2017 IEEE 7th International Advance Computing Conference

Voice control module for Low cost Local-Map navigation based Intillegent wheelchair

Sateesh Reddy Avutu
Research Scholar
Biomedical Engineering Dept
North-Eastern Hill University
Shillong, Meghalaya ,India
sateeshreddy.eldt@gmail.com

Dinesh Bhatia
Associate Professor and Head
Biomedical Engineering Dept
North-Eastern Hill University
Shillong, Meghalaya ,India
bhatiadinesh@rediffmail.com

B. Venkateswara Reddy
Research Scholar
Dept of ECE
JNTUH, Hyderabad
Telangana, India
bheemireddyv@gmail.com

***Abstract* — Wheelchair usage in India is increasing rapidly due to higher ageing population, road accidents etc. In case of nuclear families, the elder people those who are unable to do their regular activities without assistance, are left alone at their home. Manual propelling is difficult for aged population for which they employ use of the motorized wheelchair, but their costs is quite high making it difficult for the middle income people to procure it. In order to overcome the above challenge, we proposed a novel wheel chair design "low cost local-map navigation based intelligent wheelchair". In this paper, we implemented design of voice control module for a motorized wheelchair which works based on the speech processing technique and local map navigation system. In speech processing, we employed the concept of mel – frequency cepstrum coefficients (MFCCs) and mean squared error (MSE). The simulation result shows the robustness of the voice recognition module. A microphone is provided with the wheelchair, if the user utters the destination i,e, where he/she wants to go, the wheel chair automatically takes to the destination. Based on the result we ensure that the proposed wheelchair will improve the quality of life of this population and helping them in achieving their daily routine tasks with comfort.**

Keywords— Wheel chair, voice control , MFCCs, mean square error, Local map navigation

I. INTRODUCTION

Wheelchair usage in India is increasing rapidly because of ageing of the people, road accidents, and miss-happenings at home etc. Due to surge in nuclear families globally and particularly in India, the elder people are left alone during the day time and majority of them belong to the lower middle class. According to the Census of India in the year 2011, out of total Indian population approximately 20.3% people are disabled [1] due to various mobility impairments and the figures are higher when compared with the year 2001. The people with mobility impairments need an appropriate wheel chair, but in case of aged persons and children's with quadriplegia, spinal cord injuries (SCI) and amputation, require motorized wheelchair rather than the manual wheelchair. It is very difficult to control and maintain the joystick controlled motorized wheelchair which is available in the Indian market for people with quadriplegia and its cost are quite high.

In our project, the development of low-cost local-map navigation based intelligent wheel chair capable of doing movement according to the user desire and it can clearly improve the quality of life for elders and disabled persons. The low cost Local-Map navigation based intelligent wheelchair with voice controlled module discussed in the paper has been designed in order to meet the following requirements.

- ✓ It can be designed for people with different mobility impairments, lifestyle disorders, life roles and backgrounds.
- ✓ It should be customized based on the user requirements and maintained and repaired easily at low costs by the user or local mechanic.
- ✓ It should be operated always in the safe mode and durable for long period.

The rest of this paper is organized as follows: Section II explains an overview of existing work and an overview of the speech processing techniques employing MFCCs and mean squared error (MSE). The detailed description of the proposed methodology is presented in Section III. Simulation results of the proposed work are presented in Section IV, the overall conclusion and future work is given in Section V.

II. RELATED WORK

A. *Overview of the existing work*

Speech recognition is used in [2, 3]. In [2] the wheel chair has two possible driving modes named as 1) follow a wall and 2) get into an elevator. The user switches between modes using the voice control. The execution time is 6sec; it is slightly higher when compared with our proposed method. The method presented in [3] employs speech recognition to control the movement of the wheelchair in different directions (back, forward, left, right); here they used HM 2007 commercially available IC, which is unable to process the speech with the length more than 0.96 sec.

Google Speech Recognition service and Microsoft SAPI are used in [4] combination with a 2D face tracking to control a mobile robot. The speech recognition is used to

978-1-5090-1560-3/17 $31.00 © 2017 IEEE
DOI 10.1109/IACC.2017.120

switch between the different modes like forward and backward. In this case the user needs internet facility at his movable environment. Voice recognition kit is used in [5]; which employ HM 2007 commercially available IC, in which it is not possible to get the angular movement. In real time speaker recognition system using MFCC and vector quantization technique [6] used LBG algorithm along with MFCCs for pattern matching which will take more time for execution.

B. Overview of mel – frequency cepstrum coefficients (MFCCs) and meansquare error

i) Mel – frequency cepstrum coefficients (MFCCs)

The short-term power spectrum of a sound, based on a linear cosine transform of a log power spectrum on a nonlinear mel – scale of frequency is called as mel – frequency cepstrum (MFC) and its coefficients is called as MFCCs and these are used to extract the features of audio signal[7]. In MFC, the frequency bands are equally spaced on the mel scale which approximates the human auditory system's response more closely then the linearly-spaced frequency bands used in the normal cepstrum. This frequency warping can allow for better representation of sound [7]. The block diagram representation for MFCCs calculation is as shown in Fig1.

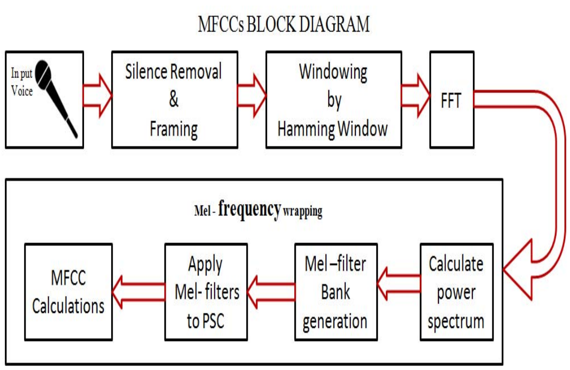


Figure – 1 (MFCCs BLOCK DIAGRAM)

ii) Mean squared error

The mean squared error (MSE) of an estimator measures the average of the squares of the errors. In this paper, the MSE is taken between the MFCCs of utter word by the user i,e where user what to move and MFCCs of the database. Minimum value of the mean squared error indicates the corresponding signal to the uttered word. The mathematical formula to calculate the mean is as shown in equation 1.

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \text{------}(1)$$

III. PRAPOSED METHODALOGY

Voice command for uttered word was recorded from 5 male and 5 female subjects mean age for male 55±2, for Female 50±2, Mean height for male 158±5, for female 152±3 and mean weight for male 68±6, for female 54±5 in the department laboratory. All subjects provided written informed consent to record their voice signals and participate in the study. Each command word is recorded 10 times with different pitch at varied time intervals and stored as separate files for each subject and repetitive task. This was repeated for all the command words. In this paper, we used 7 command words namely as Room 1, Room 2, Hall, Kitchen, Lawn, Dining and Bathroom. The command words and corresponding wheel chair action is as shown in Table – 1.

.

VOICE COMMAND AND WHEELCHAIR ACTION	
Room 1	Wheelchair will takes the user to bed-room 1
Room 2	Wheelchair will takes the user to bed-room 2
Hall	Wheelchair will takes the user to hall
Kitchen	Wheelchair will takes the user to kitchen
lawn	Wheelchair will takes the user to lawn
Dining	Wheelchair will takes the user to dining
Bathroom	Wheelchair will takes the user to bathroom
Stop	Wheelchair will stop instantaneously

Table – 1

Our voice recognition module comprises of two steps: 1) Speaker identification 2) Speech processing to perform desired task. Fig – 2 shows the flow chart of the proposed algorithm. The command voice was collected by using the microphone, subsequently it was preprocessed to remove noise and silence gaps. It was done by making the signal into frames. The nature of the waveforms is as shown in Fig – 3.

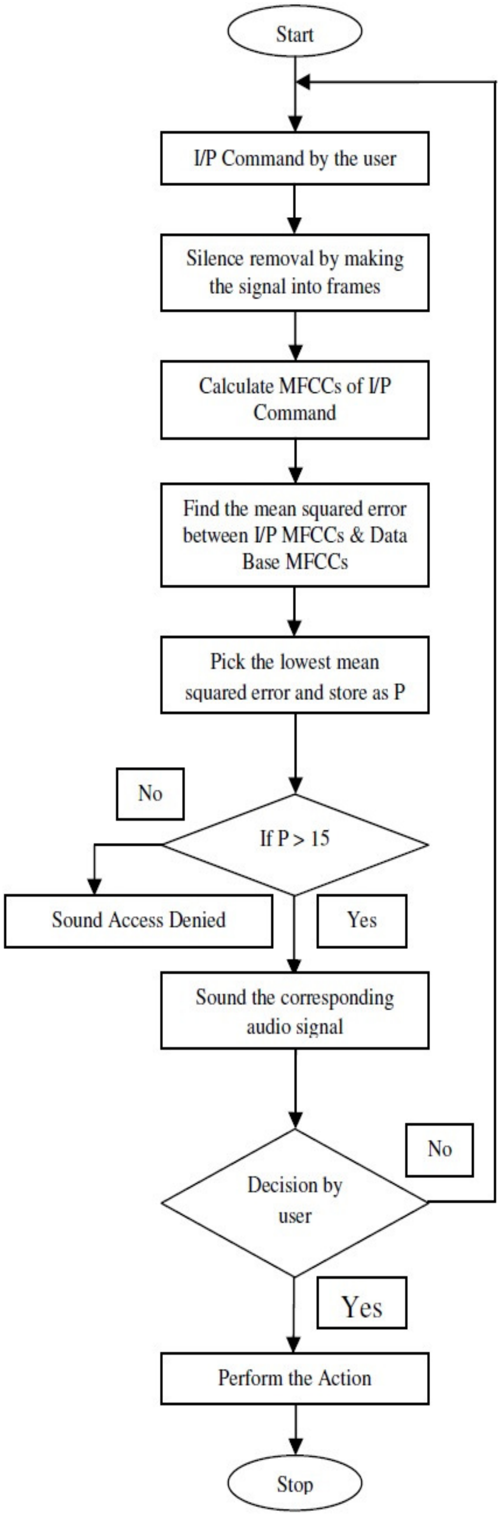
In order to find the MFCCs, the pre-processed signal is segmented into frames and windowing is done by using Hamming window. The mathematical representation of hamming window is shown below.

$$S_k = \sum_{n=0}^{N-1} s_n e^{-j2\pi kn/N}, \quad k = 0, 1, 2, \ldots, N-1 \quad \text{----------------} (2)$$

The fast Fourier transform (FFT) was used to extract the feature vectors from the frequency spectra of the windowed signal frames. The power spectrum of each windowed signal was calculated by using N point (FFT), in this paper we took N=256. The power spectrum is applied for Mel – frequency wrapping.

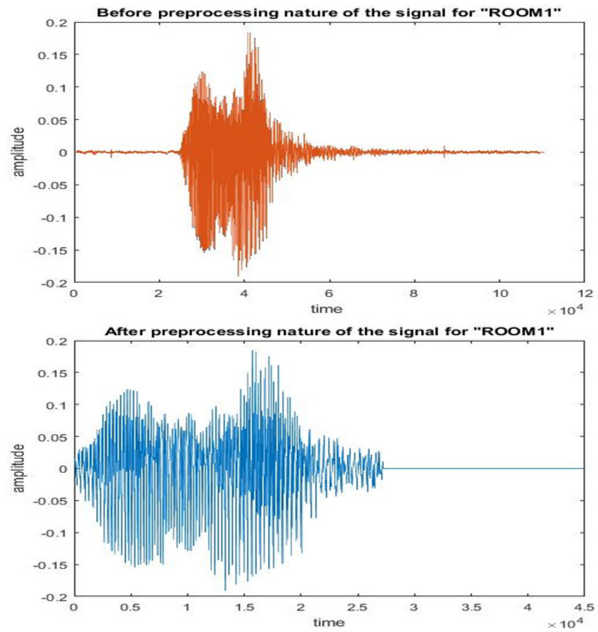
The Mel frequency filter bank is a series of triangular band pass filter, these are connected as the lower boundary of one filter which starts at the center frequency of the previous filter and the upper boundary ends at the center frequency of the next filter. The filter bank which is based on a non-linear frequency scale is called as Mel-scale. The Mel Scale relates perceived frequency of a pure tone to its actual measured frequency[7]. The mathematical relationship between the Mel-scale and normal frequency is shown below in equation number 3

$$M(f) = 1125 \ln(1 + f/700) \text{ --------------} (3)$$



Flow chart of the proposed algorithm

Figure – 2



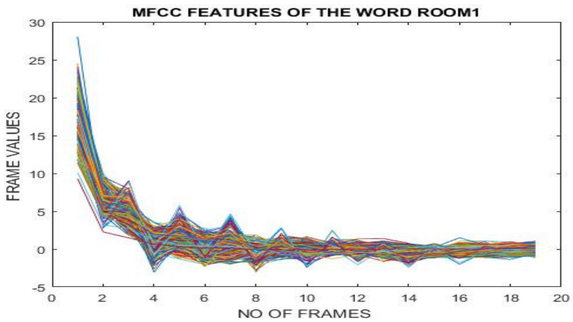
Nature of the wave form after the silence removed

Figure – 3

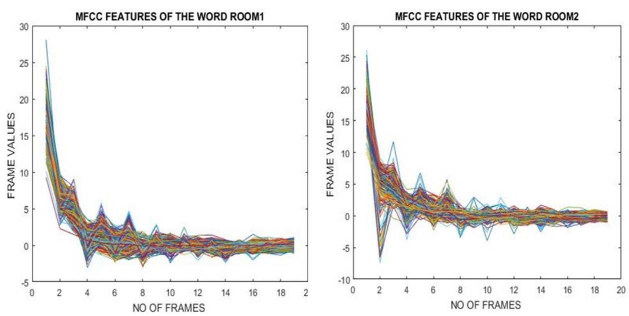
The log value of Mel frequency spectrum was found. The discrete cosine transform was done for logarithmic Mel frequency spectrum than by finding the Inverse FFT of the resultant signal we get the MFCCs of the corresponding input signal and the plots obtained are shown in Fig - 4.The mathematical equation for Discrete Cosine Transform is represented by equation number 4.

$$c_n = \sum_{k=1}^{k} (\log S_k) \cos\left[n\left(k - \frac{1}{2}\right)\frac{\pi}{K}\right], \quad n = 0, 1, \ldots, K-1 \qquad \text{-----------} \quad (4)$$

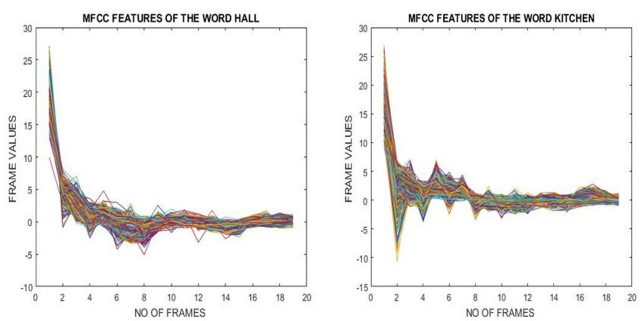
The MFCCs of all database signals are calculated then the minimum squared error is calculated between the MFCCs of the input command word and MFCCs of individual signals of the database. This value is stored in a variable say “P”, if P>15 then the system will sound the output as Access Denied else it will sound the corresponding voice command. After listening the output, the user will take the decision, if he say “yes” then the corresponding local map will get activated to maneuver the wheelchair as per input command, else it will stop.



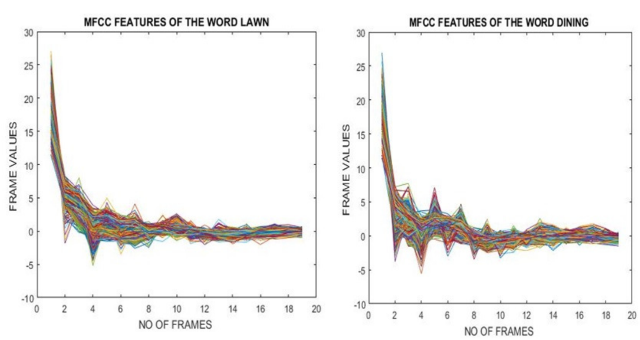
(a)



(b)



(c)



(d)

Figure – 4: a) MFCCs features of input command word b) MFCCs features of Room 1 and Room 2 c) MFCCs features of Hall and Kitchen d) MFCCs features of lawn and dining.

IV. SIMULATION RESULTS

The experiment was carried out by using male and female subjects as mentioned earlier. The voice is collected in the open environment in the department laboratory of the University. Same microphone was used to give input commands and for controlling the wheelchair through these command words. In this paper, we used seven command words which were repeated ten times with different pitch and stored as a database in the system. The database is maintained for particular user. The output of the system is shown in the Table – 2 when the wheelchair was used by the authorized person and if the wheelchair was used by the unauthorized person that output of the system is shown in Table – 3. From the tables it can be seen that……

Table – 2: Wheelchair employed by authorized person

Voice command by the user	Values of mean Squared errors when compared with the data base						
	commands						
	Room 1	Room 2	Hall	Kitchen	Lawn	Dining	Bath room
Room 1	**08.94**	14.09	14.79	20.86	12.08	14.29	15.58
Room 2	17.54	**09.64**	17.33	20.37	14.48	12.16	19.25
Hall	14.62	17.86	**10.63**	23.45	14.58	15.37	18.65
Kitchen	33.60	26.61	29.76	**10.29**	24.93	25.92	23.35
Lawn	12.77	17.00	13.66	22.58	**11.85**	14.15	20.56
Dining	12.58	13.29	13.16	17.97	12.72	**09.12**	15.23
Bath room	20.25	16.22	18.53	14.07	17.29	19.54	**11.57**

Table – 3: Wheelchair employed by unauthorized person

Voice command by the Unknown user	Values of mean Squared errors when compared with the data base						
	commands						
	Room 1	Room 2	Hall	Kitchen	Lawn	Dining	Bath room
Room 1	24.40	23.79	22.86	21.78	24.30	24.10	29.65
Room 2	16.01	16.65	17.26	22.67	18.50	17.96	22.36
Hall	18.84	24.59	20.72	19.57	22.35	27.65	26.95
Kitchen	27.00	27.60	28.85	21.63	26.73	25.78	28.34
Lawn	24.77	28.29	20.41	24.48	19.79	22.35	26.98
Dining	25.00	27.54	23.25	25.51	25.31	26.11	29.65
Bath room	26.84	28.88	22.98	23.53	26.29	27.55	26.38

V. CONCLUSIONS AND FUTURE WORK

1) Conclusion

Form Table – 2, we see that the diagonal elements has the minimum MSE in their respective rows, by this we can conclude that if an authorized person uses the wheelchair they would reach their destination safely.

From Table – 3, we see that no value is less than 15, which shows that if it is used by unauthorized person the system would represent “Access denied” and wouldn’t respond to the user commands.

2) Future work

In this paper we presented the simulation results by employing English language to pronounce the command words. Implementation of this algorithm using hardware set-up and other languages with emphasis on local dialects would make it more customer friendly and easy to use.

REFERENCES

[1] www.languageinindia.com/jan2014/disabilityinindia2011data.pdf

[2] Akira Murai; Masaharu Mizuguchi; Takeshi Saitoh; Tomoyuki Osaki; Ryosuke Konishi “Elevator available voice activated wheelchair” RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication Year: 2009 Pages: 730 – 735.

[3] Mohammed Faeik Ruzaij; S. Poonguzhali “Design and implementation of low cost intelligent wheelchair” Recent Trends In Information Technology (ICRTIT), 2012 International Conference on Year: 2012 Pages: 468 – 471.

[4] R. Berjón; M. Mateos; A. L. Barriuso; I. Muriel; G. Villarrubia “Alternative Human-Machine Interface system for powered wheelchairs “Serious Games and Applications for Health (SeGAH), 2011 IEEE 1st International Conference on Year: 2011 Pages: 1 – 5.

[5] M. Senthil Sivakumar; Jaykishan Murji; Lightness D Jacob; Frank Nyange; M. Banupriya “Speech controlled automatic wheelchair” Information Science, Computing and Telecommunications (PACT), 2013 Pan African International Conference on Year: 2013 Pages: 70 – 73.

[6] J. Kathirvelan; R. Anilkumar; Zachariah C. Alex; A. Fazul “Development of low cost automatic wheelchair controlled by oral commands using standalone controlling system Computational Intelligence & Computing Research (ICCIC), 2012 IEEE International Conference on Year: 2012 Pages: 1 – 4.

[7] Roma Bharti; Priyanka Bansal “Real Time Speaker Recognition System using MFCC and Vector Quantization Technique” International Journal of Computer Applications (0975 – 8887) Volume 117 – No. 1, May 2015.