INDIAN SIGN LANGUAGE RECOGNITION SYSTEM FOR DEAF AND DUMB USING IMAGE PROCESSING AND FINGERSPELLING: A TECHNICAL REVIEW

Mr. Rakesh Savant, Ms. Amrutha Ajay

Abstract – In deaf and dumb communication, they use sign language to communicate with each other. Communication is the way to convey our thoughts, message or information. Hearing and speech disabled people faces many problems when they communicate with normal people. They uses sign language person, which includes hand gesture, facial expressions, and head movement to convey their message. Using the image processing techniques, it is possible that system will be develop to help those disabled people for effective interaction with normal people. This paper compares and discusses various techniques that are used worldwide for region wise sign languages and proposed an idea that may applicable to develop a communication system for Indian Sign Language. An image processing based system can be made that will work on smartphones and will recognize sign perform by deaf-dumb and generate text or audio output for normal person and vice-versa communication.

Keywords -- Sign Language, deaf and dumb people, communication, Image processing, smartphones, hand gestures.

I. INTRODUCTION

The sign language recognition system is an application area of Human-Computer Interaction. Human-Computer Interaction (HCI) researches on the design and the use of computer technology in the communication. HCI includes many other fields like computer science, behavioural science, media studies etc. Its focus is on how hardware and software mechanism can be implemented for supporting Human-Computer Interaction [5].

Deaf and dumb people interact using sign language. There are many sign languages available across the world like American Sign Language, Arabic sign language, Brazilian Sign Language, Austrian Sign Language, Indian Sign language etc. American Sign Language is accepted as the standard sign language by many countries [1]. Still some countries use the sign language according to their culture, like in India Indian Sign Language (ISL) is used in India.

The sign language can be any hand movements or gesture, facial expressions etc. The normal people do not know the sign language and they do not even bother to learn it. Because of this, there is a communication gap between deaf dumb and normal person. To bridge this gap the translator can be used. However, finding a well-experienced translator and keeping them for the lifetime is neither

convenient nor affordable. Hence, we need a system to make communication easier between deaf dumb and normal person. The system takes hand gesture as an input from real-time video and processes them using image-processing techniques to extract the sign. The sign recognized are converted into text or speech. In addition, for the text input, the corresponding sign image is shown to a deaf dumb person. This can also be useful to teach deaf-dumb person, as there are not enough learning centers.

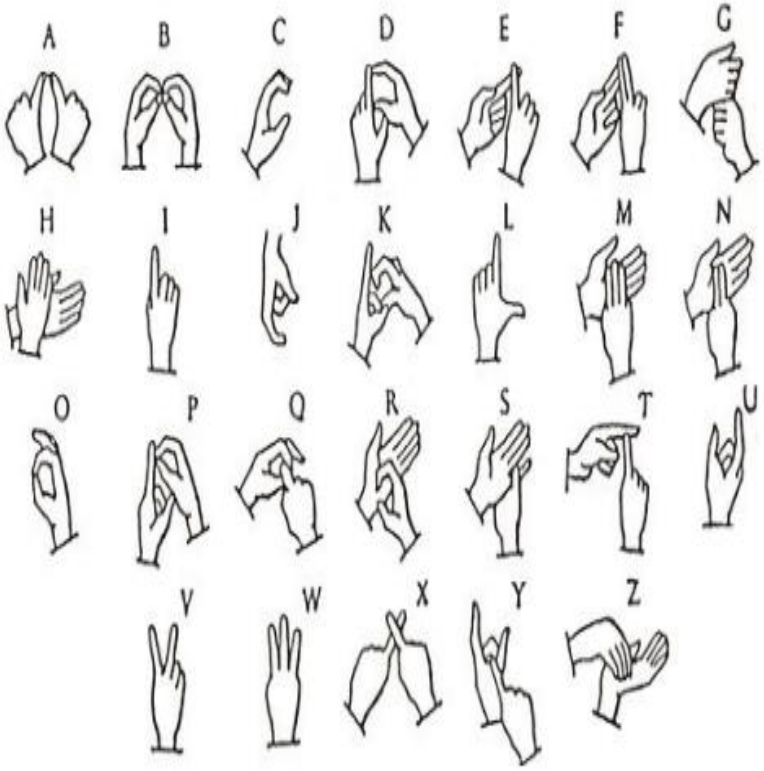

Figure 1: Indian Sign Language Alphabets [2]

In ISL, we can use word level and letter level (fingerspelling) gestures [2]. Using the word level gesture, we can interpret the words, but it is not always handy as not all the words can be included in the dataset. For word, level gestures we need to create a dataset of the single word in different forms as the position and the size of the hand of all people will not be same.

In addition, for the accuracy of the result returned it is very necessary to have a large option to match with. Not only that but for many words, the sign may not even exist or are not known to the individual. Fingerspelling can be useful in such cases. We can use letters to form the words. Hence, the recognition of the letters has significant importance in the sign language recognition. We will have 26 alphabets and 10 numbers (0-9) for the sign recognition. The figure 1 and figure 2 shows the representation of alphabets and numbers in Indian Sign Language [2]. Here we are going to use the fingerspelling.

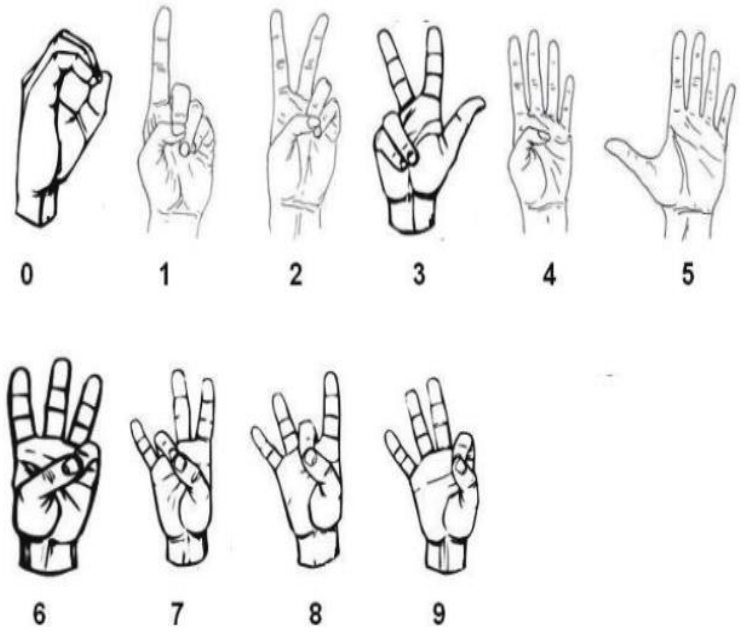

Figure 2: Indian Sign Language Numbers [2]

There are many systems proposed for deaf-dumb communication that can be divided into two parts: - hardware based and vision based system. The hardware system includes the wireless gloves, data gloves [13] that can be used to extract the hand signs. These gloves have sensors, Arduino circuit board, and Accelerometer in them to recognize the gestures [6] [13] [14]. The accelerometer is used to sense the tilt; it can be static as well as dynamic. The sensors that can be used are the flex, tactile etc.

The drawback of the hardware-based system is that we cannot always carry the hardware with us. While in the vision-based system, we do not need any extra

hardware, as smartphones are capable of doing initial processing. It uses the digital image processing techniques and even the machine learning techniques to extract the features for recognizing the sign. It takes the video as input and creates the frames from video. These frames can then be used to extract the sign. Then the output is given in the form text or audio.

II. LITRATURE REVIEW

In [2] Adithya, Vinod, and Usha proposed a method for automatically recognizing the Indian Sign language and the output is given in the textual form. In this method, the digital image processing techniques and the artificial neural network are used for recognizing the sign. For identifying the sign, the features are extracted from the hand image based on shape as a feature. Skin color based segmentation is used to extract the hand region from the image. This method has the accuracy rate of 91.11%.

Anchal Sood and Anju Mishra [1] use webcam for capturing the image. Then image acquisition, segmentation, morphological erosion and feature extraction are used to recognize the gesture. The output is given in both the textual and audio form. However, we cannot use both the hands in this technique.

For gesture recognition, the state of the finger is used [3]. The alphabets are recognized using the features like the angle between fingers, a number of fingers that are fully opened, fully closed or semi-closed and identification of the fingers used in the recognition.

Surbhi and Ujwalla proposed a full duplex way of communication for deaf and dumb i.e. gesture to text and speech and vice versa [4]. Pre-processing, Segmentation, Feature Extraction and Classification are the main steps performed for recognizing the dynamic hand gestures. For segmentation skin, a color filtering technique is used. Feature Extraction is done using Eigen vectors and Eigen values [29] techniques. Eigen value weighted Euclidean Distance [27] based classifier is used for the classification [8]. Both the hand can be used for the gesture.

Malu S Nair, Nimitha A P, and Sumam Mary Idicula presented a machine translating system for translating Malayalam text to Indian Sign Language [7]. HamNoSys (Hamburg Notation System for Sign Languages) structure is used as an intermediate representation for signs [15]. The Malayalam words are given as input to the system and the output is delivered as a 3D character animation. We can also input new words into the system using interactive sign editor that will transcribe sign into HamNoSys structure.

Shubhankar, Karan, and Tanmay use scale-invariant feature transform (SFIT) algorithm for gesture recognition of American Sign Language [8]. The delay time is reduced and the accuracy is increased using the SFIT algorithm.

Emad and Ebaa present an android application for communication with disabled users and vice versa [9]. The key feature of the application is that it employs the Arabic language as the medium of communication for learning the sign language terms. This system was tested on the real deaf, dumb and normal people, and they found this application valuable.

Asha Thalange and Dr. S. K. Dixit proposed two new feature extraction techniques of Combined Orientation Histogram and Statistical (COHST) Features and Wavelet Features to recognize the static sign of numerals in American Sign Language [10].

The performance of the system is measured by extracting the four different features of Orientation Histogram, Statistical Measures, COHST Features and Wavelet Features to train and recognize the American Sign Language numbers individually using a neural network. Observation states that COHST method forms stronger feature than the Orientation Histogram and Statistical Features. For static image based ASL numbers recognition wavelet features based system gives the best performance average recognition rate of 98.17%.

Kumud Tripathi, Neha Baranwal and G. C. Nandi [11] propose recognition of the continuous Indian Sign Language gesture done by both the hands. The problem of recognizing the continuous gesture was solved by using the gradient-based key frame extraction method. Orientation Histogram (OH) is used to extract the features of ore-processed gestures with Principal Component Analysis (PCA) [27][28][29] which is applied to reduce the dimensions of features obtained from OH.

The continuous Indian Sign Language dataset was created using the Canon EOS camera in Robotics and Artificial Intelligence laboratory (IIIT-A) for the experiments. Probes were being tested using various types of classifiers like Euclidean distance, Correlation, Manhattan distance, city block distance etc. Comparative analysis was performed on the various classifiers, from which Correlation and Euclidean distance were found to give better accuracy.

Rajaganapathy used Microsoft Kinect a motion capture device. S, Aravind. B, Keerthana. B and Sivagami. M [12] for converting the human sign language to voice.

In [15] Khushdeep Kaur and Parteek Kumar designed a system in which HamNoSys is generated for the Indian Sign Language words. The HamNoSys is then converted into XML form known as SiGML. The HamNoSys to SiGML Conversion System will reduce the dependency on eSIGNEditor that contains the database of words for the American and British Sign Language, which are different from Indian Sign Language. A virtual human tool in JA SiGML URL APP will take input as SiGML and will generate the signs for the corresponding words.

M.K. Bhuyan, Chaitanya Narra, and Darsha Sharath Chandra presented a method for regenerating the original hand gesture sequence with the help few frames, which are the keyframes [16]. A Hu moment is used to extract the key frames.

To recognize the Thai sign Language hand gesture in complex background Chana Chansri and Jakkree Srinonchat developed a technique using the fusion of depth and color video [17]. Histograms of Oriented Gradients (HOG) are used for the image feature extraction and the backpropagation of neural network is used to recognize the hand gesture. The system has good robustness with an accuracy rate of 84.05%. The system is able to recognize 24 hand gestures that are representing 42 alphabets.

Amrutha C U, Nithya Davis, Samrutha K S, Shilpa N S and Job Chunkath discussed the development of a mobile application for the persons with a sensory deficit for the language acquisition [18]. This system will convert the text to Indian Sign Language images and videos.

In [19] Pranali Loke, Juilee Paranjpe, Sayli Bhabal and Ketan Kanere introduced a sign language converter using the android application. It will convert the hand gesture of Indian Sign Language to natural language. The images are captured through Android application and then sent to the server. From the server, it is sent to MATLAB [27][29] where the image is mapped using the neural network and will convert it to text. This text is then sent to the server that in turn is displayed on the android screen.

Washef, Kunal, and Soma [20] represented an algorithm that recognizes the hand gesture by using Dynamic Time Warping methodology. This system has three modules that are real-time detection of face region and two hand regions, tracking the hand's trajectory in terms of direction among consecutive frames as well as distance from the center of the frame and gesture recognition based on analyzing variations in the hand locations along with the center of the face.

This system overcomes the limitations of vision-based approach concerning the illumination, background complexity and distance from the camera that is up to 2 meters.

In [21] Prashant, Kshitija , Tejaswini, and Pramod proposed a handy mobile application that will act as a two-way communicator. This system has two modules. First, it draws the ISL gestures from the real-time video and translates it to human understandable speech. Then in the second module, the natural language will be taken as input giving appropriate ISL gesture as output. The mapping of images is done using the Correlational based approach. Then Google's Text-to-speech and Speech-to-text APIs are used for the relevant audio and text generation.

Sidig , Luqman, and Mahmoud [22] have investigated on various distance transformation techniques for the extraction and description of the features of the sign images. They have shown the performance of three techniques, namely Fourier, Hartley, and Log-Gabor transforms. They have also tested and compared different classification schemes. The accuracy of the overall system reached over 99% using Hartley transformation technique.

In [23] Mrim and Srini have conducted a survey on the communication barrier faced by the deaf people and the communication applications (apps) available for deaf people. Then they recommended the best app among the existing app and identified the best design and the features of those apps for the developers. They examined 55 apps but only 6 were found to be specifically designed as the communication app for deaf people.

Rajam and Balakrihnan [24] proposed the basis for the development of Sign Language recognition system for one of the south Indian Language. A set of 32 signs each representing the Binary UP and DOWN positions of the fingers were defined in this system. The method was developed with respect to a single user in both training and testing phase. The 10 number of static images were trained that were pre-processed using figure point extraction.

Image processing techniques were used to identify the fingertip to convert the images into text. During the testing phase, the dynamic images were identified in this method. The system had the accuracy of 98.15% when trained with 320 images and tested with 160 images.

Sign sequence segmentation and sign recognition is two main problems in continuous sign language recognition. In [25] Wenwen, Jinxu, and Zhongfu used the hidden Markov Model (HMM) to calculate the similarity between the sign

model and testing sequence and a fast algorithm was proposed for computing the likelihood of HMM to reduce the computation complexity.

To improve the recognition rate grammar constraint and sign length constraint were employed. The coarse segmentation method was proposed to provide a maximal level number. The method showed superior recognition performance and lower computation compared to other techniques when experimented with KINECT dataset of Chinese language containing 100 sentences each composed of 5 sign.

Leap motion sensor is used in [26] to recognize the Indian Sign Language. Leap motion sensor is sensor device that can be used to recognize the gesture more specifically hand gestures and provides an only specific number of relevant points. The device extracts the significant charades from the dataset then system maps all the frames recognized by the sensor to attributes. The Random Forest (RF) classifier model is used then for feature classification. A block-list classifier is used to remove all non-gesture frames then the RF is used to form the meaningful words. This meaningful word is then displayed on an LCD.

The objective of [30] was to develop a real-time system for hand gesture recognition. This system recognizes the gestures of the hand, features like peak calculation and angle calculation then it will convert the image into voice and vice versa. The system is designed and implemented using the artificial intelligence, image processing, and data mining concepts to take input as hand gestures and generate recognizable output in form of text and voice with an accuracy of 91%.

III. CONCLUSION

The above study illustrates various techniques used in vision-based system, along with some hardware-based techniques. In vision, based technique the accuracy depends on many factors like the position of the hand, the distance from the camera, the background of image, illumination, dataset etc.

The handsome amount of sample images per sign in dataset will lead for batter and accurate results. Basically two methods namely fingerspelling based and word spelling based are used to represents a sign, most of the researchers uses word spelling based method but due to unavailability of proper dataset this method is not up to the mark.

Only a few systems have recognized both alphabets and numbers for specific regional sign languages like American Sign Language, Arabic Sign Language, etc.

The real-time images can be used for the sign recognition but capturing and interpreting sign from real-time images is not up to the mark hence most of the methods have used the static images.

As various sign languages used worldwide and various vision based sign language interpreters developed by researches, we can also propose a sign language interpreter that used for Indian Sign Language and to help Indian deaf-dumb people to communicate with normal people. In the proposed system, real-time images can be captured through held held device like smart phone and recognize sign. The fingerspelling can be used instead of word spelling thus making it more efficient.

REFERENCES

[1] A. Sood and A. Mishra, "AAWAAZ: A communication system for deaf and dumb," 2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2016.

[2] V. Adithya, P. R. Vinod, and U. Gopalakrishnan, "Artificial neural network based method for Indian sign language recognition," 2013 IEEE Conference on Information and Communication Technologies, 2013.

[3] R. Shangeetha., V. Valliammai., and S. Padmavathi., "Computer vision based approach for Indian Sign Language character recognition," 2012 International Conference on Machine Vision and Image Processing (MVIP), 2012.

[4] S. Rathi and U. Gawande, "Development of full duplex intelligent communication system for deaf and dumb people," 2017 7th International Conference on Cloud Computing, Data Science & Engineering - Confluence, 2017.

[5] "Human–computer interaction," Wikipedia, 23-Jan-2018. [Online]. Available: https://en.wikipedia.org

[6] B. Rajapandian, V. Harini, D. Raksha, and V. Sangeetha, "A novel approach as an AID for blind, deaf and dumb people," 2017 Third International Conference on Sensing, Signal Processing and Security (ICSSS), 2017.

[7] M. S. Nair, A. P. Nimitha, and S. M. Idicula, "Conversion of Malayalam text to Indian sign language using synthetic animation," 2016 International Conference on Next Generation Intelligent Systems (ICNGIS), 2016.

[8] S. Chattoraj, K. Vishwakarma, and T. Paul, "Assistive system for physically disabled people using gesture recognition," 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), 2017.

[9] E. E. Abdallah and E. Fayyoumi, "Assistive Technology for Deaf People Based on Android Platform," Procedia Computer Science, vol. 94, pp. 295–301, 2016.

[10] A. Thalange and S. Dixit, "COHST and Wavelet Features Based Static ASL Numbers Recognition," Procedia Computer Science, vol. 92, pp. 455–460, 2016.

[11] K. Tripathi and N. B. G. Nandi, "Continuous Indian Sign Language Gesture Recognition and Sentence Formation," Procedia Computer Science, vol. 54, pp. 523–531, 2015.

[12] S. Rajaganapathy, B. Aravind, B. Keerthana, and M. Sivagami, "Conversation of Sign Language to Speech with Human Gestures," Procedia Computer Science, vol. 50, pp. 10–15, 2015.

[13] N. Harish and S. Poonguzhali, "Design and development of hand gesture recognition system for speech impaired people," 2015 International Conference on Industrial Instrumentation and Control (ICIC), 2015.

[14] D. Abdulla, S. Abdulla, R. Manaf, and A. H. Jarndal, "Design and implementation of a sign-to-speech/text system for deaf and dumb people," 2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA), 2016.

[15] K. Kaur and P. Kumar, "HamNoSys to SiGML Conversion System for Sign Language Automation," Procedia Computer Science, vol. 89, pp. 794–803, 2016.

[16] M. Bhuyan, C. Narra, and D. S. Chandra, "Hand gesture animation by key frame extraction," 2011 International Conference on Image Information Processing, 2011.

[17] C. Chansri and J. Srinonchat, "Hand Gesture Recognition for Thai Sign Language in Complex Background Using Fusion of Depth and Color Video," Procedia Computer Science, vol. 86, pp. 257–260, 2016.

[18] C. Amrutha, N. Davis, K. Samrutha, N. Shilpa, and J. Chunkath, "Improving Language Acquisition in Sensory Deficit Individuals with Mobile Application," Procedia Technology, vol. 24, pp. 1068–1073, 2016.

[19] P. Loke, J. Paranjpe, S. Bhabal, and K. Kanere, "Indian sign language converter system using an android app," 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA), 2017.

[20] W. Ahmed, K. Chanda, and S. Mitra, "Vision based Hand Gesture Recognition using Dynamic Time Warping for Indian Sign Language," 2016 International Conference on Information Science (ICIS), IEEE 2016.

[21] P. G. Ahire, K. B. Tilekar, T. A. Jawake, and P. B. Warale, "Two Way Communicator between Deaf and Dumb People and Normal People," 2015 International Conference on Computing Communication Control and Automation, IEEE 2015.

[22] A. A. I. Sidig, H. Luqman, and S. A. Mahmoud, "Transform-based Arabic sign language recognition," Procedia Computer Science, vol. 117, pp. 2–9, 2017.

[23] M. Alnfiai and S. Sampali, "Social and Communication Apps for the Deaf and Hearing Impaired," 2017 International Conference on Computer and Applications (ICCA), IEEE 2017.

[24] P. S. Rajam and G. Balakrishnan, "Real time Indian Sign Language Recognition System to aid deaf-dumb people," 2011 IEEE 13th International Conference on Communication Technology, 2011.

[25] W. Yang, J. Tao, and Z. Ye, "Continuous sign language recognition using level building based on fast hidden Markov model," Pattern Recognition Letters, vol. 78, pp. 28–35, 2016.

[26] P. Chavan, T. Ghorpade, and P. Padiya, "Indian sign language to forecast text using leap motion sensor and RF classifier," 2016 Symposium on Colossal Data Analysis and Networking (CDAN), IEEE 2016.

[27] N. S. Soni, M. S. Nagmode, and R. D. Komati, "Online hand gesture recognition & classification for deaf & dumb," 2016 International Conference on Inventive Computation Technologies (ICICT), 2016.

[28] S. Upendran and A. Thamizharasi, "American Sign Language interpreter system for deaf and dumb individuals," 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), IEEE 2014.

[29] S. N. Sawant and M. S. Kumbhar, "Real time Sign Language Recognition using PCA," 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies, IEEE 2014.

[30] S. S. Shinde, R. M. Autee, and V. K. Bhosale, "Real time two way communication approach for hearing impaired and dumb person based on image processing," 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), IEEE 2016.

AUTHORS' PROFILE

Mr. Rakesh Savant is working a teaching assistant at BMIIT. He has more than four years of experience in academics. His interest area includes artificial intelligence, Machine learning, digital image processing and computer networks. He has attended many seminar and workshops He can be contacted at rakesh.savant@utu.ac.in

Ms. Amrutha Ajay Kunnath is studying in semester 8 of 5 Years Integrated M.Sc. (IT) programme at Babu Madhav Institute of Information Technology affiliated to Uka Tarsadia University-Bardoli.