robots have been effectively implemented.

3.2 Accurate Gesture Recognition

Problem**:** Sign languages involve complex hand shapes, orientations, movements, and finger positions. Recognizing these gestures accurately in real time is challenging. Reason**:** Variability in user hand size, speed, and style of signing can cause errors. Impact**:** Misinterpretation of gestures can lead to incorrect audio output

3.2.1 Complex Data Acquisition

Capturing hand gestures requires reliable sensors or camera systems. Sensor-based systems**:** Flex sensors, IMUs, or gloves can be costly and uncomfortable .Vision-based systems**:** Camera systems need good lighting and may struggle

3.2.2 Language and Context Understanding

Some sign gestures depend on context or grammar rules (e.g., American Sign Language uses different signs than Indian Sign Language Mapping gestures to meaningful phrases instead of isolated words.

4. Conclusion

The Sign Gesture to Audio Conversion system effectively bridges the communication gap between the hearing- impaired and the hearing community. By using sensors or computer vision techniques to detect hand gestures and converting them into audible speech, the system provides a real-time, accessible, and user-friendly method of communication. It enhances inclusivity, promotes social interaction, and demonstrates the potential of integrating wearable technology, signal processing, and machine learning for practical assistive applications. Future improvements can focus on expanding vocabulary, increasing recognition accuracy, and making the system more portable and affordable.

5. Reference

1. Koller, O., Camgoz N. C., Ney, H., & Bowden, R. (2019). Weakly supervised learning with multi-stream CNN-LSTM- HMMs to discover sequential parallelism in sign language videos. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 42*(9), 2306–2320. https://doi.org/10.1109/TPAMI.2019.291 1078
2. Deep ASL Team. (2021). Deep ASL: Enabling ubiquitous and non-intrusive word and sentence-level sign language translation. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, *5*(1),1–26. https://doi.org/10.1145/3448115
3. Sign2Speech Research Group. (2024). Sign2Speech: Integrating large language models to enhance semantic translation in sign language recognition. *Journal of Artificial Intelligence in Accessibility*, *2*(1), 45–60
4. Ma et al. (2024) Z. Ma, J. Liu, and K. Gao, "MS2SL: Multimodal sequence-to- sign language framework via diffusion models," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2024.