IJCSN International Journal of Computer Science and Network, Volume 4, Issue 1, February 2015
ISSN (Online) : 2277-5420 www.IJCSN.org
Impact Factor: 0.417

Hand Gesture Recognition to Speech Conversion in Regional Language

[1] **B.D.Jadhav,** [2] **Nipun Munot,** [3] **Madhura Hambarde,** [4] **Jueli Ashtikar**

[1]Assistant professor, Department of Electronics & Telecommunication,
JSPM's Rajarshi Shahu College of Engineering, Pune, India

[2,3,4] UG Student, Department of Electronics & Telecommunication,
JSPM's Rajarshi Shahu College of Engineering, Pune, India

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. We propose to develop a device which can convert the hand gestures of a deaf-mute person into speech. This methodology provides a map for developing a Digital wireless glove which is fitted with Flex sensors and accelerometer. These sensors sense the gestures of a person in the form of bend of fingers and tilt of the hand fist. This system includes a voice playback IC to give the real time speech output in regional 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 in regional language (here Marathi). So this device acts as a communicator as well as a translator providing more flexibility in communication.

Keywords - *Gesture Recognition, Sign Language, Flex Sensors, Regional Language, Accelerometer.*

1. Introduction

The development of the most popular devices for hand movement acquisition, glove-based systems started about 30 years ago and continues to engage a growing number of researchers. Communication involves the exchange of information, and this can only occur effectively if all participants use a common language. Sign language is the language used by deaf and mute people that uses gestures instead of sound to convey or to express fluidly a speaker's thoughts. A gesture in a sign language is a particular movement of the hands with a specific shape made out of them. The main aim of this paper is to present a system that can efficiently translate Sign Language gestures to both auditory voice and text. Several languages are being spoken all around the world. So this system aims to give the voice output in various regional languages (here Marathi).

2. Literature Survey

Sign language recognition system mainly have two well known approaches viz. Image processing technique and another is microcontroller and sensor based data glove [5]. These approaches are also known as vision based and sensor based techniques. In the image processing technique camera is used to capture the image/video, in this static images are analyzed and recognition of the image carried out using algorithms that produce sentences in the display. The algorithms used in vision based sign language recognition system are Hidden Markov Mode (HMM), Artificial Neural Networks (ANN) and Sum of Absolute Difference (SAD) [4]. The disadvantage of vision based techniques includes complex algorithms for data processing. Visual based mostly techniques use camera chase technologies, whereby usually the user wears a glove with specific colors or markers indicating individual parts of the hands, specially the fingers. The cameras record the ever-changing image and position of the hand because the user signs and also the pictures are then processed to retrieve the hand form, position and orientation. Another challenge in image and video processing includes variant lighting conditions, backgrounds and field of view constraints and occlusion.

Another approach is using a portable Accelerometer and tactile sensors used to measure the hand gesture [1]. Accelerometer is used to capture movement information of hand and arms. EMG sensor placed on the hand, it generates different sign gesture [3]. Sensor output signals are fed to the computer process to recognize the hand gesture and produce speech/text. In the instrumented approach [2] of sign language recognition instrumented part of the system combines an AcceleGlove and a two-link arm skeleton. The existent systems have used a text-to-speech conversion for voice output. This brings a language constraint into picture as TTS output is in English only.

3. System Architecture and Implementation

The system architecture comprises of two primary sections: Transmitter and Receiver section. The transmitter section is responsible for recognizing the hand gestures of a person. This is done by wearing the gloves which are fitted with flex sensors each of 4.5" on the 4 fingers of the glove whereas the thumb is fitted with a 2.2" flex sensor. The output of flex sensors is in the form of variation in resistance in accordance to the bend of fingers. Also along with this an accelerometer is mounted on the palm side of the glove to sense the tilts of the hand. The outputs of the flex sensors and the accelerometer are directly given to the ADC channels of the microcontroller. The ADC output of each channel is processed and a specified value is assigned to each gesture. To each specified value of the gesture, a digital value is fed to the encoder which converts the parallel data to serial and transmits it through the RF transmitter.Once the serial data is received at the RF receiver this data is converted to parallel form by the decoder and given to the microcontroller. In accordance with the digital value received the microcontroller gives designated SPI commands to the voice playback IC. This voice playback IC produces the pre-recorded voice outputs through the speaker. In this way the hand gestures of a person are converted to the speech form where the speech output is in the form of regional language. The block diagram of the system is shown in Fig. 1 and Fig. 2.

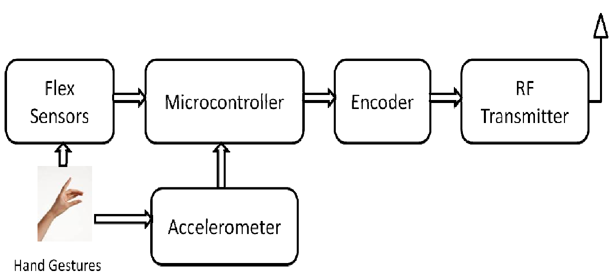

Fig. 1 Block diagram implementation of transmitter section

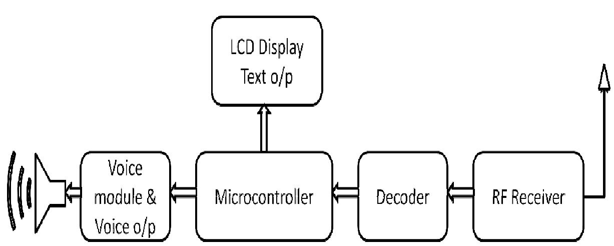

Fig. 2 Block diagram implementation of receiver section

2.1. 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 10kΩ to 30kΩ. A global organization flexed device has 10kΩ resistance and once bent the resistance will increase to 30kΩ at 90°.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 as shown in Figure 3. The electrical device and flex forms a potential divider that divides the input voltage by a quantitative relation determined by the variable and glued resistors.

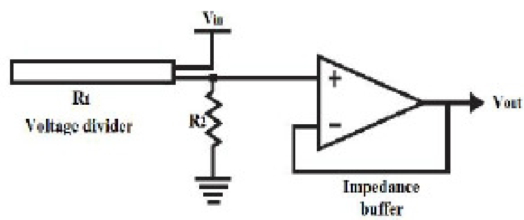

Fig. 3 Equivalent circuit of flex sensor

The characteristics of flex sensors can be studied by the bend angle vs. resistance & bend angle vs. voltage graph which is drawn according to the table below:

Table 1: Variation of voltage & resistance with respect to bend angle

Bend Angle (in degrees)	Resistance (in Kohms)	Voltage (in Volts)
0	9.77	1.93
10	10.85	2.06
20	11	2.11
30	12	2.15
50	13.1	2.3
60	13.7	2.4
90	16	2.5
120	18	2.9
150	19.5	3
180	20	3.1

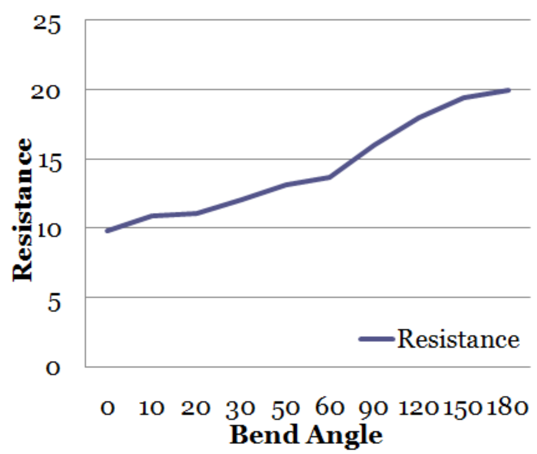

Fig. 4 Resistance vs. Bend angle

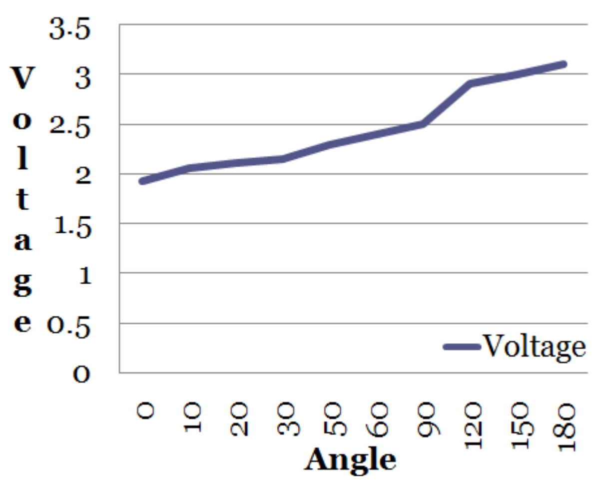

Fig. 5 Voltage vs. Bend angle

2.2. Accelerometer

An Accelerometer is a kind of sensor which gives an analog data while moving in X,Y,Z direction or may be X,Y direction only depends on the type of the sensor. The ADXL335 is a small, thin, low power, complete 3-axis accelerometer with signal conditioned voltage outputs. The product measures acceleration with a minimum full-scale range of ±3g. It can measure the static acceleration of gravity in tilt sensing applications, as well as dynamic acceleration resulting from motion, shock, or vibration.

2.3. Microcontroller (PIC18F4520)

In this project we are using PIC18F4520 microcontroller. As we require ADC (analog to digital converter) to digitize all analog signals from the sensors, we use PIC18F4520 as it has inbuilt ADC and multiplexer. It supports both serial and parallel communication facilities. Some of the key features of this microcontroller are:

- Flash Program Memory: 32 kbytes
- EEPROM Data Memory: 256 bytes
- SRAM Data Memory: 1536 bytes
  A/D Converter: 10-bit Thirteen Channels
- Enhanced USART: Addressable with RS-485, RS-232 and LIN Support
- MSSP: SPI and I²C Master and Slave Support
- External Oscillator: up to 40MHz

2.4. Encoder-Decoder

The output from the PIC microcontroller is encoded by using ST12 encoder manufactured by Sunrom technologies. Each address/data input can be set to any one of the two logic states. The programmed address/data are transmitted together with the header bits Via an RF. It is used to correct the error at the receiver end, if any error had occurred. In the receiver it is decoded by using ST12 decoder. The same IC can be used as an encoder as well as a decoder.

2.5. 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. A command is an instruction given to LCD to do a predefined task like initializing it, clearing its screen, setting the cursor position, controlling display etc. The data register stores the data to be displayed on the LCD. The data is the ASCII value of the character to be displayed on the LCD.

2.6. Voice Module

Once the gesture made by a person is matched with the database, its output should be in the form of voice audio. To bring this voice output we use MP3-TF-16P (DFPlayer Mini module). This module is a very effective alternative for voice recording and playback ICs. It has various features such as it provides perfect integrated MP3 and WMV hardware decoding. The software supports TF card driver, as well as FAT16, FAT32 file system. It has various control modes such as serial mode, AD control mode. It also has a built in 3-W amplifier so there is no need to connect an external amplifier for sound output through speakers. As here we use a microcontroller the serial mode of operation is used. Due to this only 2 ports of serial communication (Rx and Tx) are used. The pre-recorded voice files are stored in the TF card and each voice track is named according to serial numbers (001, 002, 003…). For each gesture a specific serial command is sent to the mp3 module and it playbacks the pre-recorded voice file from the TF card. This module provides flexibility for modification of voice outputs according to different languages according to different regions.

4. System Flow

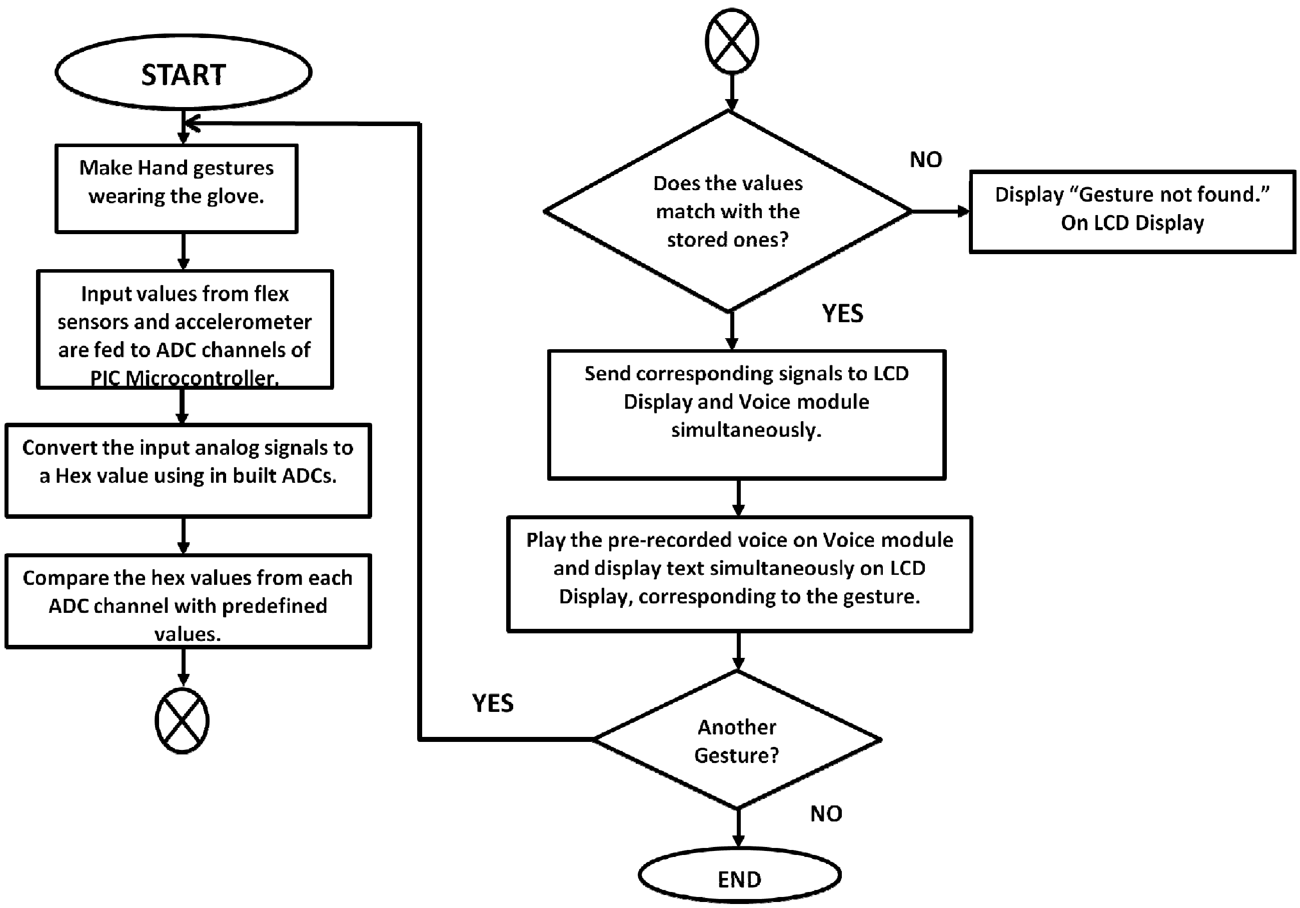

Fig. 6 Flow chart of proposed system

5. Results and Discussions

In this prototype system, the user forms a gesture and holds it for around 2 seconds to ensure proper recognition. Each gesture comprises of bending of all fingers in certain angles accordingly. Every bend of the sensor (finger) produces an analog voltage that is fed to the ADC channels of PIC microcontroller. Each bend of the sensor gives unique ADC value so that when different hand gesture is made, different ADC values are produced. Taking such ADC values for around 5 different users, a table of average ADC values for each finger (sensor) is maintained.

Using a comparison technique for each channel output value the microcontroller gives specific serial commands to the voice module and LCD display simultaneously to give the speech output and text display respectively. Using this concept of ADC count, more and more no. of hand signs can be used as per user and accuracy can be increased with a little change in the ADC count.

The hand signs taken in the prototype can be easily modified according to the user convenience. At the same time the voice output can be changed easily to gives a flexibility in selection of language according to different regions. System accuracy, user configurability, portability, its immunity to noise and environmental disturbances make it to be a better choice over the other products available in the market.

Fig. 7 Gesture-A made on glove & corresponding ADC outputs on LCD.

Fig. 8 Gesture-B made on glove & corresponding ADC outputs on LCD.

The table shown below gives the average ADC values for 5 fingers (sensors) recorded by 5 different users:

Table 2: ADC values of each sensor (finger) for different gestures

Gesture	Thumb	Index	Middle	Ring	Pinky
A	<620	<440	<445	<450	<445
B	<675	<447	<515	<515	<443
C	<670	<515	<447	<445	<443
D	<630	<515	<515	<515	<505
E	<620	<443	<515	<515	<505
F	<620	<445	<515	<515	<505
G	<620	<515	<515	<515	<445
H	<680	<515	<515	<515	<505
I	<620	<445	<445	<445	<505
J	<620	<450	<450	<515	<505
K	<680	<515	<515	<515	<505
L	<680	<450	<450	<515	<505
M	<680	<450	<515	<515	<505
N	<680	<450	<450	<445	<445
O	<680	<515	<515	<515	<445

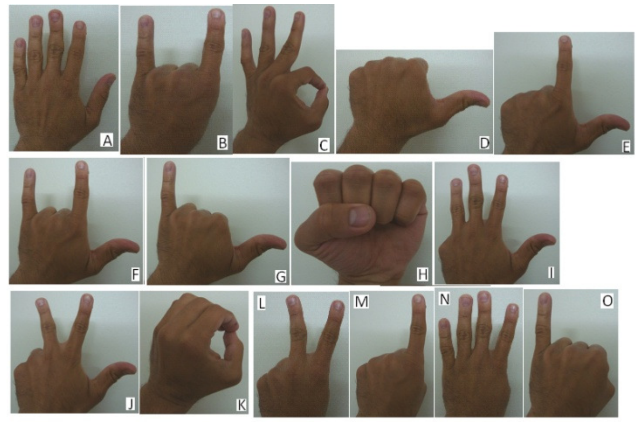

Fig. 9 Hand Gestures used in the system

The gestures shown in the above Fig. 9 are given names A, B, C,…. for naming purpose. Each of these gestures can be given respective sound output in regional languages for example in Marathi it can be done as:
Gesture A: “Thaamb” (wait)
Gesture B: “Ikde Ye” (come here)
Gesture C: “Changla ahe” (It is nice)
And so on for different gestures.

6. Conclusion

This system aims to lower the communication gap between the deaf or community and the normal world. The project proposes a translational device for deaf-mute people using glove technology. The proposed technique has enabled the placement of five flex sensor and an accelerometer on to a glove to detect the gestures of a person. As this system is having its voice output in regional language, it can be used as a translator to communicate with people of different regions with ease. To expand the systems capability, two or more accelerometers can be attached. Camera for facial detection can also be installed for better communication.

References

[1] Hernandez-Rebollar, J. L. Kyriakopoulos, N.Lindeman R.W, “A new instrumented approach for translating American Sign Language into sound and text” , Proceedings of the Sixth IEEE International Conference, 2004.

[2] Mehdi S.A., Khan Y. N., "Sign language recognition using sensor gloves", Proceedings of the 9th International Conference, Volume:5, IEEE Conference Publications, 2002.

[3] Kunal Kadam, Rucha Ganu, Ankita Bhosekar, Prof. S. D. Joshi, "American Sign Language Interpreter", Proceedings of the IEEE Fourth International Conference on Technology for Education, 2012.

[4] Satjakarn Vutinuntakasame, "An Assistive Body Sensor Network Glove for Speech- and Hearing- Impaired Disabilities", Proceedings of the IEEE Computer Society International Conference on Body Sensor Networks, 2011.

[5] Anbarasi Rajamohan, Hemavathy R., Dhanalakshmi M. , "Deaf-Mute Communication Interpreter", International Journal of Scientific Engineering and Technology (ISSN : 2277-1581) Volume 2 Issue 5, pp : 336-341, 1 May 2013

[6] K.C. Shriharipriya and K. Arthy, "Flex sensor based hand gesture recognition system", Proceedings of International Journal Of Innovative Research And Studies (IJIRS), Vellore, India, May 2013.

[7] Bhavina Patel, Vandana Shah, Ravindra Kshirsagar, "MICROCONTROLLER BASED GESTURE RECOGNITION SYSTEM FOR THE HANDICAP PEOPLE", in Proceedings of Journal of Engineering Research and Studies, December 2011.