INTERNATIONAL JOURNAL OF ADVANCE RESEARCH AND DEVELOPMENT
IJARnD

(Volume2, Issue3)
Available online at www.ijarnd.com

Conversion of Sign Language to Text and Speech

Girish U. R[1], Srinidhi Madhyastha [2], Varun A. M [3], Poornima B. G[4]

[1]UG Student, Computer Science and Engineering, Vidyavardhaka College of Engineering, Karnataka, India
[2]UG Student, Computer Science and Engineering, Vidyavardhaka College of Engineering, Karnataka, India
[3]UG Student, Computer Science and Engineering, Vidyavardhaka College of Engineering, Karnataka, India
[4]Associate Professor, Computer Science and Engineering, Vidyavardhaka College of Engineering, Karnataka, India

ABSTRACT

Deaf people rely on sign language to express their own thoughts and feelings. It becomes the major communication barrier between the deaf and other people. Sign Language has evolved as one of the major areas of research and study in computer vision. Researchers in sign language recognition used different input devices such as data gloves, web camera, depth camera, color camera, Microsoft's Kinect sensor, etc. to capture hand signs. In this paper, we display the importance of Sign Language and proposed technique for classification and their efficient results. A sign language looks up the manual communication and body language to convey meaning, as opposed to acoustically conveyed sound patterns, which involve a simultaneous combination of handshapes, orientation, and movement of hands. The signs are captured using a new digital sensor called "Leap Motion Controller". LMC is 3D non-contact motion sensor which can track and detects hands, fingers, bones and finger-like objects. The Leap device tracks the data like point, wave, reach, grab which is generated by a leap motion controller. The system implements Dynamic Time Warping (DTW) for converting the hand gestures into an appropriate text.

Keyword: Sign Language, Leap Motion Controller, Bone API.

1. INTRODUCTION

For more than 250–300 million people around the world, hearing loss and disability to speak presents everyday challenges-some large, some small. To build an application for the deaf and dumb community has become of utmost importance today. With the changing world in terms of technology innovations and education opening new areas of opportunities, it becomes necessary for everyone to be ahead of the others in this race. In order to enable the deaf and dumb community to create recognition and also to give them a standard platform to communicate and express their opinions with every other individual, this application is being created. A sign language is a communication method for impaired people. By using a sign language as an input interface to Information and Communications Technology (ICT) devices, it becomes possible for impaired people to hear, something which is hard to perform by using a conventional keyboard or touchpad. A sign language uses visual information associated with finger acts. At the same time, several fingers acts are used with a part of the face such as a line of sight and mouth.

In this paper, we propose a hand gesture recognition approach using Leap Motion Controller. This latter has skeletal tracking that recognizes the framework of fingers to obtain a highly accurate data such as the index finger, the position of finger bones and the degree of the thumb.

1.1 Leap Motion Controller

The Leap Motion controller is a small USB peripheral device which is designed to be placed on a physical desktop, facing upward. It can also be mounted onto a virtual reality headset. Using two monochromatic IR cameras and three infrared LEDs, the device observes a roughly hemispherical area, to a distance of about 1 meter. The Leap device is as shown in Fig 1 and Fig 2.The LEDs generate pattern-less IR light and the cameras generate almost 200 frames per second of reflected data. This is then sent through a USB cable to the host computer, where it is analyzed by the Leap Motion software using "complex maths" in a way that has not been disclosed by the company, in some way synthesizing 3D position data by comparing the 2D frames generated by the two cameras.

Fig 1: Internal structure of Leap

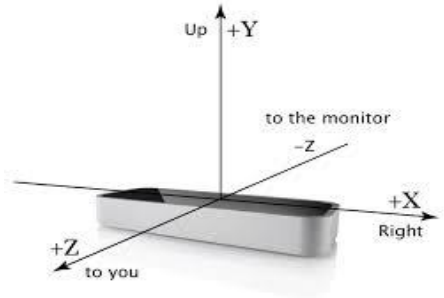


Fig 2: Leap Motion Controller

1.2 Bones of human hand and Bone API

Bone API has some attributes that we can use to get the position of bone in a limited space above the leap motion controller. Bone API has some attributes such as the center, direction, and length which returns the center, direction of finger and length of the finger respectively

The human hand has five fingers and 27 bones, not including the sesamoid bone, 14 of which are the phalanges (proximal, intermediate and distal) of the fingers. The metacarpal bones connect the fingers and the carpal bones of the wrist. Each human hand has five metacarpals and eight carpal bones. Among humans, the hands play an important function in body language and sign language.

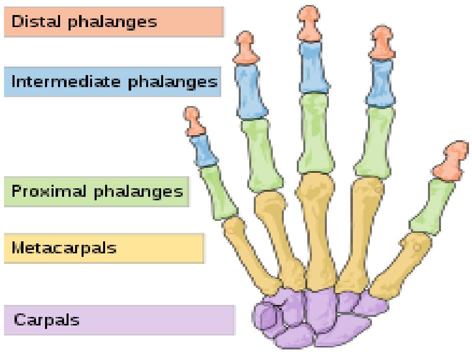


Fig 3: Bones of Human Hand

2. LITERATURE SURVEY

A Sign Language Recognition (SLR) system has to be designed to recognize a hand gesture. Gestures in sign language are defined as specific patterns or movements of the hands, face or body to make out expressions. The sign capturing methods are vision based sign extraction, a data glove, and Electromyography. The continuous sign recognition is very complex as it cannot be separated as the speech is separated. Hence continuous sign recognition system uses Hidden Markov Model (HMM) and Electromyography (EMG) segmentation. Sign Language, the natural communication medium for a deaf person, is difficult to learn for the general population. The prospective signer should learn specific hand gestures in coordination with head motion, facial expression, and body posture. Sign tutor helps the user by providing them with sign videos, text-based description, pictures of hand gestures and 3D animated avatar. A vision-based static hand gesture recognition algorithm consists of three stages: pre-processing, feature extraction and classification. The pre-processing stage involves following three sub-stages: segmentation which segments hand region from its background images using a histogram based thresholding algorithm and transforms into binary silhouette; a rotation that rotates segmented gesture to make the algorithm, rotation invariant; filtering that effectively removes background noise and object noise from the binary image by a morphological filtering technique. A localized contour sequence (LCS) based feature is used here to classify the hand gestures. A k-mean based radial basis function neural network (RBFNN) is also proposed here for classification of hand gestures from Leap controller based feature set. Microsoft Kinect sensor plays a vital role in the sensing and robotics communities. However, the applications are programmed are generated in C alone. To overcome this issues, the VU-Kinect Simulink, an integration of Kinect and simulink has been developed which provides easy access to the sensor's camera and depth image streams. The VU-Kinect block is used to track a 3-D object. Vision-based recognition system consists of three main components: hand gesture modeling, hand gesture analysis, and hand gesture recognition. Gesture model describes how the hand gesture is to be represented. The analysis is performed to compute the model parameter from the image feature. The analysis phase is followed by recognition phase that classifies the model prepared. Hand gesture recognition consists of feature detection, hand localization, feature extraction and parameter computation. Classification technique deals with the Euclidean distance metric. Translated ISL is displayed with the help of a 3D virtual human avatar. The input to the system is the clerk's speech which is in English. The speech recognition module recognizes speech and makes a text output. The output from the parser is given to an eliminator module which performs a reduction task by eliminating unwanted elements and further the root form of verbs are found using the stemmer module. The structural divergence of English and ISL is handled by a phrase reordering module using ISL dictionary and rules.

In the current research for Sign Language Recognition (SLR), the image recognition of colored images, depth images, and hand shapes are used. Since it must be taken with colored gloves, the glove worn is not suitable. The image recognition requires long calculation time to detect the hand and the fingers. Therefore; it takes relatively a long interval to attain the final recognition result.

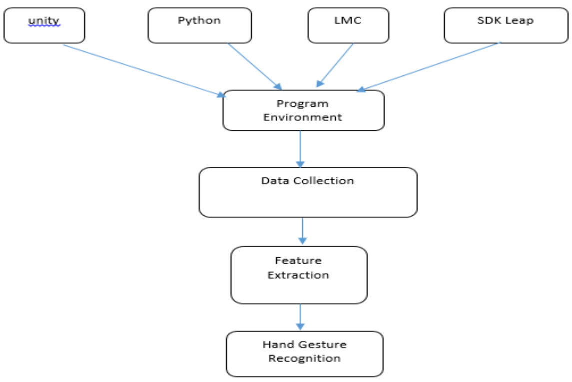
3. PROPOSED APPROACH

The overall workflow of our proposed system is composed of 5 steps as illustrated in the flow diagram. The first step consists of setting up the programming environment, including Unity, SDK Leap, LMC, and Python. Step two consists of collecting data to create a library of gestures. The third step concern the features extraction, these features are the introduced as a vector to a classifier, which generates a model for each gesture. Finally, the last step uses the trained classifier to recognize users' gesture. B. Data collection At this level, data is extracted from the LMC using the software development kit (SDK) associated with Python. The LMC returns data representing the geometry of detected hand around its vicinity. The data contains information describing the overall motion of the hand. For our approach, we use the LMC to collect hand and finger data while the sign is performed in the LMC field of view. Among the features of the leap motion controller, we mention the tracking of the hand gestures by giving the 3d coordinate "x, y, z" of the palm and tip of each finger.

We use the Leap Motion Controller to capture gestures from the user. Initially, the captured gestures are used to generate a training set wherein we train the computer to recognize the various symbols used in symbolic language. This training set is then used to compare with the actual user input and when a match is found, it outputs the corresponding letter or word along with the speech.

We create multiple instances of the same letter or word for better accuracy. Further, machine learning will be used to increase the accuracy.

3.1 Flow Diagram:



CONCLUSION

Thus we conclude that Leap Motion Sensor will overcome the limitations of previous systems and will be an alternative to low-cost hardware and real-time response property guaranteeing the system translation accuracy under different environmental factors.

REFERENCES

[1] Recognition Letters, vol. 36, pp. 221–227, 2014.

[2] X. Liu and K. Fujimura, "Hand gesture recognition using depth data," IEEE Intl. Conf. on Automatic Face and Gesture Recognition, 2004.
[3] J. Molina, M. Escudero-Vinolo, and A. Signoriello, "Real-time user independent hand gesture recognition from time-of-flight camera video using static and dynamic models," Machine Vision and Applications, 2011.
[4] S. Oprisescu and E. Barth, "3D Hand Gesture Recognition using the Hough Transform," Advances in Electrical and Computer Engineering, 2013.
[5] M. Betke, J. Gips, and P. Fleming, "The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People With Severe Disabilities," IEEE Transactions On Neural Systems and Rehabilitation Engineering, 2002.
[6] www.lifeprint.com
[7] https://developer.leapmotion.com/documentation/python/index.html
[8] X. Chai, G. Li, Y. Lin, Z. Xu, Y. Tang, X. Chen, and M. Zhou, "Sign Language Recognition and Translation with Kinect," IEEE Intl. Conf. on Automatic Face and Gesture Recognition, 2013.
[9] P. Breuer, C. Eckes, and S. Muller, "Hand gesture recognition with a novel in time-of-flight range camera – a pilot study," Proceedings of the 3rd Intl. Conf. on Computer vision/computer graphics collaboration techniques, 2007.

