Voice Controlled Wheelchair Using Feature Extraction Techniques Through MATLAB

Saumya Sah

Department of Electronics and Communication Engineering, Birla Institute of Technology, Mesra Ranchi, India. saumya12493@gmail.com

Abstract – Modern lifestyle and the desire to achieve more is the main cause of increase in number of person suffering from mobility impairments with no hands to take care. Hence there is a need of development of such mobility vehicles which ease the lifestyle of people suffering from mobility impairments. Many wheelchairs are there in the market which do ease their life to an extent but there is still a need of innovation so that they can freely move on their own. So a voice based wheelchair with high efficiency is proposed in this paper which will navigate according to the command given by the user and hence will require no amount of physical work. In this paper, a model is proposed for voice controlled wheelchair using MATLAB with the help of different recognition techniques. The recognized speech is sent to Arduino microcontroller which decides the motion of the designed bot. There are six different commands that are recorded for training purpose of the system. Different combinations of feature vectors and classifiers were experimented for recognition efficiency. For testing this model, speech is fed to the system. Once a word is successfully recognized, the bot is navigated in desired direction through the Arduino microcontroller interfaced with MATLAB. An overall efficiency of 99.1% is found for the system.

Keywords – voice recognition, ANN, GMM, MATLAB, MFCC, LPC

I. INTRODUCTION

Wheelchair is an important mobility equipment for the people who are suffering from different degree of mobility impairments. Several development on the wheelchair have been made for providing an easy going lifestyle to these kind of people. These modification are made so that the flexible demand of the user according to their needs like sports, stairs climbing, avoiding obstacle, tank chair etc. could be achieved. Recliner wheelchair is also constructed so that the user can adjust the wheelchair according to their comfort [1]. There are mainly two kinds of wheelchair which are most popular. These are hand-operating and joystick based control. These systems require at least some amount of muscular strength to generate the desired motion. This need of muscular strength is difficult for old aged and handicapped person. Hence there is a great need for intelligent wheelchair which can be controlled via speech or gesture for comfortable navigation and better safety of the user. [2], [3].

Latest works in this field focus on the use of different gestures and body movement to control the direction of motion. Some of them are voice [4], [5], face [6], eye [7] or electromyography from neck and face muscles[8], [9] and tongue movement [10], and by using electrooculography (EOG) [11], by head tilt [12], mouth shape or tongue drive [13], also the use of gesture and touchpad [14]. These gestures are useful for navigation of a handicapped person but using them for wheelchair movement in long run can cause unnaturalness in the user. The wheelchair becomes very costly when multiple features are added like control and monitoring systems with EMG, EEG techniques and navigation with assistant. Speech, on the other hand is very natural way of communication and easy for usage by all old aged and paralyzed person. Various voice based wheelchair has been developed that made the movement of these people easy and according to their will. However, problems with these voice based wheelchair still exist. Speech has several application frameworks and one of those frameworks could be this project work which will provide the physically handicapped persons an opportunity to move freely all by themselves and that too without any external help. The powered wheelchair is integrated with the voice recognition system and hence it will help in navigation as well as the propulsion of the wheelchair.

In this paper, I report an attempt to build a Voice controlled wheelchair module. Features like MFCC and LPC were tested for accuracy for both isolated and continuous text with GMM and ANN for designing the model. This paper is organized as follows: Section II describes the hardware structure of the models. Section III gives an overview of the different development stages. Section IV presents the simulation and results from

 DOI : 10.18231/2454-9150.2018.1308

different combinational models and Section V finally concludes the paper.

II. HARDWARE STRUCTURE

This section discussed about the hardware structure of the proposed model for voice based wheelchair. The elements of the models is the bot chassis, a replica of the actual wheelchair along with the Arduino board required for the control of the bot movement. Fig.1. shows the block diagram of the proposed model of the system. The system also include MATLAB based speech recognition system which will be discussed later.

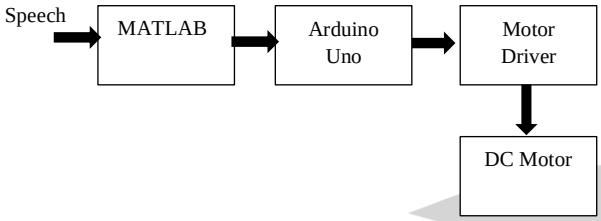

Fig.1.Block diagram for offline model

Table I briefs the hardware components required to build the offline model.

TABLE I. LIST OF COMPONENTS

Sl. No.	Components	Purpose
1	Laptop	Run MATLAB application
2	MATLABR2014a	Used to build models to convert speech to text.
3	Arduino package for MATLAB	Used by MATLAB to communicate with Arduino
4	Arduino microcontroller	Use to run decision algorithm depending on recognized word
5	L293D	Used as motor driver
6	DC Motor	Used to drive the bot

ARDUINO BOARD

Arduino Uno R3 microcontroller Board is used to prepare the system which is based on ATmega328P. It contains everything that needed to support the microcontroller. Figure 2 shows the Arduino Uno Board. It simply needed to connect it with a computer via USB cable or power it with an AC-to-DC adapter or battery to get it started. The programming for the Arduino Uno board is done through the Arduino IDE software where the program is written on the sketch and uploaded it to the Arduino Uno board. The board hence programmed with the instruction that needed for the movement of the bot.

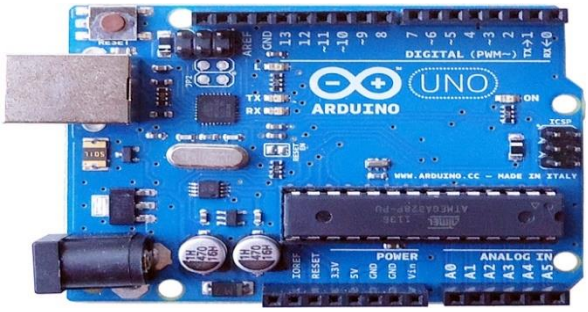

Fig.2. Arduino Uno Board

MOTOR DRIVER CIRCUIT

L293D module is used to drive the motor that is used to design the replica of wheelchair. It uses H- Bridge circuit. Figure 3 shows working of Motor Driver.

Specifications of the board are:

•Medium power motor driver for driving DC Motors and

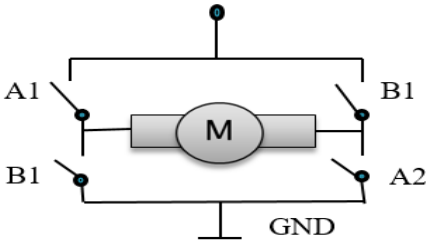

Fig.3. Working of Motor Driver

Stepper Motors

•Uses L293 motor driver IC.

•It can drive 4 DC motor on/off or 2 DC motor with direction and speed control.

•Maximum motor supply voltage: 36V

•Maximum motor supply current: 600 mA per motor

DC MOTOR

Fig.4. Battery Operated DC Motor

In this proposed system, DC motors which are four in number are used. These DC motors are of battery operated type with specification of 100 rpm for high torque and of 12 V power supply. Figure 4 shows the battery operated Dc Motor.

The voice recognition is done with the help of MATLAB. The system is trained with the previously prepared speech models of different words which are saved in the system. Input files are fed to this system and the recognized words obtained from the system. These outputs are used to generate different motor commands by sending digital signals to the Arduino board via USB cable. This is done by using Arduino package for MATAB as a user interface. Fig. 5. Shows the circuit diagram of the system.

 DOI : 10.18231/2454-9150.2018.1308

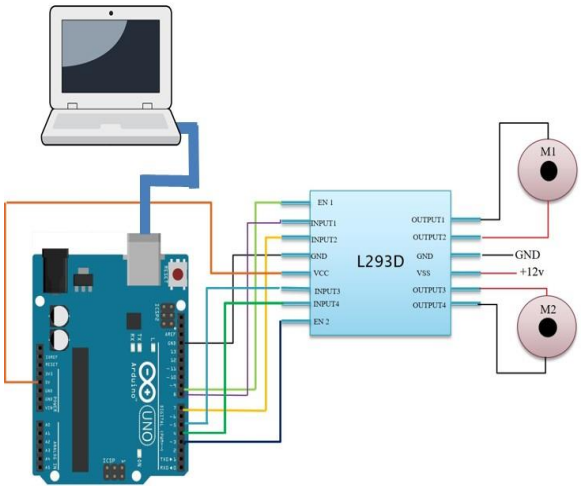

Fig.5. Circuit Diagram of System

III. SYSTEM DEVELOPMENT

This section describes the system development of offline mode in detail. This requires preparation in different stages like preparing the raw data then pre-processing the data so that the data should become as it is required for preparing the model, feature extraction etc. The feature vectors are extracted through the pre-processed data which is done by different feature extraction techniques. These extracted feature vectors are then go through comparison with the stored data with the help of classifiers. According to the comparison result the mostly closed feature vector is decided as the recognized word and hence the recognized word is obtained. The section is divided into three parts: Initially, data preparation method is discussed which is followed by feature extraction techniques and classifiers.

A. Data Preparation

System is prepared to recognize six commands for wheelchair control. These commands are "START", "FORWARD", "BACKWARD", "LEFT", "RIGHT", and "STOP". Each of these commands are recorded for 100 times making a total of 600 wav files from a single speaker. Mono channel recording is done at a sampling rate of 44100 samples per seconds with a resolution of 16 bits per sample in a quite lab environment. Out of these 100 files, 80% is used for training the system and rest 20% is used for testing.

B. Feature Extraction

For this experiment, two traditional features were used. Initially, LPC (Linear Prediction Filter Coefficients) feature was extracted. 12th order filter is used to modeled the Vocal tract. The pre emphasis speech signal is divided into frames and over these frames windowing process is performed. The Auto correlation function is used over the resulting frames of windowing speech sample signal in order to obtain the LPC parameter. Fig.6 shows the block diagram of LPC technique. The major drawback of this type of feature extraction technique is that it does not perform well in noisy environment and can give great results in clean environment.

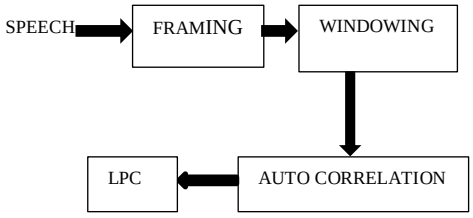

Fig.6. Block Diagram of LPC Feature Extraction Technique

The other feature used for this experiment was MFCC (Mel Frequency Cepstral Coefficient). Fig.7. shows the block diagram of MFCC technique. The MFCC technique is considered because it is the most popular and effective feature extraction technique for speech recognition. It approximates the human system response more closely than any other system. MFCC is used as the acoustic features of human voice. It considers the human voice pitch in the form of frequencies and equally scale them on the Mel scale, these extracted feature are unique to others.

In order to perform MFCC, the signal is divided into small overlapping frames. These frames are then entered into windowing process which is done by using hamming window. After this, fast Fourier transform (FFT) is use to extract the feature vector and the power spectrum of each windowed signal is calculated. The power spectrum hence obtained is applied to Mel frequency wrapping using bank filters. The logs of the powers are taken at each of the Mel frequencies. Then the resultant signal is transformed using an inverse DFT into Cepstral domain. The lower order coefficients are selected as the feature vector whereas the higher coefficients are discarded as it contains less specific information about speaker. Then the coefficients are uniformly spaced and used as output feature vector for that speech frame

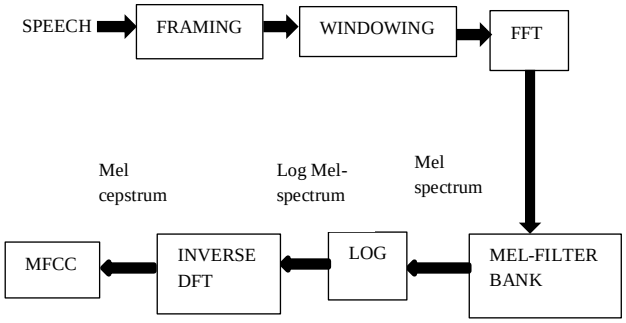

Fig.7. Block Diagram of MFCC Feature Extraction Technique

 DOI : 10.18231/2454-9150.2018.1308

C. Classifiers

Gaussian Mixture Model (GMM) can be explain as a classifier that is used to compare the features that are extracted from the MFCC technique with the stored samples of voice commands. Gaussian Mixture Model is a probabilistic model represented by its Gaussian distribution. In this all the data points generated are derived from a mixture of a finite Gaussian distribution which has unknown parameters. Each Gaussian distribution is calculated by its mean, variance and weighted sum of Gaussian components densities. Once the system is trained with the extracted features of the voice command samples, the evaluation takes place by the use of log likelihood algorithms. In evaluation process the log likelihood distance is calculated between the voice query and the voice command samples. The best result is the voice query which is closest to the trained voice command samples.

ANN is an information processing model consisting of a number of connected processing units or nodes called neurons. Each neuron accepts a weighted set of inputs and produces an output. These outputs from one neuron signals are transmitted to the other neuron connected to it. A neural network is a dense interconnection of simple, nonlinear, computational elements and hence it is a parallel distributed processing model. These artificial neurons are organized in number of layers. These different layers perform different types of transformations on its inputs. ANNs has many advantages due to its parallel distributed processing, distributed memories, error stability, and pattern learning and distinguishing ability. Among its several advantages, the most advantageous one is the fact that it can actually learn from observing data sets and hence it take data samples rather than entire data sets.

D. Back Propagation Algorithm

It is the training or learning algorithm. It learns by example. If you submit to the algorithm the example of what you want the network to do, it changes the network's weights so that it can produce desired output for a particular input on finishing the training. Back Propagation Networks are ideal for simple Pattern Recognition and Mapping Tasks.

E. Block Diagram of Model Building

Fig.8. describes the block diagram of the isolated text model. Both training and testing data are isolated words of approximately 1 second of duration. Before feature extraction, these wave files are preprocessed for amplitude normalization to remove the loudness effect. These processed wav files are fed to the feature extraction block where different features like MFCC and LPC is computed. These features are then fed to GMM and ANN Classifier. Depending on the outcome probability, a word is recognized.

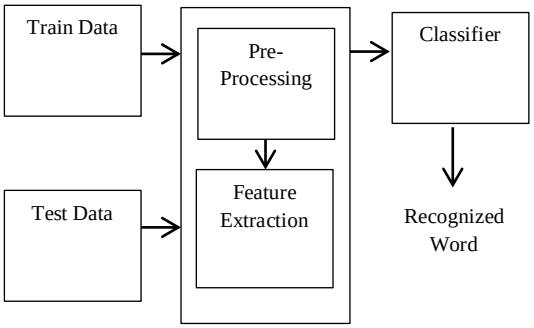

Fig.8.Block diagram of text model for system

IV. PREPARATION OF SPEECH RECOGNITION SYSTEM

The voice recognition system is done with the help of MATLAB. This can be performed in different stages. The first step includes the feature extraction of the voice command sample which was recorded and stored in early stages. The feature extraction process could be done through different techniques. MFCC and LPC techniques are being used for feature extraction which was already discussed in the earlier section of this thesis. The second step includes the training of the system with the extracted feature vectors of voice commands. This could be done through different modelling techniques like GMM and ANN. The system is trained to obtain accuracy for each command coded into the microcontroller. The third step is the recognition part where the system recognize the voice by matching voice sample of the user with the stored voice samples in the database. The output of voice recognition is then fed to the digital input ports of the microcontroller so that the wheelchair can navigate.

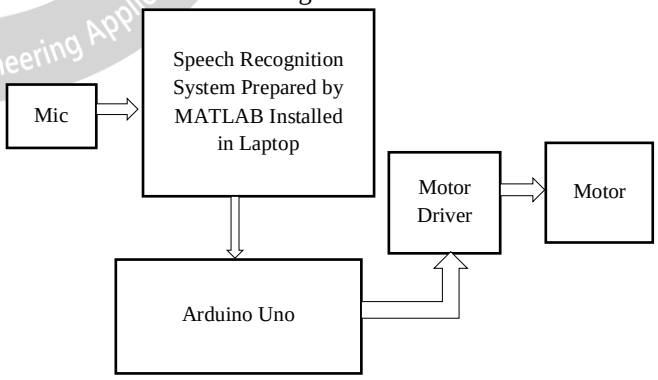

Fig.9. Block Diagram of System for voice based wheelchair

Figure 9 shows the block diagram of system for the voice based wheelchair bot. This block diagram includes one Mic which can be either externally fitted to the system or the internal microphone of the laptop could be used. The Voice recognition system is prepared in the Laptop using MATLAB. The Arduino board will be connected with the

 DOI : 10.18231/2454-9150.2018.1308

laptop to transmit the command to the motor driver to move the motor for the propulsion of the bot.

In this section, the key tools and steps that are used in preparing the speech recognition system over the MATLAB R2015a are discussed in brief. These are few steps of preparing the system and some toolbox installed in the MATLAB.

A Neural Network Toolbox

MATLAB has inbuilt ANN toolbox which helps us to implement the complex design to perform various kind of problems. Feed Forward type of neural network is performed in this toolbox. The Neural Network Toolbox provides algorithms, functions, and apps to create, train, visualize, and simulate shallow and deep neural networks. One can perform classification, regression, clustering etc. with the help of this neural network toolbox. ANNs has many advantages due to its parallel distributed processing, distributed memories, error stability, and pattern learning and distinguishing ability. Figure 10 shows the neural network toolbox of the MATLAB.

The toolbox includes convolutional neural network and auto-encoder deep learning algorithms for image classification and feature learning tasks like Input-Output Curve fitting, Pattern Recognition, and Classification, Clustering and Dynamic Time Series

B Arduino Support Package

The MATLAB Arduino Support Package is a support package that has to be installed in the MATLAB for transmitting the data over Arduino from the MATLAB based speech recognition system. This support package is functional for R2014a and beyond the version of MATLAB. This package enables us to use MATLAB to communicate with an Arduino board. We can read and write sensor data through the Arduino and immediately see the results in MATLAB without having to compile.

The task that this package enables us to perform are:

1. It can acquire analog and digital sensor data from your Arduino board
2. It can control other devices with digital and PWM outputs
3. It can drive different types of motors and also supports Adafruit Motor Shield.
4. It helps in Communicating with an Arduino board over a USB cable

C Preparation of the System

The steps involved in preparing the speech recognition system for the offline mode of the proposed system are described in this section. The voice recognition system is prepared through MATLAB R2015a software. These steps are given below

1. The database preparation is the first step in this mode. Six commands were recorded each for 100 times making a total of 600 command voice samples.
2. Feature extraction techniques are performed over these voice samples through which the feature vectors are obtained.
3. After extraction of features from the voice samples of commands. The 80% of voice samples are used to train the model where as 20% of the voice samples are used for testing the models.
4. The training of the model is done through different classification techniques.
5. When the train model is prepared successfully the testing is done and hence the recognised word is obtained.

By considering the above step, we can say that the voice recognition system is prepared successfully. The transmission of the commands to the Arduino Board is transmitted with the help of the MATLAB support package known as MATLAB Support Package for Arduino Board.

V. SIMULATION RESULTS

This section discusses the simulation results for the proposed model. The experimental setup for the system is shown in Fig.11. which consists of laptop, Arduino board, motor driver, bot and power supply. The system was tested with the help of the Speech recognition system prepared on the MATLAB. In order to obtain better test results, all the six commands are being tested one by one. For testing, the commands are being spoken over the laptop inbuilt microphone and the recognized command are shown in the MATLAB based voice recognition system. If the bot moves according to the command spoken, the test is successful and hence the requirement of the system is acquired.

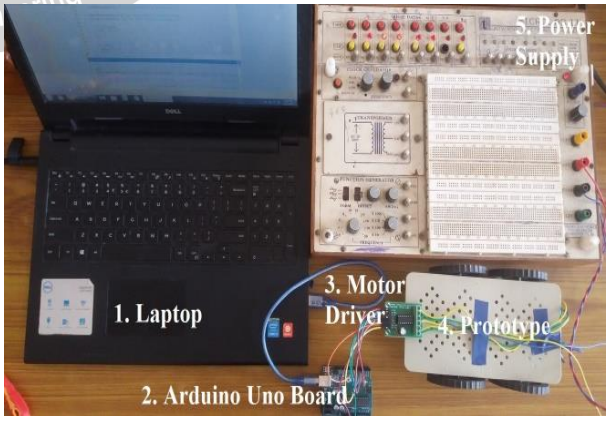

Fig.11.Showing Experimental Setup for system

Different combinations of feature vector and classifiers were used for modeling the system. The recognition efficiency is tabulated in Table II.

TABLE II EFFICENCY OF SYSTEM

 DOI : 10.18231/2454-9150.2018.1308

Sl. No.	Techniques	Efficiency
1.	MFCC + ANN	99.1%
2.	MFCC + GMM	98.3%
3.	LPC + ANN	72.5%
4.	LPC + GMM	62.5%

The performance of the system is highest to the model for MFCC feature when used with ANN. Table III shows the different parameters of the proposed system. . By observing the above four combination, one can state that the model prepared by the combination of Mel- frequency Cepstral Coeffcient and Feed Forward Neural Network is much better than the system prepared by the other combinations. The results are so because the main advantages of using Artificial Neural Networks is that it can handle a large amount of data sets and it has the ability to implicitly detect complex nonlinear variables whereas MFCC feature extraction technique has an advantage of better human auditory aspect i.e. proper and better representation of the speech signal, and also it has the ability to resists more to noise.

TABLE III PARAMETERS OF THE MODEL	
Sl. No.	**Parameters**
1.	The voice recognition is done through the models designed in MATLAB.
2.	The highest accuracy of this mode is 99.1% with MFCC+ANN
3.	It does not require Bluetooth module.
4.	The system is idle for rural places as it doesn't require an internet connection.
5.	Low hardware complexity
6.	User dependent system

From table 3 we can see that the operation of system should be preferred for designing an intelligent voice controlled wheelchair in places with no or poor internet access as it does not need an internet connection which is needed in most of the previous research work. The system is a user dependent model, as it is trained from data from a single speaker. It is less complex model and obtain highest accuracy of 99.1% with MFCC and ANN model.

VI. CONCLUSION

This paper develops a model for voice based wheelchair system. The model is prepared from different feature extraction techniques which provides different efficiency results. From the results we can say that the most efficient prepared model is with MFCC technique with ANN technique that is 99.1%. So the final model is prepared with the combination of MFCC and ANN techniques. Hence the purpose of the paper is achieved and the voice controlled wheelchair for physically handicapped person to ease their life is prepared.

ACKNOWLEDGMENT

I thank my colleagues, juniors and friends from Birla Institute of Technology, Mesra who provided insight and expertise that greatly assisted the research, although they may not agree with all of the interpretations/conclusions of this paper.

REFERENCES

[1] Jesse Leaman and Hung Manh La A Comprehensive Review of Smart Wheelchairs: Past, Present, and Future, IEEE TRANSACTIONS ON HUMAN-MACHINE SYSTEMS, VOL. 47, NO. 4, 2017, pages 486-499

[2] Sateesh Reddy Avutu, Dinesh Bhatia, B. Venkateswara Reddy, "Voice control module for Low Cost Local-Map navigation based Intelligent wheelchair". IEEE 7th International Advance Computing Conference, 2017, pages 609-613

[3] Richard C. Simpson and Simon P. Levine Voice Control of a Powered Wheelchair, IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 10, NO. 2, JUNE 2002, pages-122-125

[4] G. Pires and U. Nunes. A wheelchair steered through voice commands and assisted by a reactive fuzzy-logic controller. Journal of Intelligent and Robotic Systems, 34(3):301–314, 2002.

[5] K. Komiya, Y. Nakajima, M. Hashiba, K. Kagekawa, and K. Kurosu. Test of running a powered-wheelchair by voice commands in a narrow space. JSME Journal (C), 69(688):210–217, 2003. (in Japanese).

[6] L.M. Bergasa, M. Mazo, A. Gardel, R. Barea, and L. Boquete. Commands generation by face movements applied to the guidance of a wheelchair for handicapped people. In ICPR, volume 4, pages 4660–4663, 2000. (a) (b) (c) (d) (e) (f) Fig. 10 Running scenes.

[7] Y. Matsumoto, T. Ino, and T. Ogasawara. Development of intelligent wheelchair system with face and gaze based interface. In IEEE Int. Workshop on Robot and Human Communication, pages 262–267, 2001.

[8] K. Choi, M. Sato, and Y. Koike. Consideration of the embodiment of a new, human-centered interface. IEICE Trans. Inf. & Syst., E89-D(6):1826–1833, 2006.

[9] K. H. Kim, H. K. Kim, J. S. Kim, W. Son, and S. Y. Lee. A biosignal-based human interface controlling a power-wheelchair for people with motor disabilities. ETRI Journal, 28(1):111–114, 2006.

[10] Y. Ichinose, M. Wakumoto, K. Honda, T. Azuma, and J. Satou. Human interface using a wireless tongue-palate contact pressure sensor system and its application to the control of an electric wheelchair. IEICE Trans. Inf. & Syst., J86-D-II(2):364–367, 2003.

[11] R. Barea, L. Boquete, M. Mazo, and E. Lopez. Wheelchair guidance strategies using eog. Journal of Intelligent and Robotic Systems, 34(3):279–299, 2002.

[12] J.S. Ju, Y. Shin, and E. Y. Kim Intelligent wheelchair using head tilt and Mouth shape ELECTRONICS LETTERS 13th August 2009 Vol. 45 No. 17

[13] Jeonghee Kim, Xueliang Huo,, Julia Minocha, Jaimee Holbrook, Anne Laumann, and Maysam Ghovanloo∗ Evaluation of a Smartphone Platform as a Wireless Interface Between Tongue Drive System and Electric-Powered Wheelchairs IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING, VOL. 59, NO. 6, JUNE 2012, pages 1787-1796.

[14] Shraddha Uddhav Khadilkar, Narendra Wagdarikar Android Phone Controlled Voice, Gesture and Touch Screen Operated Smart Wheelchair. International conference on pervasive computing, 2015

 DOI : 10.18231/2454-9150.2018.1308