Sign Language Translation Systems for Hearing/Speech Impaired People: A Review

Yuvraj Grover
Amity School of Engineering and Technology
Amity University, Noida
Uttar Pradesh, India
yuvi.grover1999@gmail.com

Riya Aggarwal
Amity School of Engineering and Technology
Amity University, Noida
Uttar Pradesh, India
riya.agg72@gmail.com

Deepak Sharma
Amity School of Engineering and Technology
Amity University, Noida
Uttar Pradesh, India
dsharma15@amity.edu

Prashant K. Gupta
Amity School of Engineering and Technology
Amity University, Noida
Uttar Pradesh, India
pkgupta@amity.edu

2021 International Conference on Innovative Practices in Technology and Management (ICIPTM) | 978-1-6654-2530-8/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICIPTM52218.2021.9388330

***Abstract* – Sign language is a basic tool for those who haven't received the gift of speaking and listening. Normal people generally do not learn sign language to communicate with hearing/speech impaired people. Also, every time these differently abled people cannot keep a human translator with them. This makes a communication gap between normal and hearing/speech impaired people. So to bridge this gap, various sign language translators and recognition techniques have been developed which are discussed in this paper. Also, through various researches done in this field, a general classification of these sign language translators and recognition techniques have been made, as per the taxonomy discussed in the paper, in a tabular form along with the literature review.**

Keywords – speech processing, image processing, sign language, ISL, Indian Sign Language, vision based, sensor based

I. INTRODUCTION

Studies revealed that one out of every five deaf people in the world is from India.According to the 2011 census, approximately 7 million people in India are deaf and mute and 1.9 million have speech disability. Even then, most of the time hearing/speech impaired people find it very difficult to lead a normal life as any other person. They face a whole lot of challenges from the moment they are born. The major challenge is - Communication. They cannot communicate with people who can speak/hear because nobody learns their language i.e. Indian Sign Language (ISL). Only hearing/speech impaired's closest family members learn ISL and that too in a superficial manner. This barrier often impacts such people, on so many levels as education, and doing and understanding daily chores. They are often not well versed in reading and writing due to lack of learning. 2011 census data reveals that among deaf and mute people, 63% are not employed and 30% have never attended school. Now, consider the case when these people want to travel solo? Want/have to live all by themselves? Or simply want to go and buy some medicines/ want to do shopping/ have a haircut/ take a rickshaw/ want to eat at a restaurant? They just cannot do all of this without being accompanied by a hearing and speaking person. This makes them feel dependent on their parents, siblings or friends all of the time. This is very stressful for them. So, to help such people, a lot of work has been done in the field of sign language detection and translation with the intention to remove this barrier of communication.

Sign Language is a language that uses visual-manual modality to convey meaning. It comprises of gestures ranging from facial expressions to using body parts to communicate. It can be classified as Isolated Sign Language and Continuous Sign Language. If a single gesture corresponds to a single word, it is termed Isolated Sign Language. Whereas, if a series of gestures are used to produce a meaningful sentence, it is termed as Continuous Sign language. A complete sign language uses 3 main parts:

1. Fingerspelling (letter by letter)
2. Word level sign vocabulary (majorly used)
3. Non-Manual features (Facial Expressions, Body positions, etc.)

Sign Languages tend to be inspired from local culture and hence there is no universally accepted form of it. Each country has developed their own. For example, Indian Sign Language (ISL), American Sign Language (ASL), Chinese Sign Language (CSL), Japanese Sign Language (JSL), and many more. This contributes to difficulties in development of a generic sign language translation system. Even the Indian Sign Language is decentralized i.e., it varies geographically throughout India. The gesture for the words 'Long' and 'Flag' in ISL has been shown in Fig. 1. A description of English alphabets in ISL are depicted in the Fig. 2.

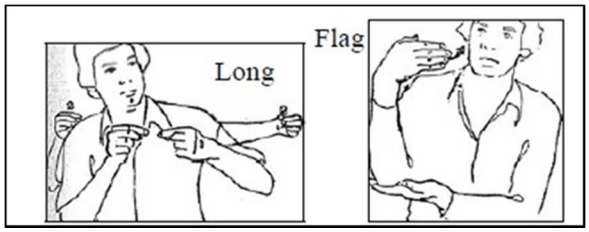

Fig. 1. "Long" and "Flag" in ISL

978-0-7381-3289-1/21/$31.00 ©2021 IEEE

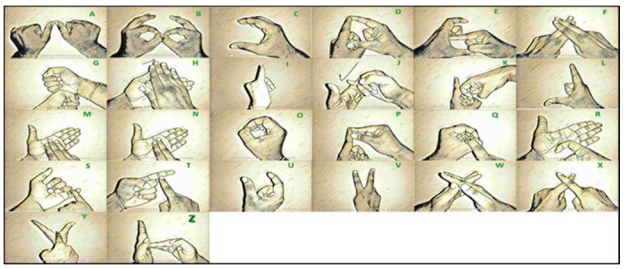

Fig. 2. Double handed ISL Alphabet

This paper provides insights captured from various publications and classifies these publications into a hierarchy of methods and types of work done in the field of sign language detection and translation. These insights are explained in chapter II, Classification of Sign Language Translation. Further, Chapter III provides literature review for these publications along with a tabular representation for the same. Chapter IV includes conclusion followed by the references.

II. CLASSIFICATION OF SIGN LANGUAGE TRANSLATION

Over the past decade, a lot of work has been done in the field of sign language recognition and translation. The work can be classified into sub categories for a better understanding as shown in Fig. 3.

A broader classification is done for translation of sign language to spoken language and spoken language to sign language. These are further classified based on their methodologies.

Work in the field of translation from normal language to Sign Language, can be achieved by either converting text form to Sign Language or Speech to Sign Language. Similarly, work in the field of sign language recognition also splits into two categories, namely, Vision Based Recognition and Sensor Based Recognition.

A. Translation to Sign Language

In this category, input is in normal spoken language and output is the preferred sign language. Here sign language animations, as shown in Fig. 4, are generated using appropriate notation systems or technologies. It may further be classified into two sub types –

a. *Text to Sign Language*

As the name suggests, here input is given in the form of text and then it is converted into sign language for hearing/speech impaired people. Text input may be given by those who cannot speak.

b. *Speech to Sign Language*

Here input is given in the form of voice using a microphone or some similar input device. The input speech is then processed using various NLP techniques such as tokenization and information extraction so as to convert it into an efficient sign language animation.

B. Translation from Sign Language

In this category, a sign language input is given by those who cannot speak. We can also say that this category is for sign language recognition. Sign language recognition can also be further classified on the basis of how the sign language input is given.

a. Sensor Based Recognition

In this approach, the input is given using hardware sensors in the form of gloves (Fig. 5) or smartphone internal sensors by tracking the movement of the hands and body. The sign language detection process gets simplified and faster when using sensor based input. In gloves, multiple sensors are embedded so as to track each and every movement of the fingers, palm and their location. Sensor based gloves can be connected using bluetooth. Smartphone based sensors have recently came in the queue however they have a lot of limitations such as not able to focus on the double handed signs. They give input in the form of 3D orientation of the smartphone (eventually hand) in which accelerometer continuously reads data from the three axis in space. Sensor based recognition has an advantage over the vision based recognition as no camera has to be kept in front of the person conveying his/her thoughts using sign language [15].

b. Vision Based Recognition

In this approach, the input is given using camera in the form an image or a video (continuous gesture or static gesture). Computer vision techniques such as image processing, object detection, video processing, skin segmentation and image segmentation are used to process this type of input. Then the processed inputs can be fed into neural networks such as CNN (Convolutional Neural Network), LSTM (Long Short-Term Memory Neural Networks), RNN (Recurrent Neural Network),etc. for translation to appropriate text/speech output. This approach has a drawback when compared to sensor based approach due to various challenges faced during video/image processing such as lighting conditions, brightness, background noise, camera angle etc.

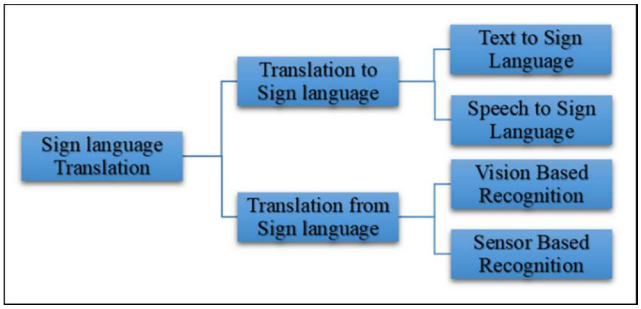

Fig 3: Classification of sign language translation

Fig 4: Sign Language animation for the word 'bad' [8]

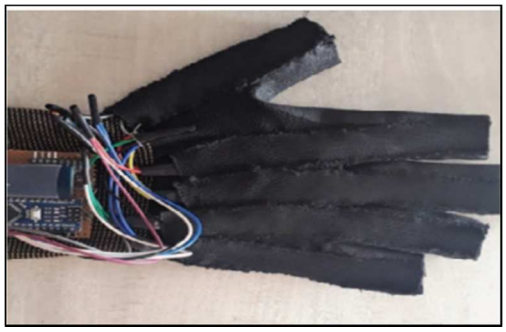

Fig 5: Sensor-Based Glove for Sign Language Recognition [12]

III. LITERATURE REVIEW

In recent times, sign language translation and recognition is a wide area of research. Sign language is the sole medium through which hearing/speech impaired people can communicate their thoughts and feelings to others. [1] has made an integrated two-way ISL translation that uses a camera and a microphone. The image from the camera was binarised and further converted to text and for the other way round, the speech was converted to text and then to the appropriate sign. Video processing based sign language recognition system was developed by [2] using a 3-layer artificial neural network. Continuous input was taken from a selfie-based video which was preprocessed, filtered and then pattern extraction was done. Finally, speech was given as the output. A generic prototype for text-to-sign machine translator was proposed by [3] which took English text as input, performs syntactic analysis on the text, and then generates the matching syntax for Indian Sign Language. This paper also gives an account of the issues that are present in Indian Sign Language because it uses two-hand sign also. Another two way communicator model was developed by [4] that used Google TTS and STT API for speech-text translations. Real time video was converted into relevant speech and real time speech was converted into animated gesture videos. A system that converts speech to Indian sign language was proposed by [5]. The paper used HamNoSys for generating signing gestures from speech. It was developed based on 100 basic words of daily routine. The system covered one-handed and two-handed signs. [6] Compared different approaches for converting speech to sign language. The approaches were mainly artificial neural networks, vision based recognition, automatic speech recognition and hidden markov models. It was found that each approach was equally powerful. A similar comparison was done by [7]. It compared different notation systems namely Stokoe, SignWriting, HamNoSys and Gloss for ISL. Notation systems are used to transcribe a spoken/written language to sign language script. It was found that HamNoSys was the most efficient notation system. Gurumukhi script was used in [8] to convert Punjabi language to signing animation using HamNoSys notation. It covered all the basic words from daily routine. [10] Developed a web application to convert Hindi text to Indian Sign Language. It used a Hindi dependency parser to perform syntax analysis and POS (parts of speech) tagging on the input Hindi text. Also, Hindi Word net API was used so as to increase the efficiency of the web application because Word net allowed replacing words whose gestures were not present in the database with their respective synonyms. The system was implemented using JSP and HTML. A portable sensor based glove was designed by [11] which could translate real-time signs into speech using LSTM (Long-Short Term Memory) networks. It had a testing accuracy of 98%. The designed glove having wireless capabilities (bluetooth) used a low-cost hardware making it affordable for many of its users. It could interpret 26 static and dynamic gestures. The 3D sensor data from glove was converted into 2D data in the form of CSV files which was then fed into the LSTM network. On a similar note, [12] designed a multi-lingual glove for conversion of ASL (American Sign Language) and ISL (Indian Sign Language) to speech. It had accuracy of 100% with ISL database and 98.91% with ASL database on SVM classifier.
A tabular representation of the some publications along with their classification in the taxonomy described in this paper has been given in table 1.

Table 1: Classification of various sign language translators and recognition techniques

Reference No.	Year of publication	Type of publication	Dataset Used	Algorithm/ Notation Used	Advantages	Limitations	Classification
Dasgupta T. et.al. [3]	2008	Conference	208 sentences collected from 'A' level introductory course in Indian sign language workbook	Text analytics, syntactic parsing, LFG representation, ISL Generation, HamNoSys	Proper natural language processing has been done	One-way translator	Speech-to-sign
Anand M.S. et.al. [1]	2013	Journal	Real-time input taken from camera and microphone	Image classification, speech recognition	Two-way translation without devices like gloves	Only fingerspelling based	Finger Spelling, vision based, speech-to-sign, sign-to-speech
Ahire et.al. [4]	2015	Conference	Real-time input taken	Region Growing, Mapping using correlation	Efficient two-way communication in real time	No such limitation observed	Gesture based, vision based, speech-to-sign, sign-to-speech
Verma A. et.al. [8]	2015	Journal	235 Punjabi words	HamNoSys, SiGML	Good accuracy, signing animations generated	One-way translator	Gesture based, vision based, speech-to-sign
Kaur S. et.al. [5]	2015	Conference	Simple 100 Punjabi words	HamNoSys, SiGML	Generation of signing animations	One-way translator with a very limited dataset	Gesture based, vision based, speech-to-sign
Dutta K.K. et.al. [14]	2015	Conference	125 signing images in database	Minimum EigenValue Algorithm	Double handed signs are used, speech output in various languages	One-way translation	Gesture based, vision based, sign-to-speech & sign-to-text
Vij P. et.al. [10]	2016	Conference	Real-time text input in Hindi language	Dependency parser, POS (Parts of Speech) tagger, WordNet, HamNoSys	Wordnet helps in expanding the signing database	One-way translator with only text as input	Text-to-sign
Rao G.A. et.al. [2]	2017	Journal	18 signs in continuous sign language recorded with 10 different signers	Gaussian Filtering, Euclidean distance, normalized Euclidean, mahalanobis distance	Sign video capture done using selfie camera and stick	One-way translator	Gesture based, vision based, sign-to-speech
Abraham E. et.al. [11]	2019	Conference	40 samples each for 26 ISL Gestures containing readings of accelerometer, gyroscope and flex sensors	LSTM (Long Short-Term Memory)	High accuracy of 98% using LSTM, low cost hardware	One-way, no record of relative-orientation of both the hands	Gesture based, Sensor Based, Sign-to-Speech
Chandra M.M. et.al. [12]	2019	Conference	22 ASL(American) Gestures and 11 ISL(Indian) Gestures	Signal Processing, SVM (support vector machine)	Can recognize both ISL and ASL	One-way translator, only single-handed gestures allowed	Gesture Based, Sensor-based, Sign-to-Speech

Zhou Z. et.al. [16]	2020	Journal	11 ASL Symbols (A, B, C, I, L, Y, 1, 2, 3, 8, I Love You)	PCA (Principal Component Analysis), multi-class SVM algorithm	Effective wearable, machine learning assisted translator	One-way translation	Gesture Based, Sensor Based, Sign-to-Speech

IV. CONCLUSION

Thus, we have surveyed various techniques and technologies employed for sign language translation along with the type of datasets used. For the survey the major focus was on Indian Sign Language but was not only limited to it because mostof the already developed systems are in American or British Sign Language. Translation systems in languages such as Chinese Sign Language, Russian Sign Language were also observed. Mostly all the translation systems were working one-way (either translating from sign language to text/speech or translation speech to sign language generation). However, a complete two-way sign language system in Indian Sign Language was not observed. Already existing systems are either domain specific or are one-way translators. They aren't of much help to all the hearing/speech impaired people in general. They expect a system which could give them all kinds of features i.e. to be able to translate to and from sign languages that too in real time. Also, it should be a generic system and not domain specific. This would help them to convey their thoughts and feelings in a better and deeper manner. Hence, a two way Indian Sign Language translation system is needed to fully bridge the communication gap between hearing/speech impaired and other people which could follow the maximum portion of the classification proposed.

REFERENCES

[1] Anand, M.S., Kumaresan, A. and Kumar, N.M., 2013. An Integrated Two Way ISL (Indian Sign Language) Translation System--A New Approach. *International Journal of Advanced Research in Computer Science*, *4*(1).

[2] Rao, G.A. and Kishore, P.V.V., 2018. Selfie video based continuous Indian sign language recognition system. *Ain Shams Engineering Journal*, *9*(4), pp.1929-1939.

[3] Dasgupta, T. and Basu, A., 2008, January. Prototype machine translation system from text-to-Indian sign language. In *Proceedings of the 13th international conference on Intelligent user interfaces* (pp. 313-316).

[4] Ahire, P.G., Tilekar, K.B., Jawake, T.A. and Warale, P.B., 2015, February. Two way communicator between deaf and dumb people and normal people. In *2015 International Conference on Computing Communication Control and Automation* (pp. 641-644). IEEE

[5] Kaur, S. and Singh, M., 2015, September. Indian Sign Language animation generation system. In *2015 1st International Conference on Next Generation Computing Technologies (NGCT)* (pp. 909-914). IEEE.

[6] Fatima, S., Agarwal, A. and Gupta, P., 2016, March. Different approaches to convert speech into sign language. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)* (pp. 180-183). IEEE.

[7] Dhanjal, A.S. and Singh, W., 2019, February. Comparative Analysis of Sign Language Notation Systems for Indian Sign Language. In *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)* (pp. 1-6). IEEE.

[8] Verma, A. and Kaur, S., 2015. Indian sign language animation generation system for Gurumukhi script. *International Journal of Computer Science and Technology*, *6*(3), pp.117-121.

[9] Ghanem, S., Conly, C. and Athitsos, V., 2017, June. A survey on sign language recognition using smartphones. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments* (pp. 171-176).

[10] Vij, P. and Kumar, P., 2016, August. Mapping Hindi Text To Indian sign language with Extension Using Wordnet. In *Proceedings of the International Conference on Advances in Information Communication Technology & Computing* (pp. 1-5).

[11] Abraham, E., Nayak, A. and Iqbal, A., 2019, October. Real-Time Translation of Indian Sign Language using LSTM. In *2019 Global Conference for Advancement in Technology (GCAT)* (pp. 1-5). IEEE.

[12] Chandra, M.M., Rajkumar, S. and Kumar, L.S., 2019, October. Sign Languages to Speech Conversion Prototype using the SVM Classifier. In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)* (pp. 1803-1807). IEEE.

[13] Grif, M. and Manueva, Y., 2016, June. Semantic analyses of text to translate to Russian sign language. In *2016 11th International Forum on Strategic Technology (IFOST)* (pp. 286-289). IEEE.

[14] Dutta, K.K. and GS, A.K., 2015, December. Double handed Indian Sign Language to speech and text. In *2015 Third International Conference on Image Information Processing (ICIIP)* (pp. 374-377). IEEE.

[15] Bairagi, V.K., 2014. Gloves Based Hand gesture recognition using Indian Sign Language. *International Journal of Latest Trends in Engineering and Technology,* Vol.(8)Issue(4-1)(pp.131-137)

[16] Zhou, Z., Chen, K., Li, X., Zhang, S., Wu, Y., Zhou, Y., Meng, K., Sun, C., He, Q., Fan, W. and Fan, E., 2020. Sign-to-speech translation using machine-learning-assisted stretchable sensor arrays. *Nature Electronics*, *3*(9), pp.571-578.