GESTURE TO SPEECH WEARBLE GLOVE

J COMPONENT PROJECT REPORT

by

CHATHURA.N.J

HARIKA.K

SAI VARSHITHA.R

SAI SRINIDHI.K

School of Electronics

Vellore Institute of Technology

Vellore

November 2020

Declaration by Authors

This is to declare that this report has been written by us as part of our coursework. No part of the report is plagiarized from other sources. All information included from other sources has been duly acknowledged. We aver that if any part of the report is found to be plagiarized, we shall take full responsibility for it.

HARIKA .K (18BSW0070)
CHATHURA .N.J (18BSW0017)
VARSHITHA.R (18BSW0077)
SAI SRINIDHI K (18BSW0007)

Place: VIT, Vellore

Date: 28th of October, 2020

TABLE OF CONTENTS

1. ABSTRACT

The intention behind proposing this project is to help the paralyzed people ask for assistance with mere hand movement. The device proposed in this project could convert hand gestures to voice commands and also notify the caretakers who are some distance away when the patient is in need. This project is a novel idea because there is no room for error since they are hard coded messages. The patients may or may not be speech impaired. The target audience is the people with Paralysis and needs to bring the attention of the caretakers whenever they are not nearby. The proposed project promises high reliability and speed

2. INTRODUCTION

Paralysis is a loss of strength in and control over a muscle or group of muscles in a part of the body. Most of the time, this is not due to a problem with the muscles themselves. It is more likely due to a problem somewhere along the chain of nerve cells that runs from the body part to your brain and back again. These nerve cells deliver the signals for your muscles to move.

Gesture recognition is the process by which gestures made by the user are used to convey the information. In everyday life, physical gestures are a powerful means of communication for a paralyzed person. A set of physical gestures may compose an entire language, as in sign languages. They can efficiently convey a rich set of facts and feelings. This project makes the modest suggestion that gesture-based input is such a beneficial technique to convey information.

S. Patel, U. Dhar, S. Gangwani, R. Lad, and P. Ahire proposed a system that uses an intrinsic mobile camera for gesture recognition and acquisition; gesture acquired is processed with the help of Algorithms like HSV model-(Skin Color Detection), LargeBlob Detection, Flood Fill and Contour Extraction. The system can recognize one-handed sign representation of the standard alphabets (A-Z) and numeric values (0-9) [1]. R. R. Itkarkar and A. V. Nandi proposed a system consisting of a camera attached to a computer that will take images of hand gestures. Image segmentation & feature extraction algorithm is used to recognize the hand gestures of the signer. According to recognized hand gestures, corresponding pre-recorded soundtracks will be played [2].

S. Vigneshwaran, M. Shifa Fathima, V. Vijay Sagar, and R. Sree Arshika proposed a project non-vision-based technique will be used. Most of the dumb people are deaf also. So the normal

people's voice can be converted into their sign language. In an emergency, messages will automatically be sent to their relation or friends[3].

P. Vijayalakshmi and M. Aarthi proposed a flex sensor-based gesture recognition module to recognize English alphabets and few words and a Text-to-Speech synthesizer based on HMM is built to convert the corresponding text [4]. L. Anusha and Y. U. Devi proposed a system consisting of MPU6050 for sensing gesture movement, Raspberry pi for processing, a three-button Keypad, and a speaker. It is implemented by using a trajectory recognition algorithm for recognizing alphabets. Raspberry pi generates voice output for the text in multiple languages using voice RSS and Microsoft translator [5].

H. S. Kala, S. R. S., S. Pal, U. S. K., and S. Chakma worked on a project that mainly focuses on removing the barrier of communication between the mute community and the people not familiar with the concept of sign language so that the messages that a dumb person is trying to relay is understandable to a person with no knowledge of sign language. The design of the device is based on embedded systems. Flex sensors and microcontrollers are the key components [6].

N. Harish and S. Poonguzhali have proposed a project using flex sensors and an accelerometer. The movements included during gesture representation are rotation, angle tilt, and direction changes. The flex sensor and the accelerometer are incorporated over fingers and wrist respectively to acquire their dynamics, these sensors are fitted over the data glove. These voltage signals will then be processed by the microcontroller and sent to the voice module, where the words voice outputs are stored and play backed equivalent to each word values to produce the appropriate voice words with the help of the speaker [7].

The novelty of our project is in the fact that our proposed system is a hybrid of the previously quoted techniques. It has simplicity and definite results. This device helps paralysis patients. Due to paralysis, patients are not able to move their body parts, and also it is very difficult for them to talk with other people for their need for help. When in need, hand movements will pass on the information to the caretakers. Hence our project will help them to convey their messages to doctors or family members. This project is the solution to the problem because the message is conveyed with the least effort of the finger movement.

This project can also be extended for an instance of an emergency. Here we propose a smart speaking system that helps them in conveying the message to regular people using hand motions or gestures.

3. METHODOLOGY

3.1. BLOCK DIAGRAM

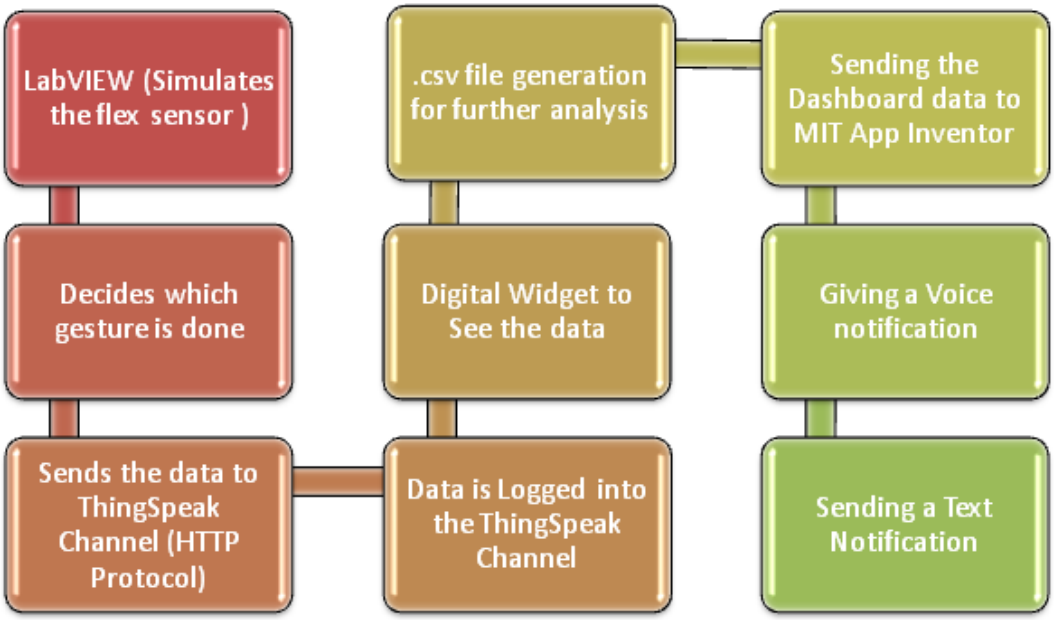


Fig 1: Block Diagram

3.2. WORKING

3.2.1.Flex Sensor

FLEX SENSOR is a variable resistor whose terminal resistance increases when the sensor is bent. So this sensor resistance increases depending on surface linearity. So it is usually used to sense the changes in linearity.

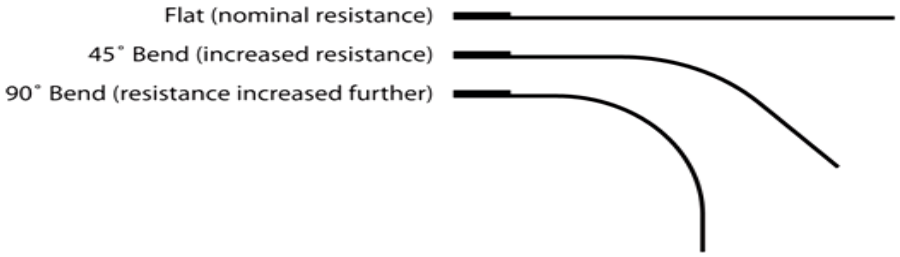


Fig 2: Different Linearity with Flex Sensor

As shown above figure, when the surface of FLEX SENSOR is completely linear it will be having its nominal resistance. When it is bent 45º angle the FLEX SENSOR resistance

increases to twice as before. And when the bent is 90º the resistance could go as high as four times the nominal resistance.

In this project the usual voltage divider is simulated using a case structure , taking into consideration only 3 angles (0,45, 90).

3.2.2 LabVIEW

LabVIEW is a framework plan stage and development environment for a visual programming language.

Instead of using a Voltage divider in the hardware, the software substitute is summation. According to the Sensor Datasheet, the 0º is encoded as 25KOhms, 45º is encoded as 45KOhms and 90º is encoded to be 125KOhms.

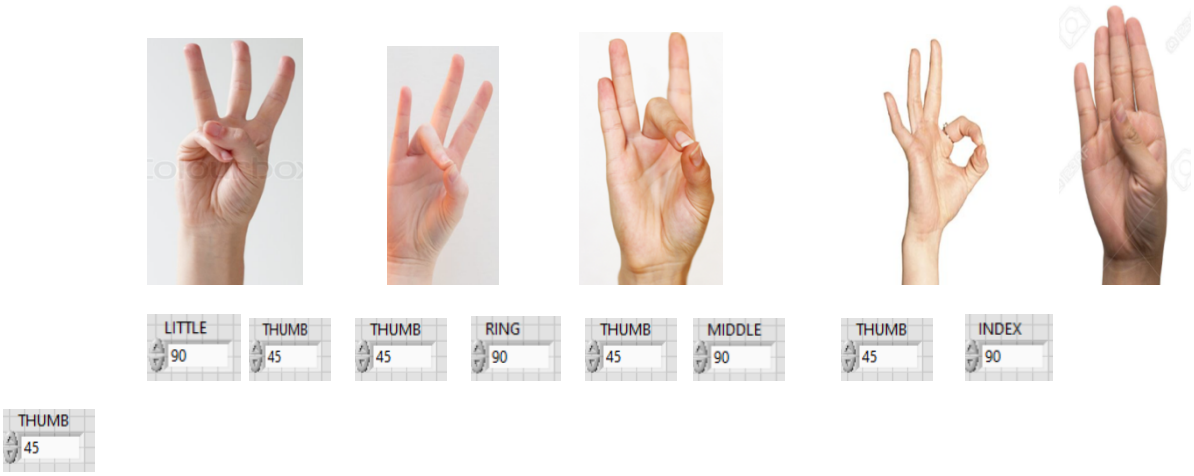


Fig 3: a) The gesture to ask for food. B) The gesture to ask for medicines c) The gesture to ask to go out d) The gesture to ask to use the restroom e) The gesture for an emergency need

In this project, we initially simulate the five flex sensors for all the fingers in LabVIEW. In this software, we have given the input controls as angle values. We have done this project currently for five gestures. The first gesture is for “Food and Water” when the little and thumb fingers are at 90 and 45. The second gesture is for “Medicines” when the ring and thumb fingers are at 90 and 45. The third gesture is for the “Restroom” when the middle and thumb fingers are at 90 and 45. The fourth gesture is for “To go out” when the index and thumb fingers are at 90 and 45. The fifth gesture is for “Emergency” when the thumb is at 90 while all other fingers are at 0. The data is sent at the rate of 1 data

point per second. We have coded the LabVIEW block diagram and sent the data to Thingspeak using the HTTP protocol.

3.2.2 ThingSpeak

IoT application platform that offers a wide variety of analysis, monitoring and counter-action capabilities is 'ThingSpeak'. ThingSpeak is a platform providing various services exclusively targeted for building IoT applications. It offers the capabilities of real-time data collection, visualizing the collected data in the form of charts, ability to create plugins and apps for collaborating with web services, social network and other APIs.

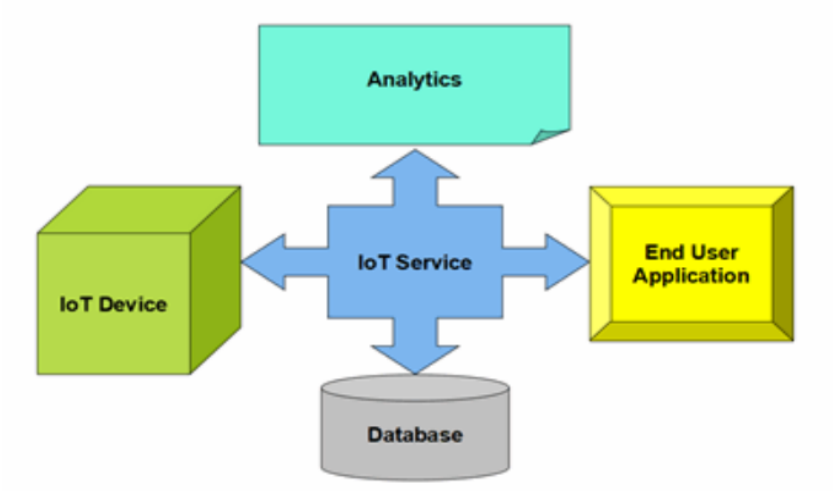


Fig 4: The ThingSpeak architechture

The data reaches the ThingSpeak Channel. Here there is a Digital Numeric Widget that shows the last passed on value. The .csv file is also generated from the ThingSpeak Channel. This can be used for the further usage to train a Machine Learning Model.

3.2.3 MIT App Inventor

MIT App Inventor is an intuitive, visual programming environment that allows everyone to build fully functional apps for smartphones and tablets

The data is sent from the ThingSpeak Channel to the MIT App Inventor. The IOT Dashboard is displayed to the App using a WebViewer Module.

The TextToSpeech Module is used to give an Audio Output. This is implemented into the project because it requires an immediate reaction. Text messages may go unnoticed , but

a Voice Message with high pitch and repetition is hard to miss. Texting Module of the MIT App Inventor is used to send Text messages to a given number.

4. RESULTS AND DISCUSSION

4.1. LabVIEW

For instance giving 45 in RING and 90 in THUMB, generates 150 in the combination of LITTLE and THUMB, MIDDLE and THUMB, INDEX and THUMB and 170 in the combination of RING and THUMB.

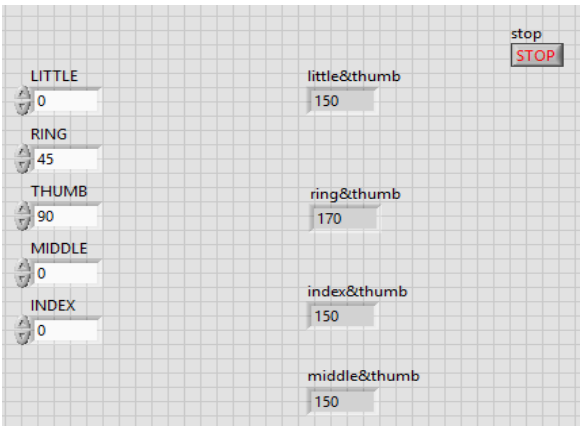


Fig 5: The Input given to the LabVIEW in the form of degrees

This is the output in the front panel of LabVIEW. The fingers are the angle inputs that are given manually. If the project is to be done in hardware the movement of the fingers at a particular angle is taken as the input. These mapped values are summed to give a result. This resultant is again passed on to a case structure to recognize the gesture made by the user.

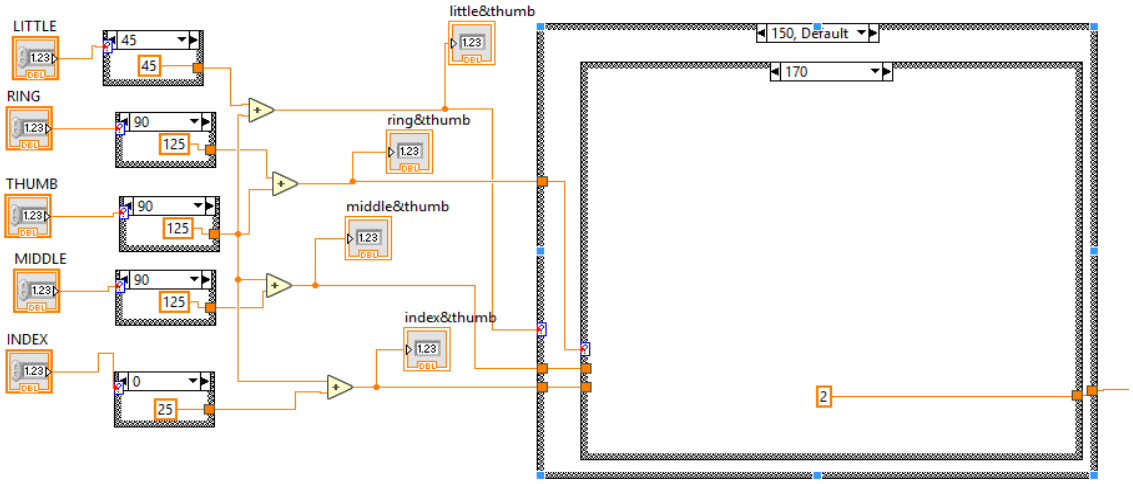


Fig 6: The Case Structure Predicting the gesture value to be 2, since the summation of RING angle and THUMB angle is 170.

This generates the value 2 in the case structure, denoting that the user needs to have his/her "Medicines".

These actions are run inside a while loop with a delay of 1 second. The latency was set to be that because in real time a person can react at least with a time difference of a second. This is sent using the HTTP Protocol to the ThingSpeak Channel.

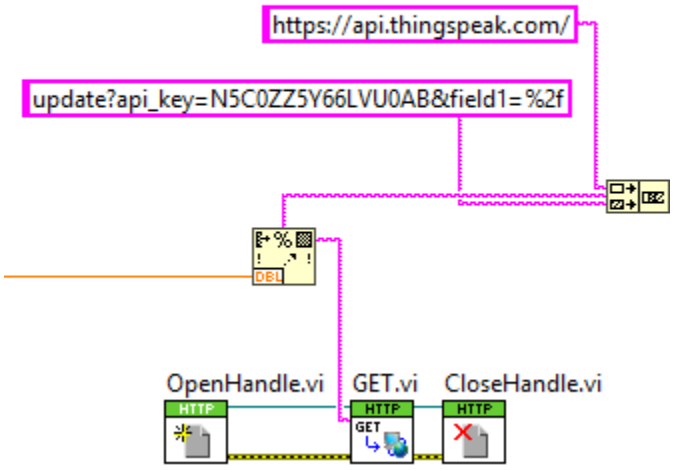


Fig 7: The HTTP protocol in play to send the data to the respective ThingSpeak Channel.

4.2. ThingSpeak

This reaches the ThingSpeak channel which is read as 2 in the Digital Numeric Widget. This is also logged into chart

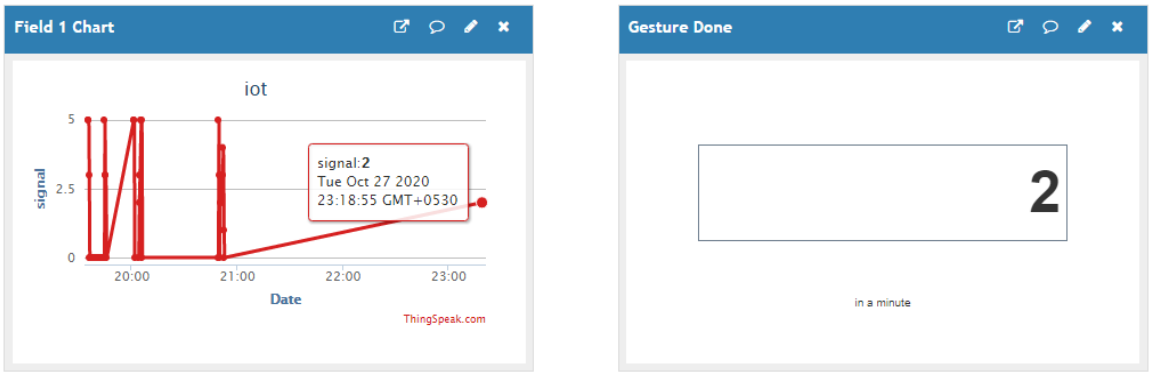


Fig 8: Data Logging graph

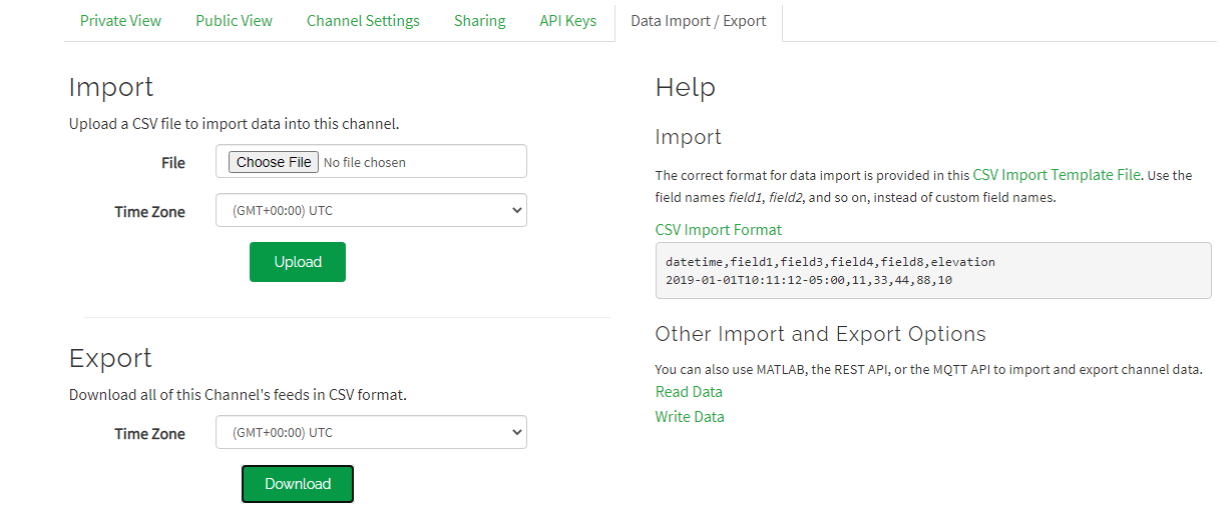


Fig 9: Exporting the .csv file

From the export option , the .csv file can be downloaded in an interval. When the app is used by many users, this data comes in handy to get the.csv file to do data analysis on which action is being done frequently.

A	B	C	D	E	F	G
created_at	entry_id	field1	latitude	longitude	elevation	status
2020-10-25 06:49:41 UTC	1	0				
2020-10-25 06:49:56 UTC	2	1				
2020-10-25 06:50:12 UTC	3	1				
2020-10-25 06:50:28 UTC	4	3				
2020-10-25 06:50:43 UTC	5	3				
2020-10-25 06:50:59 UTC	6	3				
2020-10-25 06:51:15 UTC	7	3				
2020-10-25 06:51:30 UTC	8	3				
2020-10-25 06:51:46 UTC	9	3				
2020-10-25 07:15:53 UTC	10	3				
2020-10-25 07:16:08 UTC	11	3				
2020-10-25 07:16:23 UTC	12	3				
2020-10-25 07:16:38 UTC	13	3				
2020-10-25 07:16:54 UTC	14	2				
2020-10-25 07:17:09 UTC	15	2				

Fig 9: The .csv file generated from the with the records from the day of creation

4.3. MIT App Inventor

HOMEPAGE

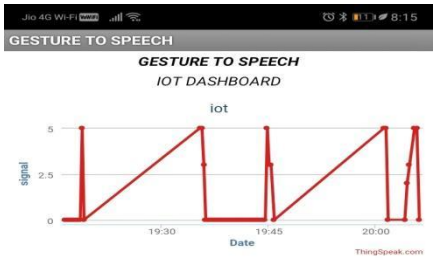


Fig10: Home Page of the App created using MIT App Inventor

The Home Page of the App looks like this with the IOT Dashboard that is previously stored. The Outputs from the App looks like this. Here each time the dashboard gets updated.

SCREENSHOTS OF THE ALERTS

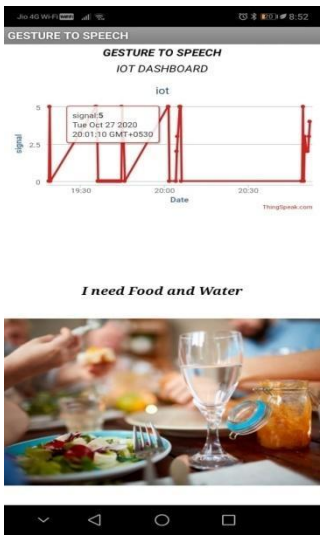


Fig 11: Alerts from the App

For the example quoted the action is “Medicines Required”. That image pops up and The Audio Message “I need Medicines” speaks up.

Fig 12: Example for the instance

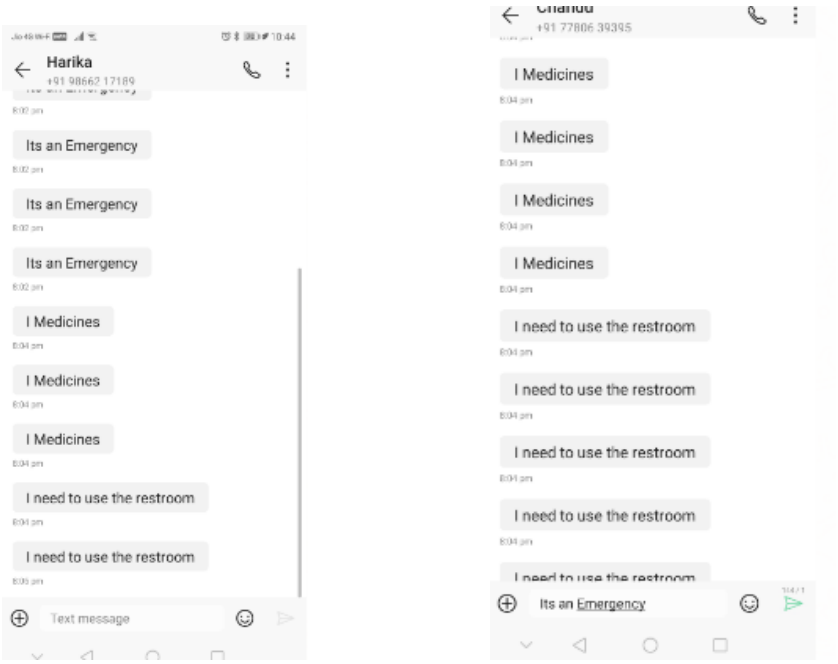


Fig 13: Message Notification

Whenever the ThingSpeak channel is updated with the gesture, this is captured by the MIT APP Inventor. Along with the speech output, there is a message alert also sent using the Texting Module. The module has been updated by the MIT APP Inventor according to the Google Play Store rules. The message output is included into the project because the caretaker can check on it it even with the requirement of intenet connectivity. He/She can have a log of data of the requirements of the patients

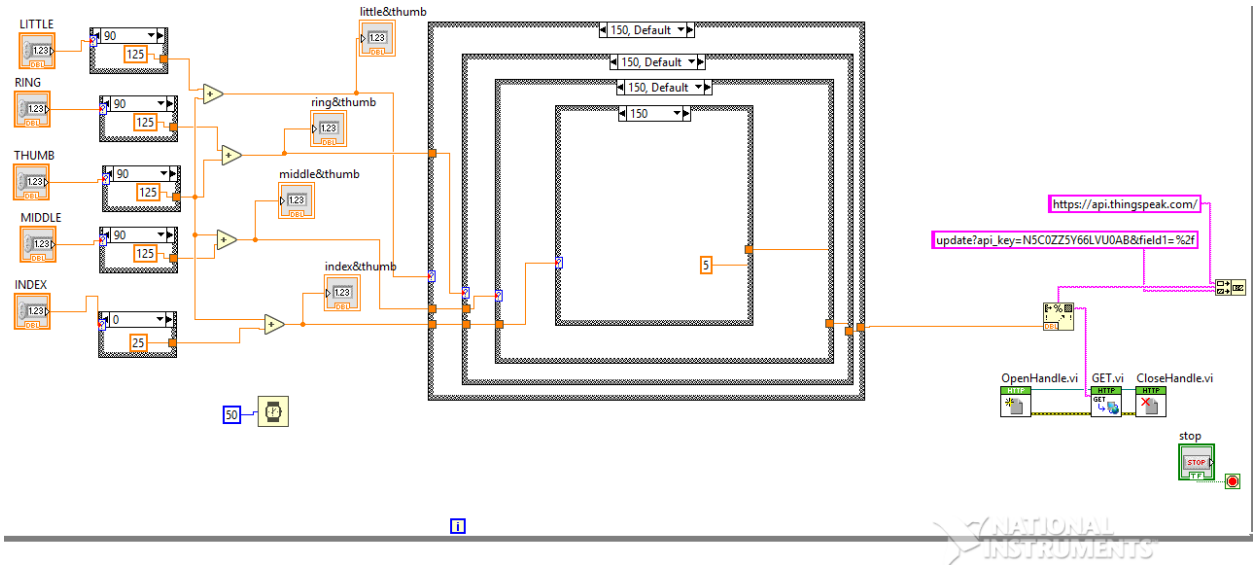
5. CONCLUSION AND RECOMMENDATIONS

ThingSpeak can allow data logging for every 15 seconds. This is an allowable latency in real time, because a need can be sent as an alert with 15 seconds of reaction time. It can be concluded that the project achieves in successfully delivering the message of a paralyzed person to a remote caretaker. This is a useful project because 0.6% of the Indian population alone suffers from Paralysis (Movement disorder) according to 2011 statistics [8]. This will be of great help to them as well as an additional help to Women Safety. If in danger mere movement of hand will send the emergency alert and we can also encorporate the Location data, that will be of extreme help.

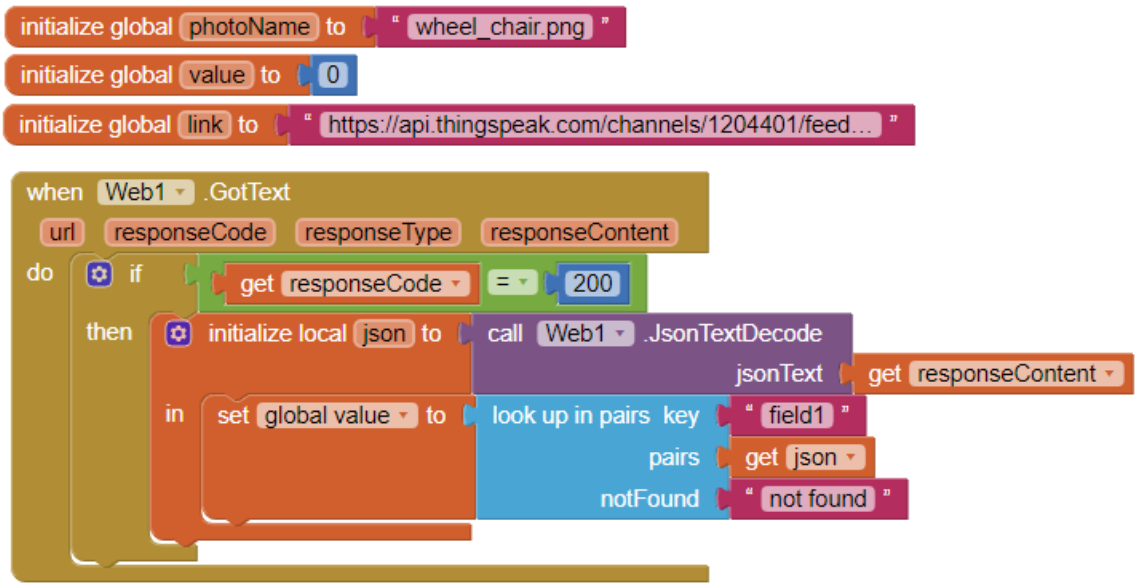
Recommendation can be made to use MQTT Protocol instead of HTTP because the speed of Data Transfer. This can be converted into a hardware project using the real flex sensors with any microcontroller to give the same results. But there the angles may not be accurate as the software simulation, thus an offset of +/- 5 to 10 degrees should be set (according to the datasheet). The .csv file generated from ThingSpeak Channel can be used to apply any ML algorithm to do data analysis on the obtained information. The app being used by many users, will generate a lot of data. This can then be logged to analyse the most required action at a particular moment of time. The number of gestures is currently limited to 5, though it can be incremented to about 10, if the person to extend this project wishes to have the same hard coded outputs. Else if the person wishes to have dynamic output, then he/she needs to add an IMU sensor to get the spatial coordinates of the hand movement, and can have different movement for each alphabet. This will increase the efficiency and versatility of the current project. We will to take this up as the future work to enhance our project and publish a paper on the same.

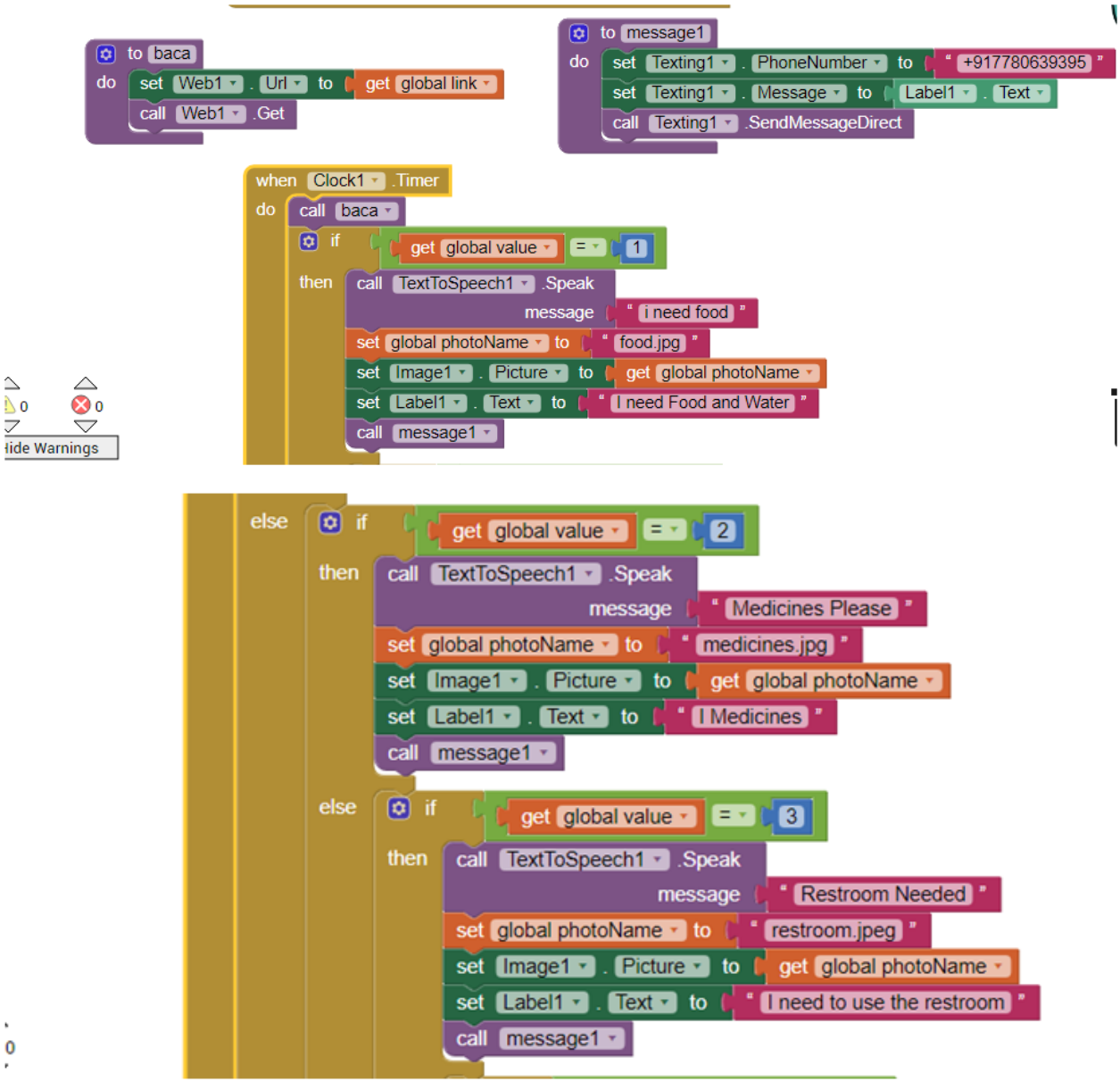
6. APPENDIX

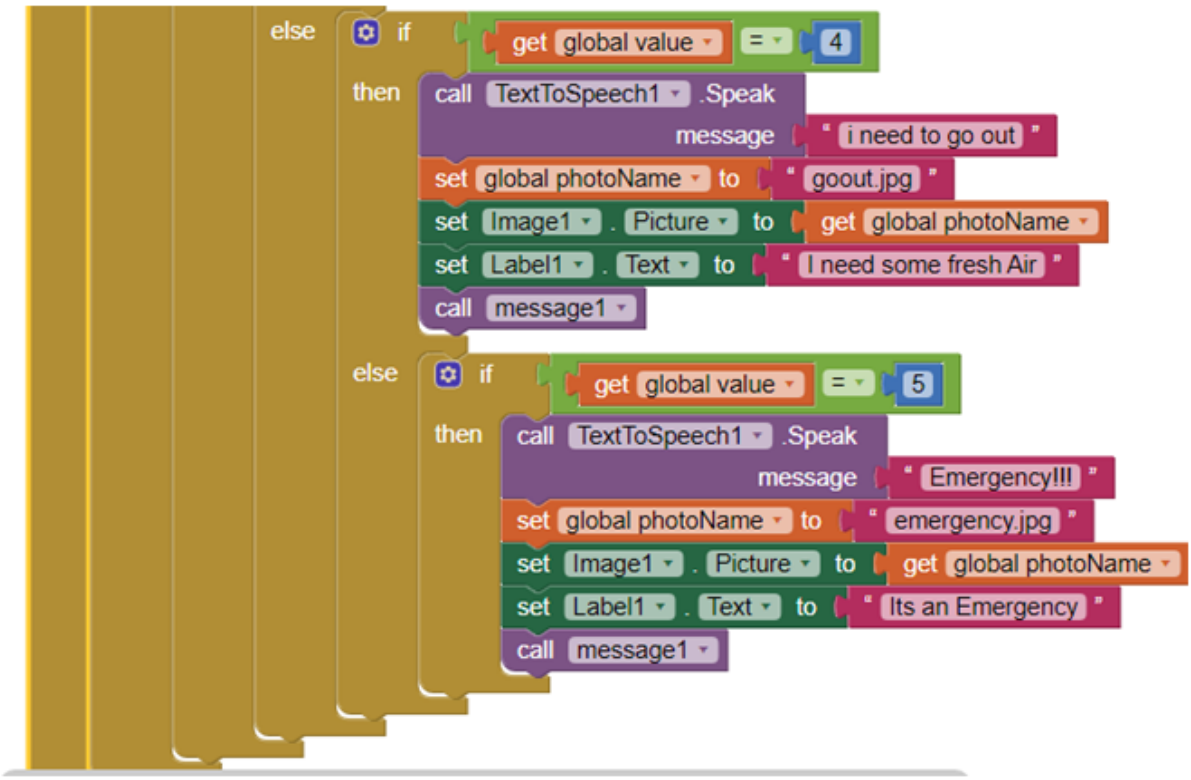
LabVIEW CODE



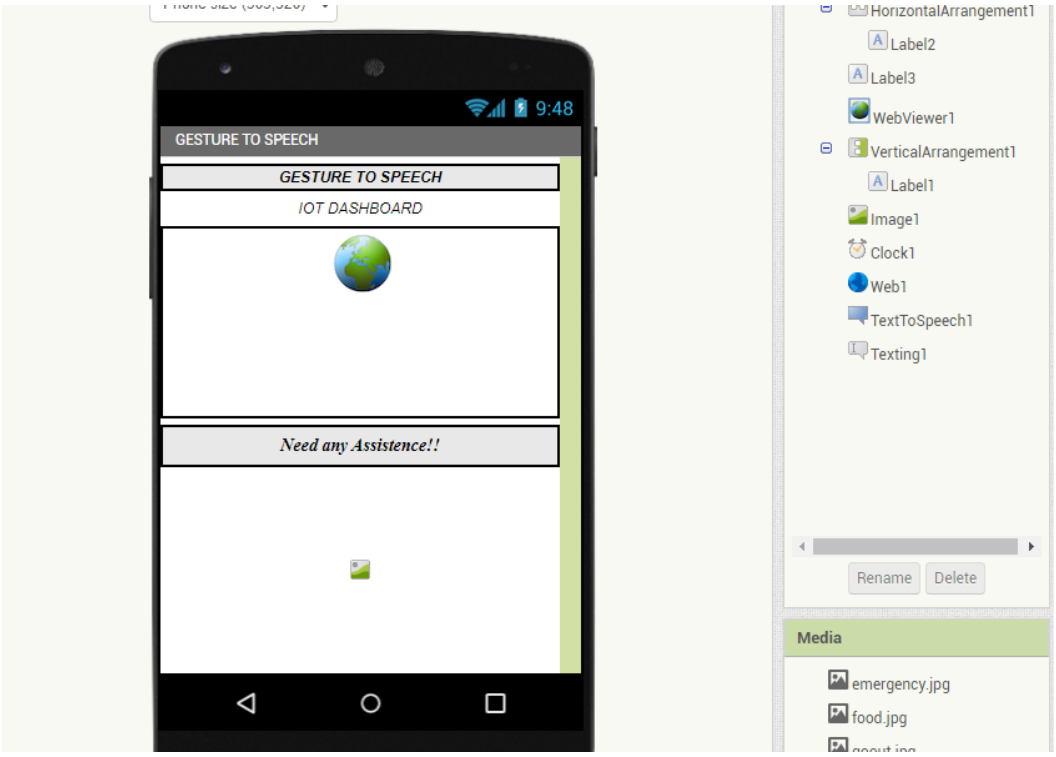
MIT App Inventor CODE




to baca
do set Web1 . Url to get global link
call Web1 .Get
to message1
do set Texting1 . PhoneNumber to " +917780639395 "
set Texting1 . Message to Label1 . Text
call Texting1 .SendMessageDirect
when Clock1 .Timer
do call baca
if get global value = 1
then call TextToSpeech1 .Speak
message " i need food "
set global photoName to " food.jpg "
set Image1 . Picture to get global photoName
set Label1 . Text to " I need Food and Water "
call message1
0
0
ide Warnings
else if get global value = 2
then call TextToSpeech1 .Speak
message " Medicines Please "
set global photoName to " medicines.jpg "
set Image1 . Picture to get global photoName
set Label1 . Text to " I Medicines "
call message1
else if get global value = 3
then call TextToSpeech1 .Speak
message " Restroom Needed "
set global photoName to " restroom.jpeg "
set Image1 . Picture to get global photoName
set Label1 . Text to " I need to use the restroom "
call message1
0



MIT App Inventor Layout



7. REFERENCES

1. S. Patel, U. Dhar, S. Gangwani, R. Lad and P. Ahire, "Hand-gesture recognition for automated speech generation," *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, 2016, pp. 226-231, DOI: 10.1109/RTEICT.2016.7807817.
2. R. R. Itkarkar and A. V. Nandi, "Hand gesture to speech conversion using Matlab," *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)*, Tiruchengode, 2013, pp. 1-4, DOI: 10.1109/ICCCNT.2013.6726505.
3. S. Vigneshwaran, M. Shifa Fathima, V. Vijay Sagar, and R., Sree Arshika, "Hand Gesture Recognition and Voice Conversion System for Dump People," *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, Coimbatore, India, 2019, pp. 762-765, DOI: 10.1109/ICACCS.2019.8728538.
4. P. Vijayalakshmi and M. Aarthi, "Sign language to speech conversion," *2016 International Conference on Recent Trends in Information Technology (ICRTIT)*, Chennai, 2016, pp. 1-6, DOI: 10.1109/ICRTIT.2016.7569545.
5. L. Anusha and Y. U. Devi, "Implementation of gesture-based voice and language translator for dumb people," *2016 International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, 2016, pp. 1-4, doi: 10DOI09/CESYS.2016.7889990.
6. H. S. Kala, S. R. S., S. Pal, U. S. K. and S., Chakma, "Development of Device for Gesture to Speech Conversion for the Mute Community," *2018 International Conference on Design Innovations for 3Cs Compute Communicate Control (ICDI3C)*, Bangalore, 2018, pp. 97-99, DOI: 10.1109/ICDI3C.2018.00029.
7. N. Harish and S. Poonguzhali, "Design and development of hand gesture recognition system for speech impaired people," *2015 International Conference on Industrial Instrumentation and Control (ICIC)*, Pune, 2015, pp. 1129-1133, DOI: 10.1109/IIC.2015.7150917.
8. https://censusindia.gov.in/Census_And_You/disabled_population.aspx