ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 14, March 2016

VOICE RECOGNITION USING K-NN CLASSIFIER

Dr. B. Vijayakumari [1], A.Johnson[2], S.Pitchai Arasu[2], K. Sadeesh Kumar[2], C.Vairamuthu[2]
[1] Associate professor, [2] UG students,
Department of Electronics and Communication Engineering,
Kamaraj College of Engineering and Technology, Virudhunagar-626001.
E-mail ID: vijayakumariece@kamarajengg.edu.in,johnsonamalraj@outlook.com, sparasucse@gmail.com, sadeeshkumark1994@gmail.com,cvairamuthu1995@gmail.com

ABSTRACT

Home automation and home monitoring is very essential now a days. By using signal processing techniques, the home will be automated and monitored. This project includes both audio processing and embedded system. By the use of signal processing, voice recognition is performed . The home appliances will be controlled by the embedded system. Using the basic feature extraction algorithm, the speech signal features are extracted and are stored. The voice signals are compared with the already available speech signals in the database by using KNN (k-nearest neighbor) algorithm. Finally, the results obtained shows that this approach will be helpful to the society for automatic home appliances control.

Keywords:Automatic Speech Recognition, Feature extraction, K-nearest neighbour method

1.Introduction

Speech is one of the ancient ways to express ourselves. Today these speech signals are also used in biometric recognition technologies and communicating with machine [5]. These speech signals are slowly timed varying signals (quasi-stationary). When examined over a sufficiently short period of time (5-100 millisecond), its characteristics are fairly stationary. But, for a period of time if the signal characteristics changes, it reflects to the different speech sounds being spoken [1]. A voice analysis is done after taking an input through microphone from a user. The design of the system involves manipulation of the input audio signal. At different levels, different operations are performed on the input signal such as Pre-emphasis, Framing, Windowing, feature extraction analysis and Recognition (Matching) of the spoken word [6]. The voice algorithms consist of two distinguished phases. The first stage is training sessions, whilst, the second stage is referred to as operation session or testing. The fundamental difficulty of speech recognition is that the speech signal is highly variable due to different speakers, and speaking rates, contents and acoustic conditions. The feature analysis component of an ASR system plays a crucial role in the overall performance of the system.

2.Methodology

The proposed methodology the consisting two processes, the initial process is creating the database that is known as the training process. Another one is testing process called as feature matching process.

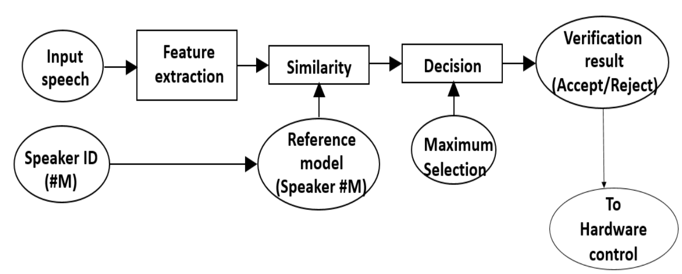


Fig1.Block Diagram of General methodology

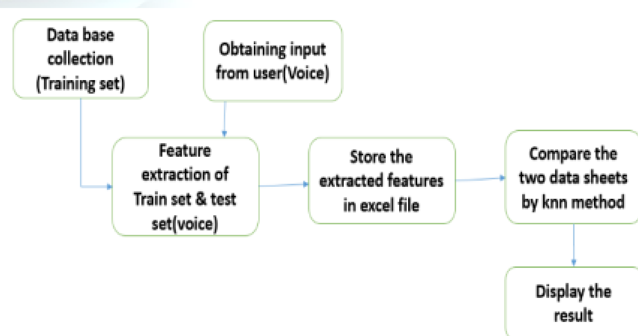


Fig 2. Block Diagram of Proposed Methodology

3. Feature Extraction

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 14, March 2016

The preferred speech signals are obtained from the users. That signals are stored and some features like energy, spectrum, variance, length, and entropy of the signal are extracted [1]. The FFT has to be performed at the time, the sampling rate of the signal has to be obtained. The above features are stored in a excel sheet like matrix format. For the database creation the set off speech signals are extracted and then stored in the excel sheet.

The Energy and some other features are extracted. The average energy of the signal has to be extracted. For the accurate matching the sampling rate will be obtained from the FFT process [10] and applying the hamming window to that signal. The up sampling is performed for that signal to improve the signal quality.

The Extracted Features are

Energy (min & maxvalues)

$$E_i = \sum_{k=N1}^{N2} x_n^2$$

Variance

$$S^2 = \frac{\sum(X-\bar{X})^2}{n-1}$$

Entropy

$H = -\sum P_i(log_2 P_i)$

Length- Sampling rate

Log energy

4. Pre-emphasis

Noise has a greater effect on the higher modulating frequencies than the lower ones. Hence, higher frequencies are artificially boosted to increase the signal-to-noise ratio. Pre-emphasis process performs spectral flattening using a first order finite impulse response (FIR) filter [7].

The filter applies to the signal to be determined by the setting stop and pass band frequency for our convenience. For applying windowing technique to that signal the length of the signal will be calculated for the particular frequency range as given below:

```
num=-20*log10(sqrt(rp*rs))-13;
dem=14.6*(fs-fp)/f;
n=ceil(num/dem);
n1=n+1;
xx=hamming(n1);
b=fir1(n,wp,xx);
x=upfirdn(a,b,1,2);
ny=length(x);
```

fp, fs = Pass band and stop band frequencies,
b= windowing range, ny= sampling rate

4.1. FIR Filter

The FIR filters are always stable. The window chosen for truncating the infinite impulse response should have some desirable characteristics.

The central lobe of the frequency response of the window should contain most of the energy and should be narrow [11].

The highest side lobe level of the frequency response should be small.

$$H(z) = 1 - \alpha z^{-1}, 0.9 \leq \alpha \leq 1.0 \quad (1)$$

4.2. Hamming window

Discontinuities at the beginning and end of the frame are likely to introduce undesirable effects in the frequency response. Hence, each row is multiplied by window function. A window alters the signal, tapering it to nearly zero at the beginning and the end. We use Hamming window as, it introduces the least amount of distortion [9]. Our implementation uses Hamming window of length 256. Equation shows the discrete time domain representation of Hamming window function. Its generates the lesser oscillations

Y(n)=X(n)×W(n) (2)

W(n)=0.54−0.46 $cos\left[\frac{2\pi n}{N-1}\right]$ $0 \leq n \leq N-1$(3)

5. K-NN Algorithm

The current speech signal is obtained from the user and the features are extracted and then stored in a excel sheet. By the use of KNN algorithm the setoff extracted feature values are compared with the already stored values. Christo Ananth et al. [8] discussed about a system, GSM based AMR has low infrastructure cost and it reduces man power. The

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
ernational Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
l. 3, Special Issue 14, March 2016

system is fully automatic, hence the probability of error is reduced. The data is highly secured and it not only solve the problem of traditional meter reading system but also provides additional features such as power disconnection, reconnection and the concept of power management. The database stores the current month and also all the previous month data for the future use. Hence the system saves a lot amount of time and energy. Due to the power fluctuations, there might be a damage in the home appliances. Hence to avoid such damages and to protect the appliances, the voltage controlling method can be implemented.

positions (i,j)=(data1(i,j)-test1(i,j));

i,j are the rows and columns of the feature values stored in excel sheet

The above equation returns the column in the excel sheet where the minimum values presents. From the return values the command will be displayed.

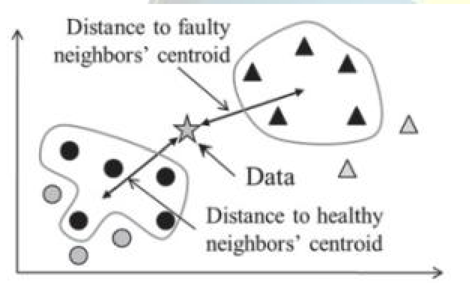


Fig 1. K-NN Algorithm Pictorial representation

The neighbours are taken from a set of objects for which the class (for *k*-NN classification) or the object property value (for *k*-NN regression) is known. This can be thought of as the training set for the algorithm, though no explicit training step is required.A shortcoming of the *k*-NN algorithm is that it is sensitive to the local structure of the data.

6. Result and Discussion

The sample database of different command signals are collected and stored. Fig 2 shows the signals of 6 different commands like turn off fan, turn on light, turn off light, etc. The extracted features of these signals are showed in excel sheet. A snapshot of this is shown in Fig 3. Fig 4 is represent the turn off light 1 command. Fig 5 represent the corresponding Matlab window that shows the features of the testing signal. By using K-NN classifier the signal recognition is performed. Hence the matched signal will be displayed as message box in the Matlab window. This is shown in the Fig 6. It will indicate the corresponding signal is "turn off light 1"

The algorithm is developed using MATLAB R 2015-a version. Initially the database is obtained from the user by recording his/her speech. From the speech signals the required features are extracted and then they are stored in the database.In the speech recognition side the current user speech is collected and then the features are extracted. That features are compared with the database. After recognition of the command this will be loaded to the embedded system to control the home appliances.

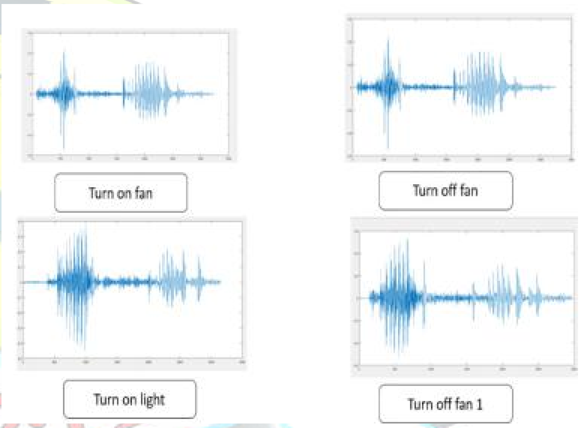


Fig2. Various Speech signal – Training Set (Databasesignals)

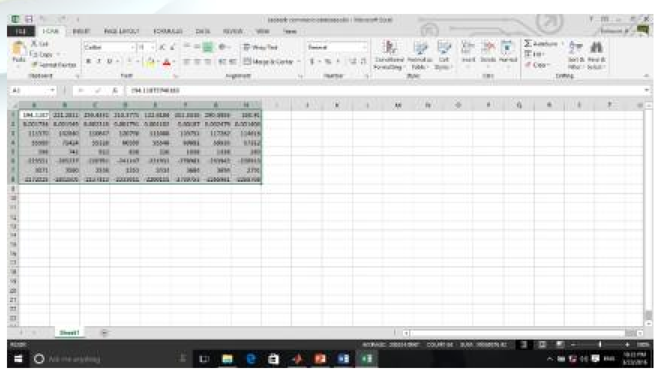


Fig 3. Various Speech signal extracted features in excel sheets (data base-training)

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
ernational Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
l. 3, Special Issue 14, March 2016

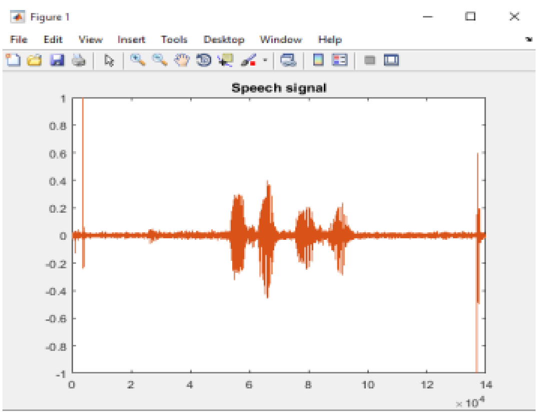


Fig 4. Input signal- Testing signal(turn off fan 1)

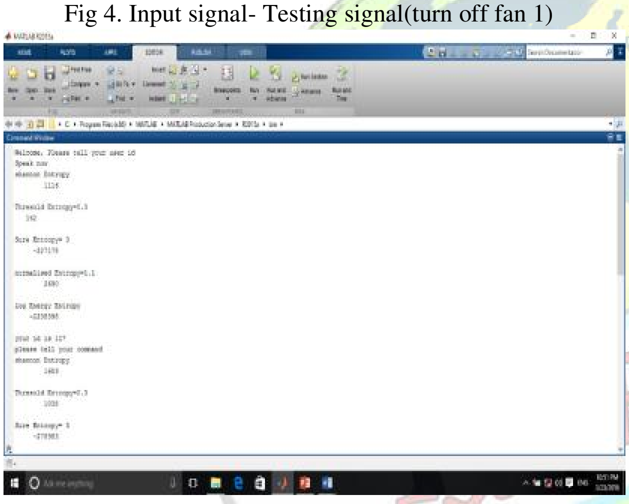
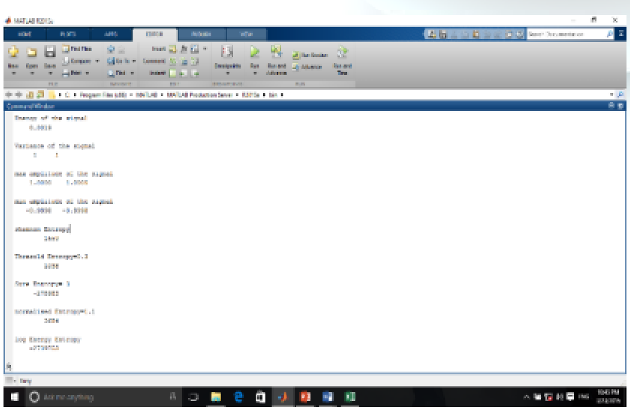
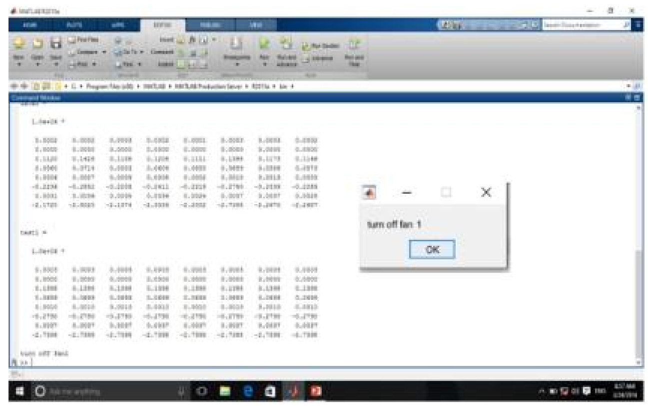


Fig 5. Display the features of the speech signal



(A)



(B)

Fig6 (A & B). Result windows that shows the features of the training and testing speech signal and corresponding output that matched with the database

7. Conclusion

By the above method is used to testing 15 commands which is daily used in the home. So this will help to make the home automation kit. This is the simple method and less expensive. The algorithm is developed using Matlab R2015A version. The K-nn classifier is used to classify the different users. The basic feature extractions techniques are used here. This will help to make the voice based home automation kit

But the only disadvantage is recording the user command as noiseless one. Otherwise the output of the programme will be accurate one.In future the signal processing concept like Kalman filtering can be applied to produce noise free commands before extracting the feature.

8. References

[1]. Bhupinder Singh, Rupinder Kaur, Nidhi Devgun, Ramandeep Kaur "The process of Feature Extraction in Automatic Speech Recognition System for Computer Machine Interaction with Humans", Journal of IJARCSSEVolume 2, Issue 2, February 2012 ISSN: 2277 128X

[2]. Siddhant C. Joshi, Dr. A.N.Cheeran “MATLAB Based Feature Extraction Using Mel Frequency Cepstrum Coefficients for Automatic Speech Recognition”, Journal of IJSETR, Volume 3, Issue 6, June 2014

[3]. Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi “Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques” Journal Of Computing, Vol.2, ISSUE 3, MARCH 2010, ISSN 2151-9617

[4]. Prabhakar V. Mhadse1, Amol C.Wani2“Speaker Identification Based Automation System through Speech Recognition”, Journal of IJETAE ISSN 2250-2459, Volume 3, Issue 1, January 2013

[5]. R.Anandan, Mr.B.Karthik, Dr.T.V.U.Kiran Kumar “Wireless Home and Industrial Automation Security System Using Gsm” Journal on Volume 4, No. 4, April 2013 Journal of Global Research in Computer Science

[6]. Selina Chu, Shrikanth Narayanan, and Jay Kuo, Environmental Sound Recognition With Time–Frequency Audio Features Ieee Transactions On Audio, Speech, And Language Processing, VOL. 17, NO. 6, August 2009

[7]. MostafaHydari, Mohammad Reza Karami, EhsanNadernejad, Speech Signals Enhancement Using LPC Analysis based on Inverse Fourier Methods, Contemporary Engineering Sciences, Vol. 2, 2009, no. 1, 1 - 15

[8]. Christo Ananth, G.Poncelina, M.Poolammal, S.Priyanka, M.Rakshana, Praghash.K., “GSM Based AMR”, International Journal of Advanced Research in Biology, Ecology, Science and Technology (IJARBEST), Volume 1,Issue 4,July 2015, pp:26-28

[9]. J.B. Allen. How do humans process and recognize speech IEEE Trans. on Speech and Audio Processing, 2(4):567–577, October 1994.

[10]. J.B. Allen. Consonant recognition and the articulation index. *J. Acoust. Soc.Am.*, 117(4):2212–2223, April 2005.

[11].Bryce E. Lobdell, Mark A. Hasegawa-Johnson, Jont B. Allen “Human Speech Perception and Feature Extraction”

[12]. Urmila Shrawankar, Dr. Vilas Thakare “Techniques for feature extraction in speech recognition system”

[13]. Rabiner, Juang, Fundamentals of Speech Recognition, 1993

[14]. www.simulink.com Matlab official website

[15]. www.stackoverflow.com website