Gesture To Speech Conversion using Flex sensors, MPU6050 and Python

Vaibhav Mehra, Aakash Choudhury, Rishu Ranjan Choubey

Abstract: *Communicating through hand gestures is one of the most common forms of non-verbal and visual communication adopted by speech impaired population all around the world. The problem existing at the moment is that most of the people are not able to comprehend hand gestures or convert them to the spoken language quick enough for the listener to understand. A large fraction of India's population is speech impaired. In addition to this communication to sign language is not a very easy task. This problem demands a better solution which can assist speech impaired population conversation without any difficulties. As a result, reducing the communication gap for the speech impaired. This paper proposes an idea which will assist in removing or at least reducing this gap between the speech impaired and normal people. The research going on this area mostly focuses on image processing approaches. However, a cheaper and user-friendly approach has been used in this paper. The idea is to make a glove that can be worn by the speech impaired people which will further be used to convert the sign language into speech and text. Our prototype involves Arduino Uno as a microcontroller which is interfaced with flex sensors and accelerometer, gyroscopic sensor for reading the hand gestures. Furthermore, to perform better execution, we have incorporated an algorithm for better interpretation of data and therefore producing more accurate result. Thereafter, we use python to interface Arduino Uno with a microprocessor and finally converting into speech. The prototype has been calibrated in accordance with the ASL (American Sign Language).*

Keywords : Arduino Uno, Flex sensors, MPU6050, Python, Text to speech.

I. INTRODUCTION

The ability to speak or the power of speech is something we take for granted. It is the most effective and powerful way of sharing our thoughts and emotions. However, a large fraction of the world population is not blessed with this ability. According to the census of India, approximately 1.6 million people in India are speech impaired [1] and a very small fraction of the entire population are comfortable in communication through sign language. The speech-impaired people have a difficult time conveying their thoughts to others. There are millions of speech impaired people and only a handful of sign language translators.

Our project aims at removing these barriers and providing the speech impaired people a method of communication so that they can communicate their thoughts and feelings effectively and independently and not just in the presence of a translator. The proposed model is based on a flex sensor mounted on a glove and interfaced with an Arduino Uno.

Extensive research has gone into the development of technology to achieve gesture to speech conversion. Gesture to speech conversion can be thought of as a two-pronged approach. The first method is using flex sensors and wearable technologies and the second approach is using Image Processing and training algorithms. MATLAB has been used to perform image processing by using image segmentation and with the help of feature extraction algorithm, the system recognizes the image of the hand gesture captured by the camera and converts it to speech from a list of pre-recorded tracks [2]. After image processing, Machine Learning and Template Matching can be used to get the desired text to speech conversion [3]. In another work, speeded up robust features algorithm has been used in combination with a kinetic sensor to perform static gesture recognition. This work has been done mainly for Mandarin and Tibetan languages in accordance with specified semantic rules [4]. The Mandarin-Tibetan bilingual speech synthesizer has also been used in [5], along with a Deep Neural Network Model and a Support Vector Machine to classify facial expressions along with hand gestures so that the system can add emotion to the speech that it generates. A recognition rate of 80.4% was achieved in [6] which used Hidden Markov Models along with Taiwanese Sign Language to implement recognition of 51 different fundamental postures and 6 different orientations. Contour Analysis and feature extraction can be used with the OpenCV library [7]. Velostat which is a low price packaging material has been used in [8] to make sensors to measure bend of fingers. The data from the sensor is then mapped onto a character set using a Minimum Mean Square Error machine learning algorithm. Aishwarya V proposed a system which works with accelerometer and gyroscope to measure the orientation of fingers in hand gestures and combined it with the Hidden Markov model for use mainly for Tamil language [9]. A PIC microcontroller can be used with flex sensor and accelerometer along with an amplifier to boost the voltage inputs from a flex sensor, and then finally a speaker to produce sound [10]. The design of a wearable band has been discussed in [11]. The band houses EMG sensors which measure the biopotential in the muscles moving the hands. Gyroscope and accelerometer have also been used to measure the orientation of the hands. Finally, machine learning and fuzzy logic have been used to identify the gesture and convert to speech. Flex sensor in combination with tactile sensors and accelerometer has been used to measure the orientation of fingers and an HMM-based model is used for final conversion to speech in [12].

Another approach involves the usage of gloves incorporated with sensors and interfaced with a Bluetooth module and smartphone app to achieve gesture to speech conversion [13]. IMU and flex sensors have been used in [14] to track the movement of fingers in three-dimensional space. The data is fed to Raspberry Pi which gives input to the speaker. The particular sound or text to be pronounced can be also be kept in a database [15]. A similar approach has been used in [16] but with an ARM7 microcontroller, LCD and speaker. Another software-based approach is by using Microsoft Kinect v2. It takes the image of the person performing the gestures and compares them against pre-stored gestures in the database and mapping a keyword associated with the best match [17-18].

II. METHODOLOGY

The paper proposes a method to read the sign language and convert it into text and speech. The prototype developed included various components such as Arduino Uno, flex sensors, MPU6050, etc

A. Arduino Uno

Arduino Uno is a microcontroller based on the chip ATmega328. It works on an operating voltage of 5 V with the cut off limit for input voltage being 6-20V. It consists of 14 digital I/O pins and 6 analog input pins. The 5 analog input is connected to the five different flex sensors placed on the fingers of the glove, we can get the reading which varies continuously with the help of these pins. It has a clock speed of 16Mhz which assist us in processing the information without facing any difficulties.

B. Flex sensors

Flex sensors is a variable resistor that can be used to detect the variation in resistance as per the bend in the sensor. The flex sensors are mounted on the fingers and thumb of each hand. These are directly connected to the microcontrollers on both the arm and are the base of gathering information regarding the human gesture at a particular moment.

C. MPU6050

MPU6050 is an accelerometer gyroscope sensor which is used to detect the alignment of an object. It is also connected to the microcontroller and assist in differentiating words that have the same gesture but different alignment as per the American Sign Language(ASL) shown in fig 1. For example, the alphabets H and U cannot be differentiated with the help of flex sensors alone.

Fig1. American Sign Language [19]

The following steps are involved in the algorithm:

Step 1 - Data from the MPU and flex sensor will be continuously taken for 5 loop cycles for a particular gesture.

Step 2 - Values received will be averaged to get the mean value of each flex sensor. Then standard deviation is calculated and the range for a particular gesture will be set as (average value - (standard deviation/2)) to (average value + (standard deviation/2)).

Step 3 - The output of the gloves gives values for each flex sensors which are rechecked with the database and it is then converted into text form if the values are accurate otherwise we provided a feature of predicting words but it is only viable when the person carries a smart-phone at the particular moment.

Step 4 - The microcontroller then sends the data to the microprocessor by serial communication. The data is then stored in the kernel and is hard coded in python to get the desired output.

Step 5 - Google Text To Speech (GTTS) library in order to attain a maximum efficiency of the model. As the data is stored in the kernel we now access the speaker of the device and convert the text to speech using GTTS and the output is achieved in the form of speech.

Step 6 – The translated sentence or words can be sent as an SMS using an online service named way2sms. The person's mobile number needs to be given as an input for the process to advance.

Multiple iterations of calibration tests were organized to attain maximum efficiency. Table I and fig 3 shows the growth in the overall efficiency of the model. Initially for the first 10 iterations the attained accuracy was low which increased gradually for higher number of iterations.

Table- I: Percentage accuracy

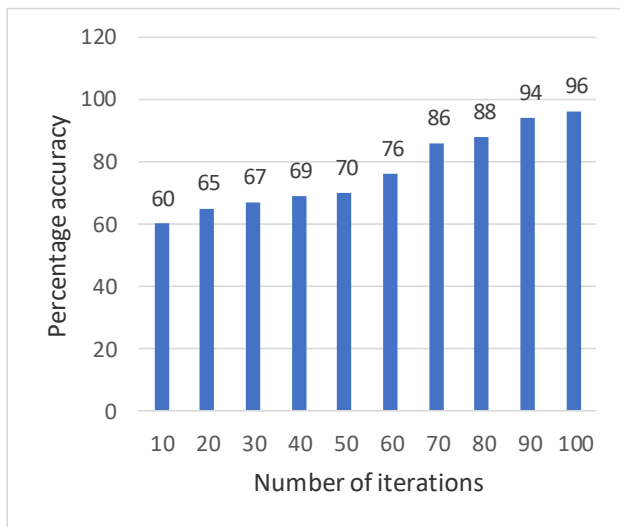| No. of iterations	Results		
	Correct	*Incorrect*	*% accuracy*
10	6	4	60
20	13	7	65
30	20	10	67
40	28	12	69
50	35	15	70
60	46	14	76
70	60	10	86
80	70	10	88
90	85	5	90
100	96	4	96

Fig 2. Flowchart

D. Flowchart

The whole hardware is integrated on a single glove. The values of the original sign language are stored after the calibration of flex sensors, these values have been tested many times in order to maintain high accuracy and persistence. But as for different individual the values will be different because of the extent by which they can flex the sensor. So, to overcome this we thought of an algorithm which would initially take input from the user for the words/letter they want to express via the glove. Fig 3 shows the flowchart of the designed system.
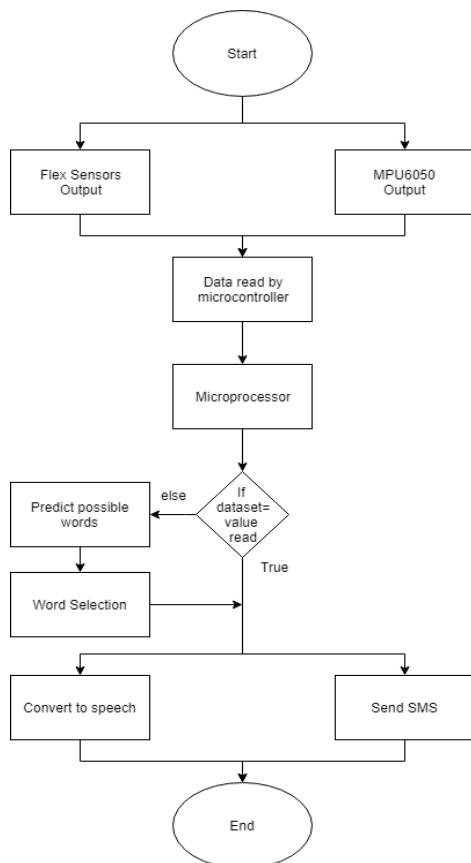

Fig 3. Plot of accuracy vs number of iterations

III. RESULT

The bending of the fingers causes a change in resistance and the MPU computes the orientation of the hand. Depending on calibration, alphabets are displayed on the Arduino Serial Monitor as shown in fig 4. The readings from the flex sensor and MPU are shown in fig 5.

The output is in the form of a sound and the accent can be varied as per the user demand. In case of misunderstanding between the machine and the user, we employed a method of predicting text for the user based on the commonly used words worldwide making it easy to communicate.

The output is also sent to the receiver in the form of a text message. This allows the person to communicate with people who are deaf and not fluent in sign language, therefore our prototype assists in eradicating multiple problems with a single device. The screenshot of the message received is shown in fig 6.

Fig 4. Serial Monitor Output

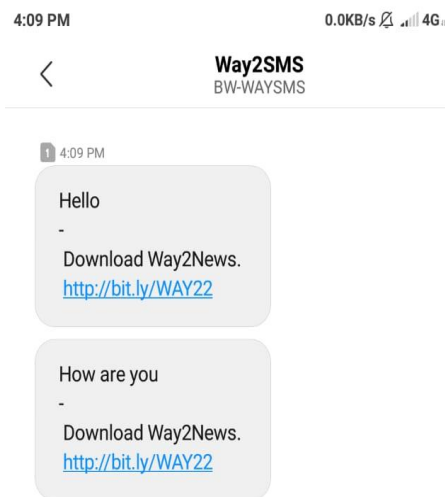**Fig 5. Flex sensor and MPU outputs**

Fig 6. Screenshot of SMS
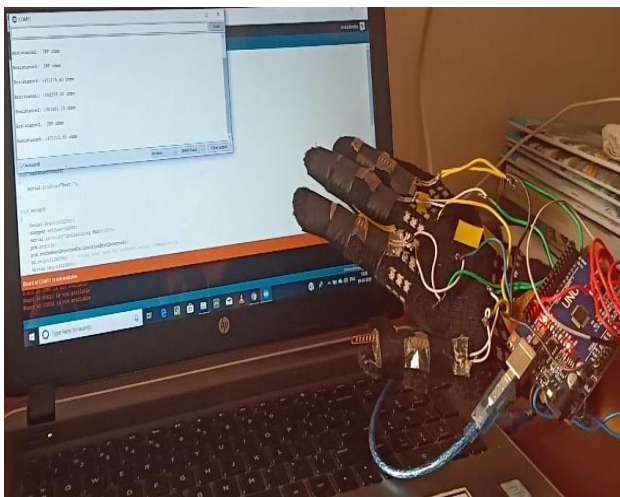
IV. DISCUSSIONS

Fig 7. Final Prototype

The devices existing in the market are generally based on hand gesture recognition using image processing but developing such a model would cost a lot and is not very convenient to use as it involves the use of a camera and someone always has to hold up a camera for that method to work. Compared to that our prototype provides a more convenient and more efficient method which assists in eradicating the problem for both the hearing and speech impaired people. It is a smart prototype that can be used by people in their daily life hassle-free.

V. CONCLUSION

The prototype allows the speech impaired people to communicate with the world giving them an opportunity to express themselves to the world without any difficulties. The flex sensor in combination with the MPU is successfully and accurately able to translate ASL to speech. By mounting these sensors on a glove, a very convenient to use wearable is made which is not only efficient but also comfortable to use in our daily lives. By using python a reliable and accurate system is developed and the messaging service makes the system even more user friendly. It provides an efficient method of

alleviating the problems of the speech impaired community. It empowers such people with the power of speech and allows them to express themselves better.

Not only can the glove be used to convert gesture to speech, but by carefully tracking the movement of the palm and fingers with even more accuracy it can also be used to perform some delicate tasks in locations inaccessible to humans but can be accessed by robots. The human specialist can wear the glove and move his fingers as if the perform the task and the robot can mimic the movements. It can be concluded that the existing methods assisted in the development of our prototype but inclusion of new technology will surely lay a major impact in this field.

REFERENCES

1. http://censusindia.gov.in/Census_And_You/disabled_population.aspx
2. Itkarkar, R.R. and Nandi, A.V., 2013, July. Hand gesture to speech conversion using Matlab. In 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT) (pp. 1-4). IEEE.
3. Pramada, S., Saylee, D., Pranita, N., Samiksha, N. and Vaidya, M.S., 2013. Intelligent sign language recognition using image processing. IOSR Journal of Engineering (IOSRJEN), 3(2), pp.45-51.
4. Yang, H., An, X., Pei, D. and Liu, Y., 2014, September. Towards realizing gesture-to-speech conversion with a HMM-based bilingual speech synthesis system. In 2014 International Conference on Orange Technologies (pp. 97-100). IEEE.
5. Song, N., Yang, H. and Wu, P., 2018, May. A Gesture-to-Emotional Speech Conversion by Combining Gesture Recognition and Facial Expression Recognition. In 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia) (pp. 1-6). IEEE.
6. Liang, R.H. and Ouhyoung, M., 1998, April. A real-time continuous gesture recognition system for sign language. In *Proceedings third IEEE international conference on automatic face and gesture recognition* (pp. 558-567). IEEE.
7. Manikandan, K., Patidar, A., Walia, P. and Roy, A.B., 2018. Hand Gesture Detection and Conversion to Speech and Text. arXiv preprint arXiv:1811.11997.
8. Preetham, C., Ramakrishnan, G., Kumar, S., Tamse, A. and Krishnapura, N., 2013, April. Hand talk-implementation of a gesture recognizing glove. In *2013 Texas Instruments India Educators' Conference* (pp. 328-331). IEEE.
9. Aiswarya, V., Raju, N.N., Joy, S.S.J., Nagarajan, T. and Vijayalakshmi, P., 2018, March. Hidden Markov Model-Based Sign Language to Speech Conversion System in TAMIL. In 2018 Fourth International Conference on Biosignals, Images and Instrumentation (ICBSII) (pp. 206-212). IEEE.
10. Harish, N. and Poonguzhali, S., 2015, May. Design and development of hand gesture recognition system for speech impaired people. In 2015 International Conference on Industrial Instrumentation and Control (ICIC) (pp. 1129-1133). IEEE.
11. Vishal, D., Aishwarya, H.M., Nishkala, K., Royan, B.T. and Ramesh, T.K., 2017, December. Sign Language to Speech Conversion. In 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)(pp. 1-4). IEEE.
12. Vijayalakshmi, P. and Aarthi, M., 2016, April. Sign language to speech conversion. In 2016 International Conference on Recent Trends in Information Technology (ICRTIT) (pp. 1-6). IEEE.
13. Heera, S.Y., Murthy, M.K., Sravanti, V.S. and Salvi, S., 2017, February. Talking hands—An Indian sign language to speech translating gloves. In 2017 International conference on innovative mechanisms for industry applications (ICIMIA) (pp. 746-751). IEEE.
14. Bhaskaran, K.A., Nair, A.G., Ram, K.D., Ananthanarayanan, K. and Vardhan, H.N., 2016, December. Smart gloves for hand gesture recognition: Sign language to speech conversion system. In 2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA) (pp. 1-6). IEEE.
15. Padmanabhan, V. and Sornalatha, M., 2014. Hand gesture recognition and voice conversion system for dumb people. International Journal of Scientific & Engineering Research, 5(5), p.427.

16. Potdar, P.R. and Yadav, D.D., 2014. Innovative Approach for Gesture to Voice Conversion. International journal of innovative research and development, 3(6), pp.459-462.
17. Ahmed, M., Idrees, M., ul Abideen, Z., Mumtaz, R. and Khalique, S., 2016, July. Deaf talk using 3D animated sign language: A sign language interpreter using Microsoft's kinect v2. In 2016 SAI Computing Conference (SAI) (pp. 330-335). IEEE.
18. Rajaganapathy, S., Aravind, B., Keerthana, B. and Sivagami, M., 2015. Conversation of Sign Language to Speech with Human Gestures. Procedia Computer Science, 50, pp.10-15.
19. https://www.teacherspayteachers.com/Product/American-Sign-Language-Alphabet-Poster-3563143

VI. AUTHORS PROFILE

Vaibhav Mehra is a BTech. EEE student, School of Electrical Engineering, VIT Vellore, India. Publications include Automated Hospital Management System implementing IOT and self-travelling robot using Image Processing and RFID Technology in IJITEE and RFID Based Secured, Remotely Accessible Personal Medical Data Base Including the Medicinal History in IEEE. He is a member of SEDS project team in VIT and has completed many projects.
Hack a thon winner organised by the School of Electrical Engineering at VIT
He has research interests in embedded systems, internet of things, VLSI and FPGA.

Aakash Choudhury is a BTech. Electrical and Electronics Engineering student, School of Electrical Engineering, VIT Vellore, India. He has completed his schooling from The Calcutta Boys' School. Publications include Automated Hospital Management System implementing IOT and self-travelling robot using Image Processing and RFID Technology in IJITEE and has presented paper on GSM based Accelerometer Mounted Accident Detection with Location Tracking and Survivor's Condition Monitoring System at IEEE Discover Conference 2019, which is pending publication. He was a member of Instrument Society of India, VIT branch. He has research interests in embedded systems, internet of things, VLSI and FPGA.

Rishu Ranjan Choubey is a BTech. EEE student, School of Electrical Engineering, VIT Vellore, India. A sports enthusiast for whom learning is exploring the world around and becoming a better person. Interested towards Automation, Control Systems, Robotics and innovation.