International Journal of Research in Engineering, Science and Management
Volume-1, Issue-9, September-2018
www.ijresm.com | ISSN (Online): 2581-5782

Sign Language Interpretation and Implementation Using ORB Algorithm and Machine Learning

Omkar Pokale[1], Shrikant Ganoje[2], Aikya Banerjee[3], Anant Agrawal[4]
[1,2,3,4]*Student, Department of Compuetr Engineering, Sinhgad Academy of Engineering, Pune, India*

***Abstract*—In the modern world, Human Gesture Recognition software built into a computer system could work as a fortune for the handicapped or the people who cannot speak. Indian sign language (ISL) has a total of about 5000 signs in its vocabulary. Each sign is to be viewed at different angles and is recognized by the images stored in the repository. In this paper, by understanding the need, we are going to obtain Human Gesture recognition through a webcam built into the computer and convert the text generated to speech, which will then be transformed by the amazon virtual assistant ALEXA in real time operation. Hand gestures are obtained in the form of a continuous video format with each sign forming a separate image in its own, that is done by the gradient based key-frame extraction method. Each image will be stored in the database, after which, when the user wants to perform the execution , images from the database will be compared to the user's sign through the use of image processing and the Oriented FAST and Rotated BRIEF(ORB) algorithm. It performs with nearly a 95% accuracy as for recognizing the sequence of signs. The project presents itself as a go-to for the deaf and dumb persons proving to be useful in numerous research areas.**

***Index Terms*—Gesture Recognition, Indian Sign Language, Camera, Deaf and Dumb**

I. Introduction

Introduction- This project consists of various aspects, the first and foremost being the camera, which is primarily essential for the implementation. It is an optical instrument that scans 2D and 3D images from the face-up and finally delivers the scan for further processing and reduction of feature components. Here we use the ORB algorithm for comparison. Although, in the modern world we have many input devices such as keyboards, mouse, microphone, etc. sign language is worthy of use for broadening the perspectives and making it much more convenient for those who have been equipped with sign language as their primary source of communication. This project could be hand-held as an innovation in many educational and industrial fields. The project also tries to explore the area for conversion of human gesture into functioning by removing the gap between action and response. It uses ALEXA as one of its main components and not only helps in the interaction but can also use by investigations bureau to expose and catch theft when special signs are performed and stored. These can be utilized in many ways as per user convenience.

One of the mesmerizing features about this project is that once the software has been installed, it can provide for fast communication through the use of ALEXA's speech engine enhancing the idea of speech rather than having to read or write from the keyboard which is much slower in comparison. All said and done, there are yet many more aspects that can be added such as sentiment analysis through human emotion or more importantly face-recognition, sign language interpretation can be taught to other humans for effectively communicating with the handicapped, etc.

The project is also extremely efficient in the sense that people will be operating according to their own requirements and will be provided with the flexibility to manage it.

II. Objectives of the Work

- To build a smart system that recognizes sign language based on key-point extraction.
- To provide real time conversion of signs to speech with image processing and artificial intelligence.
- To provide faster communication with the help amazon's virtual assistant ALEXA.
- To make an impact on the modern world about usage of speech in place of other input devices and output devices.
- To promote and help impaired persons develop a faster adaptation to computer technology.
- To revolutionize and combine innovative technologies.

III. Literature Survey

Gesture Recognition is quite an eye-catching field from its stand-point and this paper contains study pertaining to the machine learning view of gesture recognition. In the initial days considering only the hand gesture recognition, kinetic sensors were used to sense the motion of the hand and accordingly it was further processed.

After that came the SIFT and SURF algorithm which was again used for gesture recognition or mainly image processing in OpenCV. Following that was the emerging field of deep

learning in which neural networks were used to obtain and recognize the hand motion.

Ultimately, in the present, the convolutional networks have set foot in the picture. Convolutional Networks are used to detect the emotion on the face and gestures of the hand.

But with the neural network the main problem associated is that sometimes they try to memorize the task rather than recognizing the pattern and actual detection.

Now the algorithm that we are using is simple enough and helps in proper detection of the sign sequences. It is named as ORB algorithm.

Through the stream of speech synthesis, this is again one of the areas focused widely. Thus there are a variety of ways that can be used for the text-to-speech conversion engine.

The first one being the Freetts library available for conversion. Then the Marytts library is also a choice available for text to speech conversion. OpenCV provides a lot of functionalities related to these libraries.

There are numerous ways through which we can convert text to audio, one of them being, the conversion of text into an audio file and then playing it as per requirement.

For the audio conversion, many speech synthesizers can be used, while one of the best and most efficient is the Google Speech Synthesizer which provides all the desired functionalities required for the software. As for the execution part, Amazon Alexa is a marvelous product widely used and admired all over the globe.

IV. Requirements

The system is designed so as to keep in check that it is feasible for the people and requirements are tried to be kept as simple as possible. For developing the system the software module requirements include:

1. JVM
2. NETBEANS/ECLIPSE
3. WINDOWS 32/64-bit
4. OPENCV
5. GOOGLE SPEECH SYNTHESIS

The (min) Hardware requirements for the system are:

1. WINDOWS 7/8/10
2. INTEL I3/I5/I7 DUAL CORE PROCESSOR
3. INBUILT CAMERA
4. 2 GB RAM

V. System Architecture

The system architecture is a conceptual model of the software designed to provide more insight on the designing the components of the software. The architecture is then implemented step by step through conversion of the model to code. The structure of the system solely depends on the basic components of the software and the relationship (e.g. behaviour) between them. Consequently, the design of the product can then be procured with the help of the overall architecture. The architecture of the software is simple, the human gesture component is first scanned and read by the the camera component which delivers the record to the pre-processing model where most of the features are extracted into templates, so as to match it with the signs in the database.

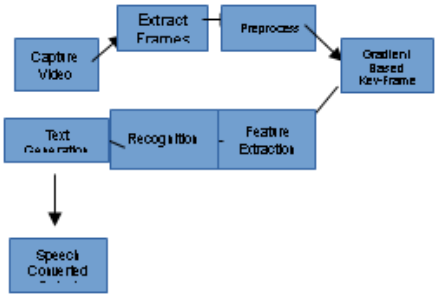

Fig. 1. System architecture

VI. Algorithm Analysis

In this system for comparing two images we use the ORB algorithm. However, there are other algorithms which are used for comparison. E. Rublee et al. introduced Oriented FAST and Rotated BRIEF (ORB) in 2011. ORB algorithm is a blend of modified FAST (Features from Accelerated Segment Test) detection and direction-normalized BRIEF (Binary Robust Independent Elementary Features) description methods. FAST corners are detected in each layer of the scale pyramid and cornerness of detected points is evaluated using Harris Corner score to filter out top quality points. As BRIEF description method is highly unstable with rotation, thus a modified version of BRIEF descriptor has been employed. ORB features are invariant to scale, rotation and limited affine changes.

ORB (1000) has the least feature-matching cost.
Quantitative Analysis Shown that:

1. ORB detects the highest number of features
2. BRISK is the runner-up as it detects more number of features than SIFT, SURF, KAZE, and AKAZE.
3. SURF detects more features than SIFT.
4. AKAZE generally detects more features than KAZE.
5. KAZE detected least number of feature-points.

VII. Conclusion

Sign Language recognition system paves the way for a revolution of digitized technology for the handicapped. This kind of technology can thus be binded with innumerable domains such as machine learning and artificial intelligence where the computer can be taught new signs just by weaving them to the camera. With close-to-perfect precision and accuracy in terms of gesture recognition, the system could be promoted and built into a portable system of low costs, such that deaf and dumb people could use it for communication anytime and anywhere.

REFERENCES

[1] Hand Gesture Recognition Using an Adapted Convolutional Neural Network with Data Augmentation
[2] Deep Learning for Hand Gesture Recognition on Skeletal Data
[3] ORB: an efficient alternative to SIFT or SURF, Ethan Rublee,Vincent Rabaud,Kurt Konolige,Gary Bradski
[4] Robust Part-Based Hand Gesture Recognition Using Kinect Sensor, Zhou Ren, Junsong Yuan.
[5] Real-time Vision-based Hand Gesture Recognition Using Haar-like Features, “Qing Chen, Nicolas D. Georganas, Emil M. Petriu
[6] Vision Based Hand Gesture Recognition, Pragati Garg, Naveen Aggarwal and Sanjeev Sofat(2009)
[7] Real Time Hand Gesture Recognition System, Neethu P S, R Suguna,Dr. Divya Sathish.
[8] Deep Convolutional Neural Networks for Sign Language Recognition, G.Anantha Rao, K.Syamala, P.V.V.Kishore, A.S.C.S.Sastry.