

Automatic Translate Real-Time Voice to Sign Language Conversion for Deaf and Dumb People

Prof. Abhishek Mehta^{1*}, Dr. Kamini Solanki², Prof. Trupti Rathod³

¹ Research Solar at Department of Computer and Informative Science, Sabarmati University, Ahmadabad, Gujarat, India. And Assistant Professor, Faculty of IT and Computer Science, Parul University, Vadodara, Gujarat, India

² Accosiate Professor, Parul Institute of Computer Application, Parul University.

³ Assistant Professor, Vidyabharti Trust College of MCA, Gujarat Technological University, Bardoli, Gujarat, India.

Abstract—Sign Language Recognition is one of the most growing fields of research area. Many new techniques have been developed recently in this area. The Sign Language is mainly used for communication of deaf-dumb people. In this paper, we propose design and initial implementation of a robust system which can automatically translates voice into text and text to sign language animations. Sign Language Translation Systems could significantly improve deaf lives especially in communications, exchange of information and employment of machine for translation conversations from one language to another has. Therefore, considering these points, it seems necessary to study the speech recognition. Usually, the voice recognition algorithms address three major challenges. The first is extracting feature form speech and the second is when limited sound gallery are available for recognition, and the final challenge is to improve speaker dependent to speaker independent voice recognition. Extracting feature form speech is an important stage in our method. Different procedures are available for extracting feature form speech. One of the commonest of which used in speech recognition systems is Mel-Frequency Cepstral Coefficients (MFCCs). The algorithm starts with preprocessing and signal conditioning. Next extracting feature form speech using Cepstral coefficients will be done. Then the result of this process sends to segmentation part. Finally recognition part recognizes the words and then converting word recognized to facial animation. The project is still in progress and some new interesting methods are described in the current report. The system will perform the recognition process through matching the parameter set of the input speech with the stored templates to finally display the sign language in caption of video on the screen of computer/mobile etc. So, Deaf and Dumb people or students easily learn the subject through the online YouTube video.

Index Terms—*image processing, sign language, speech recognition, spectral parameter, Deaf Human, Sign Language Translation Systems, Humatronics, Automatic Speech Recognition*

INTRODUCTION

India, a nation with a populace of 1.3 billion individuals, almost a fifth of the total populace [7], is evaluated to have individuals with hearing loss of the request for 5 million [8]. As per the Government of India Disabled Persons Statistics Survey 2016 [35], 32.5% of this number is comprised of kids. In the review, for the age-bunch 5–9 years of age, 209 of an example set of 100,000 kids and for the agegroup of 10–14 years of age, 212 of a lot of 100,000 youngsters have been discovered hearing impeded. A critical segment of this

populace, to be specific 32% of these youngsters have a significant hearing misfortune, and 39% are determined to have serious hearing misfortune [35]. A kid with hearing misfortune faces colossal hindrances in the improvement of discourse and language capacities. Hearing misfortune constrains the youngster's tutoring, advanced education and effects future expert chances.

Besides, various techniques for instructing utilized for the meeting hindered in India add to an absence of structure and approach concerning defeating this obstacle. Inside India, there exist three prime strategies for instructing the consultation hindered, to be specific Indian Sign Language, Oralism and Total Communication. Oralism [14] is the training of understudies with hearing misfortune through oral language by utilization of lip perusing, discourse and mirroring the mouth shapes and breathing examples of discourse. Complete Communication (TC) [38] is a way to deal with the training of individuals with hearing misfortune that plans to utilize numerous methods of correspondence, for example, marked, oral, sound-related, composed and visual guides, contingent upon the specific needs and capacities of the kid. Gesture based communication, albeit a favored method of correspondence around the world, is ascribed to minimal measure of utilization of the three strategies in India. Absolute Communication remains the most generally received philosophy. The way of thinking behind this method is that it gives the youngster different occurrences of modes to depend on. One of the downsides of TC is that it denies the offspring of complex language learning (English or ISL) and consolidates both while educating, which may credit to disarray [38]. A dominant part of schools additionally follow the Oralism technique that may not effectively help training in instances of significant and serious hearing misfortune.

The underlying exploration included setting up a comprehension of the communication via gestures training situation in India on the field. We connected with partners in the biological system—understudies, instructors, organization and officials—and dug into their torment focuses with the assistance of semi-organized meetings. After a complete conversation utilizing participatory plan, we chose to seek after a device which would go about as an instructing help that gives, printed and signage prompts on recordings and guides the comprehension of understudies for the subjects experienced. We at that point endeavored to approve the requirement for such a framework through A/B testing and doing an observational investigation [17] with the

understudies where they had the option to get a handle on better and hold data learned through the inscribing offered by the framework. The study hall condition was our objective, so we directed the field research at a school. Investigating different fields where innovation can help the consultation impeded may prompt more arrangements in such spaces, which would in the end make ready toward an allinclusive arrangement.

We propose an innovation upheld answer for connect the asset hole among individuals with hearing misfortune (in ISL—their favored method of correspondence). The proposed framework use previously existing instructive recordings on the web and gives sign subtitling accessible during the run of the video. The reason for our foundation is to make an interface that serves content in the essential language of the hearingimpaired network that makes it simple to relate mappings and all in all structure an effective framework for learning and assessment of youthful understudies during their language building years.

The framework is worked around a database of 3D created signs that go about as the sign inscriptions for the video. Captions or discourse handling is utilized to induce the sound substance of the video, and it is then sent to the Natural Language Processing module which has a Subject–Object–Verb rule-based syntax, and sentences are changed over to this organization; in the long run, the video is overlaid with the 3D sign inscriptions.

PREVIOUS WORK

Commercial products for people with hearing loss:

Efforts to aid the hearing impaired have long since been focused on communication through translation of signage. A few leading commercially available translation tools from around the globe are discussed here. The 'HandTalk Translator' Application [34] converts Brazilian Portuguese audio to Brazilian Sign Language. This product is market ready and available on the Google Play Store. It uses an interactive avatar with facial expressions and fluidity. However, it aids only one part of the communication from the non-signer to the signer. MotionSavvy's 'Uni' [24] leverages leap motion technology to convert audio and vice versa to American Sign Language. Currently in R&D phase, its interactions are limited only to hand movement and do not take into account facial expressions as a part of signage. Platforms and resources to ease communication with the hearing impaired have also been developed and maintained by various groups and organizations who promote Sign Language Learning. Within India, 'Talking Hands' [33].

A web-based platform provides an extensive dictionary of Indian Sign Language. It is used for educational purposes and interactional videos to develop skill sets. However, this platform lacks responsiveness and contains a limited subset of commonly used signs for language development. 'Ramakrishna Mission' [25] provided the first and most widely adopted online resource for ISL signs but presented a limited user interaction and a lacking vocabulary set of signs. 'SignTalk' [32] acted as a relay service between the signer and the interpreter and was by far the most evolved system to serve people with hearing loss in India. The platform, however, required paid interpreters and necessitated the need

for an interpreter to be present to relay the request. This service is no longer operational. Coming to the most straightforward mode of communication, to manually hire a sign language interpreter, due to the need and lack of ISL teachers still remains a constant struggle and unaffordable circumstance for most coming from impoverished neighborhoods.

Experimental/research work in sign language recognition

Endeavors to help the conference debilitated have since a long time ago been centered around correspondence through interpretation of signage. A couple of driving financially accessible interpretation devices from around the world are examined here. The 'HandTalk Translator' Application [34] changes over Brazilian Portuguese sound to Brazilian Sign Language. This item is showcase prepared and accessible on the Google Play Store. It utilizes an intelligent symbol with outward appearances and smoothness. Be that as it may, it helps just a single piece of the correspondence from the non-underwriter to the endorser. MotionSavvy's 'Uni' [24] use jump movement innovation to change over sound and the other way around to American Sign Language. Right now in R&D stage, its cooperations are restricted distinctly to hand development and don't consider outward appearances as a piece of signage. Stages and assets to ease correspondence with the consultation debilitated have additionally been created and kept up by different gatherings and associations who advance Sign Language Learning. Inside India, 'Talking Hands' [33].

An online stage gives a broad word reference of Indian Sign Language. It is utilized for instructive purposes and interactional recordings to create ranges of abilities. In any case, this stage needs responsiveness and contains a constrained subset of regularly utilized finishes paperwork for language advancement. 'Ramakrishna Mission' [25] gave the first and most broadly embraced online asset for ISL signs yet introduced a restricted client connection and a lacking jargon set of signs. 'SignTalk' [32] went about as a hand-off assistance between the endorser and the translator and was by a wide margin the most advanced framework to serve individuals with hearing misfortune in India. The stage, in any case, required paid mediators and required the requirement for a translator to be available to hand-off the solicitation. This administration is not, at this point operational. Going to the most direct method of correspondence, to physically employ a communication via gestures mediator, because of the need and absence of ISL instructors despite everything stays a consistent battle and unreasonably expensive situation for generally originating from ruined neighborhoods.

methodology used for implementation

The framework proposed in this paper impersonates the digestion procedure of the individuals with hearing misfortune in devouring predominant press outlets (TV, occasions, addresses, conversation, and so on.) as sketched out in Fig.1. Data trade happens through a sign translator going about as a medium between the two gatherings (two-way correspondence) or concurrent sign interpretation for single direction correspondence. In our paper, we center

around the second situation where each new wellspring of media should be deciphered for osmosis. It renders a plenty of substance difficult to reach by sheer excellence of absence of interpretation medium. The need of human mediation in this procedure frames a boundary to learning and correspondence; mechanizing this has seen various endeavors, however they are in no way, shape or form exhaustive. Such endeavors need to fuse multiple modalities to imitate a human translator.

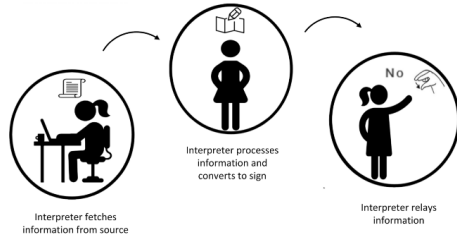


Fig -1 Operational Flow of System

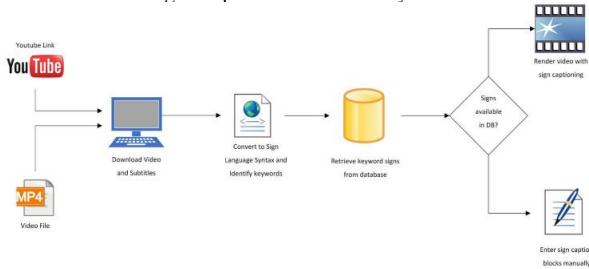


Fig -2 Logical Flow of System

As appeared in Fig.2, the contribution to the pipeline is as YouTube recordings or sound sources. The pipeline renders a smooth enlivened video at the yield end. The pipeline can be stalled extensively into these sub-modules depicted underneath.

1. Information Entry Information section alludes to the passage sub-module where the info source is surrendered to the pipeline; this information can be discourse and captions (from YouTube recordings/sound).

2. Information Processing: Information from the section submodule is prepared to increase literary data input is changed over to message and converged with caption information to acquire the best match portrayal of the information.

3. Information Understanding: The pipeline currently has text portrayal of the information. Now, we attempt to catch the specific situation and message by means of Natural Language Processing. This comprehension depends on the structure of the expected yield media. For this situation, we take a gander at the Indian Sign Language structure and concentrate components as needs be.

4. Mapping: Our data is resounding with sign lan-guage structure and would now be able to be utilized to produce relevant yield structure. The planning takes places between distinguished words and setting to 3D symbol signals put away in our database. This symbol signals database is key to our application as it goes about as an interface to trade data.

5. Interaction: Interaction is the interface between the information section and the utilization of yield. An overlay container on recordings gives a shut subtitling equalised to the end client. This connection holder can be re-sized, delayed or sped to give the greatest control of the technique

for comprehension. This element depends on an examination we led on students with hearing misfortune speaking to various methods of getting a handle on data.

The pipeline gives a total start to finish the answer for process video or sound information and is fundamental to the symbol database word reference. A full and populated database could give an enormous yield to reinforce understanding among the end clients. The association first beginnings with the client signing onto the gateway and gluing the connection of the ideal video that requires interpretation; the framework at that point merges a document of pre-recorded symbol-based signs to shape a solitary record stacked from the back end. The web-based interface likewise permits instructors to record indications of words unrecorded from past recordings, which are interpreted later by the back-end group upon approval. An intuitive application renders finishes paperwork for the general instructive purposes from the now-populated database of sign mappings.

DATA COLLECTION

A complete audio or video to sign caption pipeline allows the hearing-impaired user to understand and infer content from the input source. In the event of an important video where the user may not be familiar with the signs, content from the video allows the mapping of signs to contextual visual information from the video, helping with sign education. For a video that may not possess a plethora of visual or sign inferable cues, the sign captioning helps with understanding and learning video content. The collective performance of the system can be justified as learning and understanding.



Fig. 3 Translation of 'What a wonderful world by David Attenborough' (Rough Prototype)

CONCLUSION

Our examination advances the requirement for sign help to help with comprehension and learning for understudies with a meeting handicap since the beginning forward. From our different preliminaries, we infer that sign help enables the kids to learn, recall and comprehend the substance better. It is likewise broadly bolstered by the past exploration which diagrams the advantage of gesture based communication for comprehension and learning language and syntax. In light of these perceptions, we propose a framework that can guarantee long lasting learning by giving sign help to the consultation hindered understudies from expending broad communications, for example, any semblance of YouTube. The framework is versatile and simple to use in a study hall setting which can make a useful expansion to help the information base of understudies who in any case think that

its difficult to comprehend and learn content outside their study halls. Our examination likewise features various modules that include the framework and how every module was created dependent on understudy instructor connections and perceptions to guarantee most extreme commitment from the understudies. Our future exploration work will be led on what scale these frameworks can be at present executed and embraced as showing helps with a restricted gesture based communication database available to us.

REFERENCES

- [1] Agarwal, A., Thakur, M.K.: Sign language recognition using microsoft Kinect. In: 2013 Sixth International Conference on Contemporary Computing (IC3), pp. 181–185 (2013). <https://doi.org/10.1109/IC3.2013.6612186>
- [2] Ahmed, M.A., Zaidan, B.B., Zaidan, A.A., Salih, M.M., Lakulu, M.M.B.: A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017. *Sensors* 18(7), 2208 (2018). <https://doi.org/10.3390/s18072208>
- [3] Ahmed, S.: Real time American sign language video captioning using deep neural networks [PowerPoint presentation]. <http://on-demand.gputechconf.com/gtc/2017/presentation/s7346-syed-ahmed-real-time-american-sign-language-video-captioning.pdf> (2018).
- [4] Cannon, J.E., Fredrick, L.D., Easterbrooks, S.R.: Vocabulary instruction through books read in American sign language for English-language learners with hearing loss. *Commun. Disord. Q.* 31(2), 98–112 (2010). <https://doi.org/10.1177/1525740109332832>
- [5] Center, L.C.N.D.E.: Learning american sign language: Books, media, and classes. http://www3.gallaudet.edu/clerc-center/info-to-go/asl/learning-asl-books_media_classes.html. [Online]
- [6] Chambers, B., Abrami, P.C., McWhaw, K., Therrien, M.C.: Developing a computer-assisted tutoring program to help children at risk learn to read. *Educ. Res. Eval.* 7(2–3), 223–239 (2001). <https://doi.org/10.1076/edre.7.2.223.3863>
- [7] Chandramouli, C.: Census of India 2011, provisional population totals, government of India. http://censusindia.gov.in/2011-provisional-results/paper2/data_files/india/paper2_1.pdf (2011). [Online]
- [8] Division, S.S.: Disabled persons in India, a statistical profile 2016. http://mospi.nic.in/sites/default/files/publication_reports/Disabled_persons_in_India_2016.pdf (2016). [Online]
- [9] Easterbrooks, S.R., Huston, S.G.: The signed reading fluency of students who are deaf/hard of hearing. *J. Deaf Stud. Deaf Educ.* 13(1), 37–54 (2007). <https://doi.org/10.1093/deafed/enm030>
- [10] Fung, P.C., Chow, B.W.Y., McBride-Chang, C.: The impact of a dialogic reading program on deaf and hard-of-hearing kindergarten and early primary school-aged students in Hong Kong. *J. Deaf Educ. Deaf Stud.* 10(1), 82–95 (2005)