International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-2, July 2020

Sign Language Translation using Hand Gesture Detection

Neelam Shrivastava, Abhishek Jain, Kopal Garg, Abhishekh Pratap Singh, Abhishek Sharma

***Abstract:** This Paper demonstrate a module on a sign to speech (voice) converter for auto Conversion of American sign language (ASL) to English Speech and text. It minimizes the gap of communication between speech impaired and other humans. It could be used to understand a speech impaired person's thoughts or views which he communicates with others through its ASL gestures but failing to communicate due to a large communication gap between them. It also work as a translator for person who do not understand the sign language and allows the communication in the natural way of speaking. The proposed module is an interactive application module developed using Python and its Advanced Libraries. This module uses inbuilt camera of the system to get the images and perform analysis on those images to predict the meaning of that gesture and provide the output as text on screen and speech through speaker of the system that makes this module very much cost effective. This module recognizes one handed ASL gestures of alphabets (A-Z) with highly consistent, fairly high precision and accuracy.*

***Keywords**: ASL, Sign Language, Gesture, Image, Communication.*

I. INTRODUCTION

Sign Language is the natural way of communication for hearing and speech impaired person. There are different Sign Languages used all over the world as per the language in which the gestures or signs are being translated such as American Sign Language (ASL), Australian Sign Language (Auslan), Japanese Sign Language etc. A sign or gesture consists of actions of one or both hands, along with facial expression, which represents a specific meaning. There is a large communication gap between speech impaired persons and the whole world due to which they feel left out in our society and fails to integrate into educational, social and work environments.

Manuscript received on May 25, 2020.
Revised Manuscript received on June 29, 2020.
Manuscript published on July 30, 2020.
***** Correspondence Author
Prof. (Dr.) Neelam Shrivastava, Professor, Department of Computer Science and **Engineering**, Meerut Institute of Engineering & Technology, Meerut, India.
Kopal Garg, Student, Department of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, India.
Abhishekh Pratap Singh, Student, Department of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, India.
Abhishek Jain, Student, Department of Computer Science and Engineering, **Meerut** Institute of Engineering and Technology, Meerut, India.
Abhishek Sharma, Student, Department of Computer Science and Engineering, Meerut Institute of Engineering and Technology, Meerut, India.

The overall goal of this project is to minimize that communication gap through technology that recognize and translates sign language to text and speech. To paper presents an interactive module for automatic translation of American sign language into text and speech in English to assist the speech impaired person to communicate with others.
This sign language translator is able to translate alphabets (A-Z) to speech in English. Although facial expressions add important information to the emotional aspect of the signs but in this project work, they are excluded from the area of interest, since its analysis will complicate the already difficult problem.

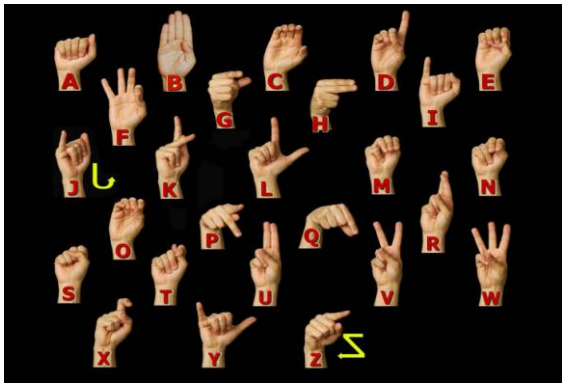


Fig 1. ASL gestures for Alphabets

II. METHODOLOGY

We present a real-time vision-based module for recognizing finger spelling continuous American Sign Language (ASL)
using a single camera to track the user's hand gestures.
This module consists of four sub parts such as: -

1. Image Capturing: here, the colored image is being captured by making use of an inbuilt camera in the system.
2. Image processing : here, the captured images are processed and converted to the required format and also extract the part of interest out of the whole image.
3. Gesture Recognition : here, the extracted part of the image is further processed in order to recognize the hand gestures in the image and make predictions on that gesture by mapping or matching it with the provided image databases corresponding to each ASL gesture.
4. Speech Conversion : here, the predicted gesture is converted to speech in English.

A.Image Capturing

The images for our Talking Hands System are captured by the inbuilt camera. The Camera is used in Video recording Mode in order to recognize the gesture in the real-time and making predictions accordingly. Capturing 30 frames per second (fps) is also sufficient as higher fps will only increase time for predict and complexity of the module. Due to increase in input data to be processed. As the images are captured at real time, so it must be efficient. The data of previously Processed frame will be deleted automatically to free the limited memory space in the Buffer.

ENVIRONMENTAL SET-UP:

Image Capturing process depends on many environmental factors such as the position of the camera or hand, lighting conditions and background conditions. The camera must be placed at sufficient distance from the user, the hand must be placed at particular position in order to capture the correct gesture. Sufficient lighting must. Be ensured to get the bright enough image to recognize the gesture easily and make more correct and efficient predictions. The more accurate result and gives the quite accurate in different lighting conditions.

B. Image Processing

Captured images are processed by using Python Open Computerized Vision 2 Library to identify the unique feature of each gesture. There are different steps that are to be performed before recognizing the gesture. Here, color thresholding method will be used to convert the colored image to YCbCr image in order to eliminate any irrelevant object in the background that may interfere in the gesture recognition and to differentiate the hand from the image due to its skin color. camera captures images at 30 frames per second (fps) due to which the difference between subsequent images will be too small. so, the images are sampled at 5 frames per second (fps). In the module, each frame is buffered and numbered sequentially.

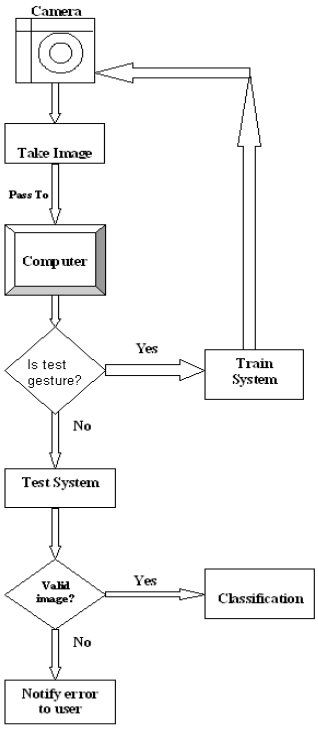


Fig 2: Flow Chart

C. Gesture Recognition

After color thresholding, the extracted images are analyzed or processed to obtain the unique features of each gesture. Each gesture has a unique combination of fingers position. This feature will help in recognition of the gesture and mapping it with the database. The recognized gesture is predicted on the basis of threshold value. The accuracy of the prediction mostly depends on the database provided to the system for training. For better results the sufficient and correct database must be provided to the system.

D.Speech Conversion

After color thresholding, the extracted images are analyzed or processed to obtain the unique features of each gesture. Each gesture has a unique combination of fingers position. This feature will help in recognition of the gesture and mapping it with the database. The recognized gesture is predicted on the basis of threshold value. The accuracy of the prediction mostly depends on the database provided to the system for training. For better results the sufficient and correct database must be provided to the system.

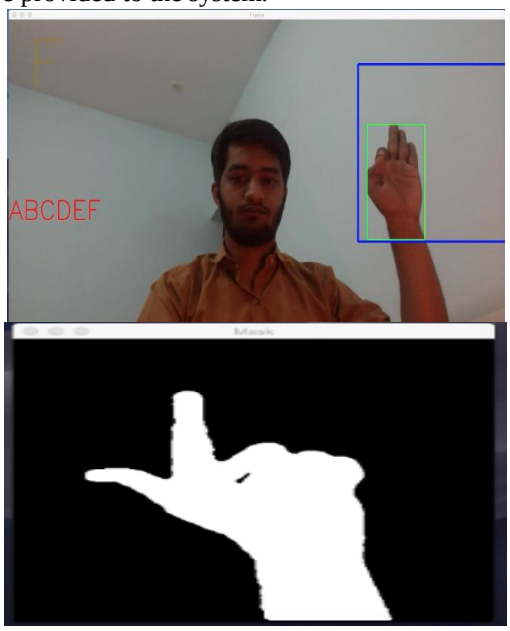
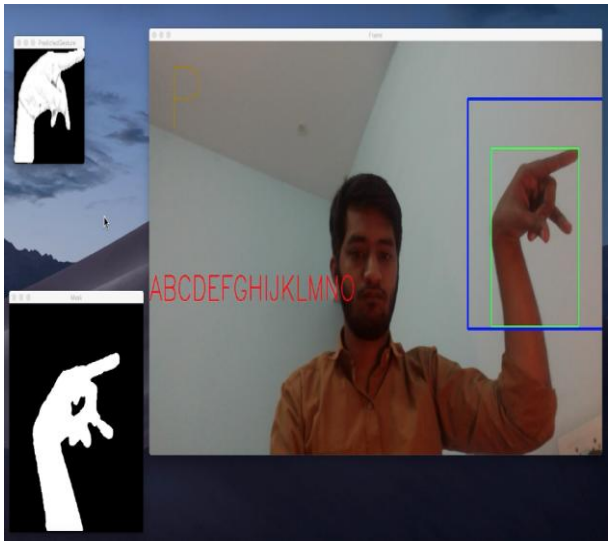


Fig. 3: Background Removal

III. RESULT

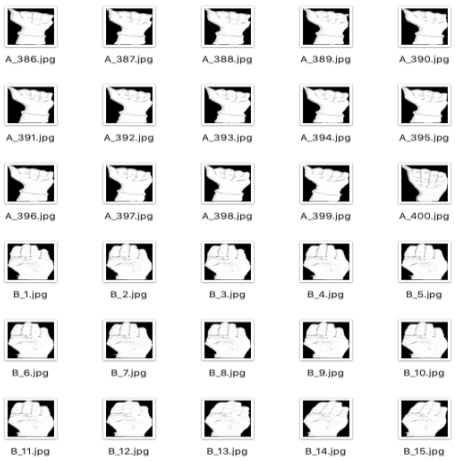
This Talking hands or sign to speech converter module is able to translate hand gestures to their respected alphabets (A-Z) with the sufficient accuracy. All the gestures can be recognized in real-time. But some hand positions may appear to be similar in posture and gesture to another alphabet's sign then that can be misinterpreted, resulting in decrease of accuracy of the system. The current system has been trained on a very small and limited database. Since there will always be variations in the hand postures and gestures from its required posture or gesture, then a larger database accommodating a larger variety of hand postures and gestures is required in order to predict more accurate sign.



IV. PERFORMANCE IMPROVEMENTS

1. To increase the performance and accuracy of the American Sign Language Translator (ASLT), the quality of images used in the training database must be enhanced to ensure that it recognizes the correct and significant characteristics in every individual sign and further improve the performance more efficiently.
2. A larger dataset will also enhance the performance in different environments, increase the robustness of the system and provide training examples for a wider variety of situations.
3. Adaptive color models and improved tracking could boost the performance of this system.
4. The ASL recognition system is one of the components of a larger system that can be hoped one day, to be an active tool for the hearing and speech impaired human beings.
5. Creating Dataset:

   This module recognizes one handed ASL gestures of alphabets (A-Z) with highly consistent, fairly high precision and accuracy. Initially, we have focused on only alphabets so that their efficient conversion would be done at faster rate. For each symbol we have trained our machine with 400 images in the dataset from different angles and under varied lightning conditions. The intensity should also be set to provide sufficient light in the camera. However, since this system is intended to be used by the consumer it would be a disadvantage if special lighting equipment were required. It was decided to attempt to extract the hand information using standard room lighting. This would permit the system to be used in a non-specialist environment. For creation of the dataset, the python code would be run under required condition, successful running of the below code would activate the camera and it can capture images at 30 frames per second (fps) due to which the difference between subsequent images will be too small. so, the images are sampled at 5 frames per second (fps). In the module, each frame is buffered and numbered sequentially.



V.EXPERIMENTAL RESULT

There are certain reasons which resulted in the adoption of this training method. This is a supervised learning algorithm. User need to train system first and then try to recognized gesture. Every gesture at least once user should have to give system for training process. The detection rate of the system achieved through this algorithm was 86%. For each type of gesture, multiple images were given to the system for training. After every training process system were tested 20 times. At the end, the system was tested with 20 images of each kind of gesture.

Here are some Results:

Input Frames	Frames Detected	Output Given
2329	1600	750
1623	759	614
1197	589	231

Table 1. Output Time Detection

VI. CONCLUSION

In this work, a sign recognition system or sign to text and speech converter is developed by using Python and its advanced libraries such as NumPy, opencv2, etc. it uses the feature vectors included in whole image frames containing all the aspects of the sign. This module has investigated the different issues of this approach to American sign language recognition using appearance-based features which are extracted from a video stream recorded with an inbuilt camera in real time. Although sign language contains many different aspects such as the position, the orientation and the shape of the hand of the user that conveys a large portion of the information of the signs.

Therefore, the different features which are extracted from the user's hand may improve the accuracy of the system to a great extent.

REFERENCES

1. Andreas Domingo, Rini Akmeliawati, Kuang Ye Chow 'Pattern Matching for Automatic Sign Language Translation System using LabVIEW', International Conference on Intelligent and Advanced Systems 2007.

2. Beifang Yi Dr. Frederick C. Harris 'A Framework for a Sign Language Interfacing System', A dissertation submitted in partial fulllment of the requirements for the degree of Doctor of Philosophy in Computer Science and Engineering May 2006 University of Nevada, Reno.
3. Jose L. Hernandez-Rebollar1, Nicholas Kyriakopoulos1, Robert W. Lindeman2 'A New Instrumented Approach For Translating American Sign Language Into Sound And Text', Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition (FGR'04) 0-7695- 2122-3/04 $ 20.00 © 2004 IEEE.
4. Helene Brashear & Thad Starner 'Using multiple sensors for Mobile Sign Recognition'. ETH- Swiss Federal Institute of Technology, Zurich,Switzerland.

AUTHORS PROFILE

Prof. (Dr.) Neelam Shrivastava is currently holding the position of Professor in Meerut Institute of Engineering &Technology, Meerut. She has more than 20 years of teaching as well as industry experience. She has been leading the way of graduates and post-graduates since years on account of her adeptness in teaching. Her hub of knowledge comprises SAP (Systems Applications and Products in Data Processing) - especially ABAP (Advanced Business Application Programming), Big Data, Machine Learning, Image Processing, ERP (Enterprise Resource Planning), Design and Analysis of Algorithm. Her publications also include many International and National papers. She is an influential person in Technology.

Kopal Garg is currently a final year Engineering student at Meerut Institute of Engineering and Technology from Computer Science and Engineering department. She has worked for more than two years in Image Processing projects and actively contributed in this project. Apart from this she has good command on almost all major programming languages including C and Python.

Abhishekh Pratap Singh is in final year of his Computer Science Engineering from Meerut Institute of Engineering &Technology, Meerut. Before doing this project he also worked on other Machine Learning and Image Processing projects which includes Driver Drowsiness Detection system and similar projects. He is also proficient in C and Python coding which he used in this project as well. Along with technical stuff he is also quite interested in management tasks. He has done training and internships in Web Development and Database Management as well.

Abhishek Jain is currently pursuing his Engineer Degree in Computer Science from Meerut Institute of Engineering & Technology, Meerut. He has work experience of his field as an intern in some startups such as Gurutva (Meerut) and Bitcs (Gurugram). In this changing world, he likes to face new and different challenges and keep himself updated with the new technologies. He has expertise in Programming with all four major languages such as C, C++, Python and Java. He has worked in the field of AI and has explored Machine Learning, Data Analysis and Image Processing as well as in the field of Web Development, he has explored different frameworks of Javascript such as React js, Next js,Meteor js,etc. He keeps himself positive and motivated towards his work and challenges.

Abhishek Sharma, belongs to Muzaffarnagar Uttar Pradesh. Currently pursuing Engineering in Computer Science and Engineering from Meerut Institute of Engineering and Technology .
He like to spend time to explore more about What is currently going on in technical area .In computer science interested areas are Data Structure, Web Developement, Database Management System .
He worked on different technologies and have a Good experience in Java, PHP, Web Development.