Sign Language Translation Using Machine Learning a Survey

[1]Sagar Achiah, [2]Sanjana k Raj, [3]Suhas C J,[4] Suha, [5]Dr Ravi kumar v Professor,

[1,2,3,4,5] *Department Of Computer Science and Engineering*

[1,2,3,4,5]*Vidyavardhaka College of Engineering*

Abstract- Sign language is a system of communication using visual gestures and signs, as used by people that have hearing and speaking disability. It is an incredible advancement that has grown over the years. Unfortunately, there exist some drawbacks that are associated with the sign language. People who can naturally and normally communicate do not know how to interpret the sign language, resulting in a communication barrier. To eradicate this barrier, we need a system that is both versatile and robust in order to convert the sign language to natural language. To overcome from these draw backs we are proposing an approach using gesture recognition and NLP with Machine Learning.

INTRODUCTION

Sign languages are languages that use the visual manual modality to convey meaning. People who have some degree of speaking or hearing disability choose to use sign language. Gestures are the movement of body parts used to convey meaningful information [1]. Gestures have been widely used for communication. Millions of people not being aware of the sign language, the main purpose of this model is to design a solution that enables communication for people with such disability. A recognition system would thus have to identify specifically the hand orientation movements. There are around 300 sign languages around the world [2]. Sign language has its own regional dialects such as American Sign Language, Indian Sign language and British sign language. We propose the design for a basic yet extensible system that is able to recognize static and dynamic gestures of American Sign Language (ASL) and the Indian Sign Language (ISL). ISL is double handed whereas ASL is single handed making ISL more intricate [7]. The existing models referenced in the related works use MATLAB, which is a general programming environment for scientific computing [5]. It is particularly popular amongst academics. Tensorflow is an open-source deep learning library for coding neural networks. The underlying language for working with Tensorflow is Python. Hence, in the model we are planning to build uses Tensorflow to obtain better and faster results compared to the existing models.

We are also planning use Natural language processing (NLP) to convert continuous gestures into a sentence that which a normal person can able to understand for better and easier communication.

RELATED WORK

Meenakshi et.al [1], proposed a hand gesture recognition system, by which gesture made by the user is made known to the system. Through the use of computer vision or machine eye, there is great emphasis on using hand gesture as a substitute of new input modality in broad range applications. They have implemented their approach by utilizing a simple web cam considering 20 fps with 7 mega pixel intensity, using K-means clustering for segmenting the hand object from rest of the background. It can identify around 45 different gestures on the basis of 5 bit binary string resulted as the output. The success rate of the chosen algorithm is about 94% approximately with average computation time of 2 second for recognizing single image in Java Kinect Library and nearest matrix matching algorithm. There is a scope for improving this time by using GPU to process the image in a much faster rate.

Anup Kumar et.al [2], proposed an approach for sign language recognition using SVM and Sphinx, that is able to perform in dynamic and minimally cluttered background with satisfactory result as it relies on skin color segmentation. For speech recognition Sphinx module is used which maps the spoken alphabet to text with high accuracy. This text is then mapped to a picture if it is a static gesture or a video if it is a dynamic gesture. System classifies the gesture as static or dynamic by measuring the distance moved by the hand in subsequent frames. For static gestures, they have used Zernike moments, a well- known shape descriptor in image processing. HSV Segmentation and Finger-tip detection showed satisfactory results in constrained environment, i.e. proper lighting and background with limited skin-colored objects. .Static Gesture recognition was carried out on a lexicon of 24 alphabets (a-y, excluding j) and with approximately 93% accuracy.

Tülay Karayılan et.al [3], proposed an approach for sign language to text by an automated sign language recognition system based on machine learning ANN with back propagation. The system was designed for ASL, with Input Layer was designed to contain 3072 neurons for Raw Features Classifier and 512 neurons for Histogram Features Classifier. Hidden Layer was designed to contain 10 neurons for each classifier. Output Layer had 3 neurons for each classifier. The system gave 70% and 85% accuracy for Raw Features Classifier and Histogram Features Classifier.

When considered other studies, the obtained results are average results. The recognition rate can be increased by improving processing image step as a future work.

Shreyashi Narayan Sawant et.al [4], developed a system to recognize a single handed gestures, accurately bare human hands using a webcam which is MATLAB interface. The aim was to recognize the gestures with highest accuracy and in least possible time and translate the alphabets of Indian Sign Language into corresponding text and voice in a vision based setup. They have considered 260 images, it includes both training and testing set. The images are captured at a resolution of 380×420 pixels. The runtime images for test phase are captured using web camera. Otsu algorithm is used for segmentation purpose and preprocessing was done using PCA is a dimensionality reduction technique based on extracting the desired number of principal components of the multi-dimensional data. The result shows that the accuracy is comparable with those of recent contributions and the proposed method gave output both in voice and text form.

Priyanka C Pankajakshan et.al [5], proposed a system based on MATLAB and ANN which does not involve any complex devices like a glove and is purely a vision based system where a user need not wear any type of cumbersome components for the recognition purpose. The system involves a web camera which is used to capture the image of the hand, a processor for the classification and recognition purpose and an output unit which can be a speaker. Here, the back propagation algorithm is used as a learning/training algorithm to determine the weights of the network. The ANN architecture used here is of a feed forward type. The image is then taken as a frame of 6×7 which makes 42 input to the artificial neural networks input layer. The ANN is then comparing the inputs with the desired outputs and then the output will be between the ranges of 0 to 1. If the output value is above 0.6 the desired output will be obtained. A database of 25 images was created for 5 types of gestures which were captured through the web camera. These were the images which were compared to the real time image to get the identified gesture output voice. The images were normal RGB images. The system proposed in this paper can be implemented in a hardware which supports image processing applications. They can be either raspberry pi or a beagle bone processor.

Pratik H. Suvagiya et.al [6], proposed a system based on Java Kinect Library and nearest matrix matching algorithm. The authors states that though Indian sign language (ISL) translation remained under examination for numerous years, still it is very difficult to implement in real-time systems. The background and brightness disturb the skeleton tracing and make the ISL translation very hard. Microsoft Kinect Xbox 360 is capable of giving in-depth vision image and color vision image of everything in front of it, created on which the skeleton body action can be tracked more precise and easier to get depth coordinate of the skeleton. The system proposed was an ISL translator system which was implemented in Java and Simple Open NI API of Java Kinect library. System cached data through Kinect IR depth sensor and IR Emitter, then nearest matrix matching algorithm was applied which matched with stored data and displayed output in string format as well as audio. The application requires a Kinect device and computer system, USB Kinect adapter. This application does not require any database provider. Text file was used for a data storage. The authors concluded that detection of the signs would require a lot of time, thus making it cumbersome to use, segmentation was also a problem and recognition of dynamic gestures was made easier.

Yellappa Madhuri et.al [7], proposed a vision based technology for recognizing and translating continuous sign-language to text. This paper presents a mobile interactive implemented in a laptop, desktop or an IOS mobile phone to operate with its inbuilt camera, processor and audio-device. The experiments use a 36 word lexicon. This system is broken down into three main parts starting with the image acquisition followed by image processing to extract features for recognition and last comes the recognition stage where signs are identified and audio output is given. The recognition of static sign is done based on the position of the finger in the bounding box. In this work, a vision based sign language recognition system using LABVIEW for automatic sign language translation has been presented. This approach uses the feature vectors which include whole image frames containing all the aspects of the sign. Adaptive color models and improved tracking could boost performance of the vision system.

Mandeep Kaur Ahuja et.al [8], proposed a bare hand gesture recognition system using a database-driven hand gesture recognition based upon skin color model approach and thresholding approach along with an effective template matching which can be effectively used for human robotics applications and similar other applications. The system makes use of Machine Learning and Statistical Procedures. Initially, hand region is segmented by applying skin color model in YCbCr color space. In the next stage Otsu thresholding is applied to separate foreground and background. Finally, template based matching technique is developed using Principal Component Analysis (PCA) for recognition. The system is tested with the controlled and uncontrolled database. The system is tested in controlled background and in different lightning conditions. The database collected in the ideal conditions has proved to be the most efficient database in terms of accuracy and gives 100% accuracy and when the lightning conditions are changed the accuracy decreases as compare to the previous one. The system shows 91.43% with low brightness images .

Neelam K. Gilorkar et.al [9], proposed a real time vision based system for hand gesture recognition for human computer interaction in many applications. The system can recognize 35 different hand gestures given by Indian and American Sign Language or ISL and ASL at faster rate with virtuous accuracy. The gestures are classified on the basis of key points. RGB-to-GRAY segmentation technique was used to minimize the chances of false detection. They proposed a method of improvised Scale Invariant Feature

Transform (SIFT) and same was used to extract features. The system is modelled using MATLAB. To design a efficient user friendly hand gesture recognition system, a GUI model has been implemented. The sign for all alphabets, numbers and string are being recognized using SIFT algorithm. Advantage of using SIFT is high processing speed in real time. The extracted features are translations, scale and rotation invariant which make the system more flexible. Experiment revealed that the system was able to recognize a subset (35 letters) of ASL, ISL with an accuracy of 92- 96% using SIFT algorithm.

Mathavan Suresh Anand et.al [10], proposed an efficient Indian Sign Language Recognition System (ISLR) is proposed for deaf and dump people using hand gesture images. The proposed ISLR sys- tem is considered as a pattern recognition technique that has two important modules: feature ex- traction and classification. The joint use of Discrete Wavelet Transform (DWT) based feature ex- traction and nearest neighbor classifier is used to recognize the sign language. In the proposed ISLR system framework before extracting features and classification, a segmentation approach is applied to segment only the hand gesture images. Then DWT based features are extracted and given to a simple KNN classifier for sign recognition. In order to implement the proposed sign recognition system effectively by using hand gesture images, a pre- processing step is applied to the original captured image. The pre-processing includes colorspace conversion, thresholding and morphological operations to segment the hand gesture from the whole image. The process of sign recognition from the hand gesture images are divided into three modules; segmentation or pre-processing, feature extraction, recognition or classification. In the pre-processing module, only the hand gesture is segmented from the whole image using Otsu thresholding and morphological operations. After segmentation, the texture features are extracted using DWT and then KNN classifier is used to classify the type of hand gesture using the extracted features. DWT sub-band energies extracted from the hand gesture images are used as features along with the area of the segmented hand gesture region. Otsu thresholding and morphological operators are used for the segmentation procedure. DWT is applied up to the 7th level for computing sub-band energy features in order to obtain the best feature set. The nearest neighbor classifier used for the classification provides 99.23% accuracy.

CONCLUSION

This paper deals with the different algorithm and techniques used for recognizing the hand gesture as seen through the previously proposed Sign Language Translation Systems. Hand gesture recognition system is considered as a way for more intuitive and proficient way of communication for the people with hearing and speaking disability. Based on this we are proposing a novel approach to ease the difficulty in communicating with those having speech and vocal disabilities. Since it follows an image-based approach it can be launched as an application in any minimal system and hence has near zero-cost.

REFERENCES

[1] Meenakshi Panwar, "Hand Gesture Recognition Based On Shape Parameters", 2012 International Conference on Computing, Communication and Applications, IEEE,22-24 Feb 2012.

[2] Anup Kumar, Karun Thankachan, Mevin M Dominic, "sign language recognition", 2016 3rd International Conference on Recent Advances in Information Technology (RAIT),IEEE, 11 July 2016.

[3] Tülay Karayılan and Özkan Kılıç, "Sign language translation", 2017 International Conference on Computer Science and Engineering (UBMK),IEEE, 02 November 2017.

[4] Shreyashi Narayan Sawant and M. S. Kumbhar, "Sign Language Translation", 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies, 26 January 2015.

[5] Priyanka C Pankajakshan and Thilagavathi B,"Sign Language Recognition System", 2015 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS),IEEE, 13 August 2015.

[6] Pratik H. Suvagiya, Chintan M. Bhatt and Ritesh P. Patel ," Indian Sign Language Translator Using Kinect",© Springer Science+Business Media Singapore 2016 S.C. Satapathy et al. (eds.), Proceedings of International Conference on ICT for Sustainable Development.

[7] Yellappa Madhuri, Anitha Anburajan, Vision-based Sign Language Translation Device-ICICES- IEEE, 2013.

[8] Mandeep Kaur Ahuja, Amardeep Singh, Hand Gesture Recognition using PCA, International Journal of Computer Science Engineering and Technology (IJCSET), 2014.

[9] Neelam K. Gilorkar, Manisha M.Ingle, Real Time Detection And Recognition Of Indian And American Sign Language Using Sift, International Journal of Electronics and Communication Engineering & Technology (IJECET), 2016

[10] Mathavan Suresh Anand, Nagarajan Mohan Kumar, Angappan Kumaresan, An Efficient Framework for Indian Sign Language Recognition Using Wavelet Transform, International Journal of science and technology(IJSETR), 2015