INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Two Way Communication Systems using Speech and Sign Language: Survey

Aishwarya Jawalkar[1], Samarth Sathe[2], Rutuja Jagtap[3], Sulaxmi Raskar[4]

[1, 2, 3, 4]Department of Computer Engineering, RMD Sinhgad school of Engineering Warje, Pune, Maharashtra.

Abstract: Sign language is complex to understand but the complete language which involves the hand's movement, facial expressions, and body postures. Sign language is core communication media to the people which cannot speak. It is not a universal language means every country has its own sign language. Every country has its own grammar for their sign language, word orders and pronunciation. The issues arise when people try to communicate using their language with the people who are unaware of this language grammar. To detect the sign using hand gesture and then convert into textual or verbal form which will can be understood by any person i.e. recognizes the result for each sign.
Keywords: Record Voice/Speech, Convert ASL Into Speech, Character Detection, Speech Processing, Text Detection.

I. HISTORY

Where Did American Sign Language Originate?
The exact beginnings of ASL are not clear. Many people believe that ASL came mostly from French Sign Language (FSL). Others claim that the foundation for ASL existed before FSL was introduced in America in 1817. Today's ASL likely contains some of this early American signing. Modern ASL and FSL share some elements, including a substantial amount of vocabulary. However, they are not mutually comprehensible.

A. 1970's Sign language slowly returns to deaf education but in English forms.
B. 1975, Sign Instructors Guidance Network. Today it is called ASL Teacher Association.
C. 1988. Now ASL was on the track of popularity because of the first Deaf President.
D. 1990 ASL accepted as a foreign language credit and course offered in many college and high school at explosive growth.
E. 1990's many states legalize ASL as a foreign language course for High Schools and college.

II. INTRODUCTION

Sign language is a way by which the gestures made by the user are used for communication. Human gestures are an efficient and powerful way of interaction.
The language allows the speech impaired people to communicate with rest of the world. But issue with sign language is that it is only confined to the people who are deprived of speech. In order to communicate, normal people also have to be aware of the sign language.
Sign language is a language through which communication happens without the means of acoustic sounds. It relies on sign patterns like orientation and movements of the arm that aids in communication between individuals. There are very subtle differences in the gestures between similar signs. The mechanics of ASL make it challenging to distinguish and capture differences between signs. The deaf and dumb people are the communication part, to convey their thought with other deaf and dumb people and with other normal people.

A. What is American Sign Language?
American Sign Language (ASL) is a complete, complex language that employs signs made with the hands and other movements, including facial expressions and postures of the body. It is the first language of many deaf all over the world, and one of several communication options available to deaf people. ASL is said to be the fourth most commonly used language in the United States.
The system is designed to visually recognize all static gestures of American Sign Language (ASL) with bare hand. Different users have different hand shapes and skin colours, making it more difficult for the system to recognize a gesture. The system combines five feature extraction algorithms for user independent and robust hand gesture recognition. The whole system works in four steps for gesture recognition such as image acquisition, pre-processing, feature extraction and feature recognition.
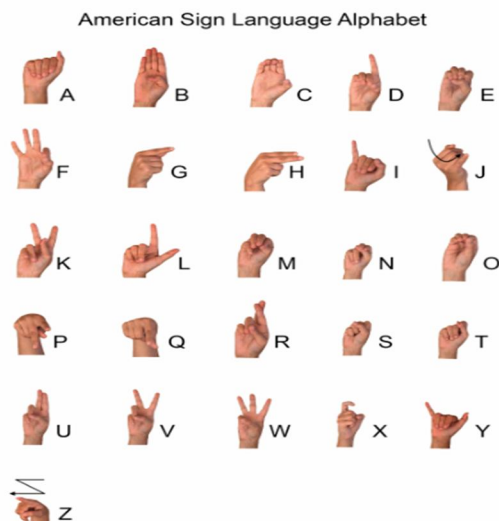
Figure: 26 alphabets of American Sign Language

We can detect ASL from the recorded speech and vice versa or detecting ASL from the user speech. The recorded speech is processed and detects the text of the speech and display. And then detected text is converting into ASL.

III. LITERATURE REVIEW

"Hiroomi Hikawa", et.al Introduced a system of hardware posture recognition based on a SOM and Hebb hybrid vector classifier. As the feature vectors used in the proposed system were invariant to location changes in input images, the recognition was robust to location changes in hand signs. This paper focused on tuning of the classifier by new training method.

"Md. Mohiminul Islam," et al they have applied different algorithms for feature extraction of hand gesture recognition system. This includes K convex hull for fingertip detection, pixel segmentation, eccentricity, elongatedness of the object. The experimental results show that K convex hull algorithm gives more accurate fingertip detection. Besides, other algorithms are applied to get greater accuracy in recognition system. Image frames taken by mobile video camera interfaced with the computer are tested by our trained ANN.

"Yifan Zhang" , et.al In this work, they have introduced up-to-date the largest dataset called Ego Gesture for the task of egocentric gesture recognition with sufficient size, variation and reality, to successfully train deep networks. The dataset is more complex than any existing datasets as our data is collected from the most diverse scenes. Compared to gesture classification in segmented data, the performance on gesture detection is far from satisfaction and has much more space to improve.

"Deepali Naglot" et al describes In this paper, the fundamental Experimentation is done on American Sign Language. American Sign Language, considered to be one of the predominant sign language for deaf communities in United States. ASL is widely used language all throughout the world. It uses only one hand to display the gestures and thus makes it easy for interpretation and understanding. It comprises of around 6000 gestures and other common words. The common words are shown with some specific gesture or spelling with the help of 26 hand gestures indicating 26 alphabets of ASL.

"Ponlawat Chophuk", et.al The paper aims to solve the problem of the first signs in American Sign Language because their recognition is not perfect due to the Euclidian distances of the fingertip to palm position used in 3D of those signs are similar, so it is difficult to recognize them. We therefore propose a system of first American Sign Language recognition with a bare hand in depth plane by using 3D non-contact motion sensor. In this system, two patterns of the polygon area between the consecutive fingertip positions with palm position in depth plane by Shoelace formula are used to identify the first sign language: 1) six triangles area, 2) one hexagon area, then a decision tree is applied to classify the alphabets. The results showed the 7 alphabets in fist ASL using the researcher's hand. The accuracy of the method proposed is approximately 96.1%. Nevertheless, N and T alphabets were not perfect because the Euclidian distances of the fingertip to palm position proposed in 3D of those signs are similar, so it is difficult to recognize them.

"Suraksha Devi" et al Presents a The primary objective of this work is to enable the speech impaired and the mute to have a seamless communication and independent living in the society. The proposed prototype and accompanying algorithm accomplished the initial objective. The glove is cost effective and is capable of translating sign gestures (conventional Indian Sign Language) into

speech-text in real time using android application on the phone. The significant recording of hand movements and feature extraction is done by applying Principle Component Analysis. The glove prototype is independent of the surrounding light or any other kind of interference. As a result, precise and accurate recognition gesture is possible in less time. This paper emphasizes on the translation of sign language with the help of the glove; however, the glove could be used for various other applications such as for the virtual reality interaction, gaming, entertainment, education technology, robotics etc. This is an ongoing work; the of the algorithm and prototype are going on.

"Fatih Erden" et al In this article multimodal hand gesture detection and recognition system is presented. Since both infrared and visible range information is used, the proposed system is more accurate than IR-only and less power consuming than camera only systems. A novel WTA code based sensor fusion algorithm is also presented for 1-D PIR sensor signal processing. The algorithm fuses the data coming from the different PIR sensors in an automatic manner to determine left-to-right, right-to-left, upward, downward, clockwise and counter-clockwise motions. A Jaccard distance based metric is used to classify the hash codes of feature vectors extracted from sensor signals.

IV. EXISTING SYSTEM

The system combines five feature extraction algorithms for user independent and robust hand gesture recognition. The whole system works in four steps for gesture recognition such as image acquisition, pre-processing, feature extraction and feature recognition. Image frames taken by video camera interfaced with the computer are tested by our trained CNN. The CNN is trained with sample images of our database and it recognizes ASL alphabets and numbers with accuracy in real time environment.

A. *Convert Speech Into American Sign Language (ASL)*
1) *Record Voice/Speech:* This is phase where we can record the voice/speech of the user for recognizing text.
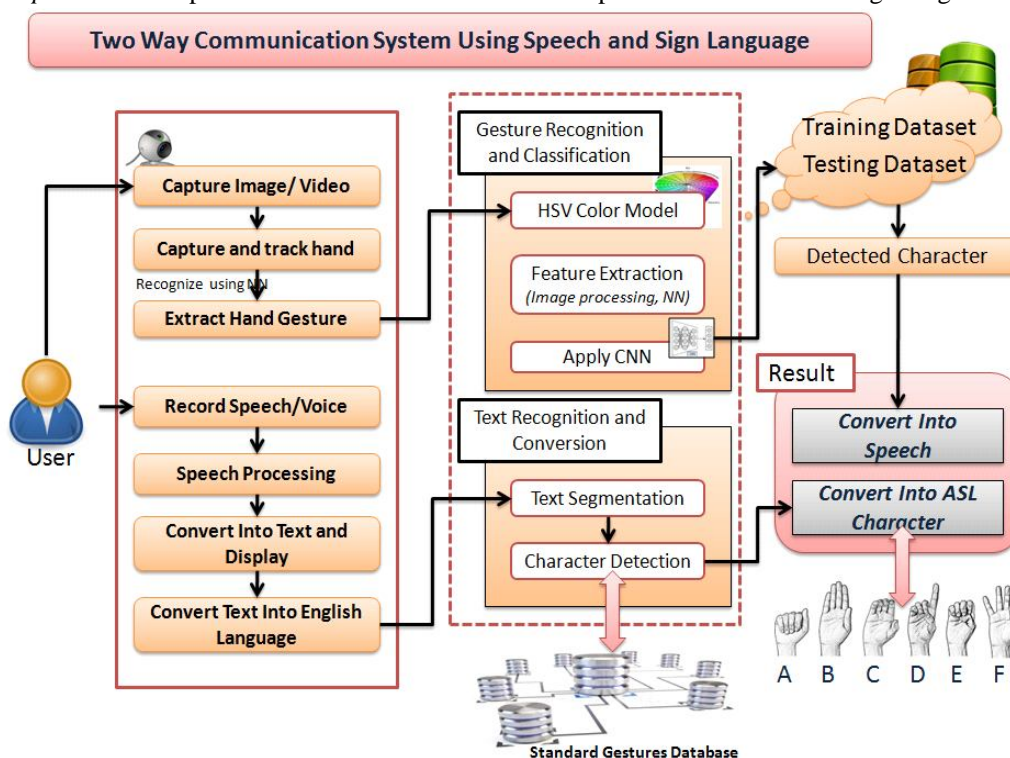

Figure: - System Architecture

2) *Speech Processing and Detect Text:* Firstly the system can process the recorded speech. After that it detect the text from processed speech and then display detected text to user.
3) *Translate Text into English:* If the text is in other than English language then the system can convert it into English language.
4) *Convert Text into ASL Character:* The system can detect character from text and then detect the ASL character. Form the text system detects each character and converts it into ASL character.

For e.g.

A   B   C   D   E   F

Detected text =

Then system detects and converts it into ASL character.

5) *Final Result:* In this phase, the system can display the detected ASL character to the user.

A   B   C   D   E   F

B. *Convert ASL Into Speech*

1) *Capture Image / Video:* This is phase where we are capturing the images/videos for recognizing hand gesture.

2) *Image Acquisition and Enhancement:* This is phase where we are getting input images and applying image enhancement techniques which is called as preprocessing phase of input image. In Image enhancement noise reduction and contrast adjustment will be performed, following is the formula for the RGB Normalization:

$$r = \frac{R}{R+G+B}; \quad g = \frac{G}{R+G+B}; \quad b = \frac{B}{R+G+B}$$

3) *Feature Extraction:* This phase is used for determining the meaning of a gesture from the image. These features should uniquely describe the gesture in order to achieve a reliable recognition. Therefore, different gestures should result in different, good discriminable features. Some of these features are presented as follows:

a) *Local Orientation Histogram:* They employing histograms of local orientation, they use the orientation histogram as a feature vector for gesture classification and interpolation.

b) *Local Brightness*: The input gesture image is divided into 25x25 blocks each of 5x5 block size, and they have calculated the local brightness of each divided block after applying colored segmentation operation using HSV (Hue, Saturation and Value) color model.

c) *Binary Object Features:* The height, area, centroid, and distance of the centroid from the origin (top-left corner) of the binary image are used as features.

4) *Character Detection:* We can apply CNN algorithm on the trained dataset to detect the character from the gesture images. By using an image processing we can easily recognize the hand gesture.

5) *Final Result:* After detecting the character from image we represent the final result represent in verbal form i.e. speech.

V.   CONCLUSION

This work survey and analyzed how to visually recognize all static gestures of American Sign Language (ASL) with bare hand. Different users have different hand shapes and skin colours, making it more difficult for the system to recognize a gesture. Sign Language Recognition is essential for the deaf and dumb people to communicate with other people.

REFERENCES

[1] Hiroomi Hikawa, Keishi Kaida, "Novel FPGA Implementation of Hand Sign Recognition System with SOM-Hebb Classifier" 2013 IEEE.

[2] Md. Mohiminul Islam,Sarah Siddiqua, and Jawata Afnan"Real Time Hand Gesture Recognition Using Different Algorithms Based on American Sign Language" 2017 IEEE.

[3] Yifan Zhang*, Congqi Cao*, Jian Cheng, and Hanqing Lu "EgoGesture: A New Dataset and Benchmark for Egocentric Hand Gesture Recognition" 2018 IEEE

[4] Peijun Bao, Ana I. Maqueda, Carlos R. del-Blanco, and Narciso García, "Tiny Hand Gesture Recognition without Localization via a Deep Convolutional Network", IEEE Transactions on Consumer Electronics, Vol. 63, No. 3, August 2017.

[5] S.H. Lee, M.K. Sohn, D.J. Kim, B. Kim, and H. Kim, "Smart TV interaction system using face and hand gesture recognition," in Proc. ICCE, Las Vegas, NV, 2013, pp. 173-174.

[6] S. Kim, G. Park, S. Yim, S. Choi and S. Choi, "Gesture-recognizing hand-held interface with vibrotactile feedback for 3D interaction," IEEE Trans. Consum. Electron., vol. 55, no. 3, pp. 1169-1177, 2009.

[7] S. S. Rautaray, and A. Agrawal, "Vision based hand gesture recognition for human computer interaction: a survey," Artificial Intelligence Review, vol. 43, no. 1, pp. 1-54, 2015.

[8] D. W. Lee, J. M. Lim, J. Sunwoo, I. Y. Cho and C. H. Lee, "Actual remote control: a universal remote control using hand motions on a virtual menu," IEEE Trans. Consum. Electron., vol. 55, no. 3, pp. 1439-1446, 2009.

[9] D. Lee and Y. Park, "Vision-based remote control system by motion detection and open finger counting," IEEE Trans. Consum. Electron., vol. 55, no. 4, pp. 2308-2313, 2009.

[10] F. Erden and A. E. Çetin, "Hand gesture based remote control system using infrared sensors and a camera," IEEE Trans. Consum. Electron., vol. 60, no. 4, pp. 675-680, 2014.

[11] S. Jeong, J. Jin, T. Song, K. Kwon and J. W. Jeon, "Single-camera dedicated television control system using gesture drawing," IEEE Trans. Consum. Electron., vol. 58, no. 4, pp. 1129-1137, 2012.

[12] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," IEEE Trans. on Pattern Anal. Mach. Intell., vol. 38, no. 1, pp. 142-158, 2016.

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)