Islamic University of Technology

Department of Computer Science and Engineering (CSE)

Vibrotactile and Visual Feedback for Deaf and Mute

Authors

MEHRAB ZAMAN CHOWDHURY (134404)

MASRUR SOBHAN SIAM (134418)

A Thesis Submitted to the Academic Faculty in Partial Fulfillment of the Requirements for the Degree of

BACHELOR OF SCIENCE IN COMPUTER SCIENCE AND ENGINEERING

BACHELOR OF SCIENCE IN COMPUTER SCIENCE AND ENGINEERING

Vibrotactile and Visual Feedback for Deaf and Mute

Authors

MEHRAB ZAMAN CHOWDHURY (134404)

MASRUR SOBHAN SIAM (134418)

Department of Computer Science and Engineering (CSE)

Islamic University of Technology (IUT)

Organization of Islamic Cooperation (OIC)

Board Bazar, Gazipur-1704

Bangladesh

Declaration of Authorship

This is to certify that the work presented in this thesis is the outcome of the analysis and experiments carried out by Mehrab Zaman Chowdhury and Masrur Sobhan Siam bearing Student No. 134404 and 134418 respectively of Academic Year 2016–2017 under the supervision of Dr. Kamrul Hasan, Associate Professor of Department of Computer Science and Engineering (CSE), Islamic University of Technology (IUT), Gazipur, Bangladesh. It is also declared that neither of this thesis nor any part of this thesis has been submitted anywhere else for any degree or diploma. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given.

Authors

Mehrab Zaman Chowdhury

Masrur Sobhan Siam

Approved by

Dr. Kamrul Hasan

Associate Professor

Department of Computer Science and Engineering

Islamic University of Technology

BoardBazar, Gazipur-1704, Bangladesh.

Acknowledgement

First of all, I would like express my heartiest gratitude to the Almighty Allah for providing me the strength to complete this thesis work. After the Almighty, it is my great pleasure to express gratitude to the people who made this thesis possible.

We would like to express our grateful appreciation to **Dr. Kamrul Hasan**, Associate Professor, Department of CSE, IUT and **Hasan Mahmud**, Assistant Professor, Department of CSE, IUT for being our advisor and mentor. Their motivation, suggestions and insights for this thesis have been invaluable. Without their support and proper guidance this research would not have been possible. Their valuable opinion, time and input provided throughout the thesis work, from first phase of thesis topics introduction, subject selection, proposing algorithms, modifications till the project implementation and finalization which helped us to do our thesis work in a proper way. We are really grateful to them.

It was our pleasure to get the cooperation and coordination from **Prof. Dr. Muhammad Mahbub Alam**, Head of CSE Department, IUT during various phases of the work. We are grateful to him for his constant and energetic guidance and valuable advice.

Our deep appreciation extends to all the respected jury members of our thesis committee for their insightful comments and constructive criticism of our research work. Surely they have helped us to improve this research work greatly. Lastly we would like to thank the faculty members of CSE Department, IUT who have helped to make our working environment a pleasant one, by providing a helpful set of eyes and ears when problems arose.

Dedication

Every challenging work needs self-efforts as well as guidance of elders especially those who are very close to our heart.

My humble effort we dedicate to our sweet and loving,

Father and Mother

Whose affection, love, encouragement and prays of day and night made us able to get such success and honor,

Along with all hard working and respected,

Teachers

Abstract

According to World Federation of the Deaf (WFD) there are approximately 70 million deaf and mute all around the world. They are the unfortunate ones who are deprived of communication properly. This creates gap between the normal people and the deprived ones. In our research, we will establish communication between both the normal people and the deprived ones. We will establish one to one and two-way communication. Our research consists of mainly two modules; first module is to generate texts from speech which will then be classified into command sets and context sets, second module is providing reply by the deaf and mute to the normal people. The commands and contexts will be mapped to the database and according to that output will be shown. It is an adaptive and multimodal approach.

Contents

Declaration of Authorship vi

Acknowledgement ii

Dedication iii

Abstract iv

List of Figures vii

1. Introduction **1**

1.1. Background …………………………………………………………………. 1

1.2. Motivation ………………………………………………………………….. 2

1.3. Problem Statement …………………………………………………………... 2

1.4. Our Approach ……………………………………………………………….... 3

2. Related Works **4**

2.1. AUDIS Wear: A Smartwatch based Assistive Device for Ubiquitous Awareness of Environmental Sounds …………………………………………… 4

2.2. The Haptic Chair as a Speech Training Aid for the Deaf ………………………... 6

2.3. Two Way Communicator between Deaf and Dumb People and Normal People ………………………………………………………………… 8

2.4. A Vibrotactile Sensory Substitution Device for the Deaf and Profoundly Hearing Impaired …………………………………………………. 12

3. Design **14**

3.1. Google Text to Speech …………………………………………………… 14

3.2. Google Speech to Text …………………………………………………… 15

3.3. Vibrotactile Approach …………………………………………………… 15

3.4. Contextual Approach ………………………………………………………….. 16

3.5. Bluetooth and Wi-Fi Communication ……………………………………………. 16

3.6. Smart Watch and Smart Phones ………………………………………………….. 17

4. Algorithms **19**

4.1. De-Syllable ………………………………………………………………………. 20

4.2. Vibration ………………………………………………………………………… 21

4.3. Google Speech to Text ………………………………………………………… 22

4.4. Text to Speech ………………………………………………………………….. 24

5. Implementation and Evaluation **25**

5.1. Implementation 25

5.2. Evaluation 29

6. Conclusion and Future work **31**

7. References **32**

List of Figures

Figure 2.1: Diagram of Sound Recognition Framework …………………………………... 5

Figure 2.2: Haptic Chair ……………………………………………………………… 7

Figure 2.3: Overview of real-time video to speech ……………………………………... 10

Figure 2.4: Overall view of speech to animated video …………………………………… 11

Figure 2.5: Vibrotactile Vest ………………………………………………………… 13

Figure 4.1: Code for De-Syllable ……………………………………………………….. 20

Figure 4.2: Code for Vibration ………………………………………………………….. 21

Figure 4.3: Code for Google Speech to Text ………………………………………………. 23

Figure 4.4: Code for Text to Speech ……………………………………………………….. 24

Figure 5.1: Overall view from speech to Vibrotactile and visual output …………………... 26

Figure 5.2: Text input by deaf and text output to normal people …………………………... 27

Figure 5.3: Vibration pattern and visual output …………………………………………….. 28

Figure 5.4: Proposed system's working procedure …………………………………………. 28

Figure 5.5: Heuristic Evaluation ……………………………………………………………. 30

Chapter 1

Introduction

1.1 Background

The current era is a zoom of technology. Each and every field has an impact of the technological advancements onto it. One such rapidly growing technical advancement is the increasing impact of mobile phones on human life. The enormous and ever increasing Internet usage along with smart phones and smart watches has proven a boon to mankind. Not surprisingly it became a popular device for applications to assist people that experience disabilities. A broad range of smartphone based applications were presented for deaf people. Like smart phones, smart watches can also help to aid the hearing disabled people or mute people as it is equipped with sensors, input and output capabilities (vibrations, touch, text writing etc.), network interfaces and considerable consumption power.

The hearing disabled and mute people cannot mingle with the social world because of their physical disabilities. Unintentionally, they are treated in an unusual manner by the rest of the society. They cannot be a part of the social events, say students cannot study in schools with normal students, elderly persons cannot work at work places, and much more. Simple activities like going and buying a commodity from the grocery shop is very complicated task for the deaf and dumb person. Despite of this large number of deaf and mute people, very less research is done in order to bridge the communication barrier using the smart devices (Smart phones, Smart watches etc.).

1.2 Motivation

Many researches have been done using sign language for deaf and mute people. But it is a troublesome approach as the normal people need to remember the sign languages. Moreover, sign languages vary from region to region. Some of the approaches are one-way communication, some devices are not portable, less available and much expensive. Some devices are not that much comforting. So our main focus is to make a one to one and two-way communication using smart devices, such as, cell phones and smart watches.

To bridge the communication gap between us, we propose a method which helps the normal and disabled people to communicate in a very effective and faster way. Vibrotactile, visual cues, multimodality, and adaptive approaches are the main focus of this research which will help to find a solution to the communication problem.

1.3 Problem Statement

Our focus is to implement a one to one and two-way communication tool for communication purpose. The input speech given by the normal people will be classified into two sectors- command sets and context sets. Vibration pattern will be generated from the given commands which will be used as output for the deaf and mute. The commands will also be available as visual output. The context sets will be shown as visual output to the smart devices.

The whole system will be adaptive and user friendly. Vibration pattern will be set according to syllable for better learning purpose.

1.4 Our Approach

Our main contribution in this research is the adaptive approach. Besides text will be classified to some commands and contexts. Vibration pattern will be set for the commands where the contexts will be shown as visual output. Some predefined answers will be set for giving reply of the deaf and mute people. Text option will also be available for reply which will then be shown to normal users' devices as text and also in voice commands. Connection between smart devices will be done with Bluetooth. This will ease the communication between normal people and deaf & mute. No sign language is needed to be remembered. Syllable based vibration pattern will be set. Words having same syllable will have same vibration pattern but the duration of vibration will be different. Besides a game will be developed for fast learning. Suppose, a vibration will start and image will be shown and then the user will have to answer the meaning given by the vibration and context image.

Chapter 2

Related Works

Many works have been done for communication with the deaf and mute people. Some of the researches are about the haptic feedback for speech based communication using smart devices. When referring to mobile phones and similar devices (smart watch), this generally means the use of vibrations from the device's vibration to denote that an event has occurred. Vibrotactile feedback allows the deaf people to sense vibration for specific set of words and enables them to interact with the physical world. The research works discussed below are related to Vibrotactile feedback for communication.

2.1 AUDIS Wear: A Smartwatch based Assistive Device for Ubiquitous Awareness of Environmental Sounds

A large set of assistive devices is available for deaf people. Besides hearing and communication aids, devices to access environmental sounds are available commercially. But the devices have two major drawbacks:

- They are targeted at indoor environments (e.g. home or work).
- Only specific events are supported (e.g. the doorbell or telephone).

Researches show that important sound can occur in all context and the drawbacks can be tackled by modern information and communication technology. The smart watch which is a form of wrist watch, offers new potential for assistive technology. Based on smart watch and algorithms from pattern recognition, a prototype [1] is discussed here that observes the surrounding environment of the user and detects the sounds. A Vibrotactile feedback that is a vibration is triggered when a sound is detected and information about the sound is appeared on the smart watch.

In the last few years' smart watch became popular among people. It allows people to install and use applications. Smart watch can be used as an assistive tool for the people having disabilities like deaf, dumb etc. Although smart phones allow several applications that provide hearing aid to the deaf, a more compact platform like smart watches also supports this feature and it is more comfortable to wear and easy to use.

A prototype has been developed in this paper that allows the use of smart watch as an assistive device. The smart watch must need to run Android Wear operating system. The UI of the application was designed based on the need of the user group.

A. Sound Processing

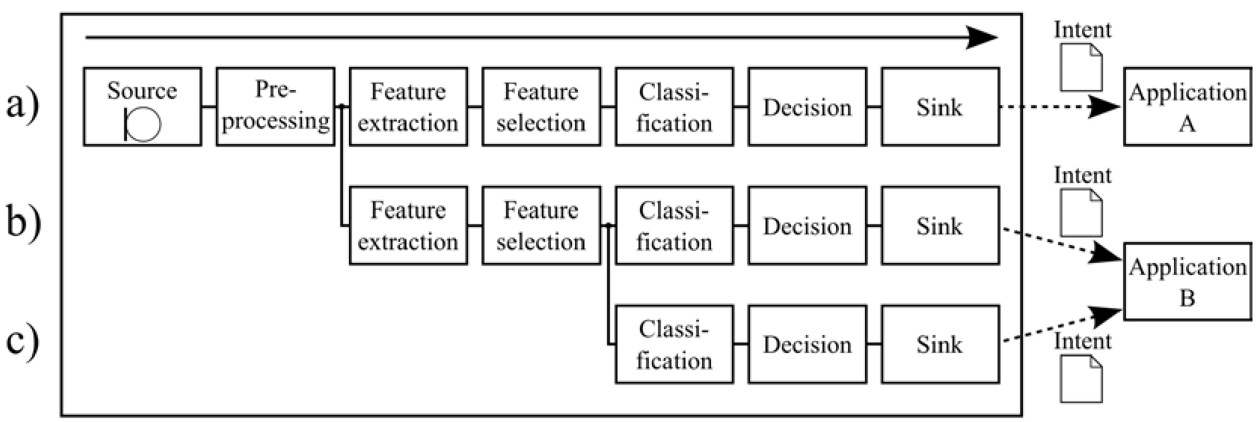

Figure 2.1: Diagram of Sound Recognition Framework

The prototype makes use of a sound recognition algorithm and is capable of detecting sound. A sound recognition pipeline is formed by seven successive processing stages, where each stage processes the results from the former stage. The data is captured in the Source stage that provides access to the microphone and organizes the data into frames. The captured data is passed on to the Preprocessing stage where the data can be preprocessed, e.g. with filters or windowing functions. In the third stage the Feature Extraction takes place, followed by a Feature Selection. The Classification stage performs the actual assignment to classes. On grounds of the classification results in the Decision stage it is determined which class is detected. Finally, the Sink stage triggers the action(s) associated with the detected sound.

B. User Interface

Before designing the UI, some measures have been taken so that people having different capabilities can use the application efficiently. Contradicting requirements are fulfilled when designing the app for a large group of users. That means Google design guidelines and the different disabilities of the users are taken into consideration. The UI of this prototype consists of one simple screen that contains a button to start or stop the service. When a sound is detected, a vibration is triggered and a visual notification is shown on the display.

Even though the deaf people accepted this prototype but it doesn't allow both way communication. The deaf people can't give feedback to the received environmental alert.

2.2 The Haptic Chair as a Speech Training Aid for the Deaf

The 'Haptic Chair' delivers Vibrotactile stimulations to several parts of the body. In this paper, the effectiveness of using the haptic chair in speech therapy sessions is explored. 20 deaf people were engaged in the study. The improvements in word clarity signifies that the haptic chair can make a significant contribution to speech therapy.

Speech therapy is a treatment for people with speech and communication disorders resulting from a wide range of medical conditions such as cleft lip, weak muscles around the mouth, to hearing impairment, etc. The therapy sessions may include oral-motor exercise, speech drills to improve clarity of voice. In a typical speech therapy session, the speech therapist and the deaf sit in front of a mirror and the deaf tries to mimic the lip movement of the therapist. Often the deaf able to mimic the lip movement but due to lack of audible feedback it is not possible to detect if he generated any sound or not. In this paper, the effectiveness of using haptic chair in speech therapy session is explored. The design of the Haptic Chair was extended according to Gault's proposed method so that users would be able to sense amplified vibrations produced by their own voice as well as others such as teachers or therapists. With this modification, we observed immediate effects on the awareness the profoundly deaf users had of whether they were matching the sound production pattern accompanying the lip movements they could see.

Figure 2.2: Haptic Chair.

The usual speech training sessions consist of a system consisting of a mini-microphone, amplifier and headphones. In this study, an additional microphone is added to the same amplifier to capture the voice of the profoundly deaf user, which has been fed back to the deaf. The speech therapist holds a microphone and the student wears headphones. The sound from the microphone is amplified and sent to the headphones. As the speech therapist speaks, the hearing impaired person hears the speech therapist's voice through the headphones and sees the lip movements in the mirror. The sound produced by both the speech therapist and the deaf were used to generate Vibrotactile feedback through the haptic chair.

In this study, a 12-week pilot user study and then 24-week follow-up study has performed. The deaf people who were given speech therapy using the haptic chair showed better performance than the people who received speech therapy without the using the haptic chair. The speech therapist give a point based on the clarity of the word pronounced by the deaf. The deaf people who received therapy using haptic chair showed better performance their pronounced word clarity is more than the pronounced word clarity of the user who got therapy without using the haptic chair.

Results suggest that the additional Vibrotactile feedback provided by the Haptic Chair had a positive impact on speech learning in this context. The user feels the vibration of the speech in his arm, legs and also in the back.

The current system doesn't process the input audio stream electronically instead delivers the entire input audio stream to each of the separate vibration system that focuses on the feet, back, arm and hand.

2.3 Two Way Communicator between Deaf and Dumb People and Normal People

Every human sees, listens and then reacts to the situation by speaking himself out. But the deaf people are deprived of this ability which creates a gap between the deaf and the normal people.

This paper tells about a communication tool that will reduce the communication gap and also tells about the communication in both ways. The system consists of two modules:

- First module is drawing out Indian Sign Language (ISL) gestures from real-time video and mapping it with human-understandable speech.
- Second module will take natural language as input and map it with equivalent Indian Sign Language animated gestures.

Processing from video to speech will include frame formation from videos, finding region of interest (ROI) and mapping of images with language knowledge base using Correlational based approach then relevant audio generation using Google Text-to-Speech (TTS) API. The other way round, natural language is mapped with equivalent Indian Sign Language gestures by conversion of speech to text using Google Speech-to-Text (STT) API, further mapping the text to relevant animated gestures from the database.

The use of mobile phones in our day to day life is increasing significantly. The usage of Internet along with smart devices has given a significant advancement to the mankind. The hearing disabled people are treated in an unusual manner. They are deprived of the social event due to lack of ability of communicating with the others. The image processing techniques are used here to build the application along with the methodologies of the human computer interactions.

The methodologies of building this application consists of two modules.

A. *Conversion of real time video (ISL) to equivalent human natural language speech.*

The process of drawing out gestures from real-time video and then converting it to human understandable language is depicted in the figure.
The following phases are followed:

- Frame formation from real time video
- Pre-processing and noise removal
- Finding ROI using Region growing
- Mapping using Correlation based approach
- Relevant audio generation using Google TTS API

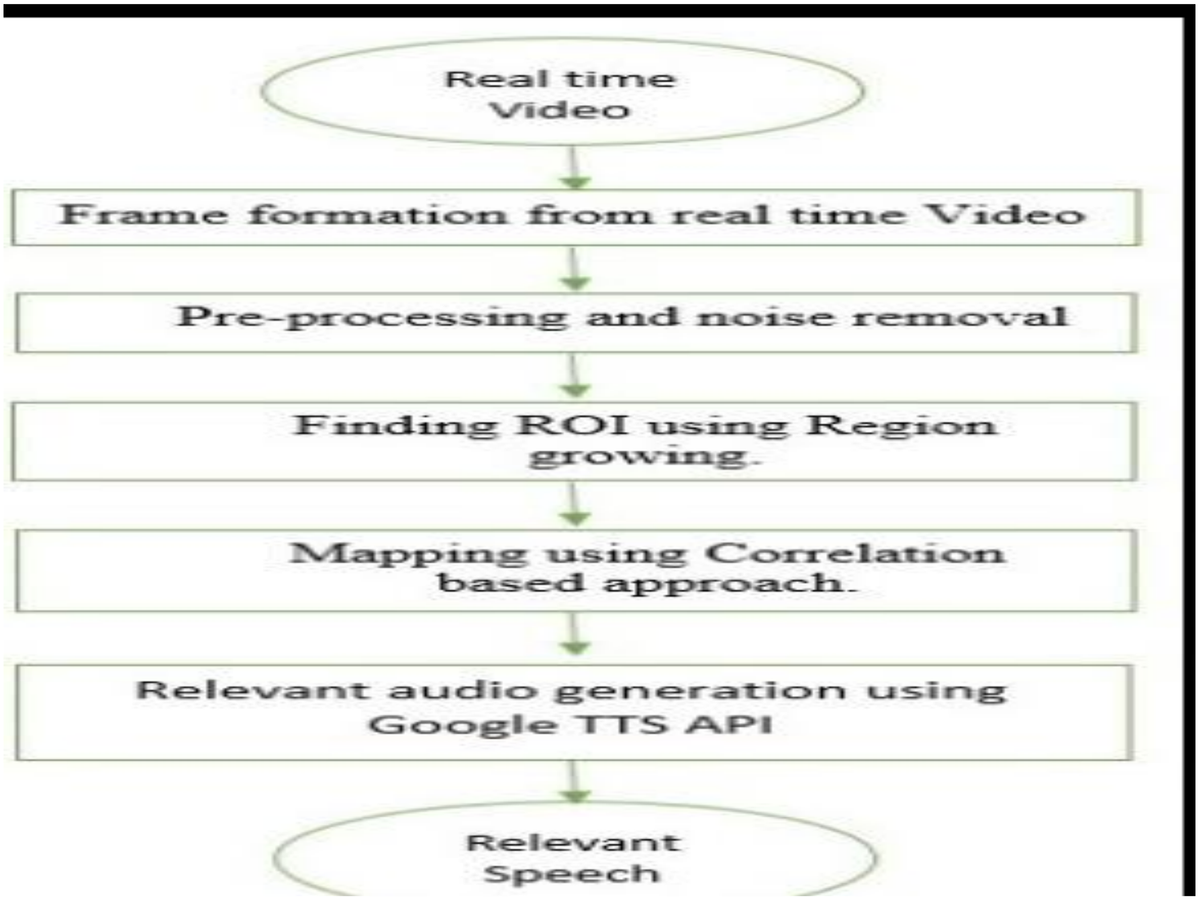

Figure 2.3: Overview of real-time video to speech

B. Mapping of human speech (natural language) to animated videos corresponding to ISL gestures.

As shown in fig.2, the process of converting natural
Language to corresponding gesture animation equivalents comprises of following phases:

- Generation of text from real time audio speech using Google STT API
- Mapping text to gesture images in database.
- Sequencing the images to form video
- Display animated gesture video.

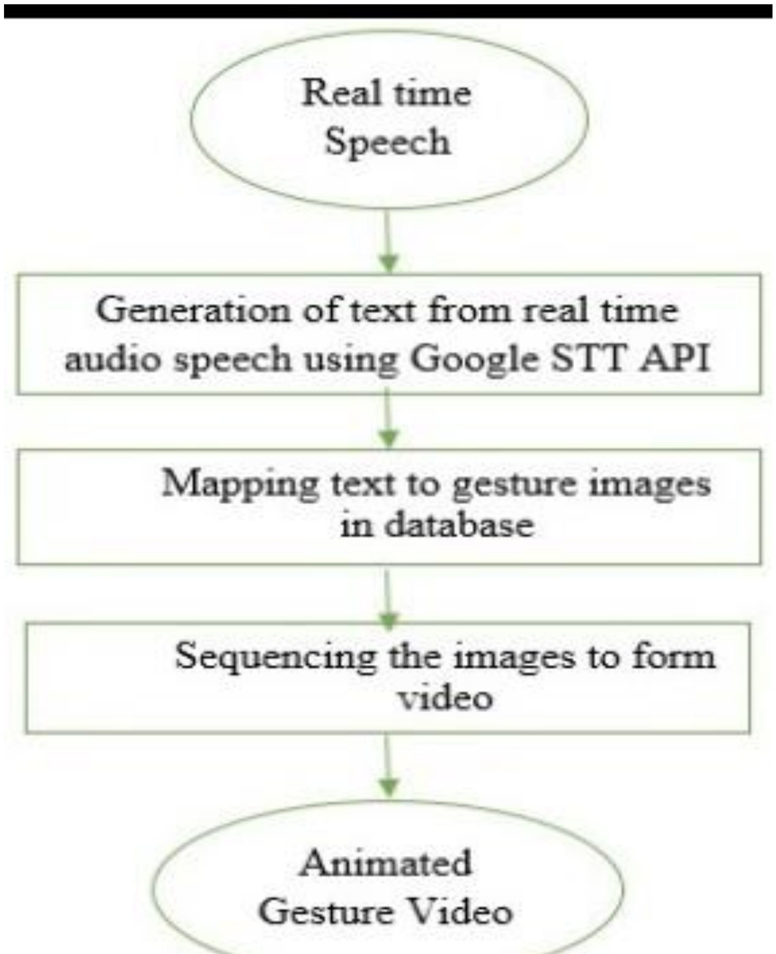

Figure 2.4: Overall view of speech to animated video

2.4 A Vibrotactile Sensory Substitution Device for the Deaf and Profoundly Hearing Impaired

The cochlear implant (CI) is an effective solution for regaining hearing capabilities for certain populations within this group, but not for all. First, CIs are expensive, ranging from $40,000 to $90,000 depending on the age of the recipient. This places CIs out of economic reach for many. Second, CIs require invasive surgery. Third, there is low efficacy of late CI implantation in early-onset deaf adults. A low-cost, non-invasive, plasticity-based solution to deliver auditory information to the brain were implanted. A "vibratory vest" was built by which auditory information is captured, digitally processed, and delivered to the skin of the torso using small vibratory motors.

Vibratory vest is used for communication purpose. It is a one-way communication tool for deaf and mute people. Here, information is sent to the brain via pattern of vibration on the screen. Our body can feel the vibration and according to the vibration the deaf people understands what are being told to them. The vest is connected with smart devices via Bluetooth and the speech is turned into pattern of vibration. There are about 32 motors attached to the jacket which are the source of the vibrations. Lights are set to the vest to watch if the motors are active while speaking. A game is developed as a learning purpose of the vibration pattern.

The vest is connected with a Bluetooth device such as smart phones. There are motors connected to the vest which makes the vibration. The vest can be worn under or over the dress. The external sound is compressed to map such
a way that for each word there will be some different vibration pattern. Vibration will be output of the external sound (input).

Figure 2.5: Vibrotactile Vest

Chapter 3

Design

This section states the description of the design of our proposed approach for communication with deaf and mute people. In our design mainly google STT, google TTS, Vibrotactile method, contextual approach, Bluetooth, Wi-Fi communication, smart phone, smart watch etc. were used.

3.1 Google Text to Speech

Google Text-to-Speech is a screen reader application developed by Android, Inc. for its Android operating system. It powers applications to read aloud (speak) the text on the screen which support many languages. Text-to-Speech may be used by apps such as Google Play Books for reading books aloud, by Google Translate for reading aloud translations providing useful insight to the pronunciation of words, by Google Talkback and other spoken feedback accessibility-based applications, as well as by third-party apps. Users must install voice data for each language.

We used Google TTS to convert the text to speech. When a mute user types and send his response to the normal people the text is then been turned to speech.

3.2 Google Speech to Text

It is just the opposite of google text to speech functionality. It is used in our application. A normal person talks and then the speech is converted to text for understanding of the deaf and mute.

3.3 Vibrotactile Approach

The majority of electronics offering haptic feedback use vibrations, and most use a type of eccentric rotating mass (ERM) actuator, consisting of an unbalanced weight attached to a motor shaft. As the shaft rotates, the spinning of this irregular mass causes the actuator, and in turn, the attached device, to shake.

Some newer devices, such as Apple's notebooks and iPhones featuring the "Taptic Engine", accomplish their vibrations with a linear resonant actuator (LRA), which moves a mass in a reciprocal manner by means of a magnetic voice coil, similar to how speaker technology translates DC electrical signals into motion of its speaker cone. LRAs are capable of quicker response times than ERMs, and thus are able to transmit more accurate haptic imagery.

Piezoelectric actuators are also employed to produce vibrations, and offer even more precise motion with less noise and in a smaller platform, but require higher voltages than the ERM and LRA implementations, and may be more fragile.

We used mobile or smart watches for vibration as an approach of alarming the deaf and mute. Vibration pattern is set according to syllable and words having same syllable can be distinguished by duration of vibration.

3.4 Contextual Approach

Some words, the contexts are shown as pictures so that the contexts are visible and understandable. Suppose a user said ‘Classroom’ and a picture of classroom will be visible in the mobile or smart watch.

3.5 Bluetooth and Wi-Fi Communication

Bluetooth is managed by the Bluetooth Special Interest Group (SIG), which has more than 30,000 member companies in the areas of telecommunication, computing, networking, and consumer electronics. The IEEE standardized Bluetooth as IEEE 802.15.1, but no longer maintains the standard. The Bluetooth SIG oversees development of the specification, manages the qualification program, and protects the trademarks. A manufacturer must meet Bluetooth SIG standards to market it as a Bluetooth device. A network of patents apply to the technology, which are licensed to individual qualifying devices.

Wi-Fi or WiFi (/ˈwaɪfaɪ/) is a technology for wireless local area networking with devices based on the IEEE 802.11 standards. *Wi-Fi* is a trademark of the Wi-Fi Alliance, which restricts the use of the term *Wi-Fi Certified* to products that successfully complete interoperability certification testing.

Devices that can use Wi-Fi technology include personal computers, video-game consoles, phones and tablets, digital cameras, smart TVs, digital audio players and modern printers. Wi-Fi compatible devices can connect to the Internet via a WLAN and a wireless access point. Such an access point (or hotspot) has a range of about 20 meters (66 feet) indoors and a greater range outdoors. Hotspot coverage can be as small as a single room with walls that block radio waves, or as large as many square kilometers achieved by using multiple overlapping access points.

We used Bluetooth and Wi-Fi communication for establishing communication between the end users. Mobile or smart watches are connected via either Bluetooth or Wi-Fi.

3.6 Smart Watch and Smart Phones

Wi-Fi or WiFi (/ˈwaɪfaɪ/) is a technology for wireless local area networking with devices based on the IEEE 802.11 standards. *Wi-Fi* is a trademark of the Wi-Fi Alliance, which restricts the use of the term *Wi-Fi Certified* to products that successfully complete interoperability certification testing.[1]

Devices that can use Wi-Fi technology include personal computers, video-game consoles, phones and tablets, digital cameras, smart TVs, digital audio players and modern printers. Wi-Fi compatible devices can connect to the Internet via a WLAN and a wireless access point. Such an access point (or hotspot) has a range of about 20 meters (66 feet) indoors and a greater range outdoors. Hotspot coverage can be as small as a single room with walls that block radio waves, or as large as many square kilometers achieved by using multiple overlapping access points.

While internal hardware varies, most have an electronic visual display, either backlit LCD or OLED or Hologram. Some use transflective or electronic paper, to consume less power. Most have a rechargeable battery and many have a touchscreen. Peripheral devices may include digital cameras, thermometers, accelerometers, pedometers, heart rate monitors, altimeters, barometers, compasses, GPS receivers, tiny speakers, and SD card (that are recognized as a storage device by a computer).

A smartphone is a handheld personal computer with a mobile operating system and an integrated mobile broadband cellular network connection for voice, SMS, and Internet data communication; most if not all smartphones also support Wi-Fi. Smartphones are typically pocketsized, as opposed to tablets, which are much larger than a pocket. They are able to run a variety of third-party software components ("apps") from places like the Google Play Store or Apple App Store, and can receive bug fixes and gain additional functionality through operating system software updates. Modern smartphones have a touchscreen color display with a graphical user interface that covers the front surface and enables the user to use a virtual keyboard to type and press onscreen icons to activate "app" features. They integrate and now largely fulfill most people's needs for a telephone, digital camera and video camera, GPS navigation, a media player, clock, news, calculator, web browsing, handheld video games, flashlight, compass, an address book, a note-taking application, digital messaging, an event calendar, etc.

We established communication between smart phones or smartphone and smart watches. As these are very popular now a days and cheap, it is much feasible using these devices for communication.

Chapter 4

Algorithms

In mathematics and computer science, an algorithm is an unambiguous specification of how to solve a class of problems. Algorithms can perform calculation, data processing and automated reasoning tasks.

An algorithm is an effective method that can be expressed within a finite amount of space and time and in a well-defined formal language for calculating a function. Starting from an initial state and initial input (perhaps empty), the instructions describe a computation that, when executed, proceeds through a finite number of well-defined successive states, eventually producing "output" and terminating at a final ending state. The transition from one state to the next is not necessarily deterministic; some algorithms, known as randomized algorithms, incorporate random input.

In our proposed methodology, we will use Google Speech to Text, De-Syllable and Vibration methods to build the system. This section deals with the algorithms used in the foundation of the prototype. Each algorithm takes an input and produces an output based on which another algorithm can produce an output. That is, the input of one algorithm depends on the output of the other algorithms. The efficiency of the algorithms is a fact here. The more efficient the algorithm is, the smaller the run time. Efficient and correct algorithms lead to produce correct output in a minimum amount of time.

4.1 De-Syllable

The code snippet given below finds out the number of syllables of a word.

```
private int countSyllables(String word) {
     //The letter 'y' can be counted as a vowel, only if it
     //creates the sound of a vowel (a, e, i, o, u).
     char[] vowels = { 'a', 'e', 'i', 'o', 'u', 'y' };
     char[] currentWord = word.toCharArray();
     int numVowels = 0;
     boolean lastWasVowel = false;
     for (char wc : currentWord) {
         boolean foundVowel = false;
         for (char v : vowels)
         {
             //don't count diphthongs
             if ((v == wc) && lastWasVowel)
             {
                 foundVowel = true;
                 lastWasVowel = true;
                 break;
             }
             else if (v == wc && !lastWasVowel)
             {
                 numVowels++;
                 foundVowel = true;
                 lastWasVowel = true;
                 break;
             }
         }
         // If full cycle and no vowel found, set lastWasVowel to false;
         if (!foundVowel)
             lastWasVowel = false;
     }
     // Remove es, it's _usually? silent
     if (word.length() > 2 &&
             word.substring(word.length() - 2) == "es")
         numVowels--;
         // remove silent e
     else if (word.length() > 1 &&
             word.substring(word.length() - 1) == "e")
         numVowels--;
     //Toast.makeText(this, "" + numVowels, Toast.LENGTH_LONG).show();
     return numVowels;
 }
```

Figure 4.1: Code for De-Syllable

4.2 Vibration

The code snippet given below generates vibration from the number of syllable. The number of vibrations will be equal to the number of syllable which can be obtained from the de-syllable algorithm.

```
private void vibration(int syllable) {
    pattern.addElement("0");
    Vibrator vk = (Vibrator) getSystemService(Context.VIBRATOR_SERVICE);

    for(int i=0;i<syllable;i++) {

        pattern.addElement("500");
        pattern.addElement("500");
    }

        int j=pattern.size();
        anArray=new String[j];
        pattern.toArray(anArray);
        longArray = new long[j];

        for (int i = 0; i < j; i++) {
            longArray[i] = Long.parseLong(anArray[i]); // <-- Fixed.
            //ohh[i] = getDateCurrentTimeZone1(Timestamp1[i]);
        }
        vk.vibrate(longArray,-1);
        pattern.clear();
    }
```

Figure 4.2: Code for Vibration

4.3 Google Speech to Text

The code snippet given below converts human speech to text.

```
 /**
 * Showing google speech input dialog
 */
private void promptSpeechInput() {
    Intent intent = new Intent(RecognizerIntent.ACTION_RECOGNIZE_SPEECH);
    intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE_MODEL,
            RecognizerIntent.LANGUAGE_MODEL_FREE_FORM);
          intent.putExtra(RecognizerIntent.EXTRA_LANGUAGE,
  Locale.getDefault());
    intent.putExtra(RecognizerIntent.EXTRA_PROMPT,
            getString(R.string.speech_prompt));
    try {
        startActivityForResult(intent, REQ_CODE_SPEECH_INPUT);
    } catch (ActivityNotFoundException a) {
        Toast.makeText(getApplicationContext(),
                getString(R.string.speech_not_supported),
                Toast.LENGTH_SHORT).show();
    }
}

/**
 * Receiving speech input
 */
@Override
      protected void onActivityResult(int requestCode, int resultCode,
  Intent data) {
    super.onActivityResult(requestCode, resultCode, data);

    switch (requestCode) {
        case REQ_CODE_SPEECH_INPUT: {
            if (resultCode == RESULT_OK && null != data) {

                ArrayList<String> result = data

  .getStringArrayListExtra(RecognizerIntent.EXTRA_RESULTS);
                //txtSpeechInput.setText(result.get(0));

                 //Modified Here
                List<ActionPair> contacts = db.getAllActionPair();
                int commandFound = 0;
                int contextFound = 0;
                String actualCommand = null;
                String actualContext = null;
                for (ActionPair cn : contacts) {
                    String command = cn.getCommand().toUpperCase();
                    String pattern = "\\b" + command + "\\b";
```

```
                Pattern p = Pattern.compile(pattern);
                Matcher m = p.matcher(result.get(0).toUpperCase());
                if (m.find()) {
                    commandFound = 1;
                    actualCommand = command;
                    break;
                }
            }

            for (ActionPair cn : contacts) {
                String context = cn.getContext().toUpperCase();
                String pattern = "\\b" + context + "\\b";
                Pattern p = Pattern.compile(pattern);
                Matcher m = p.matcher(result.get(0).toUpperCase());
                if (m.find()) {
                    contextFound = 1;
                    actualContext = context;
                    break;
                }
            }

            if (commandFound == 1 && contextFound == 1) {
                String log = actualCommand + " " + actualContext;
                txtSpeechInput.setText(log);

                this.finish();
                Intent intent = new Intent(this, Display.class);
                Bundle bundle = new Bundle();
                bundle.putString("command", actualCommand);
                bundle.putString("context", actualContext);
                intent.putExtras(bundle);
                startActivity(intent);

            } else {
                txtSpeechInput.setText("Command not recognized");
            }

            break;
        }

    }
}
}
```

Figure 4.3: Code for Google Speech to Text

4.4 Text to Speech

The code snippet given below converts text to speech.

```
t1=new TextToSpeech(getApplicationContext(), new
TextToSpeech.OnInitListener() {
         @Override
         public void onInit(int status) {
             if(status != TextToSpeech.ERROR) {
                 t1.setLanguage(Locale.UK);
             }
         }
     });

     b1.setOnClickListener(new View.OnClickListener() {
         @Override
         public void onClick(View v) {
             String toSpeak = ed1.getText().toString();
                   Toast.makeText(getApplicationContext(),
   toSpeak,Toast.LENGTH_SHORT).show();
             t1.speak(toSpeak, TextToSpeech.QUEUE_FLUSH, null);
         }
```

Figure 4.4: Code for Text to Speech

Chapter 5

Implementation and Evaluation

Implementation evaluation asks about the practical lessons that emerge from putting a new project into action. Rarely does a project go off without a hitch, and lessons learned during implementation help organizations identify if an approach may need to be modified and what critical next steps are required. In turn, these lessons can help others avoid the same pitfalls. Finally, they teach the Foundation important lessons that can help improve our grant making capacity.

Implementation evaluation is the appropriate focus for the evaluation of:

- Demonstration projects for training or service delivery where the intervention/training model is still undergoing development
- Projects that seek to replicate an existing model in one or more new settings, or with a different population
- Planning and seed grants
- Service expansion grants
- Technical assistance grants
- Advocacy and community organizing grants

5.1 Implementation

We propose one to one and two-way communication for deaf and mute with the normal people. It is an adaptive and multimodal approach. To implement the whole system natural language is needed to process to generate text from speech using google STT API. The text will then be divided into two parts- command and context. Vibration will be generated for command sets. Commands and contexts will be shown as visual output to the smart devices. As this is a two-way communication, the deaf and mute also need to be reply to the normal people. In this case some predefined answers will be set. The deaf or mute people have to tap on those answers if they want those as reply or they can write text if they want some other answers to be given to the normal people. The text will be converted to speech using Google TTS. The devices will be connected via Bluetooth or Wi-Fi.

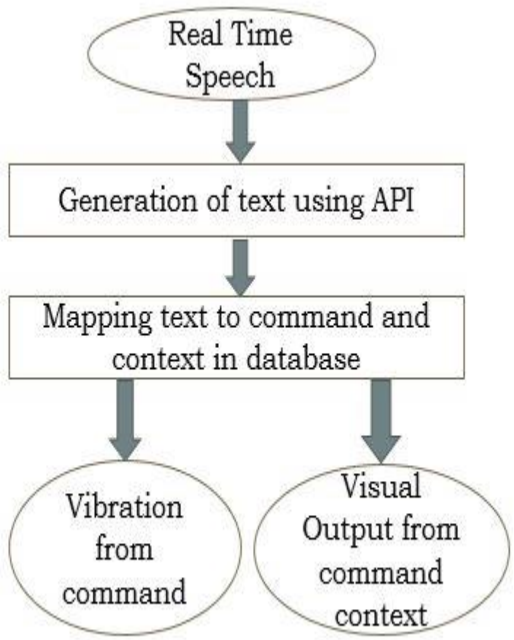

Fig: Overall view from speech to Vibrotactile and visual output

Figure 5.1: Overall view from speech to Vibrotactile and visual output

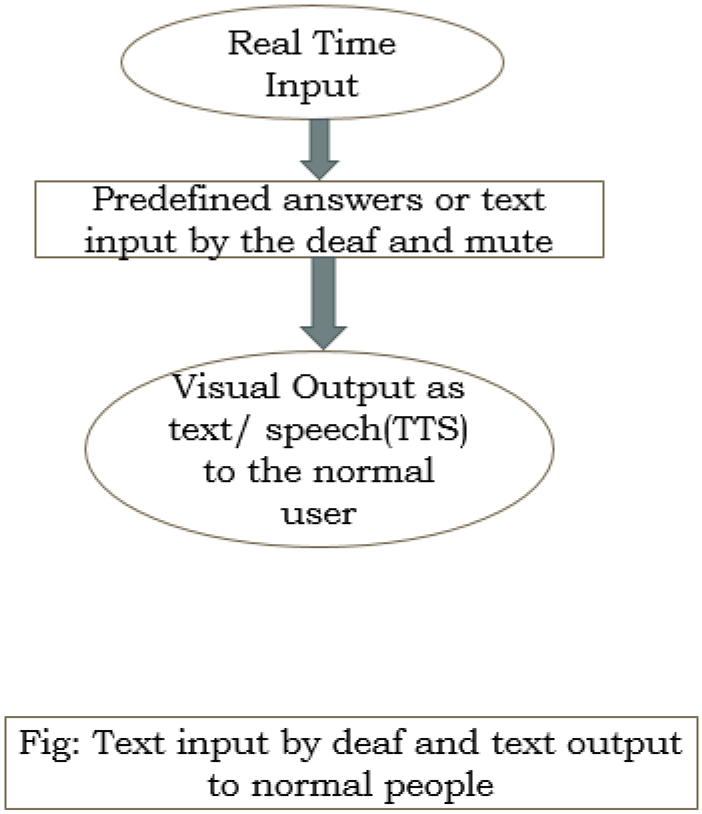

Figure 5.2: Text input by deaf and text output to normal people

Our proposed system will first find out the number of syllables of the command part. The vibration pattern depends on the number of syllable of a word. The vibration time will be proportional to the syllable. For example, if the command part consists of two syllables the then the smart watch will vibrate two time and each vibration time will be 2 seconds.

The context part will be matched with the exact picture that represents the context and will be shown visually. The figure 5.3 depicts vibration pattern and visual feedback.

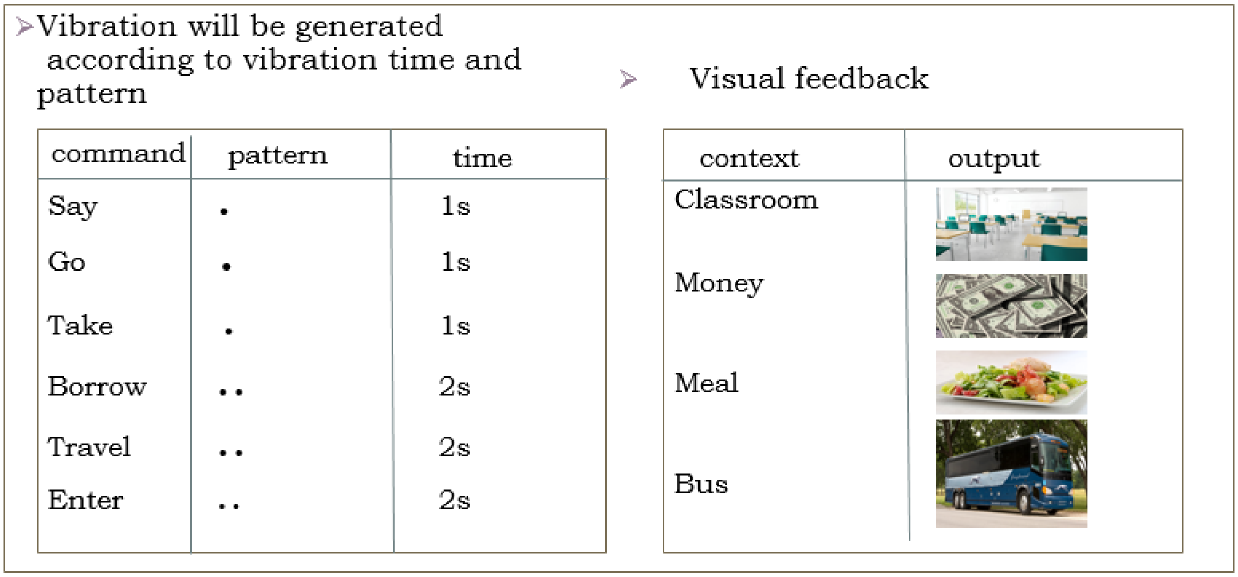

Figure 5.3: Vibration pattern and visual output

The system’s working procedure:

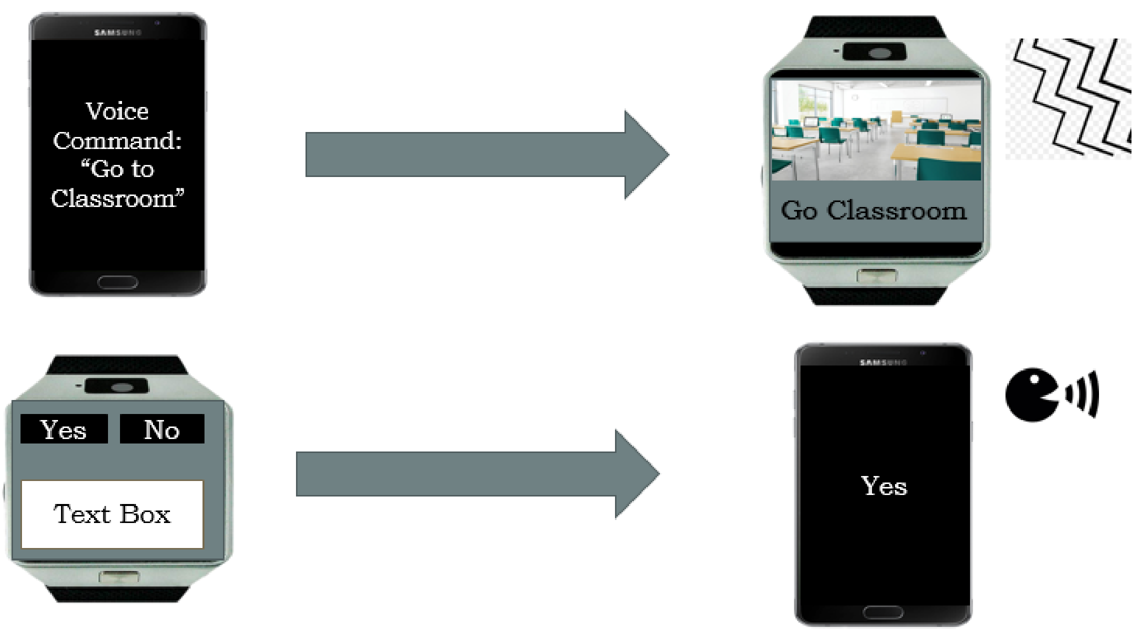

Figure 5.4: Proposed system's working procedure

5.2 Evaluation

We couldn’t evaluate our app taking information from deaf and mute but we did a heuristic evaluation on our app.

This heuristic evaluation includes visibility, match between system and real world, user control, consistency and standard, recognition rather than recall, flexibility and minimalistic design.

Visibility: This evaluation actually indicates the app visibility. How the app looks like, the color combination, proper placements of texts, button etc. are evaluated.

Match between system and real world: This evaluation refers to the matching and solving between real life problem and the app solution approach.

User control: This evaluation refers to the controlling system of the application by the user. It mainly indicates to how easy this app is to use.

Consistency and standard: This evaluation indicates the consistency of the app. It also evaluates whether this app followed the standard model of generalized app for better and easy usability.

Recognition rather than recall: This evaluation indicates whether a user has to recall the app systems to use or not and whether the user easily recognize the procedure or not.

Flexibility: This evaluation indicates the flexibility of the app. It indicates the user friendly environment of the app.

Minimalistic design: It is better to have minimal design with more information. It looks better to look at and easily understandable. This evaluation verifies this of our app.

Heuristic Checklist	Person-1	Person-2	Person-3	Person-4	Person-5	Average
Visibility	4	5	5	3	2	3.8
Match Between System and Real World	4	4	5	3	3	3.8
User Control	5	5	4	4	4	4.4
Consistency and Standard	3	4	5	4	4	4
Recognition rather than recall	5	4	3	5	3	4
Flexibility	3	5	4	3	4	3.8
Minimalistic Design	3	5	4	4	4	4

Figure 5.5: Heuristic Evaluation

Chapter 6

Conclusion and Future Works

Our proposed method is very easy to learn. It helps to create communication to both end users. Even the pictorial and Vibrotactile process help the deaf and mute to understand without learning any sign languages. We will try to make the system adaptive. We will try to make the system adaptive. New words which will not be found in database will be included in DB with the help of NLP and AI. NLP and AI is needed to differentiate the command and context set. Duration per vibration will be set if the syllable is same. For better learning and better acquaintance with the system a game is planned to be developed.

Chapter 7

References

[1] Matthias Mielke and Rainer Br¨uck \ AUDIS Wear: a Smartwatch based Assistive Device for Ubiquitous Awareness of Environmental Sounds , 2016, IEEE

[2] Suranga Nanayakkara, Lonce Wyse and Elizabeth A. Taylor \ The Haptic Chair as a Speech Training Aid for the Deaf, 2012, OZCHI

[3] Prof. Prashant G. Ahire, Kshitija B. Tilekar, Tejaswini A. Jawake and Pramod B. Warale \ Two Way Communicator Between Deaf and Dumb People And Normal People, 2015, IEEE

[4] Scott D. Novich and David M. Eagleman \ A Vibrotactile Sensory Substitution Device for the Deaf and Profoundly Hearing Impaired, 2014,IEEE Haptics Symposium (HAPTICS)

[5] R. A. D. K. Rupasinghe, D. C. R. Ailapperuma, P.M.B.N.E. De Silva and A. K. G. Siriwardana \ A Portable Tool for Deaf and Hearing Impaired People, 2014, ITRU Research Symposium

[6] Xiujuan Chai, Guang Li, Xilin Chen, Ming Zhou, Guobin Wu and Hanjing Li \VisualComm: A tool to support Communication between deaf and hearing person with Kinect, 2013, *ASSETS*

[7] Sangwook Lee, YuKyoung Lee and Yunho Kang \ LaneMate: Car Sensing Systemfor the Deaf, 2016, CHI

[8] http://www.ef.com/english-resources/english-vocabulary/top-50-verbs/