Pak. J. Biotechnol. Vol. 13 (special issue on Innovations in information Embedded and Communication Systems)
Pp. 236- 241 (2016)

INTERPRETATION OF INDIAN SIGN LANGUAGE USING STORED VOICE SAMPLED VALUES

Gargi Arya[1], Monika Mandal1[1] and Rachana Nagal[2]

[1]Amity School of Engg. and Technology, New delhi, India. [2]Amity School of Engg. and Technology, Faculty of Electronics and Communication Engineering, New Delhi, India.
Email*: gargi.arya07@gmail.com, rnagal@amity.edu

ABSTRACT

According to the census of India 2011 data on disability, around 7.4% of population is speech impaired. They use sign language as a medium to communicate but it restricts them to communicate with people who do not understand sign language. Speech impediment not only restricts them in communication but also in learning. This work make its contribution towards bringing down the barrier of communication and thus developing a compact and affordable wearable device which converts Indian sign language to voice and text. It is equipped with a accelerometer weared in index finger. For each gesture, accelerometer yields XYZ coordinates which are sensed by ATmega 2560. According to the sensed coordinates micro-controller extracts the corresponding speech file (decimal sampled sound file) from SD Card, process it and feed to audio amplifier and at the same time to LCD display. The .wav files are processed via Matlab to fetch the decimal sampling of sound files which are stored in the SD card for memory management and hence extends the potential for vast storage of words and sentences.

IndexTerms: Atmega 2560, Indian Sign language, Arduino mega 2560, accelerometer, LM 386, Text to Speech, SD card

INTRODUCTION

Hearing and auditory processing problems not only affect speech and communication but also language development and learning. The earlier these problems occur, the more it effect on the child's development. In relation with speech and language, it causes delay in the development of receptive and expressive communication skills. The language barrier affect their learning and a child's cognitive, social-emotional and academic skills. In a world where everyone has the right to be educated, language deficit lets them down. This is one and a very basic and impacting problem that they face. Lot of work has been done to improve the quality and enhance the way for learning for these people for example to improve the linguistic skills through multimedia [1], an approach to generate new tools using assistive technology for education support of children with disabilities [2], access services such as subtitles, audio-description, clean audio or sign language videos via the broadband Internet [3], use of plain English to increase web accessibility [4]. The root cause of their lower literacy and self-development is the barrier of communication. For deaf people born with no hearing and, or auditory processing problems has sign language as their first language. The lack of acceptance and usage of sign language in the society also bridge a gap with these people. According to the research, there are only 250 sign language human interpreters available across India up till now. In order to bring this barrier down, there is a need for a device which can convert the sign language into corresponding text and speech to enable people to communicate effectively and with ease. This project is an attempt to achieve the same.

On studying the English alphabet representation by Indian sign language, we find out that among 26 alphabet letters 15 are uniquely represented and the rest 11 has the index finger position as common [5,6]. As the usage of index finger as compared with other fingers made more representation of letters than others, thus accelerometer is being weared in index finger. As far as words and sentence formation is concerned, the unique representation and position lead us to program most of the words and sentences with 1 accelerometer. These words are mostly from the category of colours, computer and greetings.

There are many methods carried out to recognize and interpret the Indian sign language. The following lists some of the existing methods and approach. Using digital image processing techniques and artificial neural network for recognizing different signs proposed by Adithya, et al., [7]. A vision based dynamic gesture recognition using 3D dynamic signs corresponding to ISL words proposed by Geetha, et al., [8]. Animation system for Indian Sign Language communication using LOTS notation enabling the system to identify features like hand location(L), Hand Orientation (O) in the 3D space, Hand Trajectory movement (T), hand shapes (S) and non-manual components like facial expression, presented by Ruviansh et al., [9].

In comparison with existing methods the proposed model is user friendly, affordable, portable, works on low power and of course no skill is required to operate other than being aware of Indian sign language. It produces voice which is audible to the neighbors of the user so that they can react (or) respond for the voice message. Since the voice message is retrieved from SD card, the most important benefit of this device is regardless of time any length and language, large chunk of voice messages can be stored in a form of sampled file in the SD card and played. The approximated number of gestures which can be programmed using 1 accelerometer is 10,000 as accelerometer provides a wide range of values for each x,y and z axis. It can also serve the purpose of learning the sign language.

MATERIALS AND METHODS

A. Block Diagram

The objective of the project is to sense gestures made by person's hand and convert these gestures into corresponding text and speech using 1 accelerometer sensor [10] as shown in Fig.1. The task is carried out using 2 micro-controller's, one working as master and other as slave. Master is responsible for voice output whereas slave gets command from master and is responsible for text on LCD display.

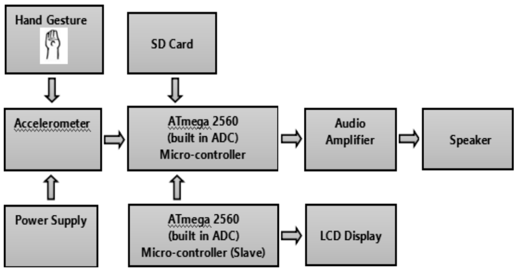

Fig. 1. Block diagram of proposed model

B. Proposed Model

- Accelerometer sense the gestures made through hand. Creates analog values/readings depending upon the changed axis of the accelerometer attached on finger and passes these values to master micro-controller.
- Master micro-controller read and convert these values into digital.
- If the values received from accelerometer matches with programmed values then master micro-controller fetch corresponding speech files from SD card and commands slave microcontroller to process for the display of text on LCD.
- Master convert digital values of speech file into analog values, fed to audio amplifier.
- Slave forwards the corresponding text to the LCD
- The microcontroller's are programmed such that voice output and text are outputted concurrently.

C. Speech Synthesis.

The files stored in the SD card module are sampled values of the recorded audio files of alphabets(a-z), words and sentences originally in the mp3 format. The process of speech synthesis requires mp3 format audio files to be converted into text file format which can be read by micro-controller in order to process them and generate voice output. Following are the steps performed for speech synthesis:

1) Step 1: Compressing audio file and converting them into the .wav format using audacity software.

- The need to change the format from mp3 to .wav [11] is that the c code requires the sampled values in the form of array of the audio file and only .wav file can provide this array.
- The original audio is in the mp3 format which is compressed using audacity software at 8kHz sample rate and 16 bits per sample.
- The need to compress the file is because samples do not fit in the micro-controller uncompressed.

2) Step 2: Output scaling of .wav format file.

- The need to perform output scaling is that the c code can only read the array existing in native format.
- The samples of the audio file in the .wav format contains value of the samples in the double format. This needs to be converted into native format. This is called Output Scaling [12].
- Output scaling is based on typical bit-widths found in a .wav file. 'Native' format uses data type unit8 (unsigned integer) and data range ranges from $0 <= Y <= 255$.
- Result is the array ranging between values $0 <= Y <= 255$. This array is copied into the c code and stored in SD card which is further read by the microcontroller.
- Micro-controller process the sampled array, convert these into analog values and feed to audio amplifier.

I. SYSTEM DESCRIPTION

A. Gesture Reading

This module consist of 1 accelerometer sensor. It is assembled such that it can be weared into the index finger and 2 other modules rest at the wrist. Whenever any gesture is made the ADXL335 3-axis accelerometer measures static acceleration of gravity when tilted. It is measured with a minimum full-scale range of ±3 g. The output signals are analog voltages proportional to acceleration. It sense 3 coordinate (X,Y and Z). Shown in fig. 2 the axes of acceleration of the coordinates. These values are received by Atmega 2560 micro-controller (master) and are converted into digital values with inbuilt ADC [13]. Since given Vs = +5v for accelerometer the output values at microcontroller ranges between 300 < 500. These value output are shown in Fig. 8, 9 and 10.

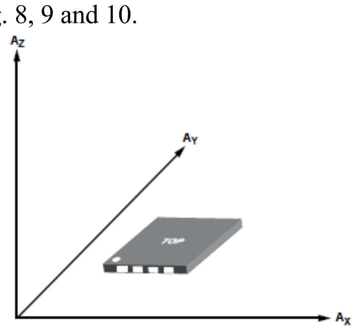

Fig. 2. Axes of acceleration

B. Text Display and Voice Process

This section consist of 2 Atmega 2560 master and slave, SD Card and LCD module. The converted digital values of accelerometer are processed by master micro-controller and are matched with re-written programmed values. If they match, master fetch the corresponding PCM.h speech file (sampled array file) from the SD card [14], process it, convert the file into analog values and forward them to audio amplifier. Meanwhile it also gives command to slave micro-controller to display corresponding text on LCD module.

C. Voice Output

Voice output is carried out by audio amplifier. The audio amplifier circuit as shown in Fig. 3 comprises of LM386 and a speaker. LM386 is a power amplifier and is used in low voltage consumer applications. The gain is internally set to 20 to keep external part count low [15], but the addition of an external resistor and capacitor between pins 1 and 8 increases the gain to any value from 20 to 200. LM 386 [16] receive analog signals from the slave micro-controller. These signals are amplified to a level at which the voice can be heard and are fed to the speaker for the voice output.

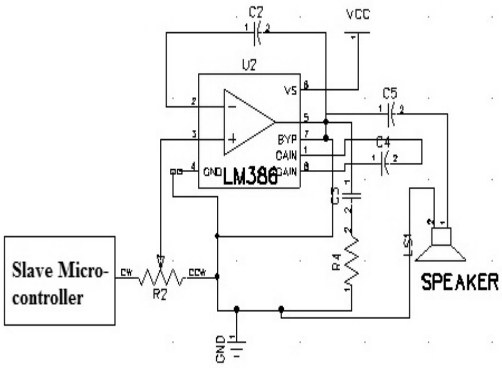

Fig. 3. Audio amplifier circuit

D. Hardware Implementation

The hardware circuitry of the proposed device includes accelerometer, arduino development board with built-in microcontroller ATmega 2560, SD card module, LCD module and audio amplifier circuit using LM386.

1) *Accelerometer (ADLX 335):* An ADLX (accelerometer) as shown in Fig. 4 is used at the tip of the index finger to get various 3 dimensional values which further decide the text and speech. The ADXL 335 complete 3-axis accelerometer. It measures static acceleration of gravity while tilted with a minimum full-scale range of ±3 g. The sensor being a polysilicon surface-micromachined structure built on top of a silicon wafer provide a resistance against acceleration forces. The deflection of the structure is measured by differential capacitor consisting of independent fixed plates and plates attached to the moving mass. Acceleration deflects the moving mass and unbalances the differential capacitor resulting in a sensor output whose amplitude is proportional to acceleration. The ADLX takes in value by the difference of values due to the different orientation in space. Supply voltage provided is +5v. The X_{out} , Y_{out} , Z_{out} are connected to analog input pins of accelerometer.

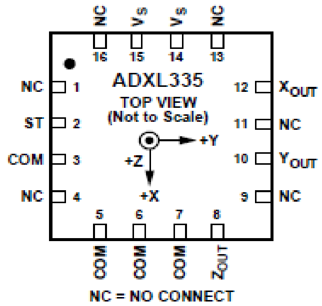

Fig. 4. Accelerometer ADLX 335

2) *Arduino mega 2560 development board:*The Arduino Mega as shown in the Fig. 5 is a micro-controller board based on the ATmega2560. It has 54 digital input/output pins of which 15 can be used as PWM outputs, 16 analog inputs, 4 UARTs, a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button. It communicates using the original STK500 protocol. It features the Atmega 16U2 programmed as a USB-to-serial converter. The two microcontrollers one being master and other as slave are connected through Rx, Tx pins, and support one way communication. The two micro-controllers are introduced for fast processing of instructions.

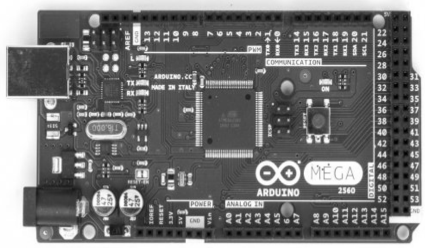

Fig. 5. Arduino mega 2560 development board

3) *SD card module:* The SD card as shown in Fig. 6 is a new mass-storage system integrated with flash memory with serial and random access capability. Accessed via a dedicated serial interface optimized for fast and reliable data transmission. SD files contains the digital sampled values in the form of array for each programmed gesture. It is connected to the master micro-controller.

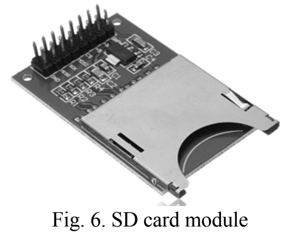

Fig. 6. SD card module

4) *LCD module:* The LCD module as shown in Fig. 7 is a 16x2 display. The register select is opted and the read/write is kept ground for all time write. No fetching of instructions is activated by keeping R/W to zero. Signal processed by microcontroller is sent to the LCD to be displayed on the 16x2 screen. The LCD is enabled at the time when the accelerometer reads in

Pak. J. Biotechnol. Vol. 13 (special issue on Innovations in information Embedded and Communication Systems)
Pp. 236- 241 (2016)

any value which is present in the library of the program. Whensoever the values are read, LCD is displayed on and the corresponding text is displayed on the screen.

Fig. 7. LCD module 16x2

5) *LM 386 power amplifier IC:* LM 386 is a power amplifier as shown in Fig. 8 and is used in low voltage consumer applications. The gain is internally set to 20 to keep external part count low, but the addition of an external resistor and capacitor between pins 1 and 8 will increase the gain to any value from 20 to 200. LM386 receives analog signals from the master micro-controller. These signals are amplified to a level at which the voice can be heard and are fed to the speaker for the voice output.

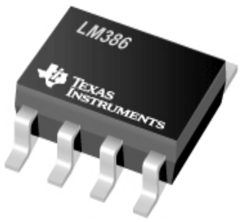

Fig. 8. LM 386 power amplifier

E. Software Implementation

Number of software's are used in the implementation of project. Arduino 1.0.6 is used for programming the micro-controller to carry out various functions such as accelerometer readings, text display and voice output. MATLAB is used for output scaling of .wav audio file. Audacity is used to compress the audio .mp3 file. The designing of the pcb is done using Diptrace software. Proteus 8 is used for the simulation of the project.

RESULTS AND DISCUSSION

The results are observed for three gestures. The following Fig. 9, 10 and 11 shows readings generated by accelerometer on arduino 1.0.6 software while calibrating and also the range of values formed for gestures. The readings are taken for X,Y and Z axis at a particular position of gesture made by the hand. They explain the way in which the readings for X,Y and Z axis is taken when a gesture is made. First column depict the X axis, second column depict Y axis and third column depict Z axis.

A. Gesture Readings by accelerometer

Figure 8 shows values when the accelerometer is oriented in Z axis: values: at the third column = 400+. Fig. 10 shows values when the accelerometer is oriented in Y axis: values: at the third column = 400+. Fig. 10 shows values when the accelerometer is oriented in X axis : values: at the middle column = 400+

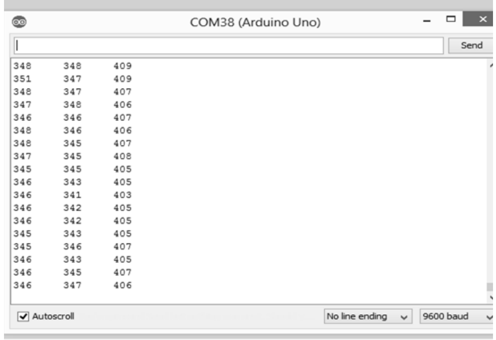

Fig. 9. Readings in z-axis

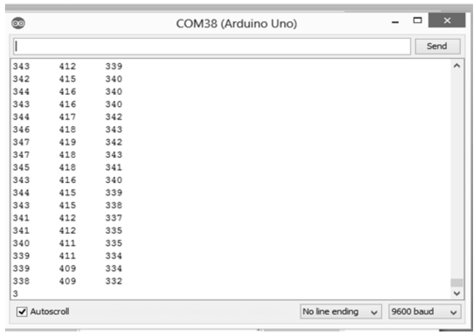

Fig. 10. Readings in y-axis

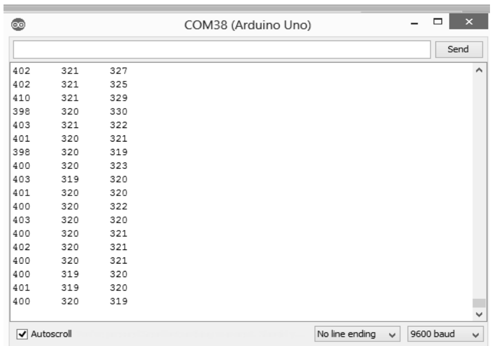

Fig. 11. Readings in x-axis

E. Operational Parameters and corresponding MATLAB gesture speech waveform

The operational axis values of the gestures at which they are programmed is shown. Shows the plot of the sound of three gestures. Fig. 12 shows the text displayed on LCD module corresponding to the gesture detected.

1) Gesture 1: Gesture 1 shown in Fig. 13 interpreting "Hi how are you". The ADLX operating parameters are as follows.

AdLx:

X axis = 320<=X<=350

Y axis = 280<=Y<=290

Z axis = 340 <=Z<=365

The corresponding speech waveform is shown in Fig. 14. The waveform represents 13793 samples with sampling frequency of 8kHz

2) Gesture 2: Gesture 2 shown in Fig. 15 interpret the word "I need help". The ADLX operating parameters are as follows.

AdLx:

X axis = 330<=X<=360
Y axis = 270<=Y<=290
Z axis = 320 <=Z<=360
The corresponding speech waveform is shown in Fig. 16. The waveform represents 16704 samples with 8kHz of sampling frequency.

3) Gesture 3: Gesture 3 shown in Fig. 17 Interpret the word "Thank you". The ADLX operating parameters are as follows.
AdLx :
X axis = 330<=X<=360
Y axis = Y>=400
Z axis = 370<=Z<=380
The corresponding speech waveform is shown in Fig. 18. The waveform represents 6912 samples with 8kHz of sampling frequency.

Fig. 12. Text display of gesture 1

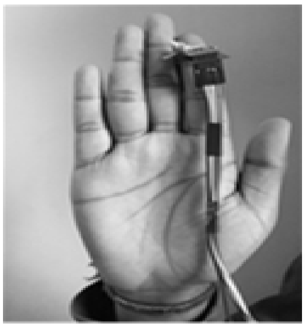

Fig. 13. Gesture 1

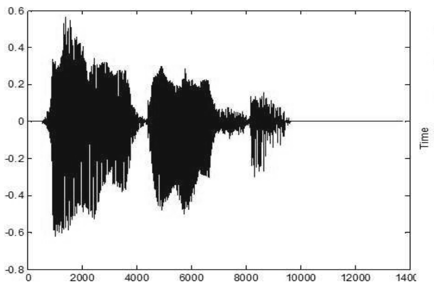

Fig. 14. Speech waveform output for gesture 1

Fig. 15. Gesture 2

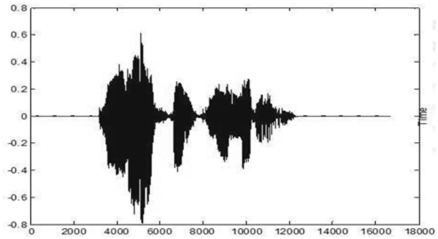

Fig. 16. Speech waveform output for gesture 2

Fig. 17. Gesture 3

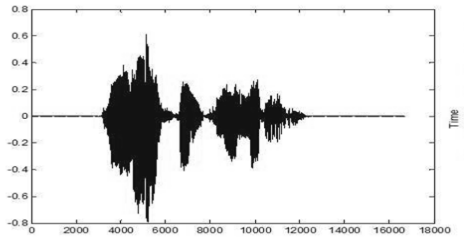

Fig. 18. Speech waveform output for gesture 3

Conclusion and Future Scope

The project aimed to interpret Indian sign language into speech and text. To enable a speech impaired person to convey his/her message easily to a second person has been designed and implemented successfully. The result is observed using three gestures, these gestures were interpreted by the micro-controller correctly, the corresponding output in the form of voice was heard from the speaker and text was displayed on LCD module. The system is portable and re-programmable. Also, it turned out to be cost-effective and handy device.

The device can be further enhanced to process a large from of lexicon. The output time from the speaker after a gesture is formed can be reduced. More words and alphabets can be formed by using more accelerometer. It can be equipped with bluetooth module or RF module to transfer the text corresponding to the gesture to other person's mobile as a text message. It can be made compatible with windows inbuild Text-to-Speech feature for voice output.

References

[1] Diego Leonardo Heredia and Manuel Perez Cota, Multimedia system for children with hearing disability. IEEE sponsored Iberian Conference of Information Systems and Technologies (CISTI)., Pp. 1-5. (2015).

[2] Vladímir Robles Bykbaev, Eduardo Pinos Vélez and Paola Ingavélez Guerra, An educational approach to generate new tools for education support of children with disabilities. IEEE sponsored International Conference of Education, Entertainment and e-Management (ICEEE), Pp. 80-83, Bali (2011).

[3] Christoph Dosch, Convergence of broadcasting and broadband Internet - a benefit for people with disabilities (and for us all). IEEE Proceedings of the ITU Kaleidoscope Academic Conference, Living in a converged world - Impossible without standards? Pp. 89-95, St. Petersburg (2014).

[4] Boldyreff, C., E. Burd, J. Donkin and S. Marshall, The case for the use of plain English to increase web accessibility. IEEE Proceedings 3rd International Workshop of Web Site Evolution, Pp. 42-48 (2001).

[5] Sudeep D. Thepade, Ashwini Kawale, Poonam Shipure, Sherin Thomas and Shruti Nathe, Novel Indian Sign Language Recognition technique using Energy Compaction of five orthogonal transforms. IEEE sponsored International Conference of Communications and Signal Processing (ICCSP) Pp. 1713-1717., Melmaruvathur (2015).

[6] Kalpana Sharma, Garima Joshi and Maitreyee Dutta, Analysis of shape and orientation recognition capability of Complex Zernike Moments for signed gestures. IEEE sponsored 2nd International Conference of Signal Processing and Integrated Networks (SPIN), Pp. 730-735, Noida (2015).

[7] Adithya, V., P. R. Vinod and Usha Gopalakrishnan, Artificial neural network based method for Indian sign language recognition. IEEE sponsored Conference of Information & Communication Technologies (ICT) Pp. 1080-1085, JeJu Island (2013).

[8] Geetha, M., C. Manjusha, P. Unnikrishnan and R. Harikrishnan, A vision based dynamic gesture recognition of Indian Sign Language on Kinect based depth images. IEEE sponsored International Conference of Emerging Trends in Communication, Control, Signal Processing and Computing Applications (C2SPCA) Pp. 1-7, Bangalore (2013).

[9] Ruviansh J. Raghavan, K. Arun Prasad, Rajani Muraleedharan and M. Geetha, Animation system for Indian Sign Language communication using LOTS notation. IEEE sponsored International Conference of Emerging Trends in Communication, Control, Signal Processing & Computing Applications (C2SPCA) Pp.1-7, Bangalore (2013).

[10] Sriram, N. and M. Nithiyanandham, A hand gesture recognition based communication system for silent speakers. IEEE sponsored International Conference of Human Computer Interactions (ICHCI) Pp. 1-5, Chennai (2013).

[11] Arduino, Arduino.cc", Simple Audio Player [online]. Available:https://www.arduino.cc/en/Tutorial/SimpleAudioPlayer

[12] Mathworks. Wavread, "mathworks", [online]. Available: http://in.math works.com/help/ matlab/ref/ wavread.html

[13] Instructables. Interfacing ADXL335 with Arduino, "Instructables", [online]. Available:http://www. instruc tables.com/id/Interfacing-ADXL335-with-ARDUINO/ step4/Serial-Monitor

[14] Bajdi. Arduino mega 2560 and SD card module, "bajdi", [online]. Available:http://www.bajdi.com/arduino-mega-2560-and-sd-card-modul/

[15] Texas Instruments., LM386 Low Voltage Audio Power Amplifier, "ti", [online]. Available:http://www.ti.com/ lit/ds/symlink/lm386.pdf

[16] Texas Instruments., Calculating Gain for Audio Amplifiers, "ti", [online].Available:http://www.ti.com/ lit/an/sloa105a/sloa105a.pdf