IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified

Vol. 6, Issue 9, September 2017

Sign to Speech

Shubham Jadhav[1], Pratik Shah[2], Ashutosh Wadhvekar[3], Parag Hoshing[4], Avinash Bagul[5]

Student, Computer Department, NBNSSOE, Pune, India [1,2,3,4]

Assistant Professor, Computer Department, NBNSSOE, Pune, India [5]

Abstract: This paper addresses the development of a hand gesture recognition system for Indian Sign Language. In this a Sensor Glove is developed which is used by speech impaired people to communicate with others. The Glove converts the hand gestures into corresponding text as well as speech. The Hand glove is composed of 5 flex sensors, a MPU 6050 which combination of both accelerometer and gyroscope, a Microcontroller which is Arduino Pro Micro and Bluetooth transceiver HC-05. The readings from the sensors send by the microcontroller to the Mobile Application using Bluetooth connectivity via Bluetooth transceiver. Mobile applications directly send these readings to the server using UDP (User Datagram Packet) Connection. On the Server side hand gesture is recognized using Artificial Neural Networks (ANNs). A Machine Learning algorithm using Multilevel Perceptron is implemented on Server, to which a input of 78 features extracted from readings are given a and output is obtained as recognition of words or letters corresponding to the gesture. This output is send back to the App where it is displayed as text as well as can be heard as a speech. Thus in this way system helps to give the voice to the speech impaired people.

Keywords: Sign Language, Flex Sensors, 3-axis Accelerometer, Microcontroller, Bluetooth Transceiver, Multilayer Perceptron.

I. INTRODUCTION

There are many people amongst us who don't have the power of speech. Our system is devoted to these speech impaired people who cannot speak. One can simply imagine how difficult it would be to interact with others when you are not allowed speak. This makes it very difficult for speech impaired people to express their ideas to others, to talk to others. Our system aims to bridge this gap between those who can speak and those who can't. Thus promoting a world where everyone equally exercises the power of speech.

As deprived of speech, Sign languages are the only medium of interaction for speech impaired people with others. Then again, problem with Sign Languages is that they are confined only to the speech impaired people. Normal people generally don't have any understanding of Sign language, which results in communication gap.

The solution to this problem is achieved by the system by eliminating the limitations of Sign languages. Sign languages are composed of different generic hand gestures, where each gesture has its own specific telling. The system recognizes these hand gestures and converts them into their corresponding speech as well as text. Thus by using the same Sign language known to speech impaired people they can interact with other people in form of speech. The gestures are recognized with the use Machine Learning, unlike various previously used pattern matching techniques.

There are different Sign languages for different countries and different languages. Each of these Sign Languages has generic set of hand gestures developed to convey that specific language. Thus these hand gestures vary from one Sign language to another. Currently the System is being developed for Indian Sign Language. The System can be used for recognition of words and letters used in Indian Sign Language.

II. SYSTEM COMPOSITION

The system can be visualized as composition of three parts namely a Hand Glove, a Mobile App and Server. The fig 1 shows a Generalized Architecture of the system.

A. Hand Glove

Hand Glove is composed of two sensors namely flex sensors and MPU 6050, a microcontroller which is Arduino Micro pro and a Bluetooth transceiver HC-05. The arrangement of these components on hand glove is inspired from [4]. The Hand Glove can be seen in fig 2. There are 5 flex sensors, one placed on every finger of hand within the pockets made in the glove along the fingers. The MPU 6050 and Microcontroller are placed on the top of the hand as it can be seen in fig 2 and Bluetooth transceiver is placed in pocket made in glove along the wrist of hand.

IJARCCE
ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

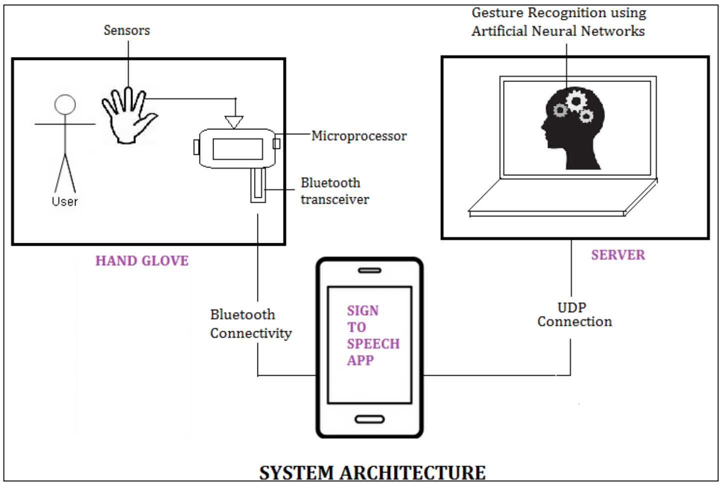

Fig. 1 Generalized System Architecture

Flex sensors are connected to 5 analog pins on the Arduino Micro Pro, MPU 6050 is connected using I2C bus and HC-05 Bluetooth transceiver is connected via Serial Peripheral Interface using Transmitter(TX) and Receiver(RX) lines. Flex sensors are used to measure the bend in the fingers, depending upon the bend, output voltage is obtained for each finger. MPU 6050 consist of both accelerometer and gyroscope but for implementation of glove only accelerometer readings are used. 3-axis Accelerometer is used to detect and measure motion of the glove. Accelerometer gives projection of vector g along X-Y-Z axis or in simple words acceleration along X-Y-Z axis. A total of 8 readings, 5 from flex sensors and 3 from accelerometer are received by the microcontroller.

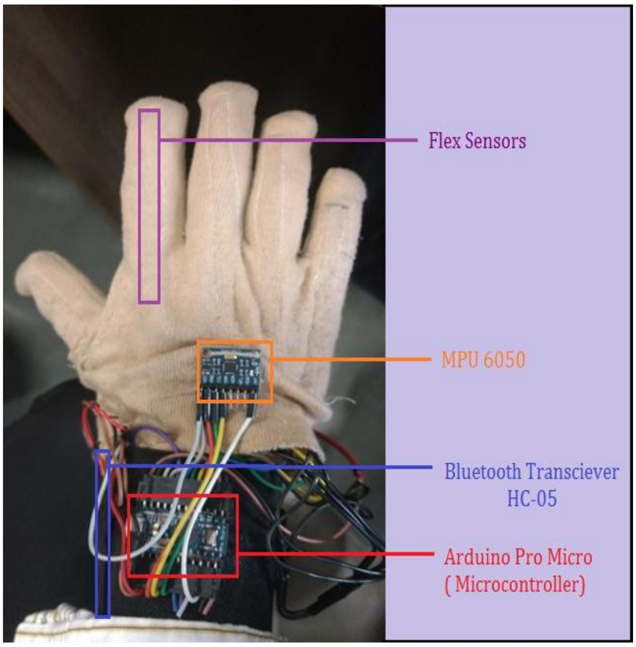

Fig. 2 Hand Glove

B. Mobile App

The data transmission medium between the microcontroller placed on Hand Glove and the Smartphone is Bluetooth. These readings from microcontroller are transmitted to the Smartphone using Bluetooth transceiver HC-05. Then the mobile App transfers these readings directly to the Server using a UDP connection as seen in System Architecture.

C. Server

Servers performs the processing on these readings received from the mobile app, recognizes the hand gesture and send the corresponding Sign language letter or word to the mobile app. Then by using Mobile app the text is displayed as well as can be heard as a speech from the smartphone.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

III. WORKING OF THE SYSTEM

The working of the System begins when hand gestures are made by the user using the hand glove. The analog reading received from the flex sensors and MPU 6050 are send to the microcontroller. At microcontroller these analog readings are converted into digital format using inbuilt ADC (Analog to Digital Converter). These 8 reading are then converted into floating point values using the code uploaded on microcontroller.

Then these 8 floating point values using Bluetooth transceiver are send to the mobile App. Currently it is an applications developed for Android users only. App receives these 8 reading from the microcontroller via Bluetooth connectivity. Once these 8 readings are received by the App they are forwarded directly as it is to the server.

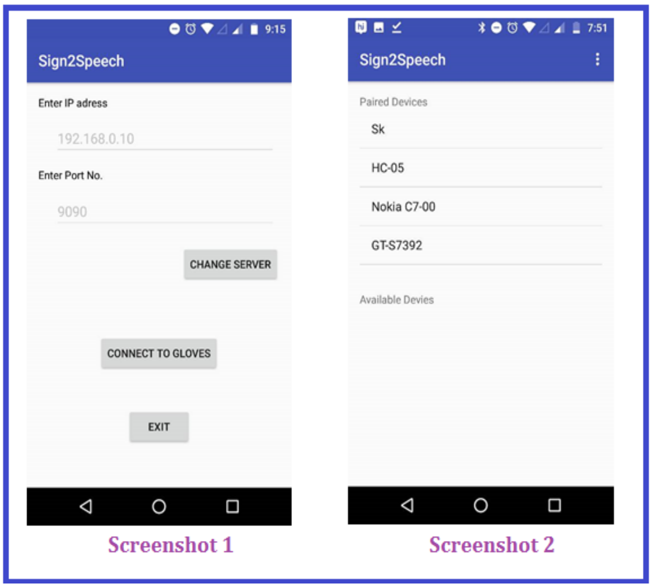

Fig. 3 Connectivity between Hand Glove and Server Using Mobile Application

The App connects to the Server via User Datagram Protocol (UDP) socket. First screenshot in Fig 3 shows the app taking IP address and port number of the server as input in order to make a UDP connection with the Server. Once the connection is established with server, the app shows the list of connected devices as can be seen in second screenshot, where HC-05 represents is the hand glove. When we select the HC -05 the flow of data begins from microcontroller to App and then App to server.

Now at the server side, input received is the same sensor data that is received by the app. We are using Artificial Neural Networks (ANN) for recognition of hand gestures on Server. On server side we are running a machine learning algorithm for Multilayer Perceptron. Code running on server side is developed in java and uses Weka Library [3]. From this library, we have imported the Multilayer Perceptron package. We have developed a model of 3 layers, which are: input layer (78 nodes), hidden layer (47), Output layer (15 nodes). When data is received by server following steps are performed for recognition of hand gesture:

A. Pre-processing

In pre-processing, based on the sensor data we try to find whether user is doing gesture or not and if he is doing gesture, then we start recording the sensor data. Now to do this we need acceleration data from 3 axes. As told in this paper [1], we calculate the value of ema as

$$Hk = \sqrt{(x_k - x_{k-1})^2 + (y_k - y_{k-1})^2 + (z_k - z_{k-1})^2}$$

$$Ema_{hk} = \alpha H_k + (1\text{-}\alpha)\, Ema_{hk\text{-}1}$$

If the value of ema goes above the threshold (based on our practical analysis we decided threshold value as 1300), it means user is doing the gesture. So we start recording the sensor data till value of ema is above 1300.

Fig 4 gives an idea about this. The green line is our ema. So when user does a gesture (portion between two red lines) ema rises above 1300 and we start recording the sensor data.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

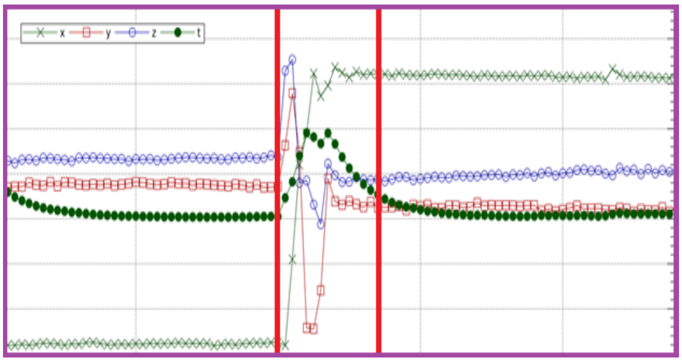

Fig. 4 Graph representing recording of gesture

B. Feature Calculation

The recorded sensor data is sent to Feature calculation Class. Here we calculate 78 features which are later given as input to our Multilayer Perceptron Model. Table shows the main features that are calculated, they are mostly the statistical and wave features.

Domain	Number of features	Name
Time Domain	8	Mean, Median, STD, MAD, Max, Min, Mean of last 5 sensor data.
Time Domain	13	Correlation
Frequency Domain	3	Energy, Entropy, Dftmean

C. Prediction

Then based on these 78 features extracted from the sensor readings, Multilayer Perceptron predicts the word or letter for the gesture. This predicted word or letter is then sent to the App again via UDP socket.

Finally the App after receiving the word displays it as shown in Screenshots of Mobile App in Fig 5. In order for word to be heard, we used Google Text to Speech library [2] to pronounce the word. In this way the word can be seen as well as heard as a speech.

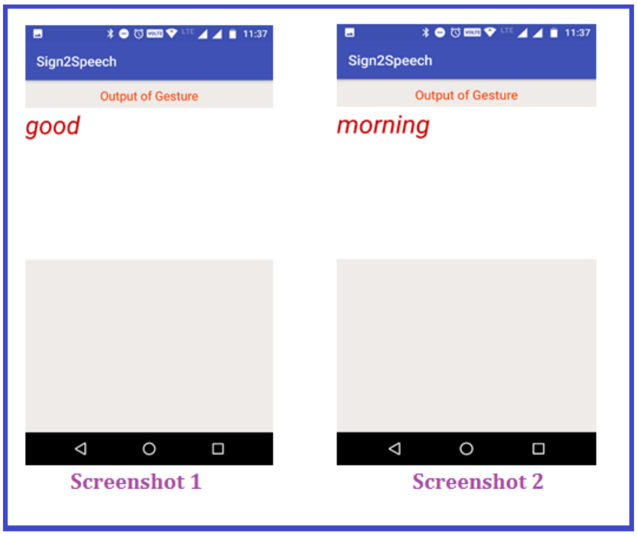

Fig. 5 Displaying output text (word) on Mobile App

IV. EXPERIMENTAL RESULTS

Currently the System has been trained for 15 generic introductory words specified in Indian Sign language, which can be seen in the Fig 6. For every word we have trained the system for 5 times. Performing the gestures for 5 times forms the training data sets for Multilayer Perceptron, from which a trained model is generated. When the gesture is performed at the real-time it is recognized by using the trained model. The generic meaning that is specified in Sign

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 9, September 2017

Language can be assigned by the user while training the system or any other meaning as they desire for a specific gesture. So that when gesture is performed a unique output speech is obtained, as specified while training.
The accuracy achieved for these 15 words is up to 100% which can be seen with help of confusion matrix in Fig 6.

```
=== Confusion Matrix ===

 a b c d e f g h i j k l m n o   <-- classified as
 3 0 0 0 0 0 0 0 0 0 0 0 0 0 0 | a = friends
 0 3 0 0 0 0 0 0 0 0 0 0 0 0 0 | b = 2
 0 0 3 0 0 0 0 0 0 0 0 0 0 0 0 | c = my
 0 0 0 3 0 0 0 0 0 0 0 0 0 0 0 | d = name is
 0 0 0 0 3 0 0 0 0 0 0 0 0 0 0 | e = acp pradyumna
 0 0 0 0 0 3 0 0 0 0 0 0 0 0 0 | f = morning
 0 0 0 0 0 0 3 0 0 0 0 0 0 0 0 | g = good
 0 0 0 0 0 0 0 3 0 0 0 0 0 0 0 | h = you
 0 0 0 0 0 0 0 0 3 0 0 0 0 0 0 | i = meet
 0 0 0 0 0 0 0 0 0 3 0 0 0 0 0 | j = nice
 0 0 0 0 0 0 0 0 0 0 3 0 0 0 0 | k = how are you
 0 0 0 0 0 0 0 0 0 0 0 3 0 0 0 | l = hello
 0 0 0 0 0 0 0 0 0 0 0 0 3 0 0 | m = thank you
 0 0 0 0 0 0 0 0 0 0 0 0 0 3 0 | n = r
 0 0 0 0 0 0 0 0 0 0 0 0 0 0 3 | o = we
```

Fig. 5 Confusion Matrix

The confusion matrix is plotted on the basis of testing performed of the system. Testing was performed by unseen dataset for each word among the 15 words. As we can see in Fig 6 only diagonal elements are present. So we are getting up to 100% accuracy.
Since the dataset consist of only 15 words, accuracy attained is up to 100%. This accuracy will vary when much more words are added to the system.

V. CONCLUSION AND FUTURE SCOPE

The System thus converts Indian Sign Language into corresponding speech as well as text, allowing speech impaired to communicate with others using Sign Language. The system currently has a vocabulary of 15 introductory words with their generic hand gestures as specified in Indian Sign Language and attains accuracy up to 100% as discussed earlier. Speech impaired people can train the system for any hand gesture and bind it to its generic meaning or any other word as per their wish. But care has to be taken that the hand gesture corresponds either to its generic meaning or to the meaning decided for it by the user. In this way a customization features is provided by the system with the use of machine learning where user can decide the meaning for any hand gesture.
The System is currently being tested and used for recognition of 15 words only. Much more words can be added to the System. This will also affect the accuracy for recognition of words. System has to be developed further in such way that it balances nature between number of words and their accuracy.
There is a little lag in between performing of hand gesture and getting its corresponding output speech due to the required processing in recognition of hand gesture. The system can be improved to minimize this lag.
Currently the System is developed for recognition of letters or words conveyed by using only one hand. Further this can developed for recognition of hand gestures made by using both the hands, thereby increasing the range of hand gestures that can be utilized by the user.

REFERENCES

[1] Lam T. Phil, Hung D. Nguyen, T.T. Quyen Suil, Thang T. Vul, A Glove-Based Gesture Recognition System for Vietnamese Sign Language, 2015 15th International Conference on Control, Automation and Systems (ICCAS 2015), Oct. 13-16,2015 in BEXCO, Busan, Korea.
[2] Vishal Pathak, Sushant Mongia, Gaurav Chitranshe, A Framework for Hand Gesture Recognition based on fusion of Flex, Contact and Accelerometer Sensor, 2015 Third International Conference on Image Information Processing.
[3] Mrs.NeelaHarish , Dr.S.Poonguzhali, Design and development of hand gesture recognition system for speech impaired people, 2015, International Conference on Industrial Instrumentation and Control (ICIC) ,College of Engineering Pune, India. May 28-30, 2015.
[4] Mohammed Elmahgiubi, Mohamed Ennajar, Nabil Drawil, and Mohamed Samir Elbuni,Sign Language Translator and Gesture Recognition.
[5] D. Xu, W. Yao, and Y. Zhang, "Hand gesture interaction for virtual training of SPG," in Proc. Int. Con! Artif. Reality Telexistence, pp. 672-676, 2006.
[6] W. Gao, J. Ma, J.Wu, and C.Wang, "Sign language recognition based on HMM/ANN/DP," Int. J Pattern Recognit. Artif Intell., vol. 14, no. 5, pp. 587-602, 2000.
[7] M.E. Cabrera, I.M. Bogado, L. Fermin, R. Acufia, and D. Ralev, "Glove based gesture recognition system, " Proc. of the 15th Int. Con! On Climbing and Walking Robots and the Support Technologies for Mobile Machines, Baltimore, MD, USA, pp. 747-753, 20 12.
[8] http://www.analog.com/en/products/mems/mems-accelerometer/adxl1345.html.