Sign Language Decipherer by Hand Behavior Image Acquistion using Leapmotion Controller

Saritha.S , Dhepak. S, Geetha Barathi.S.B
Department of ECE, Nandha College of Technology, Erode, India
sarithas.ece@gmail.com, dhepaks@gmail.com, geethabharathivio@gmail.com

***Abstract*—A sign language decipherer which recognize the hand behavior through the leapmotion controller. Proposed system used Artificial Neural Network (ANN) algorithm with Back Propagation (BP) algorithm to construct a classification model which takes feature as input. In this system, the input hand behavior is performed by verbally challenged persons and is then converted into natural audio. The various set of the hand behaviors can be converted into complete sentence and grammatically corrected using NLTK Library. In the next step mapping is done with the pre-recorded audio file which is stored in MP3 module and then the audio output which comprised of natural sentence is interpreted through Galileo board and the sequence of sentence is received as audio output.**

Keywords—Leapmotion Controller, Sign Language Decipher, Intel Galileo, Sign Language Recognition, Artificial Nueral Network, Machine Learning, Back Propogation.

I. Introduction

A Sign Language is a hand behavioral based language and that uses manual hand behavioral actions to convey the meaning of the words and to interpret the real- time sentences which can be used familiar by verbally challenged persons. But there are many obstacle between the verbally challenged persons and normal persons. This unconnected bridge can lead to an unfavorable life style with other persons and with their relationships as well. The existing and stereotype method of communication between verbally challenged and normal hearing persons can be executed traditionally through interpreters or the method of text by writing or typing. The interpreters are very expensive for the routine and daily usage and their involvement will make the verbally challenged persons to feel uncomfortable as they loss their privacy and independence. The communication method of text writing or typing is slow paced and the efficiency of instant communication will be degraded. Then the traditional system was then replaced by a wearable system for recognizing sign language and to perform the operation of interpreter with a digital circuits outcome with the audio representation that can able to bridge the gap between the verbally challenged persons and the hearing normal persons. But this additional representation cannot able to completely satisfy the verbally challenged persons in the routine world because it disturbs the routine and natural behavior of the person and it can able to reveal the identity of the person even before the meet or first talk with a normal person. This will make the verbally challenged person to feel that they are somewhat separated from the normal community in the globe.

A Sign Language Decipherer (SLD) system is the salient feature that provide the efficient communication between verbally challenged persons and normal hearing persons as this system translates the sign language into natural audio. The system functions by using the leapmotion controller which is a consumer 3D gesture sensor [7]. This leapmotion controller can be either placed in smartphones by developing a mobile application and this part is our future work. Here we have used Intel Galileo microcontroller board and MP3 Module as a hardware device to design and the independent system. Using the leapmotion controller the verbally challenged persons can able to assign by recording their easier way of a gesture by a new sign language. This proposed system can even provide the additional path to create a new easier sign language and that can be practiced and can be established in the upcoming years when our proposed system attains a standard. By Utilizing the monocular video sensor, it is hard to capture the dynamic hand behaviors called as gestures and using this system will result in a slow process execution. In the recent years, the development of the Leapmotion Controller, provides three-dimensional depth data within the user's interactive range. It can segment the objects and recognize the 3D hand behavioral gestures in an efficient and in a fastest way as the Leapmotion Controller uses Infrared based cameras that filters out the background visible light from the hand.

In this paper, we present a neural network algorithm to record the hand behavior gestures by using the leapmotion controller and send the data to the Intel Galileo board which is interfaced then the pre- recorded voices stored on the MP3 module will fetch and the respective audio is received when the sign language is given as input and the addition algorithm have been used to combine the several words into a sentence and the sentence is grammatically corrected using NLTK library in the program.

II. LITERATURE REVIEW

In [1], Jian Wu, Lu Sun and Roozbeh Jafari presented a sign language recognition system using a wearable system by using intertial sensors and sEMG sensors. In this system, the 3-D accelerometer (ACC), 3-D Gyroscope and and four channel

SEMG are preprocessed for noise rejection and synchronization purposes The sEMG-based auto-segmentation technique obtains the beginning and ending of a sign for both IMU and sEMG. As the segmentation is done, a broad set of well-established features are extracted for both IMU and sEMG signals. All extracted features are then put into one feature vector.The best feature subset is obtained using an information gain (IG)-based feature selection scheme

Cao Dong, Ming C. Leu and Zhaozheng Yin used Microsoft Kinect sensor to recognize American Sign Language alphabet. This Paper used Microsoft's Kinect sensor to obtain depth data. The Hand segmentation is done using per-pixel classification method. Random Forest (RF) gesture classifier was implemented to recognize the user signs using the joint angles.

In [2], Adithya and Vinod presented a method of recognizing the Indian sign language. Images are captured using web camera under lighting and environmental condition. Captured images are then converted to YCbCr colour space. Distance transform and Discrete Fourier Transform (DFT) is used for extract features from image and artificial neural network used for recognizing Indian sign language alphabets.

In [3], Kai-Yin Fok, Nuwan Ganganath introduced a method for recognizing digits (0-9) of American SignLanguage. The proposed system utilizes two Leap Motion sensors for collecting hand states from different viewing angle. Collected data fused using multiple sensors data fusion(MSDF).Key features of proposed system were tip to palm ratio, tip to tip ratio and tip to joint ratio. The recognition was performed using hidden Markov models (HMM).

Priyanka Mekala, Ying Gao introduced method for real-time American Sign Language recognition. A camera sensor was used to capture the video from the user. Image frames were extracted from video. Backgrounds of image frames were subtracted using Running Gaussian average method. The Discrete Fourier Transform (DFT) used for feature extraction. Combinational Neural Networks minimize the size of search space was used as classifier to detect signs [4].

Ching-Hua Chuan and Eric Regina proposed a system for ASL recognition using leap motion controller .This system uses leap motion sensor to collect data from the user. They used k-nearest neighbour and SVM to classify the 26 alphabets of American Sign Language. Average classification rate of knearest neighbour and support vector machine was 72.78% and 79.83% respectively [5].

An approach for the recognition of Arabic sign language is addressed in [6]. A.S.Elons and Menna Ahmed used Leapmotion sensor. This system works with 50 different dynamic signs from Arabic sign language. The signs were collected from 4 different persons, two signs set used for training and other two for testing. The fingers position and distance between the fingers in each frame were features sets provided to system. Proposed system used Artificial Neural Network (ANN) as classifier to recognize the gestures. Recognition accuracy of 88% was achieved.

III. PROPOSED SYSTEM

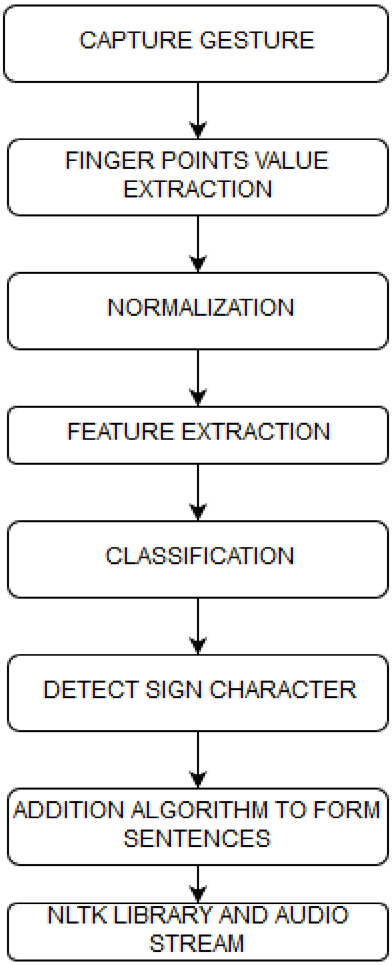

Fig.1: Architecture of Proposed System

1) Data Acquiring

The proposed system will acquire the 3D dynamic signs provided by use as input. Here the data is collected using leap motion controller. The Leap Motion controller is a small USB peripheral device which is designed to allow users to control their computers with hand gestures alone. This sensor is 3D non-contact motion sensor which can detects and tracks hands, fingers, bones and finger-like objects reporting discrete position and motion. The heart of the device consists of two monochromatic IR cameras and three infrared LEDs as shown in Fig. 2. The device has a large interaction space of eight cubic feet and viewing range is approximately 1 inch to 2 feet (60 cm) above the device. The Leap Motion system employs a right-handed Cartesian coordinate system. The origin is centered at the top of the Leap Motion Controller. The x-axis lies in the horizontal plane and running parallel to the long edge of the device. The y-axis lies in vertical plane, with positive values increasing upwards. The z-axis has positive values increasing toward the user and it is lie in the horizontal plane.

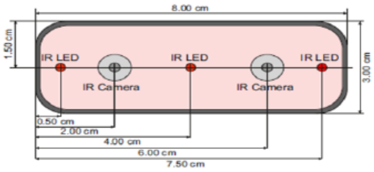

Fig. 2:Leapmotion Controller

2) Features Extraction

The Leap motion includes different types of features for hand, fingers, bone and gesture inputs.

Hand: The variations can be defined from the information acquired from the hand model, it can be left, right or both hand. From the information of velocity (in millimeter per second) and other characteristics of a detected hand, the attachment of the arm and the lists of the fingers associated with the hand.

Fingers: The features for the fingers include finger direction expressed in a unit direction vector, velocity (in millimeter), width, dip position, pip position, tip position, tip velocity, mcp position.

Gestures: The Leap motion controller can recognize movement patterns. It recognizes the motion of a finger tracing a circle in space as a *Circle* gesture, downward movement of the finger as *Key Tap gesture* as it also feature the tapping movement.

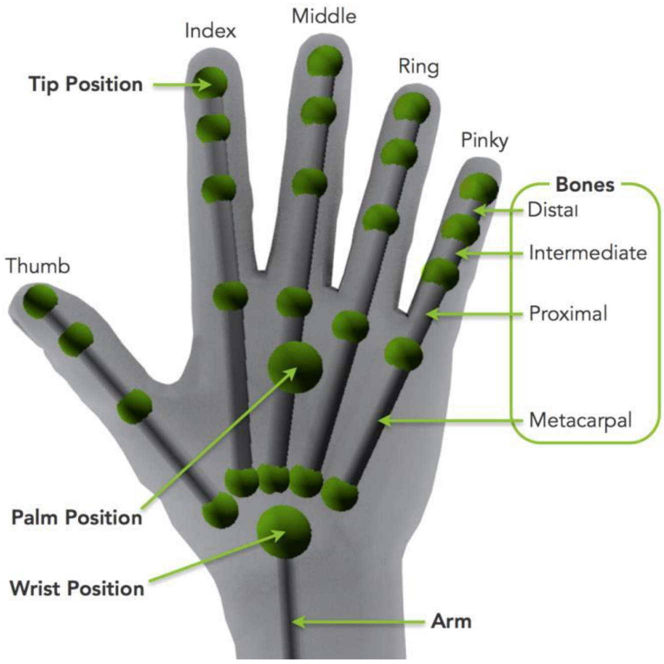

Fig. 3 : Features provided by LMC API

In this proposed system of data extraction, we select Palm and finger dataset to be the features. The features which are normalized are provided to the Artificial Neural Network (ANN). These features are used to train Artificial Neural Network are

1) The Eucledian distances between the consecutive finger tip position to palm postion.

2) The Euclidian distances between the finger tip position of each consecutive finger.

3) Classification

1) *Artificial Neural Network (ANN)*

An artificial neuron network (ANN) is a computational model based on the structure and functions of biological neural networks. Information that flows through the network affects the structure of the ANN because a neural network changes - or learns, in a sense - based on that input and output. The input layer neurons are connected to the hidden layer neurons and the hidden layer neurons are connected to output layer neurons by means of interconnecting weights as shown in fig.4. ANNs are considered nonlinear statistical data modeling tools where the complex relationships between inputs and outputs are modeled or patterns are found. An ANN has several advantages but one of the most recognized of these is the fact that it can actually learn from observing data sets. In this way, ANN is used as a random function approximation tool. These types of tools help estimate the most cost-effective and ideal methods for arriving at solutions while defining computing functions or distributions. ANN takes data samples rather than entire data sets to arrive at solutions, which saves both time and money. ANNs are considered fairly simple mathematical models to enhance existing data analysis technologies. Thus, the size input layer is set by number of feature set we want the network to process.

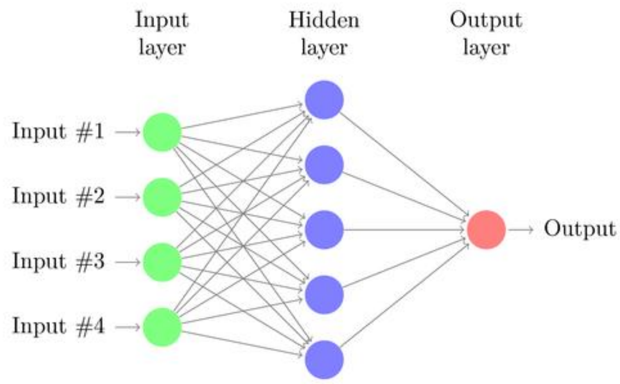

Fig.4 : Artificial Neural Network (ANN)

In this system the Networks are trained with the Back Propogation (BP) algorithm and this can be used for the training of various gestures as the sign languages.

Networks trained with the Back Propagation (BP) algorithm consist of the following steps as following Apply input to network and initialize weights and bias

- Forward Pass
  - Initially, it calculates net input to the each of the hidden layer unit. Then, it calculates output to the each of the hidden layer unit. In the next step, Calculate net input to the each of the output layer unit. Finally, it calculates net output to the each of the output layer unit
- Reverse Pass
  - It calculates the error. Then, it updates the weights of interconnection from hidden layer unit to output layer unit. Then, it calculates the error of hidden layer and finally, it updates the weight of interconnections from the input layer neurons to hidden layer neurons.

IV. RESULTS AND DISCUSSION

The proposed system was implemented using two systems, PC and Intel Galileo. Initially the data is collected from the leap motion controller connected to PC, which was running in Intel(R) Core™ i7-4510U CPU @ 2.60Ghz and the code was written using Java programming language. The python code runs on Galileo board, captures the messages sent through PC. This code also handles the tasks related to interfacing the MP3 Module to play the audio via UART communication. The another code on PC will recognizes the sign and convert them into words, this also handles the sentence to correct grammatically using NLTK library. The gesture recognizing is done through Data numpy, thus all data is converted into matrix to provide a faster accessible system. The data is collected using 3 different persons and initially 76 input samples are collected, for that input samples corresponding audio is generated using IVONA text to speech [8] and pre-defined in the system .The processing time is about 2 sec for a sentence as it was proposed by data numpy. Fig. 5. Shows the matrix form of data collection using data numpy.

```
[[-0.09443539 -0.09443531  0.29860729 -0.09761513 -0.09440866]
 [-0.09443526 -0.09443531  0.25596021 -0.10824217 -0.094422  ]
 [-0.09443524 -0.09443531  0.37198598 -0.12371693 -0.09442577]
 [-0.09443568 -0.09443531  0.30667577 -0.10257815 -0.09441752]
 [-0.09443562 -0.09443531  0.41545527 -0.06368836 -0.09441873]
 [-0.09443647 -0.09443531  0.34410876  0.00738793 -0.09440932]
 [-0.0944355  -0.09443531  0.33180906 -0.12472302 -0.09442687]
 [-0.09443587 -0.09443531  0.3643611  -0.16894118 -0.09443041]
 [-0.09443721 -0.09443531  0.43028699  0.0095     -0.09441093]
 [-0.09443846 -0.09443531  0.34737789 -0.07818481 -0.09439922]
```

Fig 5. Data Numpy collection of data

V. CONCLUSION

The Sign Language Recognition is essential for the verbally challenged people to communicate with hearing individuals. This paper presents method for the sign language recognition by capturing from Leap motion controller. Leap Motion Controller is 3D non-contact motion sensor which can detects and tracks hands, fingers, bones and fingerlike objects reporting discrete *position and motion. Some of the* advantages of Leap Motion Controller are: Robustness, Require less memory, Fast Processing. The Artificial Neural Network (ANN) is used with Back Propogation (BP) to train the system using the feature extraction. The performance of the proposed system improved in terms of time to build using the Data numpy approach. Recognition at even more faster build time can be implemented by improving the algorithm and all other features of various sign language can be adapted from the web services and proceeded over online network and these features to the system will be implemented in future.

V. REFERENCES

[1] Jian Wu, Lu Sun and Roozbeh Jafari presented a sign language recognition system using a wearable system By using intertial sensors and sEMG sensors. In this system, the 3-D accelerometer (ACC), 3-D Gyroscope and and four channel SEMG are preprocessed for noise rejection and synchronization purposes The sEMG-based auto-segmentation technique obtains the beginning and ending of a sign for both IMU and sEMG. As the segmentation is done, a broad set of well-established features are extracted for both IMU and sEMG signals. All extracted features are then put into one feature vector.The best feature subset is obtained using an information gain (IG)-based feature selection scheme

[2] Adithya, V.; Vinod, P.R.; Gopalakrishnan, U.,"Artificial neural network based method for Indian sign language recognition," Information &Communication Technologies (ICT), 2013 IEEE Conference on, vol., no., pp.1080,1085,2011-12 April,2013.

[3] K. Y. Fok, N. Ganganath, C. T. Cheng and C. K. Tse,"A Real-Time ASL Recognition System Using Leap Motion Sensors" Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 2015 International Conference on, Xi'an, 2015, pp. 411-414.

[4] P. Mekala, Y. Gao, J. Fan and A. Davari, "Real-time sign language recognition based on neural network architecture," 2011 IEEE 43rd Southeastern Symposium on System Theory, Auburn, AL, 2011, pp.195-199.

[5] H. Chuan, E. Regina and C. Guardino, "American Sign Language Recognition Using Leap Motion Sensor," Machine Learning and Applications (ICMLA), 2014 13th International Conference on, Detroit, MI, 2014, pp. 541-544

[6] S. Elons, M. Ahmed, H. Shedid and M. F. Tolba "Arabic sign language recognition using leap motion sensor," Computer Engineering & Systems (ICCES), 2014 9th International Conference on, Cairo, 2014, pp.368-373.

[7] Leap Motion, http://www.leapmotion.com

[8] IVONA Text to Speech Converter https://www.ivona.com/