Contents

Sign language recognition using image-based hand gesture recognition techniques

January 7, 2020

1 Introduction/Motivation

Sign language is one of the effective communication tools for the people who are not able to speak or hear anything. It is also useful for the person who can speak but not able to hear or vice versa. Sign language is boon for the deaf and dumb people. Sign language is the combination of different gesture, shape and movement of hand, body and facial expression. Deaf people use these different gestures to represent their thoughts. Each gesture or movement of hand, facial expression and body movement has special assigned meaning. Hand gestures are the motions of the hands to express ideas or to generally communicate. Therefore, hand gesture recognition is the interpretation of ideas expressed visually by the hands. A good example of hand gestures and their recognition is American Sign Language. We daily see that deaf and dumb people communicate with those who know their sign language, and they can't express their feelings to those who don't know the sign language. So, this is the problem we face every day. So, we created the project which helps us to understand the sign of the deaf and dumb people.

2 Market Research/Literature Survey

2.1 Cost estimation

Since our project is based on software platform, our requirements are mainly a laptop, web camera and a tool to implement, that is Python. Therefore, our budget would be less than 1000 (INR) (Since no heavy processor or machinery is required). Although in order to make a prototype we require raspberry pi, web camera, python language, which may add up to within 3000(INR).

2.2 Market survey

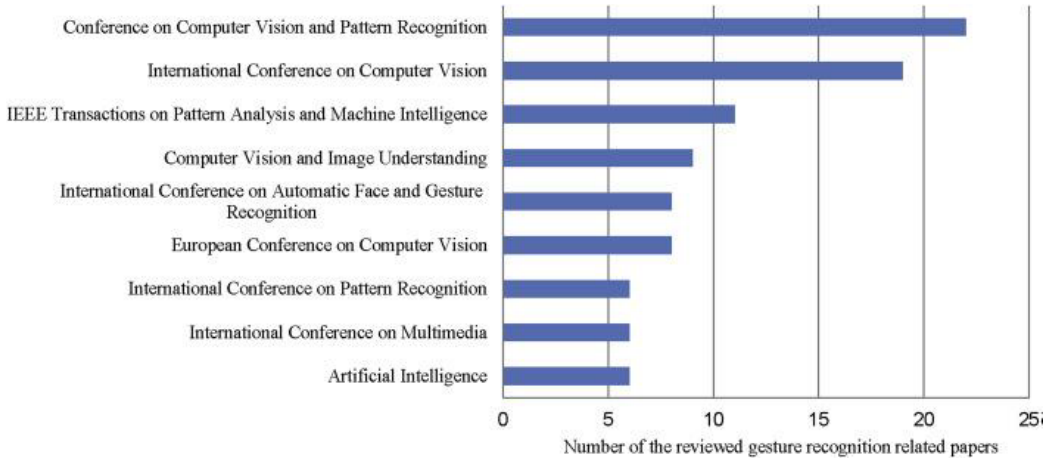

Figure 1: Number of the reviewed gesture recognition related papers

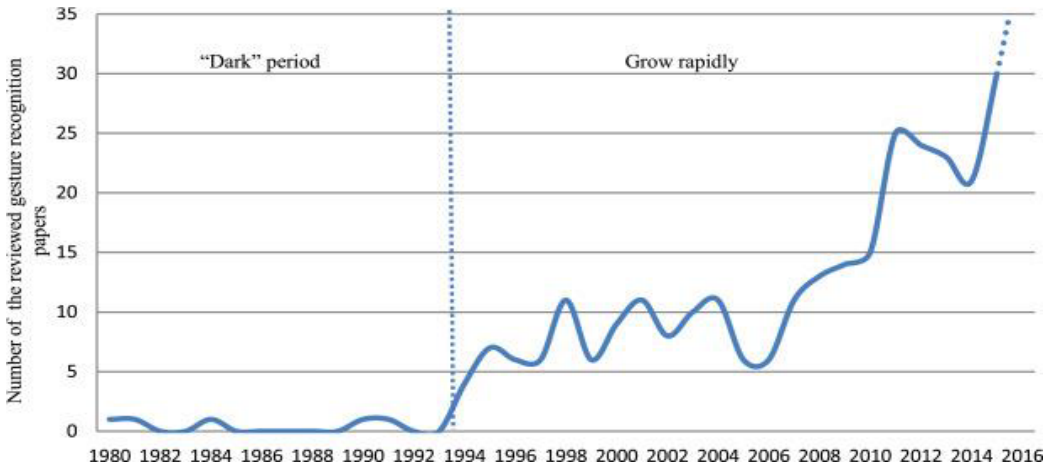

Figure 2: Yearly distribution of the reviewed gesture recognition related papers

Above figures presents a brief statistical analysis of gesture recognition technologies. As shown in Figure 1, we selected 9 different journals and conferences that each published more than 5 papers within the 285 reviewed papers related to gesture recognition. Regarding the number of papers, 65% are conferences papers, 35% are journal papers. Note that it is a common practice in computer science to publish extensively in conference proceedings. The most popular conference that published gesture recognition related papers is Conference on Computer Vision and Pattern Recognition. Figure 5 shows the yearly distributions of the reviewed gesture recognition papers. The number of gesture recognition papers has increased rapidly since 1994, indicating the growing interests in this field.

3 Hardware requirements

- Web Camera:

A webcam is a video camera that feeds or streams an image or video in real time to or through a computer to a computer network, such as the Internet. Webcams are typically small cameras that sit on a desk, attach to a user's monitor, or are built into the hardware. Webcams can be used during a video chat session involving two or more people, with conversations that include live audio and video.

4 Software requirements

- Python IDLE with video and image processing modules:

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms and can be freely distributed.

- Python IDLE with video and image processing modules:

1. **Scikit image:** Scikit-image is an open source Python package that works with NumPy arrays. It implements algorithms and utilities for use in research, education and industry applications. It is a simple and straightforward library even for those who are new to Python's ecosystem. This code is of high-quality and peer-reviewed, written by an active community of volunteers.

2. **NumPy:** NumPy is one of the core libraries in python programming and provides support for arrays. An image is essentially a standard NumPy array containing pixels of data points. Therefore, by using basic NumPy operations, such as slicing, masking and fancy indexing, we can modify the pixel values of an image. The image can be loaded using skimage and displayed using matplotlib.

3. **Scipy:** SciPy is another of Python's core scientific modules like NumPy and can be used for basic image manipulation and processing tasks. In particular, the submodule scipy.ndimage provides functions operating on n-dimensional NumPy arrays. The package currently includes functions for linear and non-linear filtering, binary morphology, B-spline interpolation, and object measurements.

4. **Opencv python:** OpenCV (Open Source Computer Vision Library) is one of the most widely used libraries for computer vision applications. OpenCV-Python is the python API for OpenCV. OpenCV-Python is not only fast since the background consists of code written in C/C++ but is also easy to code and deploy (due to the Python wrapper in foreground). This makes it a great choice to perform computationally intensive computer vision programs.

5 Implementation

Capturing the hand gesture of deaf and dumb people by camera and processing the video by converting it into frames and then processing the frames using the image processing methods and comparing the processed images with the saved database images and printing the appropriate message saved with highest compared images.

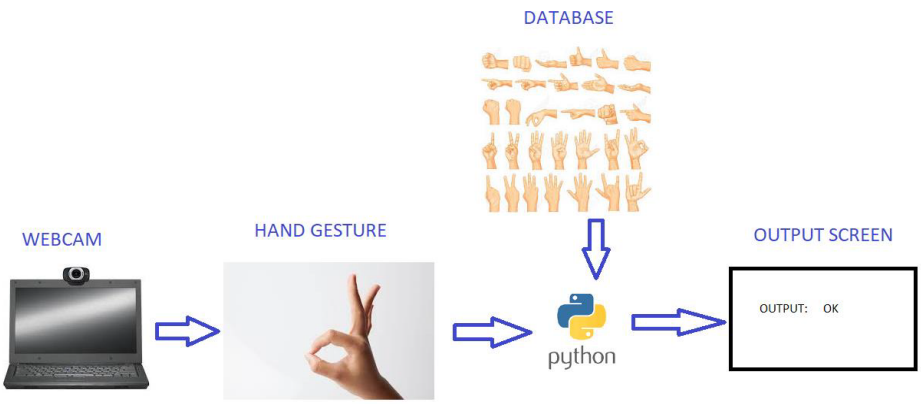

Figure 3: Flow diagram

- Video processing
- Image processing
- Image comparison between processed image and images stored in database
- First the video of the hand of deaf and dumb people is captured for a short duration of the time. He can press any key to stop the recording as an interrupt.
- Now the video is sent for processing which is converted into frames with the help of OpenCV library.
- The images are processed and the image irregularities such as blur, contrast, resize and normalization is done.

6 Feasibility:

Here we are using the video and image processing methods using python modules, so it is very effective method and highly efficient. If you go for the neural network and machine learning methods, it takes more execution time and more processing time. Since python takes more time for execution of complex programs, so it feasible to go for image processing and image comparison using simple image and video processing methods. Since we are using the web camera for capturing the hand gesture of deaf and dumb people, so we can capture the video of hand gesture of deaf and dumb people in wide angle and it covers the large portion and it helps to get better video quality hence it decreases the video and image processing time. In some research papers using the Glove based Hand Gesture Recognition, Colour Marker Based Hand Gesture Recognition, 3-D model Based Approach for Hand Gesture methods which cost effective and takes more compilation, processing and rum time. Since we are using video and image processing which ineffective and takes less compile, processing and rum time.

7 Result

Here we are using the video and image processing methods using python modules, so it is very effective method and highly efficient. If you go for the neural network and machine learning methods, it takes more execution time and more processing time. Since python takes more time for execution of complex programs, so it feasible to go for image processing and image comparison using simple image and video processing methods. Since we are using the web camera for capturing the hand gesture of deaf and dumb people, so we can capture the video of hand gesture of deaf and dumb people in wide angle and it covers the large portion and it helps to get better video quality hence it decreases the video and image processing time. In some research papers using the Glove based Hand Gesture Recognition, Colour Marker Based Hand Gesture Recognition, 3-D model Based Approach for Hand Gesture methods which cost effective and takes more compilation, processing and rum time. Since we are using video and image processing which ineffective and takes less compile, processing and rum time.

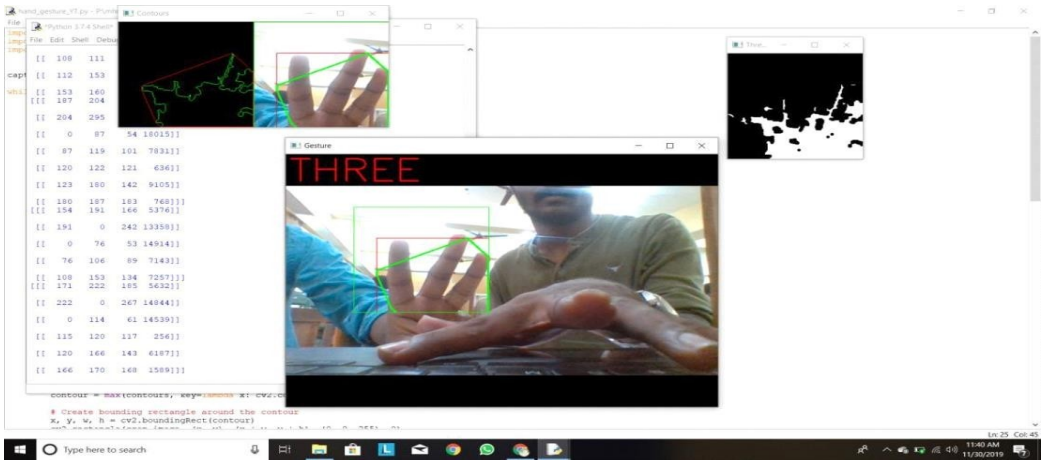

Figure 4: Hand gesture converted into corresponding number

Here we saved hand image of deaf and dumb people in database and by processing the image of hand gesture, we displayed the meaning of the hand gesture. Below image shows that, we converted hand symbol of number five into corresponding text and speech form. By finding number of edges of convex hull, we are finding the number to be displayed.

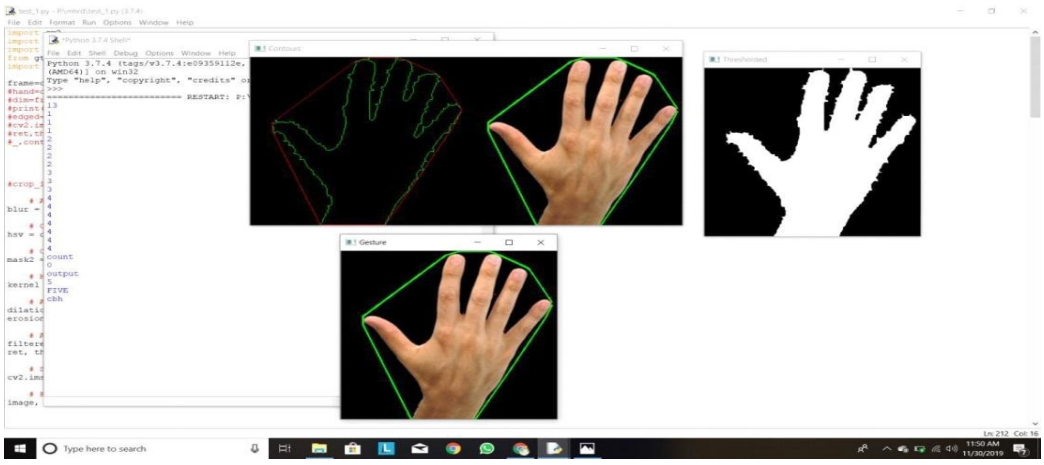

Figure 5: Database image converted into corresponding meaning

Below image shows that, converted hand symbol of alphabet B into corresponding text and speech form by counting the number of convexity defaults. If convexity defaults are 13, it indicates alphabet 'B', convexity defaults are 17 indicates alphabet 'A' and so on.

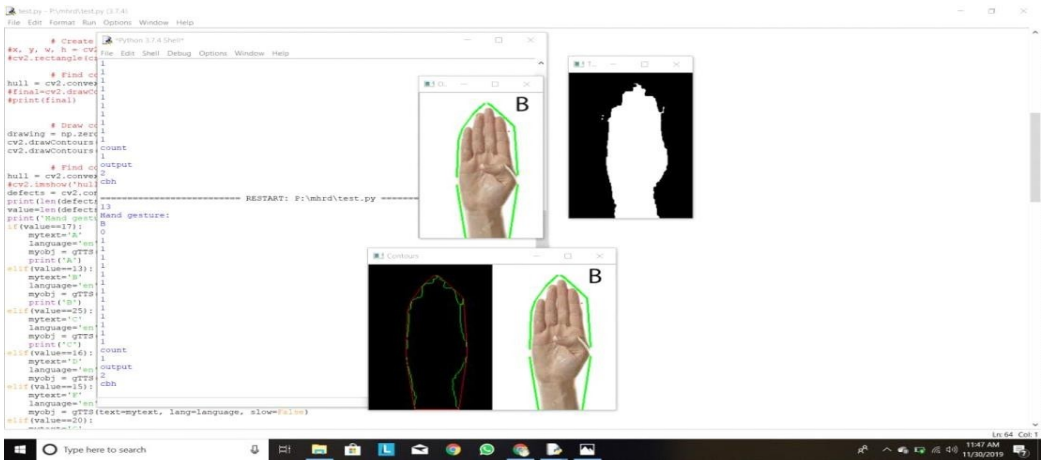

Figure 6: Hand gesture converted into corresponding alphabet

8 Conclusion

Our project aims to bridge the gap by introducing an inexpensive computer in the communication path so that the sign language can be automatically captured, recognized and translated to speech for the benefit of blind people. In the other direction, speech must be analyzed and converted to either sign or textual display on the screen for the benefit of the hearing impaired.

References

[1] Research paper : Finger detection for sign language recognition
`https://www.researchgate.net/publication/44259601_Finger_Detection_for_Sign_Language_Recognition`

[2] IEEE published paper: Sign language recognition using image-based hand gesture recognition techniques
`https://ieeexplore.ieee.org/document/7916786`

[3] Digital Image Processing (4th Edition) written by Rafael C. Gonzalez, Richard E. Woods