Hand Gesture Recognition and Voice Conversion System

6th Semester Mini Project Report

Submitted to

Department of Electronics and Communication

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, ALLAHABAD

(A centre of excellence in IT, established by Ministry of HRD, Govt. of India)

Project Team

Akash Tiwari [IEC2015020]

Naveen Kewalramani [IEC2015016]

Rachit Parashar [IEC2015081]

Project Supervisor :

Dr. Suneel Yadav

ABSTRACT

Generally deaf-dumb people use sign language for communication, but they find difficulty in communicating with others who don't understand sign language. Due to which communications between deaf-mute and a normal person have always been a challenging task. Gesture recognition refers to recognizing meaningful expressions of motion by a human, involving the hands, arms, face, head, and/or body. We have developed a device which can convert the hand gestures of a mute person into speech. This methodology provides a map for developing a Digital glove which is fitted with Flex sensors. These sensors sense the gestures of a person in the form of bend of fingers by changing the resistance. This system includes a mobile app which can convert signal into real time speech output in multiple language as well as a LCD module to display the text. The text display being in English, the voice output of this device will be multiple languages . So this device acts as a communicator as well as a translator providing more flexibility in communication.

INTRODUCTION

The hand is one of the richest sources of tactile sensory data, enabling precise, and complex manipulation.Human robot interaction systems which are operated by the hand have been actively researched for rehabilitation, virtual reality, entertainment, and so on. For the development of such systems, measurement of unconstrained hand motion should be preceded. hand motion measurement systems have not fully been exploited yet.

Loss of hearing and speech can cause people to become isolated. Sign Language is the only means of communication for deaf people. Sign Language is a well structured code gesture where every gesture has a meaning and conveys a message . With advancement of science and technology many techniques have been developed not only to minimize the problem of deaf and dumb people but also to implement it in different fields. This involves mostly the combination of shapes, orientation and movement of the hands.

We are going to design a system which will translate sign language and at the output we are having playback voice module, such that we can get output in form of sound. Flex sensors will be used to sense the hand gesture. The device designed will portable and user friendly. It will be flexible to any common person.

Methodology:

Stage 1: The data glove is fitted with flex sensors and accelerometer sensors along the length of each finger and the thumb. The flex sensors outputs stream of data that varies with degree of bend.

Stage 2: Degree of bend determines the perceived value of flex sensor data. The voltage drop across flex sensor is mapped to values from 0 to 100. If value is less than 25, it is taken as a ‘0’; if the value is from 26 to 65, it is taken as a ‘1’; if value if greater than 65 , it is taken as ‘2’. In this way a flex sensor data can be interpreted as 3 different values.

Stage 3: We have used 8 flex sensors. 4 for sensing the flex from fingers, other 4 for mode selection and for addition of other variable extension.
Total number of phrases possible = (3)^4 = 81
Used number of phrases = 54
Possible combination sentences = (54) C (2)

Stage 4: The output data stream is fed from sensors to Arduino Uno to convert the analog signal to digital signal.

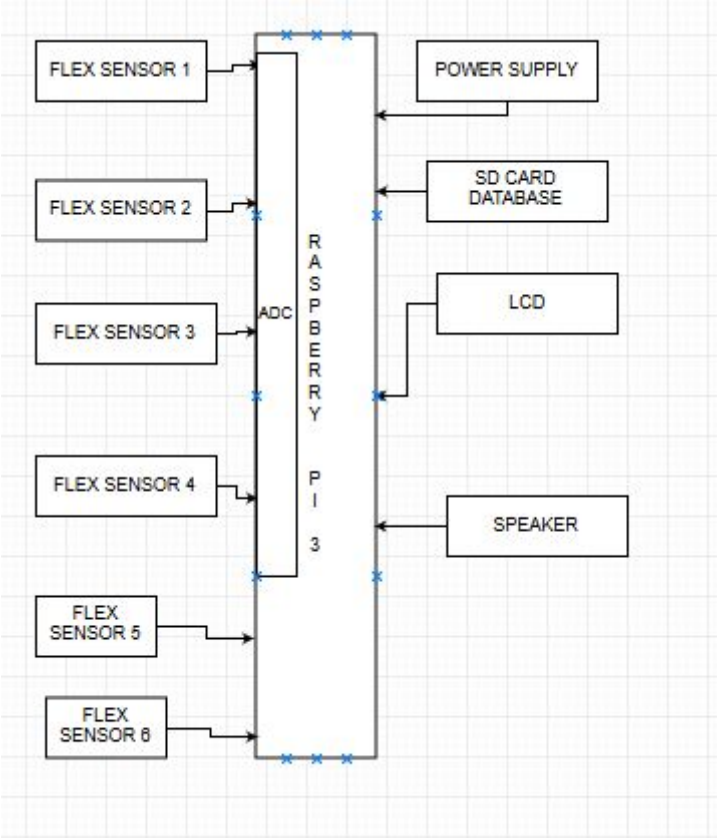
Stage 5: The output data stream from Arduino Uno is fed to Raspberry Pi3 Microcontroller board using USB cable cord.

Stage 6: The microcontroller will compare these readings to the lookup table stored in internal lookup table memory or the database, if it finds data to get matched to any element in dataset in the lookup table, microcontroller will select that phrase.

Stage 7: After this microcontroller will search the SD card for .wav/.mp3 file with similar name and send to speaker.At the same time the same text will display on LCD.

Stage 8: After presenting the audio and text result, system will keep taking inputs from the gloves. If the input data changes it will display it and play it. If any disconnection happens, the LCD will display “Goodbye”.

SYSTEM FLOW DIAGRAM



SOFTWARE TO BE USED

1. **Raspbian OS**
   Raspbian is a Debian-based computer operating system for Raspberry Pi.Raspbian comes pre-installed with plenty of

software for education, programming and general use. It has Python, Scratch, Sonic Pi, Java, Mathematica and more.

2. **Text TO Speech:**
   The IBM Text to Speech service provides an Application Programming Interface (API) that uses IBM's speech-synthesis capabilities to convert written text to natural-sounding speech. The service streams the results back to the client with minimal delay.The service supports voices in the following languages: English (UK and US dialects), French, German, Italian, Japanese, and Spanish .

LANGUAGES USED :

1. **PYTHON** : Python is a general-purpose interpreted, interactive, object-oriented, and high-level programming language. It was created by Guido van Rossum.

2. **SQL** : *SQL* is a standard language for storing, manipulating and retrieving data in databases.

CODE SNIPPETS

READING INPUT

```
 while True:
        read_serial=list(ser.readline().split('\t'))
          read_serial[3] = read_serial[3][0:2]
          mode1 = GPIO.input(17);
          mode2 = GPIO.input(27);
          print (read_serial)
          curs=conn.cursor()
        for row in curs.execute("SELECT
language1,language2,language3,music FROM data where
```

```
"+read_serial[0]+">=sensor1_low and
"+read_serial[0]+"<=sensor1_high and
"+read_serial[1]+">=sensor2_low and "+read_serial[1] + "<=
sensor2_high and "+read_serial[2]+">=sensor3_low and
"+read_serial[2]+"<=sensor3_high and
"+read_serial[3]+">=sensor4_low and
"+read_serial[3]+"<=sensor4_high"):
        print (row)
```

PRINTING ON LCD

```
if(int(mode1)==0):
          if(len(row[0])!=0):
             s1 = row[0][0:16]
             s2 = row[0][16:]
             lcd_string(s1,LCD_LINE_1)
             lcd_string(s2,LCD_LINE_2)
```

PLAYING MUSIC

```
 if(len(row[3])!=0):
              song = '/home/pi/project/english' + row[3];
              pygame.mixer.init()
              pygame.mixer.music.load(song)
              pygame.mixer.music.play()
              while pygame.mixer.music.get_busy()==True:
```

ANALOG TO DIGITAL

```
  const float VCC = 5.0;
  const float R_DIV = 67000.0;

  const float STRAIGHT_RESISTANCE_0 = 45000.0;
  const float BEND_RESISTANCE_0=160000.0;
```

```
 const float STRAIGHT_RESISTANCE_1 = 62000.0;
 const float BEND_RESISTANCE_1=185000.0;

 const float STRAIGHT_RESISTANCE_2 = 40000.0;
 const float BEND_RESISTANCE_2=150000.0;

 const float STRAIGHT_RESISTANCE_3= 30000.0;
 const float BEND_RESISTANCE_3=130000.0;
 void setup()
 {
  Serial.begin(9600);
  pinMode(A0, INPUT);
  pinMode(A1,INPUT);
  pinMode(A2,INPUT);
  pinMode(A3,INPUT);
  int sensor_0=0,sensor_1=0,sensor_2=0,sensor_3=0;
 }

 void loop()
 {
  int flexADC0 = analogRead(A0);
  float flexV0 = flexADC0 * VCC / 1024.0;
  float flexR0 = R_DIV * flexV0/(VCC - flexV0);
  float sensor_0 = map(flexR0, STRAIGHT_RESISTANCE_0,
BEND_RESISTANCE_0,0.0, 100.0);

  int flexADC1 = analogRead(A1);
  float flexV1 = flexADC1 * VCC / 1024.0;
  float flexR1 = R_DIV *flexV1 /(VCC-flexV1 );
  float sensor_1 = map(flexR1, STRAIGHT_RESISTANCE_1,
BEND_RESISTANCE_1,0.0, 100.0);

  int flexADC2 = analogRead(A2);
  float flexV2 = flexADC2 * VCC / 1024.0;
  float flexR2 = R_DIV * flexV2/(VCC - flexV2);
```

```
  float sensor_2 = map(flexR2, STRAIGHT_RESISTANCE_2,
BEND_RESISTANCE_2,0.0, 100.0);

  int flexADC3 = analogRead(A3);
  float flexV3 = flexADC3 * VCC / 1024.0;
  float flexR3 = R_DIV * flexV3/(VCC - flexV3);
  float sensor_3 = map(flexR3, STRAIGHT_RESISTANCE_3,
BEND_RESISTANCE_3,0.0, 100.0);

  Serial.print(sensor_0);
  Serial.print('\t');
  Serial.print(sensor_1);
  Serial.print('\t');
  Serial.print(sensor_2);
  Serial.print('\t');
  Serial.println(sensor_3);


  delay(3500);
 }
```

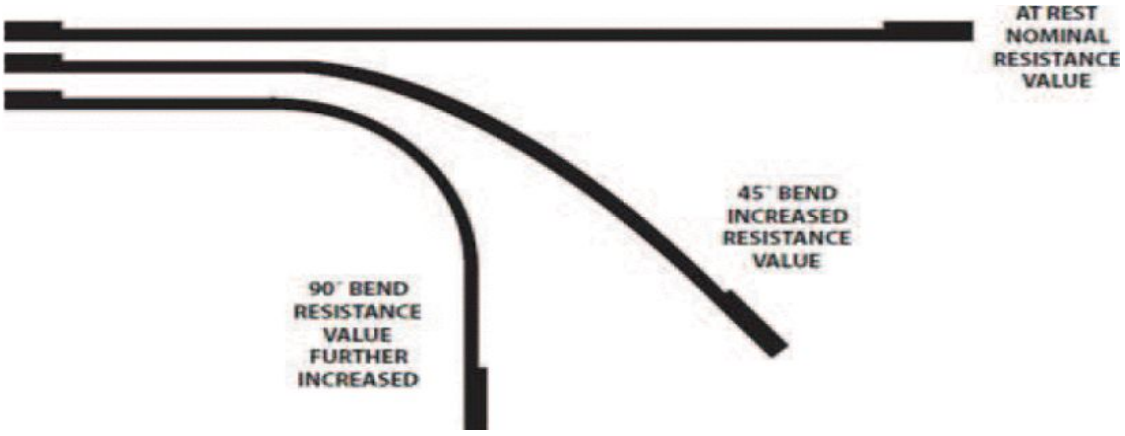
COMPONENTS USED

S.NO	PRODUCT	QUANTITY
1	Raspberry Pi 3.0 B	1
2	Flex Sensor	8
3	Speaker	1
4	Wearable Gloves	2
5	Arduino Uno	1
6	LCD(16X2 character)	1
7	Battery 9V	1
8	Micro SD Card(16GB)	1

COMPONENT DESCRIPTION

1. **RASPBERRY PI:**The Raspberry Pi is a series of small single-board computers .All models feature a Broadcom system on a chip (SoC) with an integrated ARM compatible central processing unit (CPU) and on-chip graphics processing unit (GPU).Processor speed ranges from 700 MHz to 1.2 GHz for the Pi 3; on-board memory ranges from 256 MB to 1 GB RAM. Secure Digital (SD) cards are used to store the operating system and program memory.

2. **Flex Sensors:** Flex sensors are resistive carbon parts. When bent, the device develops a resistance output correlative to the bend radius. The variation in resistance is just about 10k to 30k. A global organization flexed device has 10k resistance and once bent the resistance will increase to 30k at 90 degree. The device incorporates within the device employing a potential divider network. The potential divider is employed to line the output voltage across 2 resistors connected non-parallel.



3. **LCD Display:**A 16x2 LCD means it can display 16 characters per line and there are 2 such lines. In this LCD each character is displayed in 5x7 pixel matrix. This LCD has two registers, namely: Command and Data. The command register stores the command instructions given to the LCD.

4. **Arduino Uno:** Arduino Uno is a microcontroller board based on the ATmega328P (datasheet). It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz quartz crystal.

HARDWARE CONFIGURATION OF FLEX SENSOR

A voltage divider circuit is used with one flex sensor(variable resistance) and one fixed resistor. In case of sign detection the resistance value is-
R = 67 KΩ
Fixed resistance in case of mode conversion is R = 200 KΩ.
this gives a satisfactory variation of voltage drop in the range between 1V to 4V.
Flex sensor gives input which ranges from 0 to 1024 (mapped from 0 to 5 V).

- Maximum voltage drop = 940*(5/1024)
- Minimum voltage = 100*(5/1024)

Now we map this value from 0 - 100.

- from 0-25 , the input is considered ‘0’.
- from 26-65, input is considered ‘1’.
- from 66-100, input is considered ‘2’.

this makes it a ternary valued sensor.
Total number of signs allowed by 4 sensors = 3^4 = 81.

DATASET

	ENGLISH LANGUAGE	FRENCH LANGUAGE	SPANISH LANGUAGE	s1	s2	s3	s4
1	Stop there	Arrêtez là	Detente ahí	0	0	0	1
2	See you later!	À plus tard!	¡Nos vemos más tarde!	0	0	0	2
3	We'll catch up later!	Nous nous rattraperons plus tard!	¡Nos pondremos al día más tarde!	0	0	2	0
4	I'm having a blast.	Je m'emballe	Estoy teniendo una explosión	0	0	1	1
5	We believe in peace	Nous croyons en la paix	Creemos en la paz	0	0	2	2
6	My name is akash	Je m'appelle Akash	Mi nombre es akash	0	1	0	0
7	Excuse me	Excusez-moi	Disculpe	0	1	1	0
8	Good morning	Bonjour	Buenos días	0	1	1	1
9	Where is nearest cafeteria	Où est la cafétéria la plus proche	Dónde está la cafetería más cercana	0	1	1	2
10	Sorry for keeping you waiting!	Désolé de vous avoir fait attendre!	Disculpa por hacerte esperar	0	1	2	2
12	I am fine	je vais bien	Estoy bien	0	2	2	0
13	Help me	Aidez moi	Ayuadame	2	2	2	2
14	Turn on fan	Activer le ventilateur	Activar ventilador	0	2	2	2
15	Thank you	Je vous remercie	Gracias	1	0	0	0
16	Could you please talk slower?	Pourriez-vous s'il vous plaît parler plus lentement?	¿Podrías hablar más lento?	1	0	0	1
17	I need to pee	je dois faire pipi	Necesito orinar	2	2	2	0
18	Where is the bank?	Où est la banque?	¿Dónde está el banco?	1	0	1	1

19	Can you say it again, please	Pouvez vous répéter s'il vous plait	Por favor, puedes decirlo otra vez	2	2	1	0
20	I am from iiit allahabad	Je viens de iiit allahabad	Soy de iiit allahabad	1	1	0	0
21	I'm tired	je suis fatigué	Estoy cansado	1	1	1	0
22	You are welcome	Je vous en prie	De nada	1	1	1	2
23	What do you mean?	Que voulez-vous dire?	¿Qué quieres decir?	1	1	2	2
24	what are you doing?	Qu'est-ce que tu fais?	¿Qué estás haciendo?	1	2	0	0
25	come here	venez ici	ven aca	1	1	1	1
26	What is your name	Viens ici Quel est ton nom?	¿Cuál es su nombre?	1	2	2	0
27	You've got to be kidding me!	Dis moi que c'est une blague!	¡Tienes que estar bromeando!	2	1	1	0
28	yes	Oui	sí	2	0	0	0
29	I am Sorry	Je suis désolé	Lo siento	2	0	0	2
30	You are so precious.	tu es si précieux.	eres preciosa.	2	0	2	0
31	no	non	no	2	1	1	1
32	I am hungry	j'ai faim	estoy hambriento	1	2	2	2
33	When will you get back?	Quand reviendras-tu?	¿Cuándo volverás?	1	2	2	1
35	Come back soon.	Reviens bientôt.	Vuelve pronto.	1	1	0	1
36	I am very happy.	Je suis très heureux.	Estoy muy feliz.	2	0	0	1
38	I don't like this.	Je n'aime pas ça.	No me gusta esto	2	0	1	1
41	I am sad.	Je suis triste.	Estoy triste.	2	0	2	2

42	I am bored.	Je m'ennuie.	Estoy aburrido.	2	1	0	0
45	I am busy.	Je suis occupé.	Estoy ocupado.	2	1	1	2
46	You are right.	Tu as raison.	Tienes razón.	2	1	2	2
47	You are wrong.	Vous avez tort	Está usted equivocado.	2	2	0	0
49	How are you?	Comment allez-vous?	¿Cómo estás?	2	2	0	2
50	Are you serious?	Es-tu sérieux?	En serio?	2	2	1	1
51	I understand.	je comprends	entiendo	2	2	2	1
52	May I help you?	puis-je vous aider?	Puedo ayudarlo?	0	2	0	0
53	God bless you.	Dieu te bénisse.	Dios te bendiga.	1	0	2	0
54	Rest in peace.	repose en paix.	descansa en paz.	1	0	2	2

ACHIEVED RESULTS :

1) The dataset contains 54 phrases, combination of these can create more than 1000 sentences.
2) Minimum number of hand movements required to make a meaningful sentence.
3) Both text display and sound, lets the other people to understand properly, if there is a case of hearing or visual impairment with the receiving person.
4) 3 different language modes for user utility (English /French /Spanish) which can be opted by the user easily with left hand.

5) Easy data creation, update and deletion by the use of Database management system.
6) Real time display and sound lets the user to keep the conversation going.

REFERENCES

1) C. Fahn and H. Sun,“ Development of a Data Glove With Reducing Sensors Based on Magnetic Induction” IEEE Transactions On Industrial Electronics,Vol.No.52 ,
2) K. Tadano, M. Akai,” Development of Grip Amplified Glove using Bi-articular Mechanism with Pneumatic Artificial Rubber Muscle” IEEE International Conference on Robotics and Automation, 2010.
3) A. Karime, H. Al-Osman,” E-Glove: An Electronic glove with Vibro-Tactile feedback for wrist rehabilitation of post-stroke patients” IEEE International Conference,2011