

AUDIO TO SIGN LANGUAGE TRANSLATION USING AI

Dr. Regonda Nagaraju¹, A. Asha², A. Vinay³, A. Varun⁴, A. Preetham⁵, A. Harshitha⁶

^{1,2,3,4,5,6}Department of Artificial Intelligence and Machine Learning, Malla Reddy University, Hyderabad, India-500100

ABSTRACT

The development of technology that facilitates communication across language barriers has seen tremendous growth with the integration of Artificial Intelligence (AI). This paper presents an innovative AI-driven system for real-time translation of audio to sign language, aimed at enhancing accessibility for the Deaf and hard-of-hearing communities. Traditional methods of bridging spoken language and sign language, such as human interpreters or textual captions, face challenges in terms of availability, accuracy, and immediacy. Our proposed solution leverages advanced AI techniques, including speech recognition, natural language processing (NLP), and 3D computer vision, to provide an accurate and context-aware translation of spoken audio into sign language.

The system operates by first converting spoken language into text through automatic speech recognition (ASR) modules. NLP algorithms then analyze the text to understand context and semantics, ensuring nuanced interpretations aligned with sign language structure and grammar. Finally, a 3D avatar is used to visualize the corresponding sign language gestures, with attention to natural movement and expressiveness critical for effective communication. Our approach prioritizes accuracy in linguistic features such as finger spelling, non-manual markers (facial expressions and body language), and sentence structure, to better reflect the complexities of sign language.

1. INTRODUCTION

Sign language is communication language used by the deaf peoples using face, hands or eyes while using vocal tract. Sign language recognizer tool is used for recognizing sign language of deaf and dumb people. Gesture recognition is an important topic due to the fact that segmenting a foreground object from a cluttered background is a challenging problem.

There is a difference when human looks at an image and a computer looking at an image. For Humans it is easier to find out what is in an image but not for a computer. It is because of this, computer vision problems remain a challenge.

It is said that Sign language is the mother language of deaf people. This includes the combination of hand movements, arms or body and facial expressions. There are 135 types of sign languages all over the world. Some of them are American Sign Language (ASL), Indian Sign Language (ISL), British Sign Language (BSL), Australian Sign Language (Auslan) and many more. We are using American Sign Language data in this project. This system allows the deaf community to enjoy all sort of things that normal people do from daily interaction to accessing the information. This application takes speech as input, converts it into text and then displays the Sign Language images.

Our project aims to address the communication challenges faced by deaf individuals by developing a comprehensive communication system. This system will have the capability to convert spoken audio messages into text and display corresponding Indian Sign Language images or GIFs. It will provide a bridge between the deaf community and the general public, facilitating easier communication in various contexts such as conversations, gaming, seminars, and video conferences.

2. LITERATURE SURVEY

Title: Implementation of Indian Sign Language in educational settings.

Author: Zeshan, U. Vasishta, M. N. Sethna.

Year: 2005

Description: This article reports on several sub-projects of research and development related to the use of Indian Sign Language in educational settings. In many countries around the world, sign languages are now recognised as the legitimate, full-fledged languages of the deaf communities that use them. In India, the development of sign language resources and their application in educational contexts, is still in its initial stages. The work reported on here, is the first principled and comprehensive effort of establishing educational programmes in Indian Sign Language at a national level. Programmes are of several types: a) Indian Sign Language instruction for hearing people; b) sign language teacher training programmes for deaf people; and c) educational materials for use in schools for the Deaf. The conceptual approach used in the programmes for deaf students is known as bilingual education, which emphasises the acquisition of a first language, Indian Sign Language, alongside the acquisition of spoken languages, primarily in their written form.

Title: An Adaptable Speech to Sign Language Translation System.

Author: Shekainah Paulson and Mrs. B. Thilagavathi.

Year: 2014

Description: This project is a new version of a speech to sign language translation system with new tools and characteristics for increasing adaptability to a new task or a new semantic domain. It consists of a speech recognizer that converts spoken sentences into utterances and silences, and recognizes it as text- a sequence of words, and a video displaying the sign language interpretation of the spoken sentence. It is an adaptable system capable of reducing significantly the effort and the parallel corpus needed for adapting a speech to sign language translation system to a new domain.

Title: Mobile Translation System from Speech-Language to Hand Motion Language.

Author: Mrs.K.Rekha , Dr.B.Latha.

Year: 2014

Description: To access information anywhere, advanced technology is needed. Nowadays mobile devices and applications are developing rapidly. Deaf and hard of hearing people find difficult to use mobile devices since they are not able to access information anywhere due to lack of services. They have difficulties with reading and writing, to read and understand all information to use mobile devices due to their hearing impairment which is an invisible disability. The gain of experience collected by deaf children in four years is equivalent to the gain of one year for hearing children. All visual textual information is not accessible for this category of people with disabilities. To provide benefits for deaf people to improve their social integration and communication. In this paper, we propose a solution to this problem by a mobile based sign language translation device for automatic translation of Indian English speech language to sign language. This helps the hearing impaired people to communicate easily with the hearing people. We focus on the relationship between deaf people and the technological progress.

Title: Automatic Sign Language Translation to Improve Communication.

Author: Tiago Oliveira , Paula Escudeiro, Nuno Escueiro, Emanuel Rocha and Fernando Maciel Barbosa.

Year: 2019

Description: Over the last years, there has been an increase in hearing-impaired students who use sign language as their main form of communication attending higher education institutions around the world. The knowledge that their comprehension of texts is reduced due to sentence structure differences causes a need for more solutions to improve communication and support students in environments where they are unable to be accompanied by sign interpreters. This article details the improvements and current structure of the Virtual Sign platform, a bidirectional sign language to text translation tool that has been in development since 2015. The platform is divided into two main parts, sign to text and text to sign, and both components are described and explained. The solution has received positive feedback on several tests and a pilot experiment, and is being developed with partnerships with sign interpreters from six different European countries. Some planned improvements and future functionalities for the tool are also mentioned and detailed.

Title: Sign Language Recognition Using Template Matching Technique.

Author: Soma Shrenika and Myneni Madhu Bala

Year: 2020

Description: There is an absence of communication with deaf people in our society. To overcome this barrier the introduction of Sign Language (SL) took place. To convey meaning to normal people, sign language makes use of patterns that are visually transmitted sign patterns. Sign language is also useful for people suffering with Autism Spectrum Disorder (ASD). Normal people cannot understand the signs used by deaf, as they do not know the meaning of a particular sign. The system proposed here aims at solving this problem. This system uses a camera, which captures various gestures of the hand. Then, processing of the image takes place by using various algorithms. First, pre-processing of the image takes place. Then, determination of edges occurs by using an edge detection algorithm. Finally, a template-matching algorithm identifies the sign and display the text. As the output is text, one can easily interpret the meaning of a particular sign. This also curtails the difficulty to communicate with the deaf. The implementation of the system is by using OpenCV-Python. The system uses various libraries.

3. ANALYSIS AND FINDINGS

1. Speech Recognition Accuracy

A key component in audio to sign language translation is the initial transcription of spoken language into text using Automatic Speech Recognition (ASR). ASR accuracy depends heavily on factors like audio clarity, speaker accent, background noise, and domain-specific vocabulary. Studies show that state-of-the-art ASR systems achieve high

accuracy in controlled environments, but their performance can degrade significantly in noisy or diverse conditions. Ensuring high ASR accuracy is essential, as errors in transcription can lead to mistranslation or loss of meaning in the sign language output.

2. Sign Language Synthesis with 3D Avatars

Translating text into sign language gestures is typically achieved using 3D avatars that animate the hand signs, facial expressions, and body language. Studies reveal that high-quality, human-like avatars can significantly enhance user experience and comprehension. However, achieving realistic movements requires precise modeling of the anatomy and expressions unique to each sign language. Many systems still struggle with accurately replicating facial expressions, speed, and fluidity of hand gestures, which are essential for natural and understandable sign language. Breakthroughs in motion capture technology and machine learning have led to improvements, but fine-tuning to avoid unnatural or robotic gestures remains an active research area.

3. Real-Time Performance and Latency

Real-time processing is crucial for audio to sign language systems, especially in live environments such as classrooms or conferences. Achieving low latency is challenging due to the computational demand of ASR, NLP processing, and avatar rendering. While advancements in processing power and optimized algorithms have reduced latency, achieving seamless, real-time translation remains a technical hurdle, especially in resource-limited settings. Reducing latency without sacrificing accuracy or detail is a key focus area for ongoing research.

4. Cultural and Linguistic Adaptability

Sign languages are not universal; each has unique linguistic and cultural components. This system needs to adapt to the specific grammar, expressions, and idioms of each sign language. For instance, ASL has distinct vocabulary and structure compared to British Sign Language (BSL). Customization for various sign languages is crucial but also complex, as it requires extensive datasets and training for each variant. Efforts to build multi-language support have made progress but require further development to accommodate global sign language diversity.

5. User Feedback and System Usability

User testing and feedback are essential for refining AI-driven audio to sign language systems. Studies with Deaf and hard-of-hearing participants indicate that while the technology is promising, it still needs improvement in naturalness and accuracy to be fully effective. Feedback suggests that users appreciate the accessibility provided by the system but are sensitive to translation inaccuracies and unnatural movements that disrupt comprehension. Continuous iterations based on user feedback are essential to make the system truly user-friendly and reliable.

Findings and Future Directions

Integration of Multimodal Data: Incorporating visual and contextual cues, such as lip reading and body language detection, can improve translation accuracy by providing additional context to ASR and NLP processes.

Enhanced 3D Animation and Avatar Design: Continued improvements in avatar technology, such as integrating motion capture and deep learning-based gesture synthesis, can enhance the fluidity and expressiveness of sign language gestures.

Domain-Specific Adaptation: Training AI models for specific domains (e.g., medical, educational) can improve contextual accuracy, allowing for more accurate translations in professional settings.

User-Centric Development: Ongoing collaboration with Deaf communities to gather feedback, identify pain points, and improve usability will be critical in developing an effective and inclusive translation system.

Overall, AI-driven audio to sign language translation is a promising technology with significant potential to break down communication barriers. However, further advancements in ASR, NLP, and avatar realism, coupled with continuous user feedback, are essential to refine these systems for widespread, real-world adoption.

4. CONCLUSION

The integration of a sign language translator holds widespread applicability, offering a transformative solution for communication barriers faced by individuals with hearing impairments. Its utility spans across various sectors, including education, healthcare, travel, and legal proceedings. In educational settings, this technology fosters inclusive learning environments by facilitating seamless communication between educators and students with hearing challenges. Similarly, in healthcare, medical professionals can convey information accurately and empathetically to patients, ensuring their understanding of crucial details about their health.

Expanding this initiative to news channels, beyond the current manual interpretation on DD News, presents a groundbreaking opportunity. Proposing an automated system that incorporates facial expressions not only enhances the authenticity of sign language interpretation but also broadens the scope of accessibility within the media

landscape. Such innovation aligns with the principles of diversity and inclusion, allowing news broadcasts to cater to a wider audience, including the hard of hearing community. Collaborating with news channels to refine and implement this technology is not only a step towards a more inclusive society but also a model for other industries seeking to leverage technology for enhanced communication and accessibility. As this project progresses, its impact is poised to extend beyond news channels, becoming a benchmark for inclusive communication practices across various sectors.

5. REFERENCES

- [1] A Dictionary on Indian Sign Language (ISL) signs by FDMSE, Coimbatore. <https://indiansignlanguage.org/>.
- [2] A Dictionary of Indian Sign Language Research and Training Center (ISLRTC), Government of India. <http://www.islrte.nic.in/>
- [3] Documentation from MDN Web Docs - a Web API named as Web Speech API. <https://wicg.github.io/speech-api/>
- [4] Tim Ashwell and J.R Elam. How accurately can the Google Web Speech API recognize and transcribe Japanese L2 English learners' oral production?
- [5] Steven Bird, Ewan Klein, and Edward Loper. Natural Language Processing with Python. O'Reilly Media, 2009
- [6] Charles Sutton, Andrew McCallum. An Introduction to Conditional Random Fields. Volume 4, 2012
- [7] John Lafferty, Andrew McCallum, and Fernando Pereira. 2001. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In Proceedings of the Eighteenth International Conference on Machine Learning, pages 282–289.
- [8] Gouri Sankar Mishra, Parma Nand, Pooja. English text to Indian Sign Language Machine Translation: A Rule Based Method. International Journal of Innovative Technology and Exploring Engineering (IJITEE). Volume-8, Issue-10S, August 2019- Page 460.