International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-9 Issue-1, May 2020

Voice Controlled Vehicle Dashboard

Shridhar D. Pagar, Shivani J. Pote, Ankush S. Anmulwar, Ashwini S. Shinde

Abstract: ***Driving a vehicle or a car has become tedious job nowadays due to heavy traffic so focus on driving is utmost important. This makes a scope for automation in Automobiles in minimizing human intervention in controlling the dashboard functions such as Headlamps, Indicators, Power window, Wiper System, and to make it possible this is a small effort from this paper to make driving distraction free using Voice controlled dashboard. and system proposed in this paper works on speech commands from the user (Driver or Passenger). As Speech Recognition system acts Human machine Interface (HMI) in this system hence this system makes use of Speaker recognition and Speech recognition for recognizing the command and recognize whether the command is coming from authenticated user(Driver or Passenger). System performs Feature Extraction and extracts speech features such Mel Frequency Cepstral Coefficients(MFCC),Power Spectral Density(PSD),Pitch, Spectrogram. Then further for Feature matching system uses Vector Quantization Linde Buzo Gray(VQLBG) algorithm. This algorithm makes use of Euclidean distance for calculating the distance between test feature and codebook feature. Then based on speech command recognized controller (Raspberry Pi-3b) activates the device driver for motor, Solenoid valve depending on function. This system is mainly aimed to work in low noise environment as most speech recognition systems suffer when noise is introduced. When it comes to speech recognition acoustics of the room matters a lot as recognition rate differs depending on acoustics. when several testing and simulation trials were taken for testing, system has speech recognition rate of 76.13%. This system encourages Automation of vehicle dashboard and hence making driving Distraction Free.***

Keywords: ***Linde Buzo Gray (LBG), Mel Frequency Cepstral Coefficients (MFCC), Vector Quantization(VQ), Speaker Recognition, Speech Recognition.***

I. INTRODUCTION

As, Automation is in demand nowadays whether it will be home automation or factory automation using cutting edge technologies and products and systems are being developed which will have minimal or no human intervention in its operation. And also Vehicles nowadays are becoming technologically advanced so the dashboard is getting more and more complex with so many controls placed on dashboard. Vehicle Dashboard provides control of many functions through conventional Lever based dashboard, So it becomes difficult for the driver to drive and simultaneously control various dashboard functions by looking for the respective lever or switch while driving and increasing traffic is imposing new challenges ahead of us, Hence in order to increase safety and comfort levels in cars and taking into account the high number of new systems that could distract the attention from the driving task, Speech Recognition is becoming an essential field in the Vehicle environment.

Speech signal is made up of sequence of sounds of different frequencies coming together to form a long stream of samples when observed in broad spectrum[15]. Human speech ranges from 65- 600HZ and adult male speech frequency is around 130HZ and 210 HZ is adult female speech frequency [14]. The proposed system makes use of speech recognition for identification of verbal commands from the user and through running recognition algorithm it will identify the command spoken by the user and will execute the said operation. Speaker recognition is implemented along with Speech recognition for Authentication of user. Speech recognition makes use of Vector Quantization technique.

Fig.1.Shift from lever based dashboard to voice Dashboard

By implementing Voice based dashboard we not only ensure distraction free driving but also increases comfort and ensures vehicle safety as well as pedestrian safety.

II. SPEECH RECOGNITION

Speech recognition can be said as a case of pattern recognition. Speech recognition is process of conversion of speech to text[7]. It consists of Feature extraction from the speech signal and features extracted are Pitch, Treble. Spectrogram, Power Spectral Density. The training phase comprises estimation of parameters of the classification model using a large training dataset. In testing phase, the features of a test pattern data is matched with the trained model of each and every class data. The test pattern belongs to the class whose model matches the test pattern data perfectly. A sentence comprises of linguistic units such as words, syllables[15]. Acoustic evidences provided by the acoustic modelling is combined to form a valid and meaningful sentence. Hence, when it comes to speech recognition, the pattern matching stage can be viewed as taking place in two domains: acoustic and symbolic.

Manuscript received on April 02, 2020.
Revised Manuscript received on April 15, 2020.
Manuscript published on May 30, 2020.

* Correspondence Author
Shridhar D. Pagar*, Student ,Bachelor of Engineering, Electronics and Telecommunication, Pimpri Chinchwad College of Engineering Nigdi,
Shivani J. Pote, Student ,Bachelor of Engineering, Electronics and Telecommunication, Pimpri Chinchwad College of Engineering Nigdi
Ankush S. Anmulwar , Student ,Bachelor of Engineering, Electronics and Telecommunication, Pimpri Chinchwad College of Engineering Nigdi
Mrs. Ashwini S. Shinde, Assistant Professor, Department of Electronics & Telecommunication, Pimpri Chinchwad College of Engineering Nigdi

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

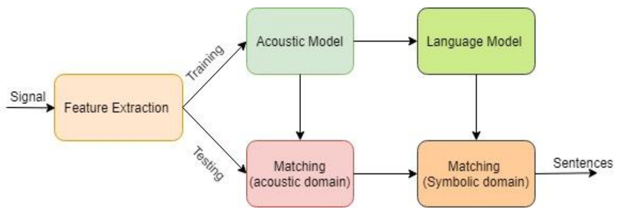


Fig.2.Typical block diagram of speech recognition.

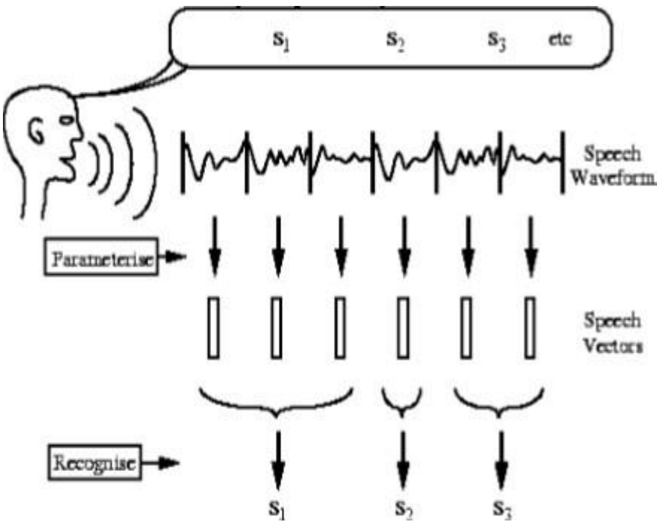


Fig. 1. Speech identifiacation from speech fragments(vectors).

The speech waveform when observed in broad spectrum and high and low peaks are clearly visible then using this we can obtain cepstral (speech) vectors through this vector information and analyzing them carefully we can differentiate the spoken words and identify them to form a meaningful sentence.

III. FEATURE EXTRACTION

A. Mel Frequency Cepstral Coefficients

Majority of Automatic Speech Recognition systems perform Mel- scale filter bank analysis and estimate cepstral (speech) coefficients called Mel Scale Filter Coefficients (MFCC). Mel filters are approximated as overlapping triangular filters having bandwidth of 1 bark. MFCC are based on variation of the human ear's critical bandwidths with frequency. Filters are spaced at low frequencies and logarithmically they are at high frequencies and they are used to capture important characteristics of speech[16]. The relation between frequency in Hz and frequency in Mel scale can be illustrated as:

$$m = 1125LN\left(1 + \frac{f}{700}\right) \qquad -(1)$$

$$f = 700\left(e^{\frac{m}{1125}-1}\right) \qquad -(2)$$

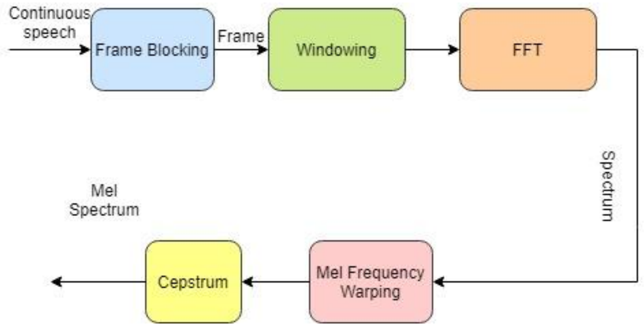


Fig.4.MFCC calculation flow

To calculate MFCCs, a schematic of this process is given in Figure 4

1. Speech signal is divided into frames of 25ms and overlap of 10ms. Each frame is multiplied with Hamming window.
2. Periodogram of each Speech frame is calculated by performing FFT of 512 samples on each frame , then power spectrum is:

$$P(K) = \frac{1}{N}|S(K)^2| \qquad -(3)$$

3. Entire frequency range is divided into n Mel filter banks.
4. To estimate filter bank energy multiply each and every filter bank with power spectrum and add the coefficients.
5. Then logarithm of these 'n' energies and compute Discrete Cosine Transform to get the final MFCC.
   Fig.8 to 11 shows step wise results obtained during mfcc extraction, and Fig.11 shows Mel filterbank.

B. Power Spectral Density (PSD)

The signal might be an electromagnetic wave or an acoustic wave. The power spectral density (PSD) of a signal is the power present in the signal as a function of frequency. PSD of the speech signal is performed on MATLAB 2016 Software. Simulation for power spectral density so as to analyze the power of a signal at different frequencies. which will be helpful in working with voice commands during training. This simulation is implemented by using the Welch function of periodogram plotting which is an extended version of the standard periodogram method. The speech signal simulated to obtain PSD contains "L-indicator" speech command recorded in my voice in .WAV file format. The Fig.14 shows the power spectral density of speech signal.

C. Pitch

Pitch is the fundamental frequency of vocal tract[16]. It represents the vibration frequency of the vocal cords during the productions of sound. So it helps in analyzing the speech signal and we have implemented the pitch-plot using Auto correlation method of plotting the pitch. The Fig.15 shows the pitch plot of speech signal. Pitch of the speech signal is simulated on MATLAB 2016 Software

D. Spectrogram

Spectrogram is Pictorial representation of the Frequency Spectrum of a signal as it varies with respect to time. The audio signal spectrograms are called sonographs. This simulation result of Spectrogram is obtained by using Praat audio software. The Fig.16shows the Spectrogram of speech signal.

Retrieval Number: A2148059120/2020©BEIESP
DOI:10.35940/ijrte.A2148.059120
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

IV. FEATURE MATCHING

Most popular algorithms for Feature Matching in speaker recognition and speech recognition are Dynamic Time Warping (DTW), Hidden Markov Model (HMM) and Vector Quantization (VQ). Here, we have used Vector Quantization. VQ is the process of vector mapping from a large vector space to a finite region in that space. Each region forms a cluster and represented by its center called a codeword. The collection of all codewords forms a codebook.
Vector Quantization technique is beneficial in data reduction. It generates optimal codebook. Fig.13 shows the plot of VQ-Codebook. In [13] they have used Vector Quantization method for speech recognition and to remove noise end point detection algorithm is used and for testing Euclidean distance is checked for each test feature and codebook.

The testing is done by evaluating all codebooks and system picks the word whose codebook is located at least distance. Vector Quantization is the process quantization of data in contiguous blocks known as vectors. Quantization is the process in which infinite vector are mapped into finite ones. Quantization has important role in signal processing. VQ has gained importance due to introduction of LBG algorithm. The performance depends on efficient generation of codebooks and are generated using training set of speech signals. Computational complexity and memory required is decreased while decreasing the number of bits for generating codebook. Codebooks have been generated in VQ using LBG algorithm which is an iterative algorithm in itself. In VQ, finite regions are generated in that particular space.

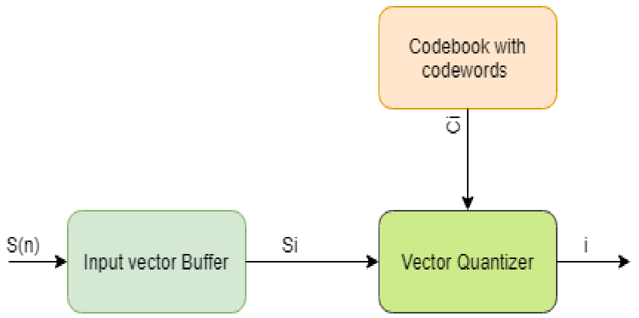


Fig. 5. Vector Quantizer Block

Linde-Buzo Gray(LBG)
The LBG algorithm is used for clustering a set of training vectors into a set of codebook vectors. Feature training involves randomly selecting feature vectors of the recorded speech samples and performs training for the codebook using the LBG vector quantization (VQ) algorithm[13].The algorithm is implemented using following procedure:

1. Design a 1-vector codebook, this is the centroid of entire training set vectors.
2. Double the size of the codebook by splitting codebook according to the rule

$$Yn^{+} = Yn(1 + \varepsilon) \quad -(4)$$

$$Yn^{-} = Yn(1 + \varepsilon) \quad -(5)$$

3. Nearest-Neighbor Search: for each training vector, find the codeword in the current codebook that is closest and assign that vector to the corresponding cell (near to the closest codeword).
4. Centroid Update: update the codeword using the centroid of the training vectors.
5. In Iteration 1, Repeat steps 3 & 4 until vector distortion for current iteration falls below the previous iterations distortion.
6. In Iteration 2, Repeat steps 2, 3 and 4 until a codebook size of K is formed.

V. PROPOSED SYSTEM METHODOLOGY

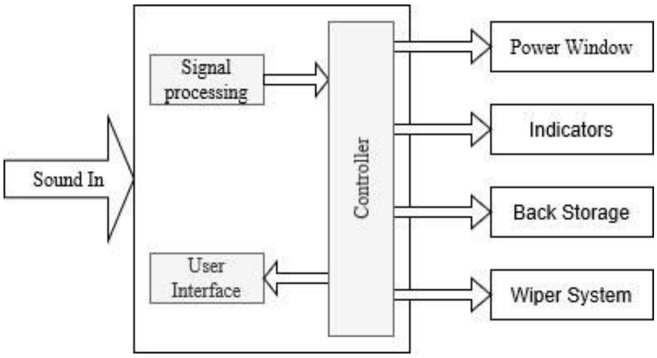


Fig .6. Proposed System block diagram.

As seen from Fig.6 the proposed system uses Raspberry Pi 3b controller and User interface is designed to display the information to user about the currently executing command. The signal processing of speech signal received from the user is carried out initially and speech command is recognized and then it is given to the controller for performing the particular operation. The proposed system has Power window control, Indicators control, Back storage opening control, Wiper system.

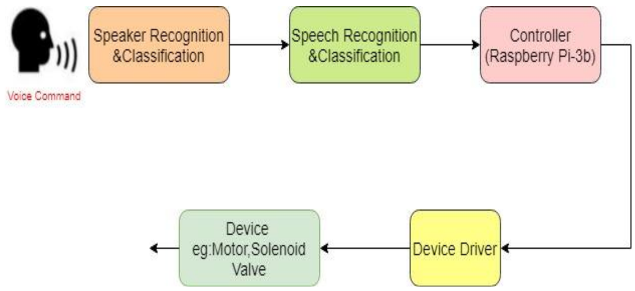


Fig.7. Process flow

The approach revolves in formation of simple block diagram structure as seen in Fig.7 with three different modules involves input speech signal and its feature extraction and recognizes in using classification techniques. The Speaker Recognition system is employed to recognize and authenticate whether the voice is of authenticated user or not. Table.1 shows the speech commands which are used in proposed system. The speech commands from 1 to 9 are speaker dependent commands for the execution of this particular commands the speech should come from the authenticated user or we can say driver of the vehicle because this commands are directly related to the driver and are important functions of vehicle with respect to safety. Hence in respect of this particular 9 commands first of all speaker recognition is done and if it authenticates the user then only this commands are executed. Rest all commands are accessible through any user inside the vehicle whether he\she might be driver or Passenger. First 9 speech commands are completely driver related functions and have no relation with passengers hence analyzing the safety and convenience aspect speaker recognition is used to provide access for operating these commands.

The speech commands which have been recognized are sent to the controller (Raspberry Pi-3b) and further controller sends signal to device driver of particular device with respect to operation for reference consider Wiper system the motor is employed hence the motor driver IC (L293D) will be activated and further Turning ON Motor.

Table 2. Speech Commands

Sr.No	Speech Commands	Operation Performed
1.	**R-INDICATOR**	Right Indicator Turns ON.
2.	**L-INDICATOR**	LEFT Indicator Turns ON.
3.	**HL UP**	Head lights are at Upper position .
4.	**HL DIP**	Head Lights are at Dipper position .
5.	**HL OFF**	Head Lights become OFF.
6.	**WIPE S1**	Wiper System Turns ON and works on First speed Level.
7.	**WIPE S2**	Wiper System is working on second speed level.
8.	**WIPE OFF**	Wiper System Turns OFF.
9.	**SPRAY ON**	Water sprayer starts spraying for specified time.
10.	**BS OPEN**	Back Storage Lock gets OPEN.
11.	**WIND-UP**	To keep Window Fully Closed.
12.	**WIND-DOWN**	To keep Window Fully Open.
13.	**WIND-STOP**	To keep window open\close as per convenient height.

VI. SIMULATIONS / RESULT

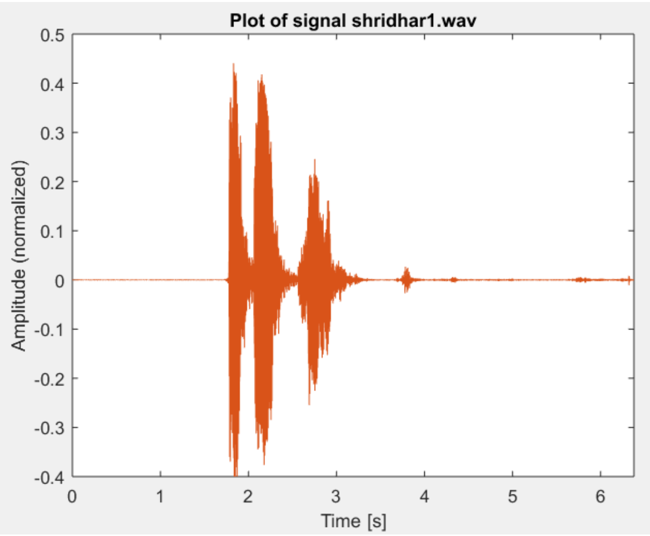


Fig.8. Plot of Signal

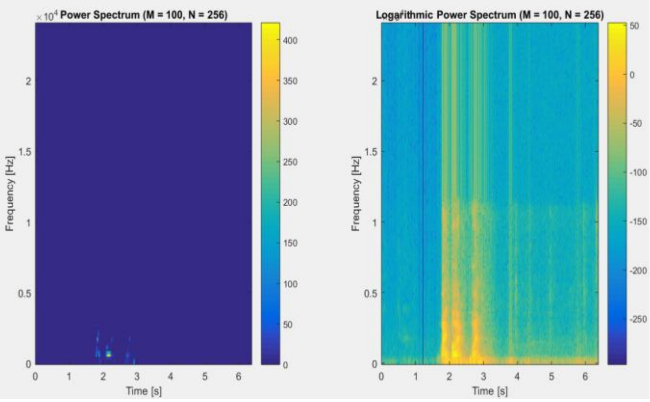


Fig.9. Linear & Logarithmic Plot

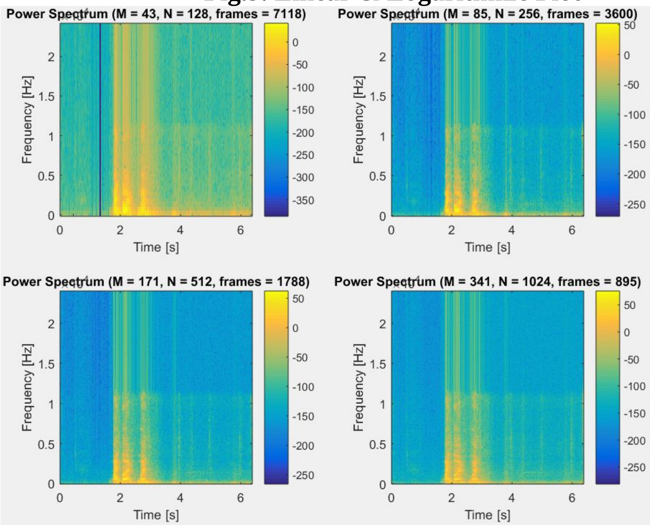


Fig.10. Power Spectrum For Different N values

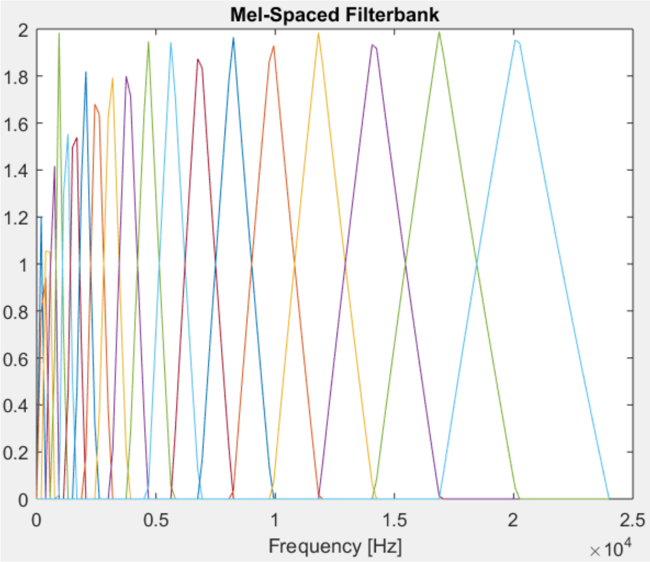


Fig.11.Mel Filter-bank.

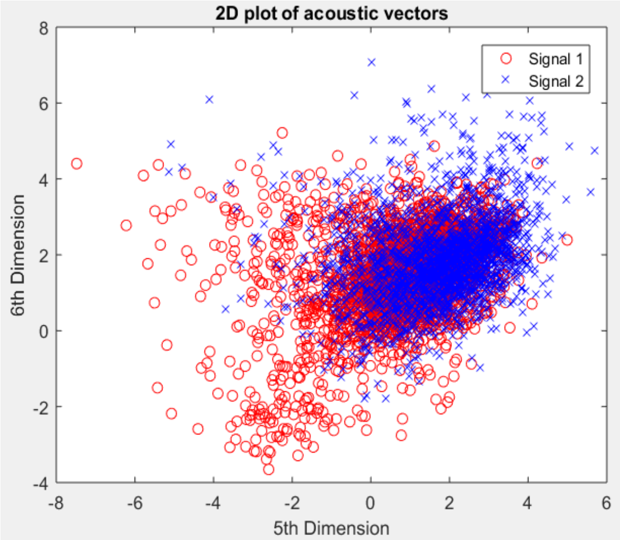


Fig.12. Vector Quantization(VQ)- Codebook

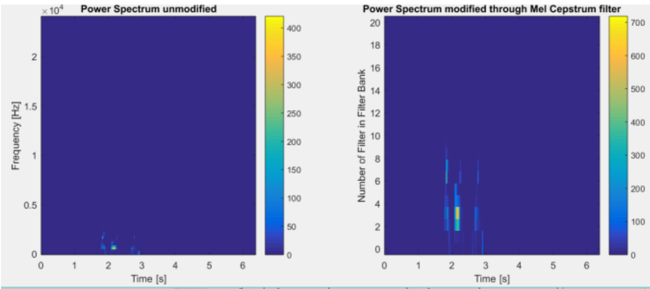


Fig.13.Modified power spectrum

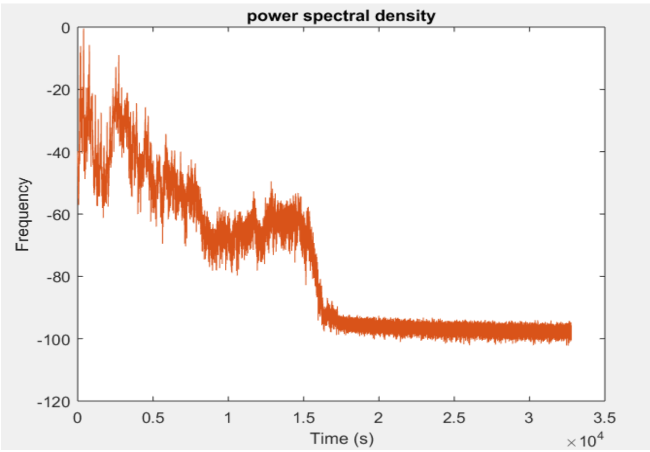


Fig.14.Power spectral density

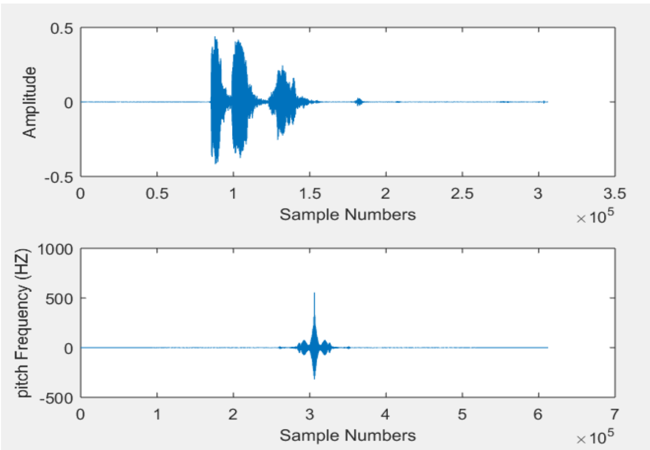


Fig.15.Pitch

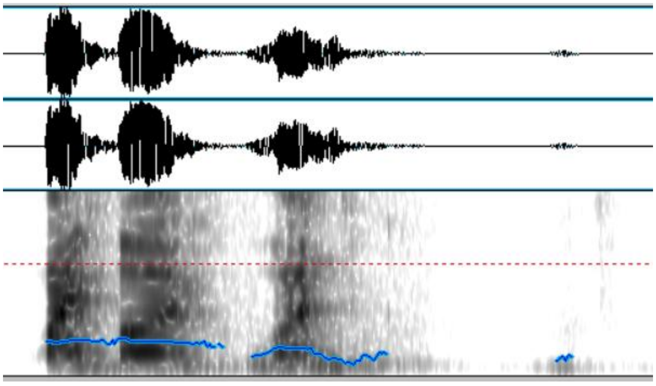


Fig.16.Spectrogram

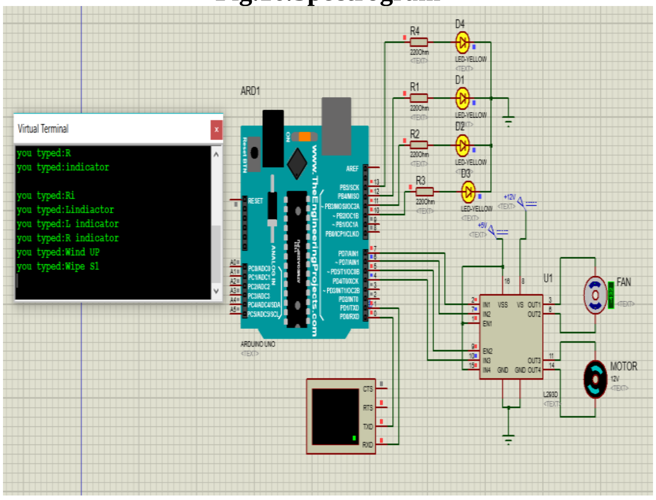


Fig.17.Hardware simulation

In Fig.17 hardware simulation is performed using Proteus Software 2018 In this simulation for testing purpose we have used Arduino Uno instead of using Raspberry Pi because Proteus was not having inbuilt package of Raspberry pi.

Table2.Difference in spoken and detected commands

Sr. No	Spoken Speech	Detected Speech	Accuracy
1	R-INDICATOR	R-INDICATOR	78.5%
2	WIPE S2	VIBE S2	61.9%
3	HL UP	HL UP	73.7%
4	HL OFF	HL OFF	74.8%
5	BS OPEN	BS OPEN	86.7%
6	WIND-UP	WIND-UP	81.2%

VII. CONCLUSION

A Voice based Dashboard system is developed to be operated on speech commands, which allows user (driver of the vehicle) to solely focus on driving hence increasing on road safety of vehicle as well as passengers. This system not only ensures safety but enhances the user experience. when several testing and simulation trials were taken for testing system system has speech recognition rate of 76.13%. When tested for 6 speech commands.

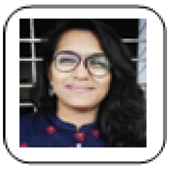
As most speech recognition suffer from addition of noise into the signal either while recording or real time then it results into decrease in recognition rate. +-Hence we have tried our best to minimize the effect of noise on to the system. The system while testing performed differently when acoustics of the room were changed from previous. But as this system designed for automotive application hence due to closed environment inside car resulted in better recognition rate.

REFERENCES

1. M. F. M. Idros, A. H. A. Razak, S A M Al Junid, A. K. Halim, N Khairudin,“Capability of Voice Recognition System for Automatic Signal in Autonomous Vehicle (AV) Application”,IEEE 5th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA 2018), November 2018.
2. Y. Zhao and L. Zhu, "Speaker-Dependent Isolated-Word Speech Recognition System Based on Vector Quantization," 2017 International Conference on Computer Network, Electronic and Automation (ICCNEA), Xi'an, 2017, pp. 133-137.
3. Jianliang Meng, Junwei Zhang, Haoquan Zhao “Overview of the Speech Recognition Technology’’Ninth International Conference on Computational and Information Sciences,pp.199-2012017.
4. Chee Yang Lon, Kai Lung Boey Kai Sze Hong “Speech Recognition Interactive System for Vehicle ’’ IEEE 13th International Colloquium on Signal Processing & its Applications, 10 - 12 March 2017
5. Shashanki Singh, Sumedha Tode, Rekha Takalkar,“MultitaskingSmart Control using Voice Command”, International Journal of Engineering Science and Computing IJESC,2017
6. V.Partha.Saradi, P.Kailasapathy, “Voice Based motion of robotic vehicle through visible light communication”, computer and electrical Engineering Science direct (Elsevier),pp.156-159,March 2019.
7. Samudravijaya K, “Automatic Speech Recognition”, Tata Institute of Fundamental Research.
8. Ashok Kumar, Vikas Mittal, “Speech Recognition :A Complete Perspective”, International Journal of Recent Technology and Engineering (IJRTE),April 2019.
9. Ande.StanleyKumar,Dr.K.MallikarjunaRao,Dr.A.BalaKrishna,“Speech Recognition System for Controlling the Robot”, International Journal of Engineering Research &Technology (IJERT),Septemb+-er 2012.
10. Orchisama Das ,“Speaker Recognition”. Center for Computer Research in Music and Acoustic, Stanford University
11. Thomas Mohan, Amrutha K, Anjana Anilkumar, Helen Johnson, Silsha K, “VOICE OPERATED INTELLIGENT LIFT”, International Research Journal of Engineering and Technology (IRJET),June 2018.
12. C. Jeeva, Anwar Naseer Khan, Junaid Azad Wani, Amit Kumar, “Voice Based Vehicle Parameter Control”, International Journal of Advanced Research in Computer Science and Software Engineering, pp.687-688 May 2016
13. H. B. Kekre, A. A. Athawale, and G. J. Sharma,“Speech recognition using vector Quantization”,In Proceedings of the International Conference & Workshop on Emerging Trends in Technology (ICWET ’11). Association for Computing Machinery, New York, NY, USA,pp.400-402,2011
14. Lu, Tzu-Chuen and Ching-Yun Chang. “A Survey of VQ Codebook Generation.” (2010).
15. Matthias Wolfel, John McDonough, “DISTANT SPEECH RECOGNITION”, A John Wiley and Sons, Ltd. Publication, ,pp.34-215, 2009
16. Homayoon Beigi, “Fundamentals of Speech Recognition”, Springer,pp.87-175, 2011.
17. Wikipedia Contributors(2020).Voice_ frequency.[online]Wikipedia. Availableat: https://en.m.wikipedia.org/org/wiki/Voice_frequency

AUTHORS PROFILE

Shridhar D. Pagar is currently in Final Year of Bachelor of Engineering Degree in Electronics and Telecommunication from Pimpri Chinchwad College of Engineering Nigdi, affiliated to Savitribai Phule Pune University(SPPU), Pune. He has been intern with Research & Development (Electrical & Electronics) Department of Force Motors, Pune. E-mail : shridharpagar@gmail.com



Shivani J. Pote, is currently in Final year of Bachelor of Engineering Degree in Electronics and Telecommunication from Pimpri Chinchwad College of Engineering Nigdi, affiliated to Savitribai Phule Pune University(SPPU), Pune. E-mail : shivanipote532@gmail.com

Ankush S. Anmulwar is currently in Final Year of Bachelor of Engineering Degree in Electronics and Telecommunication from Pimpri Chinchwad College of Engineering, Nigdi, affiliated to Savitribai Phule Pune University, Pune. E-mail : ankush.s.anmulwar@gmail.com

Mrs. Ashwini S. Shinde, has completed her Bachelor's degree (B.E) in Electronics & Communication and Masters in Engineering (M.E) with specialization in Electronics from Dr. Babasaheb Ambedkar Marathwada University, Aurangabad. She is pursuing her Doctorate degree from Savitribai Phule Pune university. Her research work speech processing, speech emotion recognition. She has more than 11 years of teaching experience and currently she is working as Assistant Professor at Pimpri Chinchwad College of Engineering affiliated to Savitribai Phule Pune University, Pune. She has more than 9 papers in international journals and conferences to her credit. E-mail : ashwini.shinde@pccoepune.org