International Journal for Innovative Engineering and Management Research

A Peer Revieved Open Access International Journal

www.ijiemr.org

IJIEMR Transactions, online available on 12th Mar 2018. Link

:http://www.ijiemr.org/downloads.php?vol=Volume-08&issue=ISSUE-03

Title: **SPEECH STRATIFICATION USING KNN METHOD**

Volume 08, Issue 03, Pages: 133–136.

Paper Authors

DR. P. V. RAMA RAJU,G. NAGA RAJU,R.DEVI PRIYA,S.KRISHNA SRI, T.S.S.PRASAD

SRKR Engineering college(A), Bhimavaram, India

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

SPEECH STRATIFICATION USING KNN METHOD

[1]DR. P. V. RAMA RAJU,[2]G. NAGA RAJU,[3]R.DEVI PRIYA,[4]S.KRISHNA SRI,
[5]T.S.S.PRASAD
[1]Professor & HOD,[2]Asst. Professor,[3,4,5]B.E Students
Department of ECE, SRKR Engineering college(A), Bhimavaram, India.

ABSTRACT:

Speech has been the most commanding and convenient method of communication. However many problems appears because of speech miscommunications such as noisy communication,language hurdle,accents of speakers etc.,. Hence,in this paper,Speech was divided by statistic method,KNN(K-Nearest Neighbour),which separates vowels from consonants.The KNN method was set up in MATLAB by analysing the phoneme data in TIMIT training database,and generating the corresponding Mel Frequency Cepstrum Coefficients for each phoneme.Later the trained KNN divider was tested using TIMIT test database.

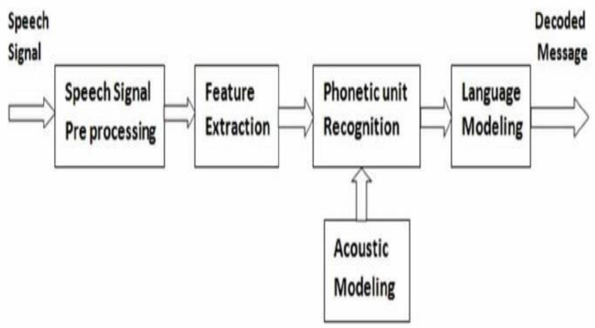
Keywords –Speech,KNN method,MFCC coefficient,phoneme,TIMIT database.

1. INTRODUCTION:

Speech is most instinctive communicative medium for humanbeings.However,many problems appears because of speech miscommunications such as noisy communication,language hurdle,accents of speakers etc.These can disturb the perception and conception the spoken words. This can be bad for normal hearing listeners,but it is even worse for the hearing loss people. Hearing Loss is one of the physical disability in the whole world people having hearing loss problem can't able to listen the normal speech and interpret the meaning. Hearing aids are devices which are used to improve the quality of hearing for hearing loss people.Loss of hearing for a person can affect the quality of life as the speech is the way of communication.Hearing aids having a problem of amplifying both desired signals as well as noise.so filters are used in aids to remove the unwanted signals.But using filters in hearing aids do not improve overall hearing. The motto of this study is to create a speech classification algorithm that will improve on the 80% accuracy of classifyingvowels and consonants of spoken words. The developed algorithm can then be implemented into digital hearing aids, so as to help thehearing impaired person to better understand spoken words.

2. FLOW CHART:

Block Diagram of Speech Recognition System:



3. SOFTWARE TOOLS:

- MATLAB

4. INFORMATION REVIEW:

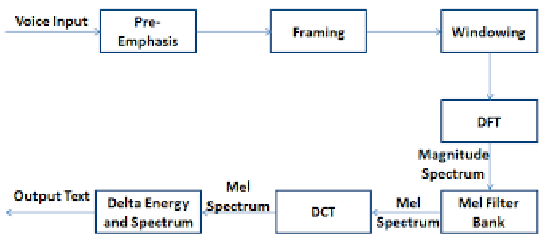
KNN:

International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

KNN is a non-parametric, it does not make any assumptions on the underlying data distribution. This algorithm is based on feature resemblance that is, KNN algorithm takes the points which are in similar with the neighbour's points or values. KNN method is used to identify vowels and consonants differently. The KNN model was built in MATLAB by passing the phoneme data in TIMIT training database and generates the corresponding MEL FREQUENCY CEPSTRUM COEFFECIENTS for each phoneme.There will be no assumptions of nonlinear data. Easy to understand.Accuracy is more adaptable. It is expensive. Requires high memory. Most of the training data is stored.For the entry of large data prediction is somewhat difficult.

MFCC:1

For speech stratification, this is the most important, **melfrequency cepstrum coefficient** (**MFCC**). MFCC takes the human perception and the frequencies into consideration.



PRE-EMPHASIS:

In this process, the voice input k(n) signal is passed through the high pass filter and the output obtained from it be k1(n). The z-transform of the filter is

$$K1(n) = k(n) - a*k(n-1)$$

$$H(z)=1-a*z^{-1}$$

The goal of pre-emphasis are to (1) high frequency has smaller magnitudes when compared to low frequency to create the balance between them pre emphasis is used (2) avoid numerical problems during the Fourier transform operation and (3) it improves the quality ofsignal to noise ratio(SNR).

FRAME BLOCKING:

The given signal is made into frames of 15-20 ms withoverlap of 50% of the frame size. In general the frame size will be equal to power of two in order for the use in the Fast Fourier Transform.

HAMMING WINDOW:

The voice signal is multiplied with the windowing signal. Time domain multiplication equals frequency domain convolution. Here the replacement will takes place, the frequencies will get replaced by the hamming window.

$$k(n)*W(n)$$

TIMIT:

TIMIT supported word,phoneme search criteria.

5. METHODOLOGY:

The methodological approach choosen included:

1.Acquiring the TIMIT database

2.Acquiring MATLAB

3: Speech Selection

4: Pre-Emphasis

5: Analysing TIMIT data into individual phonemes

6: Feature Extraction using MFCC

7: Tabulating the MFCC

8: Creating the KNN

9:Testing the new Query data

6. RESULTS:

Analysing timit sound file:

It created phonemes which resembled those of the phonetic alphabet

Feature Extraction:

Envelopes are enough to represent the difference, so we can recognize phonemes through MFCC.

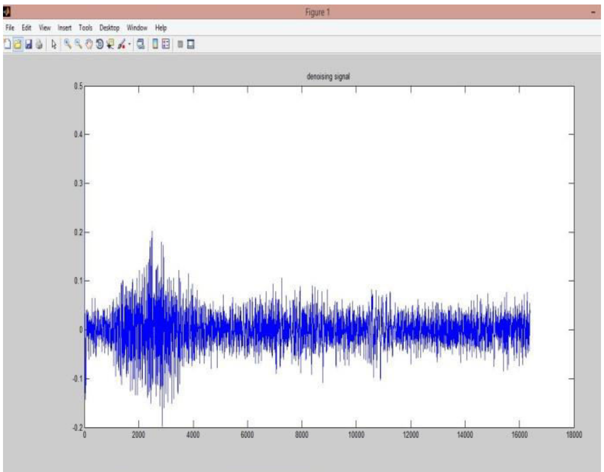
KNN Method:

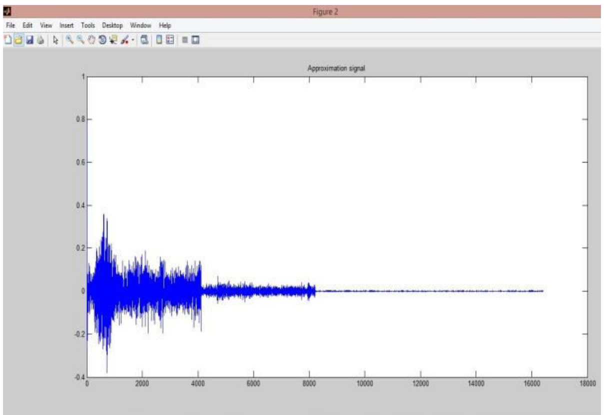
International Journal for Innovative Engineering and Management Research
A Peer Revieved Open Access International Journal
www.ijiemr.org

MATLAB software has an App that was used to generate the associated K-NN model.

TESTING:

This gives an accuracy of 83.4% for all the phoneme of the speech, but given that two of these phonemes produced an error, the actual accuracy of the K-NN model created is 89.7% for this randomly selected speech file.





7: CONCLUSION:

Based on the research done and the results obtained , these are the following conclusions we obtained:

- Using the K-NN model, developed using MATLAB and the TIMIT database, the differentiation of phoneme English vowels and consonants are done.
- The accuracy attained by our project is about 84% to 96%.
- A statistic classifier is less difficult method than known HMM classifier model as KNN used to classify speech and gain an acceptable accuracy level

8: REFERENCES:

[1] McLoughghlin I., 2009, Applied Speech And Audio Processing with MATLAB Examples, United Kingdom Cambridge University Press.

[2] Yunusova Y, Weismer G, Westbury J, Lindstrom M. Articulatory movements during vowels in speakers with dysarthria and in normal controls. Journal of Speech, Language, and Hearing Research. 2008;51:596–611

[3] B. Prasad, S.R.M. Prasanna (Eds.), Speech, Audio, Image and Biomedical Signal Processing using Neural Networks, Springer-Verleg, Berlin (2008), pp. 239-264

[4] Moore, B.C., Glasberg, B.R. Modeling binaural loudness. J Acoust Soc Am. 2007;121:1604–1612.

[5] Strom KE. The HR 2006 Dispenser Survey. Hear Rev. 2006; 13: 16–39

[6] G. Potamianos, pp. 800, 2006, ISBN 9780080448541

[7] M. Rizwan, D. V. Anderson, "Using k-Nearest Neighbour and speaker ranking for phoneme prediction", Machine Learning and Applications (ICMLA) 2014 13th International Conference on, pp. 383-387, 2014

[8] Suraj Jadhav1, Shashank Kava, (2013) Voice Activated Calculator, International Journal of Emerging Technology and Advanced Engineering

[9] Fogerty, D., and Humes, L. E. (2012). "The role of vowel and consonant fundamental frequency, envelope, and temporal fine structure cues to the

intelligibility of words and sentences," J. Acoust. Soc. Am. 131, 1490–1501

About authors:

Dr.P. V.Rama Raju

Presently working as a Professor and HOD of Department of Electronics and Communication Engineering, S.R.K.R. Engineering College, AP, India. His research interests include Biomedical-Signal Processing, Signal Processing, Image Processing, VLSI Design, Antennas and Microwave Anechoic Chambers Design. He is author of several research studies published in national and international journals and conference proceedings.

G. Naga Raju

Presently working as assistant professor in Dept. of ECE, S.R.K.R. Engineering College, Bhimavaram, India. He received BTech degree from S.R.K.R Engineering College, Bhimavaram in 2012, and MTech degree in Computer electronics specialization from Govt. College of Eng., Pune University in 2004. His current research interests include Image processing, digital security systems, Signal processing, Biomedical Signal processing, and VLSI Design.

R.Devi Priya

Presently pursuing Bachelor of Engineering degree in Electronics & Communication engineering S.R.K.R. Engineering College, AP, India.
Email ID :devipriyaragireddy@gmail.com

S. Krishna Sree Presently pursuing Bachelor of Engineering degree in Electronics &Communication engineering S.R.K.R. Engineering College, AP, India.
EmailID:krishnasree8374@gmail.com

T.Sohith Sai Prasad Presently pursuing Bachelor of Engineering degree in Electronics & Communication engineering S.R.K.R. Engineering College, AP, India.
sohithsaiprasad1@gmail.com