A Survey: Hand Gesture Vocalizer

Asha Shajee [1], Manonmani Lakshmanan [2], Lekhita Sharma [3], Jaya Kumar [4]
[1, 2, 3, 4] SRM IST, Ramapuram, Chennai, Tamil Nadu, India.

Abstract – Gestures are a very essential element of communication. This project of ours aims to create a system which will be able to detect the gestures given as input and to produce the resulting output in the form of messages or audio, mainly aiming to aid the deaf and dumb people to cope up with their difficulties and to find an efficient solution for them by converting their hand gestures to signals or voice messages. The technology being used behind this features IoT and image processing. This is a project for social purpose to try implementing a system which makes communication gap between deaf people and hearing people as less as possible. Thus, the conclusion of this project is to make a simple prototype by taking the gestures and converting them into audio and visual formats so that they're recognised efficiently, helping in communication. This project has wide applications, making it cost-efficient and reliable.

Index Terms – Flex sensors, Data Glove, Accelerometer, EMG sensors, Microcontroller, Template Matching.

1. INTRODUCTION

Communication is the most basic and important form of interaction with anyone. Thus for interacting with deaf and dumb people, we use sign language or gestures.

Gestures are the only approach of communication for the people who are suffering from disabilities like hearing and speaking.

According to the survey taken in the year 2011 by the government of India, it is reported that over 2.68 crores of people in India suffer from some form of the disabilities. Out of this, 18.9% people have speech and 7.5% people have hearing disability.

Sign language is composed of visual gestures and signs, which are used by those who are specially challenged in hearing and speaking. It is a well-structured code gesture where every sign has a specific meaning allotted to it.

There are 143 different existing sign languages all over the world, mainly American Sign Language, British Sign Language, French Sign Language, Japanese Sign Language, and Indian Sign Language.

Every country has its own language since sign language is not a universal language and also has its own grammatical and syntactical meaning as per the language that is referred, which is differs from country to country.

According to the research that has been done in sign languages, data glove (sensor) based and vision-based are the two primary approaches that are being used to recognize sign languages.

In sensor-based approach, the user has to put on a glove to the hand which carries a load of cables that is connected to the computer. Such gloves are high-priced and are not easily available everywhere. Whereas, vision-based approach works on dataset containing sign gestures images captured by the camera.

This method also gives a usual environment to the user and reduces the problems as faced in the glove-based method.

The foremost features of a person are thus the hand gestures that can be summarized by their dynamic, multi-attribute properties.

The people who are physically challenged in speaking and hearing are not able to work along with the social world.

Because of this, these people are not treated as a part of the society. They aren't able to even take part in any social events of normal people; also physically impaired students can't study with normal students in school, and many other such instances.

This creates a large gap between normal and deaf-dumb people, causing a rift that rises day-by-day. Because of this, it leads to the latter to create their own and separate society.

To reduce the isolation and the above issues, we propose a system that provides vision to the deaf and voice to the dumb people, promising them an independent life without any help of human translators.

This can be accomplished by developing a Gesture Detection Vocalizer which allows communication to occur among everyone.

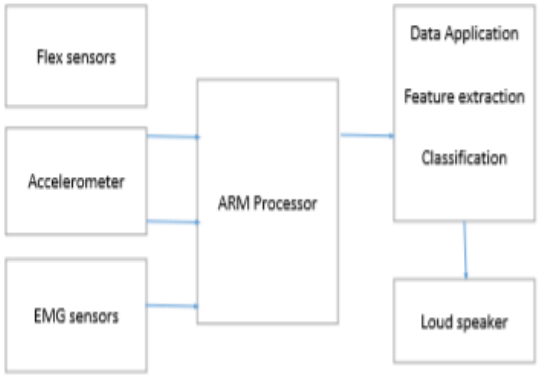

Figure 1. Block Diagram of working module of Hand Gesture

2. RELATED WORK

Hand gestures, facial expressions and body language-together these form the sign language, but majority of the information about a particular sign is revealed from the performance of the hand and thus from now on most of the research work done involves hand gestures and its detection.

On static hand gestures, majority of the work has been done but working on dynamic hand gestures is still a challenging task because sometimes one gesture represents the same meaning. For example, 'Very good' and 'Beautiful' have the same hand gesture.

- Albanian Sign Language Recognition
- Leap Motion Controller for Gestures
- Gesture Processing System
- American Sign Language Alphabet Recognition Using Microsoft Kinect
- Vision-based Gesture Detection
- Flex-sensored Hand Detection
- Android-based Gesture Detection
- Dynamic Hand Gestures
- Minimum Time Gesture Detection
- Automatic Gesture and Speech Recognition

Thus, we see that in this literature survey, it shows the related wok of the other authors and their papers for the designing of their systems. The papers of these scholars that we have taken are all dated of the years in which they were published. Hence, these 10 papers that we have taken is unique and specific from the authors.

These include the current knowledge, substantive findings, as well as theoretical and methodological contributions of each of the authors to Speech and Gesture Detection.

Thus, it explains the methods which were used in them along with all their outcomes as well as their drawbacks, which are focused on each paper.

In our Literature Survey, we have referred around 10 different papers from the year 2012-2016.

One of the example:

Gesture Processing System –

Here, the developers who developed this system used Fourier descriptors for feature extractions using Hindi text.

They then they converted it into Hindi language as output speech or voice for detection.

Hence, we intend to take these references from the methodology used in different projects of speech recognition system.

3. PORPOSED MODELLING

Thus, by the analysis of this project, we see that the recognition of hand gestures is such done that it is then converted to speech or else it can be done vice-versa as well.

Thus, in the diagram alongside, we can see how the gestures are converted into speech- monologues as well as dialogs.

Image Acquisition:

There are many input devices for image acquisition are available some of them are hand images, data gloves and markers. In this system a real time image is acquired by using 20 mega pixel web cam using MATLAB inbuilt command.

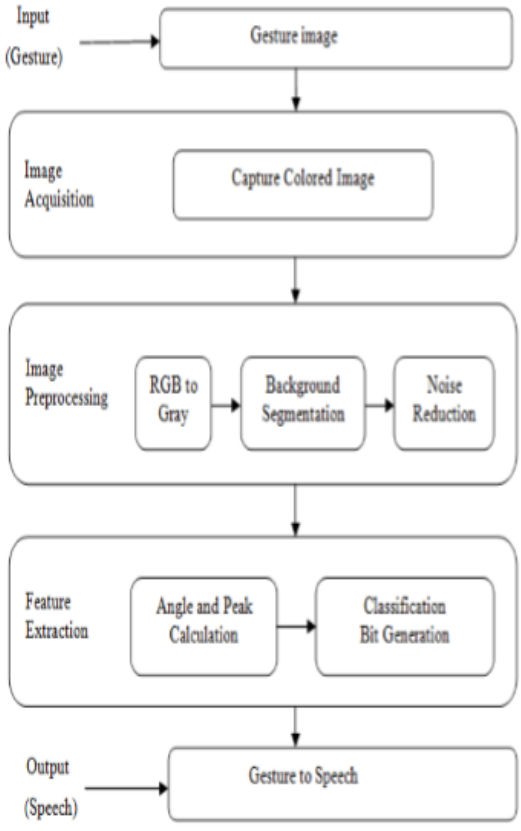

Figure 2. Flow diagram of hand gestures to speech

Image Preprocessing:

Image preprocessing is very important step for getting good results. The real time RGB color images are captured using a 20 MP webcam and converted into gray Images. Background segmentation is used to separate the hand object in the image

from its background. Noise elimination steps are applied to remove connected components.

Feature Extraction:

There are many techniques are available for feature extraction like edge detection, hidden markov model, contour tracking algorithm and Gabor filter. The proposed system uses the angle and peak calculation to extract the features of hand gestures. Peak calculation calculates the number of raised and folded fingers and to differentiate gestures we used zero, positiveand negative angles. For classification 12 bit binary sequence is generated for each hand gesture which classifies the different hand gestures, Then finally we recognize the hand gestures and then convert it into speech by using MATLAB inbuilt command.

Gesture to Speech conversion:

After recognizing the gestures the hand gesture images are converted into speech for providing better communication for hearing impaired and dumb person.

A new gesture dataset is record by the Kinect Sensor that records 10 different gestures and each gesture contains 50 samples. Each sample contains a depth image and a corresponding color image. Considering the practical application, a few pairs of similar gestures is set in the dataset, such as gesture 1 and gesture 7, gesture 2 and gesture 8. The collecting of dataset is done in a natural situation. The samples of gestures are not exactly the same as the standard gestures. Thus, the size and direction of each person's gesture may be different. Because of different personal habits, the same gestures of different people may also have nuances.

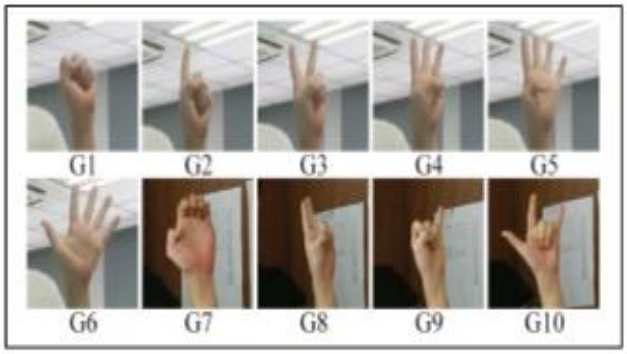

Figure 3. Hand Gestures

This is used for denoting the gesture signals that can be captured so that it can be converted into speech. By using such hand gestures to determine actions or verbal instincts, it thus justifies the purpose of the gesture vocalizer. Thus, through these such hand gestures, it first detects the hand gestures and then makes them recognizable so as to understand the gestures efficiently. There are numerous types of gestures available nowadays, which makes it more easy and more accessible to the users. Hence, this will thus help to be effective to the users in a gigantic way. The gestures thus play an important role and it can be replayed in native and other languages as wellIn the following figures, we see a few gestures and their meanings too. The gestures thus play an important role and it can be replayed in native and other languages as well.

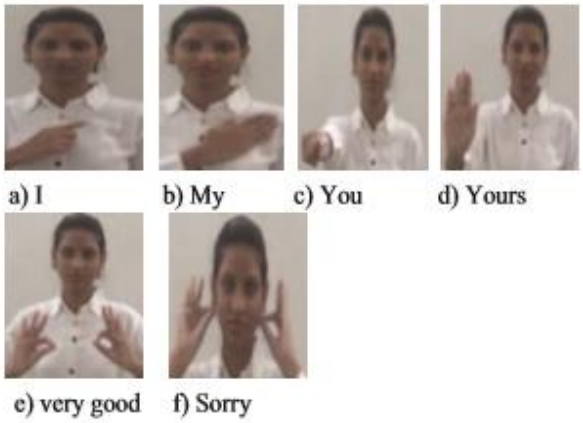

Fig. 4. Various Indian sign language words sign using single and both hands.

Also, the five fingers display the five peak filtered points in the hands wherein the sensors detect and sense the maximum.

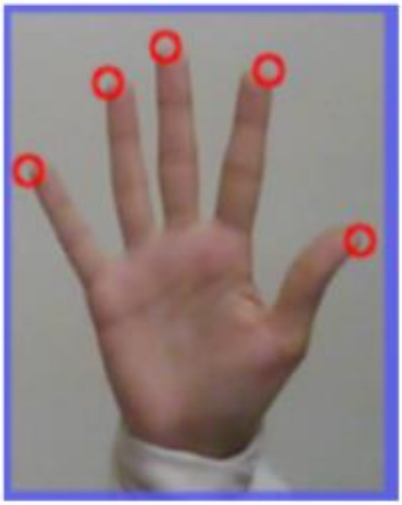

Figure 5. Five-peak filtered points of hands.

4. RESULTS AND DISCUSSIONS

With this project, the deaf-mute people can use the hand gestures to perform sign language and it will be converted into speech with accuracy 93%; and the speech of normal person is converted into hand gesture, so the communication between them can take place easily.

5. CONCLUSION

The proposed system is easy to implement as there is no complex feature calculation. This system provides us with high gesture recognition rate with accuracy 90% within minimum time. The system aims to lower the communication gap between deaf people and normal world, since it facilitates dual communications. The purpose of this device is to make conversations easy between dumb and the normal people around the world. In order to minimize their limitations and enable them to expand their boundaries, this is the best suitable device. The projected methodology interprets hand gestures into speech and vice versa. With this project, the deaf-mute people can use the hand gestures to perform sign language and it will be converted into speech with accuracy 93%; and the speech of normal person is converted into hand gesture, so the communication between them can take place easily. The proposed system converts gesture images into speech so that the normal person can know what the deaf and dumb person said and the second part consists of speech to gesture conversion; hence deaf and dumb person knows what the normal person said hence this system enables both to further connect with their society and aids them in overcoming communication obstacles created by the society's incapability of understanding and expressing sign language. By increasing databases, the efficiency of the system can be further improved.

REFERENCES

[1] Eriglen Gani , Alda Kika, "Albanian Sign Language (AlbSL) Number Recognition from Both Hand's Gestures Acquired by Kinect Sensors" International Journal of Advanced Computer Science and Applications, Vol. 7, No. 7, 2016.

[2] Miada A. Almasre, Hana Al-Nuaim, "A Real-Time Letter Recognition Model for Arabic Sign Language Using Kinect and Leap Motion Controller v2" International Journal of Advanced Engineering, Management and Science (IJAEMS), Vol-2, Issue-5,ISSN: 24541311,May 2016.

[3] Rashmi. B. Hiremath, Ramesh. M. Kagalkar, "Methodology for Sign Language Video Interpretation in Hindi Text Language" International Journal of Innovative Research in Computer and Communication Engineering, Vol. 4, Issue 5, May 2016.

[4] Cao Dong, Ming C.Leu, Zhaozheng Yin, " American sign language Alphabet Recognition Using Microsoft Kinect", Computer Vision and pattern Recognition workshop, IEEE conference, pp ,2015.

[5] Paulo Trigueiros, Ferando Ribeiro, Luis Paulo Reis, " Vision Based Portuguese Sign Language Recognition System", New Perspectives in Information System and Technologies, Volume 1,Advances in Intelligent System and Computing, Springer International Publishing Switzerland,pp-605-608,2014.

[6] Vajjarapu Lavanya, M.S. Akulapravin, Madhan Mohan, "Hand Gesture Recognition and Voice Conversion System Using Sign Language Transcription System", IJECT Vol. 5, Issue 4, Oct - Dec 2014.

[7] K. Sangeetha, L. Barathi Krishna, "Gesture Detection For Deaf and Dumb People ", International Journal of Development Research Vol. 4, Issue, 3, pp. 749-752, March, 2014.

[8] M.M.Gharasuie, H.Seyedarabi, "Real-time Dynamic Hand Gesture Recognition using Hidden Markov Models", 8th Iranian Conference on Machine Vision and Image Processing (MVIP), 978-1-4673-61842/13/$31.00,2013.

[9] Aditi Kalsh, N.S. Garewal , "Sign Language Recognition for Deaf & Dumb", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 9, September 2013.

[10] Bhupinder Singh, Neha Kapur, Puneet Kaur "Speech Recognition with Hidden Markov Model: A Review" International Journal of Advanced Research in Computer and Software Engineering, Vol. 2, Issue 3, March 2012.