ISSN: 2277-9655
(I2OR), Publication Impact Factor: 3.785

IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

# TEXT-DEPENDENT SPEAKER RECOGNITION SYSTEM IN NOISY ENVIRONMENT USING MFCC

**Mr.Youghesh Dawande*, Dr. Mukta Dhopeshwarkar**
Department of Computer Science and IT
Dr.BabasahebAmbedkarMarathwada University, Aurangabad (MS),431001 India

## ABSTRACT

This paper is about text-dependent speaker recognition system in noisy environment.The goal of this paper is to recognize the speaker using noisy environment or to identify the speaker using noised speech signal. We have used low-pass filter to remove the background noise from noised speech signal.For extracting the feature we have used standard speech parameterization method such as MFCC. After completion of parameterization of speech, Vector quantization has been used in the learning and recognition purpose.Wehave achieved 94% of recognition rate for noisy environment.

**KEYWORDS**: Recogination , Speaker, Noisy,Identity.

## INTRODUCTION

Due to the beginning of utilization, identifying a human beings has been become a crucial part of our society. Identification of human or person identification is need for different business sectors such as finance, health care, transportation, entertainment, low enforcement, security a access control, border control, government and communication.As our society becomes electronically connected to from a big global community. Now it becomes necessary to identify a person based on his characteristics.In old tradition people used password, cards to identify a person. But these mode of authentication is not so secured because password may guess or leak by unauthorized person and card may stolen or loss.In now a day's biometrics is automatic mode of identification human being. Biometric is refers to identifying human being based on his physiological and behavioral characteristics. Physiological characteristics such finger print, Iris, retina, hand geometry, face. Behavioral characteristics such as signature, gait, voice, speaker, speech, keystroke etc.The recognition characteristics cannot be stolen or loss because biometric cant not identification is shared or misplaced and they intrinsically represent the individuals bodily identity. Biometric is not only pattern recognition research problem but also if it used carefully could also be enabling technology with the potential to make our society safer reduce fraud and better convenience.[1,2,5,10]

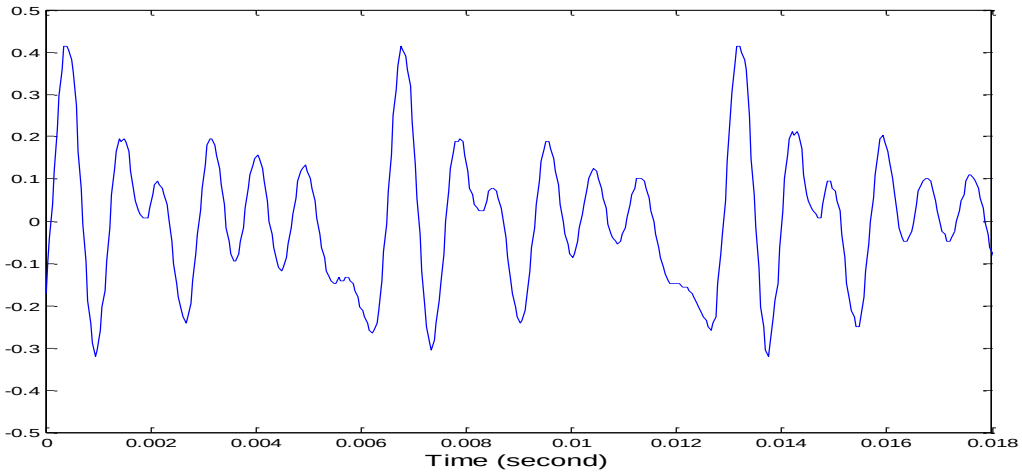
## SPEAKER RECOGNITION

Now a day's more and more attention has been paid on speaker recognition area. Speaker recognition has two major applications: a) Speaker Identification, and b) Speaker Verification. These are the technique for automatically recognizing who is speaking on the basis of individual informational included in speech waves. This technique make it possible to use the speakers voice to verify their identity and control accessing to service such voice dialing, banking by telephone, telephone shopping, database access service, information service, voice mail, security control for confidential areas, and remote access to computers.[6]

Speaker recognition is technique that can identify or verify of the pre known speakers. The recognition process is based on the parameter of speech samples. Speaker recognition is classified in to two sections Text-dependent recognition and Text-independent recognition.

## EXPERIMENTAL WORK

We have collected the sample speech from various 50 speakers of 10 utterance of each. Input speech is recorded at sampling rate of 16000Hz. The example recorded speech sampled in digital form as following diagram.




A speaker recognition system has three important parts

1) Preprocessings
2) Feature extraction
3) Pattern matching

## PRE-PROCESSING

The first step of speech processing, the system needs to remove the silence and background noise from the speech signal. Silence detection strategy is based on the energy level in way that if variation of the signal sample in a known speech frame, compare to the frame mean are big enough, the frame is considered as frame, otherwise as a silence.

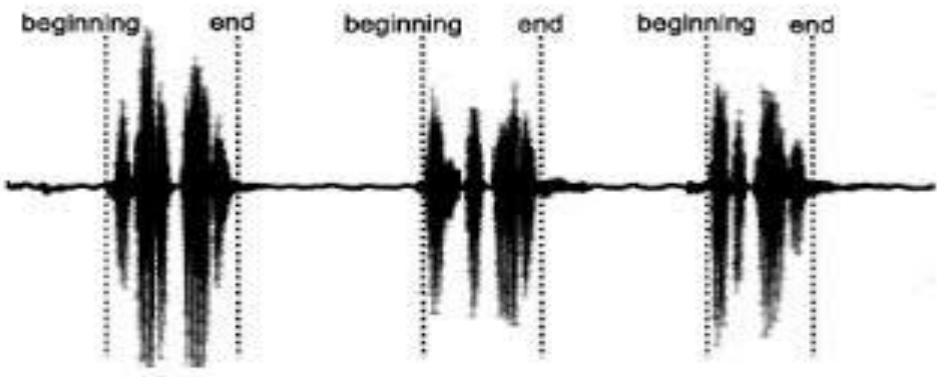



***Figure 2: silence detection***

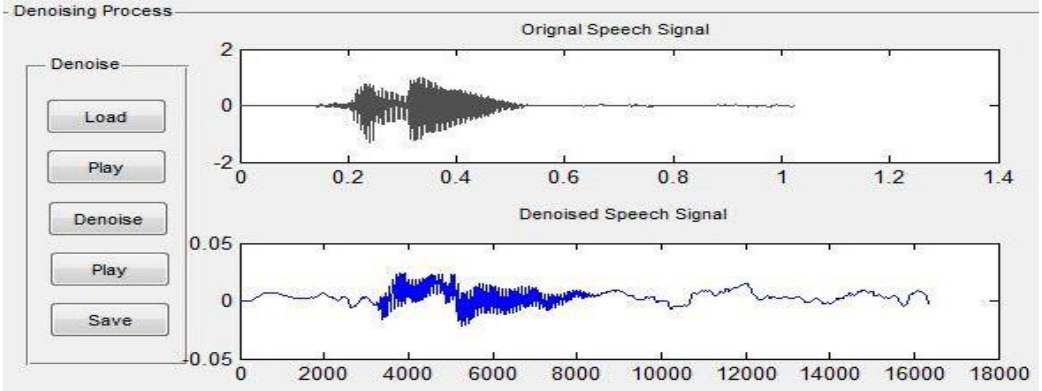



***Figure 3: noise removes***

After silence detection and noise remove from speech signal, the raw speech signal was prepared for feature extraction.

## FEATURE EXTRACTION

The most important part of speaker recognition is the extraction of features from the speech signal. This step helps us to bring the important part of data from set of samples. This make easier to find the different types of features extraction techniques. Linear Predictive Cepstral Coefficients (LPCC), Mel-frequency Cepstral Coefficients (MFCC). To extract the feature from the speech signal mel-frequency cepstral coefficients are frequently used speech parameterization in speaker recognition. The example of MFCC feature extraction from speech sample in digital form as following figure

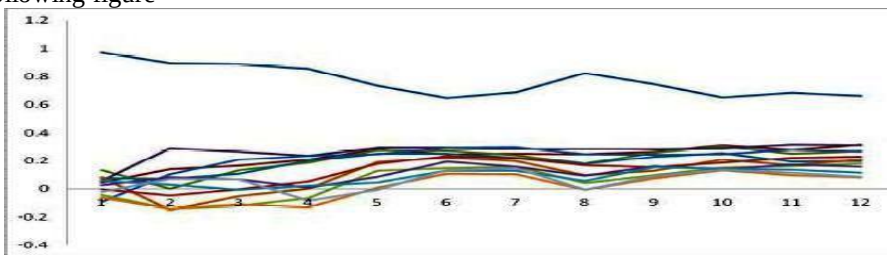


***Figure 4: Plot of MFCC 13 features for 12 frames***

The extracted feature speech samples of a speaker are quantized to number of centroids using quantization algorithm MFCC are calculated in training phase and testing phase. The speaker uttered same word in a training session and once in testing session later. Then the Euclidean distance the MFCC of each speaker in training phase to centroids of individual speaker in testing phase. The main goal is that of recognize the speaker from set of N known speaker. [8]

## FEATURE MATCHING

The feature matching techniques used in speaker recognition include Dynamic Time Wrapping (DTW), Hidden Markov Model (HMM) and Vector Quantization (VQ). We have used VQ approach to ease of implementation and high accuracy. The vector quantization also called block quantization or pattern matching quantization is often used in lossy data compression.VQ divides a large set of points into groups having approximately the same number of points close set to them. Each group is represented centroid point.

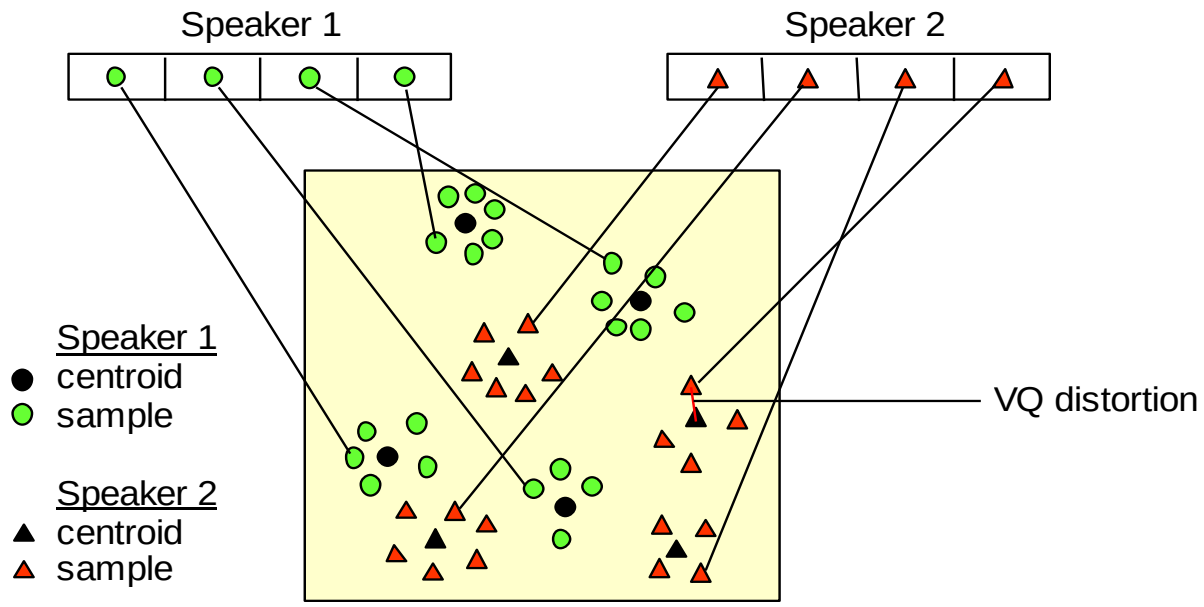



***Figure 5: Illustrating Vector Quantization***

After the enrolment session, the acoustic vectors extracted from input speech of a speaker provide a set of training vectors. The next imported step is to build a speaker specific VQ codebook for this speaker using those training vectors. [9]

## RESULT OF SYSTEM

The database consists of 50 distinct speakers including both male and female known and unknown speakers. It also contains 500 speech samples used for training and testing the speaker recognition system. Speech samples recorded in noisy environment. Recognition rate of the system defined as follows

$$RR \frac{Total\ Number\ Reconised\ speakers}{Total\ number\ of\ speaker} \times 100$$

| Sr.No | Environment | Number of sample tested | Number of sample recognized accurately | Recognition Rate (%) |
|---|---|---|---|---|
| 1 | Noisy | 50 | 47 | 94% |

***Table1: Overall recognition rate of proposed speaker recognition system***

| Sr.No | Feature Maching Technique | FeatureMatching Technique | Recognition Rate |
|---|---|---|---|
| 1 | MFCC | VQ | 94% |

***Table2:MFCC+VQ Recognition Rate of our System***

## CONCLUSION

The performance of measured on the basis of accuracy, time taken to compute the feature recognition, it was observed that the speaker recognition system performs well in noisy environment. The entire research process was carried out using MATLAB R2012a on an Intel i5powerd machine. It noticed that the recognition result 94% of the noisy environment.

## ACKNOWLEDGEMENTS

We author sincerely thank to Department of Computer Science and IT, Dr.BabasahebAmbedkarMarathwada University Aurangabad to provide facility to carry this research work.

## REFERENCES

[1] Campbell, J.P., Jr.; "Speaker recognition: a tutorial" Proceedings of the IEEE Volume 85, Issue 9, Sept. 1997 Page(s):1437 – 1462.

[2] Pradeep.Ch," Text Depandent Speaker reogination Using MFCC and LBG VQ", National Institute of Technology, Rourkela, 2007.

[3] Fu Zhonghua; Zhao Rongchun; "An overview of modeling technology of speaker recognition", IEEE Proceedings of the International Conference on Neural Networks and Signal Processing Volume 2, Page(s):887 – 891, Dec. 2003.

[4] Nakai, M.; Shimodaira, H.; Kimura, M.; "A fast VQ codebook design algorithm for a large number of data", IEEE International Conference on Acoustics, Speech, and Signal Processing, Volume 1, Page(s):109 – 112, March 1992.

[5] D. Reynolds and B. Carlson, "Text-Dependent Speaker Verification Using Decoupled andIntegrated Speaker and Speech Recognizers," In EUROSPEECH in Madrid, ESCA, pp.647-650, 1995.

[6] Mangesh S. Deshpande and Raghunath S. Holambe; Speaker Identification Using Admissible Wavelet Packet Based Decomposition International Journal of Signal Processing 6:1, 2010

[7] S. Furui., "Speaker independent isolated word recognition using dynamic features of the speech spectrum", IEEETransactions on Acoustics, Speech and Signal Processing, Vol. 3, pp. 52–59, Feb. 1986.B. S. Atal,"Text-independent speaker recognition," J.A.S.A., 52, 181 (A), 83th ASA, 1972.

[8] S. Furui, "Fifty years of progress in speech and speakerrecognition," Proc. 148th ASA Meeting, 2004

[9] Kavitha K J An automatic speaker recognition system using MATLAB World Journal of Science and Technology 2012, 2(10):36-38

[10] Philipos C. Loizou,"Speech Enhancement: Theory and Practice", CRC Press. June 7, 2007.

[11] VibhaTiwari, "MFCC and its application in speaker recognition", International Journal on Emerging Technologies, Vol. 1, No. 1, pp. 19-22 (2010).

[12] Voiers, W.D.: Diagnostic evaluation of speech intelligibility. In: Hawley, M.E. (ed.) Speech Intelligibility and Speaker Recognition, pp. 374–387. Dowden, Hutchinson & Ross, Stroudsburg (1977).