International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 04 | Apr 2019 www.irjet.net p-ISSN: 2395-0072

Sign Language Interpreter using Image Processing and Machine Learning

Omkar Vedak[1], Prasad Zavre[2], Abhijeet Todkar[3], Manoj Patil[4]

[1,2,3]*Student, Department of Computer Engineering, Datta Meghe College of Engineering, Mumbai University, Airoli, India*
[4]*Assistant Professor, Department of Computer Engineering, Datta Meghe College of Engineering, Mumbai University, Airoli, India*

Abstract - *Speech impairment is a disability which affects one's ability to speak and hear. Such individuals use sign language to communicate with other people. Although it is an effective form of communication, there remains a challenge for people who do not understand sign language to communicate with speech impaired people. The aim of this paper is to develop an application which will translate sign language to English in the form of text and audio, thus aiding communication with sign language. The application acquires image data using the webcam of the computer, then it is pre-processed using a combinational algorithm and recognition is done using template matching. The translation in the form of text is then converted to audio. The database used for this system includes 6000 images of English alphabets. We used 4800 images for training and 1200 images for testing. The system produces 88% accuracy.*

Key Words: Pre-processing, Feature Extraction, Edge Detection, Classification.

1. INTRODUCTION

Sign language is an important part of life for deaf and mute people. They rely on it for everyday communication with their peers. A sign language consists of a well-structured code of signs, and gestures, each of which has a particular meaning assigned to it. They have their own grammar and lexicon. It includes a mixture of hand positioning, shapes and movements of the hand.

The people who know sign language can communicate with each other efficiently. However, when it comes to communicating with people who don't understand sign language it causes a lot of problems. Communication is a very important part of our lives. We interact with our mates at offices, schools, hospitals and other public places. Deaf and mute people may find it difficult to express themselves in such situations because not everyone understands sign language. There are many highly talented people suffering from speech impediment. We feel that their disability should become a hindrance to achieve their goals. Adding them into the workforce will only improve the socio-economic development of the country.

Deaf and mute people usually depend on sign language interpreters for communication. However, finding a good interpreter is difficult and often expensive. Thus, a computerized interpreter could be a reliable and cheaper alternative. A system that can translate sign language into plain text or audio can help in real-time communication. It can also be used to provide interactive learning of sign language.

There is no universal sign language for deaf people. Different countries use their own sign language, although there some striking similarities among them. It is yet unclear how many sign languages exist in the world. Some languages have got legal recognition and some have not. India's National Association of Deaf estimates that there are 18 million people in India with hearing impairment. This paper discusses the implementation of a system which translates Indian Sign Language gestures to its English language interpretation.

2. LITERATURE SURVEY

Several types of researches have been done in translating Indian Sign language using deep learning. Some of them used instrument-based approach and some have used a video-based approach.

In Ref. [1] Pham The Hai uses Microsoft Kinect to translate Vietnamese Sign Language. In the proposed system, the person has to place himself with Kinect's field of view and then perform sign language gestures. It can recognize both static and dynamic gestures using multiclass Support Vector Machine. During recognition, the gesture features are extracted, normalized and filtered out based on Euclidean distance.

Purva Badhe [2] uses Fourier Descriptor for feature extraction. The system translates Indian Sign language gestures to English language. To represent the boundary points, the Fourier Series were calculated using Fast Fourier Transform (FFT) algorithm. The extracted data being too large is compressed using vector quantization. This data is then stored into a codebook. For testing purpose, the code vector generated from gestures is compared with existing codebook and gesture is recognized.

Azher Uddin [3] has proposed a system for Bangla sign language interpretation. It first converts the RGB image to HSV color space. It used Gabor filters to acquire hand features and Kernel PCA for dimension reduction of the feature vector. Finally, it uses Support Vector Machine for classification of candidate features

Muhammad Aminur Rahaman [4] uses a novel method for Bengali hand gesture recognition. The system uses cascaded classifiers to detect the hand in each frame. It captures hand gestures based on Hue and Saturation value from HIS color model. Then the images are classified using K-Nearest Neighbours Classifier.

In [5] the authors have compared two models, one using PCA and other using LDA. The paper concludes that the Neural Network Classification works well with LDA algorithm for translating Bangla Sign language.

P. Gajalakshmi [6] uses Support Vector Machine and Error Correcting Output Codes for American Sign Language recognition.

3. METHODOLOGY

The proposed system consists of two main stages: (1) Segmentation of hand (2) Recognition of hand sign. The block diagram shows the working of the proposed system. The features of a hand are an important criterion for the classifier to differentiate between the hand gestures. These characteristics must be able to adapt to different hand and gestures by different people. In this system, we have used histograms of oriented gradients (HOG) as a feature descriptor. It is better than other descriptors because it can adapt to changing illuminations and rotation of objectives. It does not consider an image as a whole, but divides it into smaller cells and then for the pixels within the cells edge or gradient direction histogram is calculated. This approach creates a bin, and clubs the histograms of different samples based on magnitude and angle. In the proposed system, we are first segmenting the hand using YCbCr color space and then processing the image through HOG and then provide it to the model. We trained the SVM classifier using 5000 images and developed a model.

3.1 Preparing the dataset

We have created a dataset of 26 English language alphabets in Indian Sign Language. Each sign gesture is performed by two different individuals with different hand structure in varied lighting conditions. The videos were recorded on a camera and then each video was broken down frame by frame to images and adjusted to 100 frames and then augmented to get about 250 images for each sign. The data was then divided into 4800 images for training and 1200 images for testing.

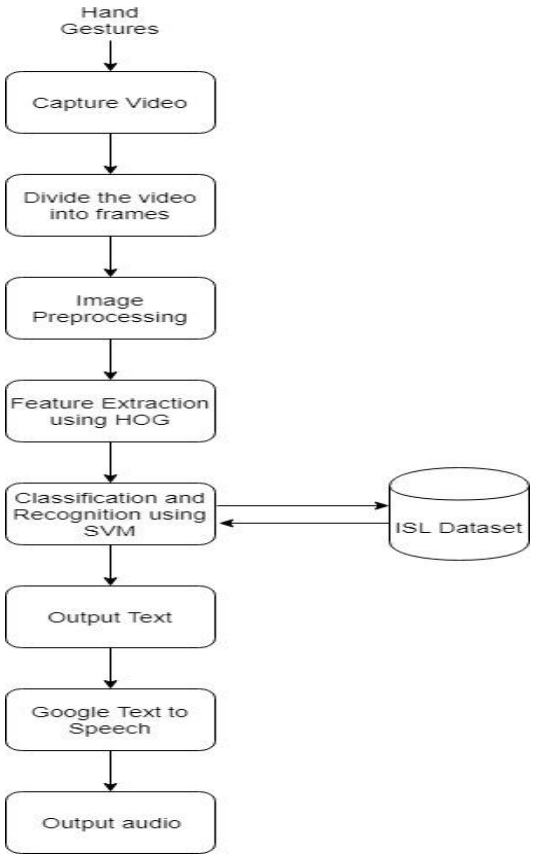

Fig- 1: Sign language interpreter flowchart

3.2 Image preprocessing and edge detection

Before performing feature extraction, the images must be processed in such a way that only the useful information is considered and the redundant, distracting noise and superficial data are neglected. The images are first converted to 100 * 100 pixel size for faster computations. The image is then converted to grayscale and finally transform into a binary image. Simultaneously, skin color is detected using YCbCr model. Finally, edge detection is performed using Canny edge detector. The process is illustrated in figure below.

(a) (b) (c)

Fig- 2: Output of skin colour detection: a) Original cropped image, b) Grey scale converted image, c) Skin colour detection output

3.3 Feature extraction

Extracting features from the data is a critical part of any object detection system. It can be performed using various methods such as Fourier Descriptor, Scale Invariant Feature Transform (SIFT) or Principle Component Analysis (PCA). Histogram of Oriented Gradients (HOG) is another successful method of feature extraction. In this paper, we have used the HOG method for feature extraction. The basic idea behind HOG is that an object or shape within an image can be represented by intensity distribution gradients or edge directions. In this method, the image is divided into small cells and a histogram of gradient is calculated for each cell. It creates a bin and combines the histogram of different samples based on magnitude and angle. For improving the accuracy, all the cells in the image are normalized so that they are not affected by variations in lighting. Finally, HOG feature vector is calculated for the entire image.

3.4 Template matching and sign recognition

The feature vector produced in the above step is fed into an image classifier. In this paper, we have used Support Vector Machine (SVM) for classification. By using SVM classifier we can maximize accuracy and avoid overfitting of data. In SVM data items are plotted in n-dimensional space where n is the number of features. Each feature is associated with a coordinate value. Then it finds a hyperplane that differentiates the classes. The model is saved for real-time sign language recognition.

3.5 Text to Speech

We have used Google's Text to Speech API for transforming the sign language into audio. It is one of the best text to speech API available. Unlike other TTS APIs, this API generates human-like voice. The sign language is interpreted using the above steps and then the result is fed to text to speech function which converts it to audio. In this system, we can see and hear the sign language translation at the same time which makes it very convenient to use.

4. CONCLUSIONS

In this paper, we have explained the implementation of a system which translates Indian Sign Language to English. We have discussed the importance of ISL translator for interacting with the deaf and mute. In this system, still hand image frame is captured using a webcam. These frames are processed to get enhanced features. Then feature extraction and classification algorithms are used to translate the sign language into English text. This translation is converted to speech using text to speech API. The system has implemented using the above algorithms to get the final output. The proposed model is evaluated by a dataset containing 26 signs from two different people. The results show the overall accuracy of the system to be 88%. Our future research will work towards implementing this model on a mobile application.

REFERENCES

[1] Hai, P. T., Thinh, H. C., Van Phuc, B., & Kha, H. H. (2018). Automatic feature extraction for Vietnamese sign language recognition using support vector machine. 2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom).

[2] Purva C. Badhe, Vaishali Kulkarni, "Indian Sign Language Translator using Gesture Recognition Algorithm", 2015 IEEE International Conference on Computer Graphics, Graphics and Information Security (CGVIS).

[3] Md Azher Uddin, Shayhan Ameen Chowdhury, "Hand Sign language recognition for Bangla Alphabet using Support Vector Machines", 2016 International Conference on Innovations in Science Engineerig and Technology (ICISET).

[4] Muhammad Aminur Rahaman, Mahmood Jasim, Md. Haider Ali and Md. Hasanuzzaman, "Real-Time Computer Vision based Bengali Sign Language Recognition", 2014 17th International Conference on Computer and Information Technology (ICCIT).

[5] Rahat Yasir, Riasat Azim Khan, "Two Handed Hand Gesture Recognition for Bangla Sign Language using LDA and ANN", The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014).

[6] P.Gajalaxmi, T, Sree Sharmila, "Sign Language Recognition for Invariant features based on multiclass Support Vector Machine with BeamECOC Optimization", IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI-2017).