

Sign Language Interpretation and Conversion to Text

Mrs. Pallavi Ghaste, Mrs.Soniya Bastwade, Rajvi Khandelwal, Shubhangi Ambapkar, Ziaurrehman Ansari and Chirag Singh Salaria.

Department of Computer Engineering
DYPCOE Akurdi, Pune

ABSTRACT

In our day to day life, communication plays important role for conveying information from one person to another person. But it becomes very difficult for the people who are deaf and dumb to communicate with normal people. Sign language is the only one way to communicate with them. But normal people are unaware of sign language. So there is only one way and that is to covert sign language into text & speech & vice versa. That is known as sign recognition. Sign language is a combination of body languages, hand gestures and facial expressions. Among those hand gestures are provides majority of the information

In this paper we will discuss the method proposed by author for Sign Language Interpretation and its conversion to text.

Index Terms- Pattern Recognition, Media pipe Framework, Image Pre-Processing, Feature Extraction.

I. INTRODUCTION

Over 5% of the total population on Planet are deaf people. 5% may seem like a small number, but that totals over 360 million-! Being specially abled of this type, it is very difficult to communicate with ordinary people. To help them communicate among themselves and others, there is a sign language through which they communicate. However, ordinary people find it hard to learn this language and hence, interaction with specially abled people is difficult.

To overcome this problem, several methods had been proposed. However, some are expensive and others rather inaccurate. Here we propose an accurate, affordable and easily accessible method to solve the problem. In the proposed method, we make use of the image recognition and deep learning, to make machine act as a translator. Through camera image of the sign alphabet will be

captured and the interface (machine) will output the sign alphabet in text.

II. LITERATURE REVIEW ON DIFFERENT SYSTEMS.

The hand gesture recognition system builds a natural, interactive and the most favorable environment for the Human-Computer Interaction. Due to its utmost importance in the present scenario, and more importantly in the future, many techniques have been thought about, developed and implemented till now. Mostly sign language recognition is done by two approaches: Image based [1].sensor-based approach.

Truong *et al.* [2] presents a work on applying AdaBoost and Haarlike classifiers for recognizing American Sign Language with complex background from live videos and translates into text and speech. The key factor of the high accuracy result was the big dataset in the training process.

Fuet *et al.* [3] proposed a wavelet enhanced image preprocessing and neural networks for hand gesture recognition approach. Collected images are processed to extract the contour of the hand gestures also provide the neural networks for hand gesture learning with wavelet-based feature extraction.

Since many proposed methods does not works on geometric properties and require any extracted features. Vishwakarma *et al.* [4] implemented method which works on geometric properties and does not require any extracted features also works well on low-resolution images which is a great advantage for developers regarding data storage.

Vishwakarma *et al.* [5] presents a novel hand gesture recognition system, for the efficient human-computer interaction. Also proposed algorithm for hand gesture recognition in the complex background in proposed based on shape and texture feature.

Kumar et al. [6] proposed an idea for feasible communication between hearing impaired and normal person with the help of machine learning approach. Instead of preprocessing techniques, like filtering and segmentation of hand patch images, features are extracted from dataset by discrete wavelet transform.

Anup Kumar et al. [7] propose the design for a basic yet extensible system that is able to recognize static and dynamic gestures of American Sign Language. Mapping of the spoken alphabet to text is done by proposed algorithm.

III. THE PROPOSED SYSTEM

Communication plays a major role in human life, but deaf and dumb people find it difficult to communicate with normal people. To minimize this communication gap, we introduce a handgesture recognition system to recognize the alphabets of Sign Language. We propose an easy-to-use and inexpensive approach to recognize single and double handed gestures accurately. We intend to develop the said project in total of three modules namely:

1. Image Input Module:

This Module will consist of capturing the image of hand to recognize the gesture within the given resolution from camera.

2. Image Processing Module:

This module will give 21 3D key points of the given image of hand. It will be essentially a hand tracking module implemented using Media pipe framework.

3. Classification Module:

With provided points from the Image Processing Module, the classifier will decide the sign alphabet by taking reference from the training set by the help of classification algorithm. The output will be generated as text corresponding to the Sign alphabet decided by Classifier.

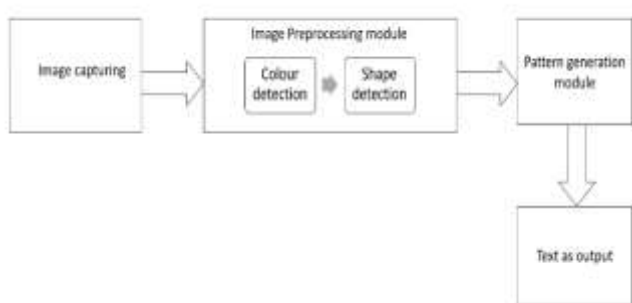


Fig 1. Architecture of the Proposed System.

Fig1 demonstrates the working of sign language interpretation system. Initially the image is captured which is then pre-pre-processed. The pre-processing consists of colour detection and shape detection module. The processed image is then given as input to pattern generation module further to which the generated pattern is converted to its equivalent text.

IV. ALGORITHM AND MATHEMATICAL MODEL

I. Sign Language Detection.

1. Capture image.
2. Retrieve key points of the hand.
3. Calculate distances between the focus of the given image and the clusters of different sign alphabets.
4. Whichever is nearest to the given image, give it as an output text.

CNN Algorithm:

Convolution operation: Prepare feature detector for the input image. Feature detector will be images of sign language. Make a corresponding feature map for it.

Rectified Linear Unit (ReLU): Apply Rectifier Function to increase the non-linearity of image.

Pooling: Apply max pooling on the image to make the network realize whichever the condition of the image, CNN must be able to identify it. Roughly 75% of information is filtered after this step.

Flattening: Flatten the 2*2 matrix to single dimensional array.

Full connection step : After Flattening feed the flattened array to the fully connected trained network and whichever class has the highest value as output, return the label of that class. (Here class represents alphabets of sign language) Mathematical model for Hand Gesture Recognition System as follows:

Let us consider S be the system such that

$$S = \{I, F, O\}$$

Where,

I = Set of Input

as a images captured by camera or by video such as raw video.

F = Set of Functions

$F = \{I, TD, IF, SS, G\}$

where,

I = Input images extracted from video or captured by camera.

$TD = \{T1, T2, T3...Tn | T \text{ is the train data used to compare the captured images.}$

$IF = \{IF1, IF2, IF3... IFn | IF \text{ is the important frames calculated by score function}$

$$F[V, S, w] = w^t f(V, S)$$

where F is the score function, V and S stands for the feature matrix of images and its candidate summary, w is the property weight

SS = Storage server

SS is used to store the extracted imported frames.

G = Generate the recognized images of Hand Gesture.

O = Output

Textual or Verbal result

V. ACTIVITIES AND EVENTS:

EVENT 1

User will make registration on SYSTEM and Storage Server. Let $f(U)$ be a function of User.

Thus, $f(U) \rightarrow \{U Ds\}$

EVENT 2

Let $f(C)$ be a capture function of System.

Thus, $f(C) \rightarrow \{I1, I2, I3In\} \in C$

EVENT 3

Pre Processing is performed on detected images.

Let $f(I)$ be a image processing function.

Thus, $f(I) \rightarrow \{I1, I2, I3In\} \in T$

EVENT 4

Let $f(FT)$ be an function used to extract the features on basis of train dataset

Thus, $f(FT) \rightarrow \{T, I\} \in FT$

EVENT5

Let $f(E)$ be the Character Detection on extracted features using CNN .

Thus, $f(E) \rightarrow \{FT\} \in T$

FLOWCHART

Object Detection using media pipe.

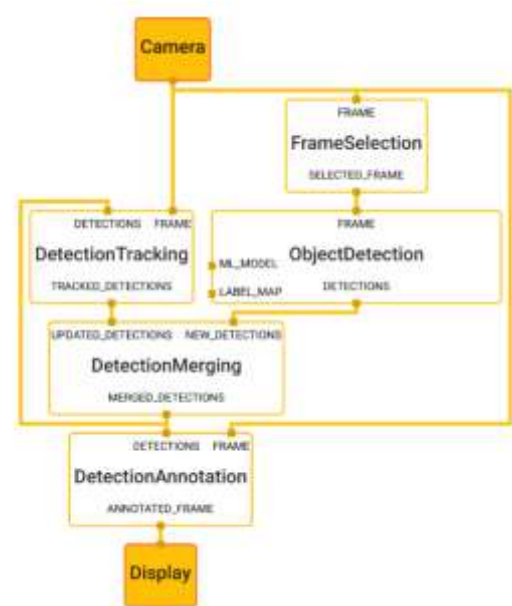


Fig 2. Object Detection using mediapipe.[8]

VI. CONCLUSION

In this paper we have proposed a sign language interpretation system and its conversion to text for deaf and dumb people. The system will make use of camera to capture hand gestures further media pipe framework will detect the hand gesture and the this gesture will be classified by matching it with the sample dataset which is used for the conversion of gesture into its equivalent text.

VII. REFERENCES

- [1] Umang Patel, Arti G Ambekar, "Moment Based Sign Language Recognition for Indian Languages" Third International Conference on Computing, Communication, Control and Automation (ICCUBEA), 2017
- [2] Xingang Fu, Ting Zhang, Chadwell H. Bonair, and Marvin L. Coats, "Wavelet Enhanced Image Preprocessing and Neural Networks for Hand Gesture Recognition" The University of Alabama Tuscaloosa, Alabama, USA, 2015

[3] D.K. Vishwakarma, Sahib Majithiax and Nikhil Kumar Mishray, "*A Static Hand Gesture Recognition System to Recognize the Total number of Fingers*" Department of Electronics and Communication Engineering, Delhi Technological University New Delhi, India, 2017

[4] Dinesh Kumar Vishwakarma, "*Hand Gesture Recognition using Shape and Texture evidences in Complex Background*" Department of Electronics and Communication Engineering, Delhi Technological University Delhi, India, 2017.

[5] Naresh Kumar, "*Sign Language Recognition for Hearing Impaired People based on Hands Symbols Classification*" PhD Research Scholar Department of

Mathematics Indian Institute of Technology, Roorkee India, 2017.

[6] Anup Kumar et al "*Sign language recognition*" 3rd International Conference on Recent Advances in Information Technology (RAIT), 2016.

[7] Anup Kumar et al "*Sign language recognition*" 3rd International Conference on Recent Advances in Information Technology (RAIT), 2016.

[8] Flowchart Available on:
<https://www.groundai.com/project/mediapipe-a-framework-for-building-perception-pipelines/1>