International Journal of Technical Innovation in Modern Engineering & Science (IJTIMES)
Impact Factor: 5.22 (SJIF-2017), e-ISSN: 2455-2585
Volume 4, Issue 7, July-2018

Speech enhancement and Speech recognition by improving PSNR Value

G. Venkat Rao[1]
[1]*M.Tech, Assistant professor, Department of ECE, Lakireddy Balireddy College of engineering, AP, India.*

Abstract— In order to improve performance robustness of Speech enhancement and speech recognition system under difficult situations, we recommend improved Peak-Signal to Noise Ratio (PSNR).It consists of two stages .In First stage we Collect speech samples from group of people and stored in database. In the second stage, we recognize the speech from the database, whether it is present in database or not. Stochastic matching criterion is used during estimation of the super vector for a new speech signal. In the first stage, aim is speech enhancement and in the second stage aim is speech recognition with minimum error. Comparing to baseline system, speech enhancement and speech recognition system by improving PSNR provides better results.

Keywords— PSNR, speech, enhancement, recognition, stochastic matching criterion.

I. INTRODUCTION

Numbers or symbols are used in the representation of digital signal. In addition to representation of digital signals, processing of signals gives us digital signal processing. Signal processing can be separated in to analog signal processing and digital signal processing. Speech and audio signal processing is one part of Digital signal processing. Some of the digital signal fields are radar and sonar signal processing, sensor array processing, spectral estimation, statistical signal processing, digital image processing for communications, biomedical signal processing, and seismic data processing[2].

1.1 Signal Sampling

Discretization and quantization are two steps involved in sampling. During sampling one of the criteria we should follow is Nyquist sampling, i.e if the sampling frequency is larger than the two times of highest frequency of the signal, it can be fully reconstructed[2].

1.2 Digital signal processing Applications

Speech signal processing, audio signal processing, audio signal compression, digital image processing, video signal compression, digital communications, Bio Medicine are major applications of digital signal processing. For high speed applications FPGAs are used. ASICs can be specially designed for high speed applications with larger usage purpose[3].

1.3 Speech Recognition

Spoken words are converted into text in speech recognition. In some cases voice recognition is referred as speech recognition. Speech recognition systems that must be trained to a particular speaker as is the case for most desktop recognition software. Recognizing the speaker can simplify the task of translating speech. Speech recognition is a broader solution which refers to technology that can recognize speech without being targeted at single speaker such as a call center system that can recognize arbitrary voices[3].

II. PROPOSE METHOD

In speech enhancement and speech recognition system, with the help of the micro phone and graphical user interface speeches are recorded and stored in database.

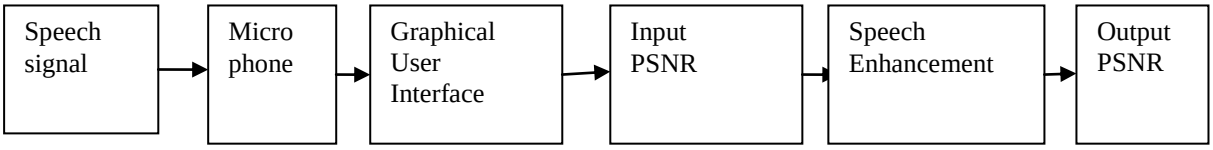

Fig 1: speech enhancement block diagram

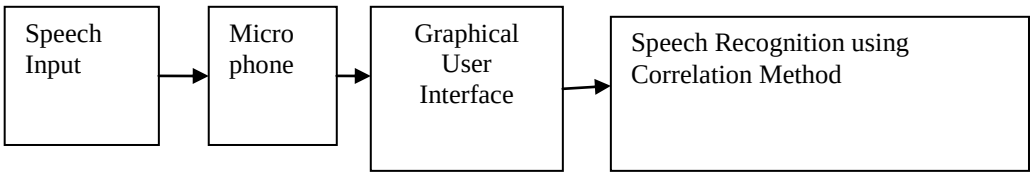

Fig 2: speech recognition block diagram

For speech recognition, speech is recorded through microphone and graphical user interface. It is correlating with database speech signals. The entire process of speech enhancement is explained through flow chart1.The process for Speech recognition is explained in flow chart2.

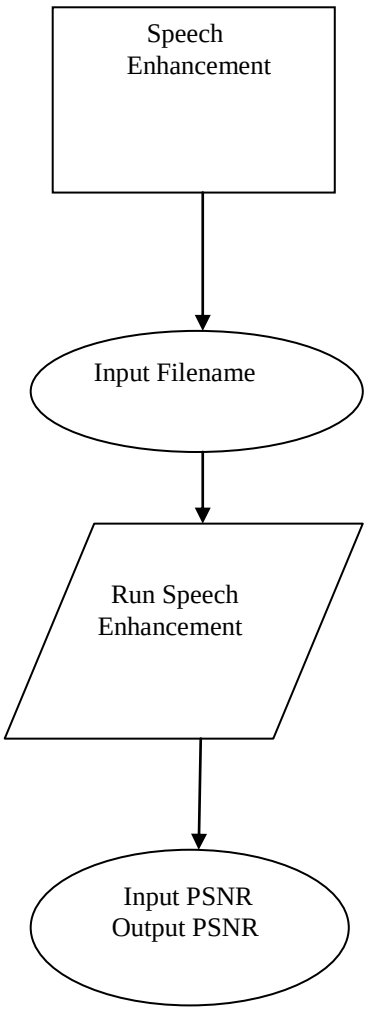

Fig -3: Flow chart1 for speech enhancement

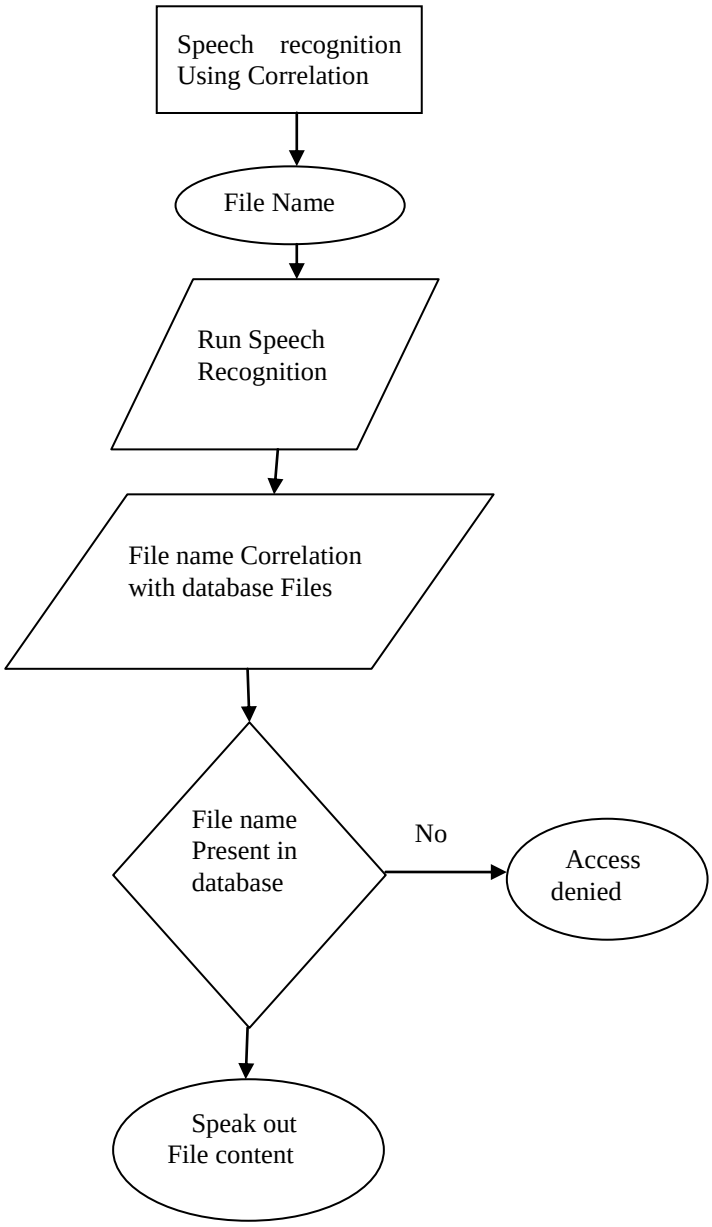

Fig -4: Flow chart2 for speech recognition

For Input speech signals PSNR value is calculated. After performing speech enhancement, output signal PSNR value is calculated. The speech is recognized by using correlation method.

The input speech PSNR and output speech PSNR values are shown in table1.

Table1: Speech PSNR Values

Input PSNR Value	84.2632
Output PSNR Value	88.5023

CONCLUSIONS

Speech Enhancement and Speech recognition system provides us recognition of Speeches from distorted and noise presented speeches. Here mentioned techniques improve the Signal PSNR value. First introduce Hamming window and FIR filtering then, propose Peak-Signal to Noise Ratio to improve environment discriminative power and tested the performance [1].

For Speech Recognition, although it requires an online cluster selection process before stochastic matching, the dimensionality of the selected subspace is smaller than the original space. The computational cost is therefore lower than the original method. Moreover, the selected subspace can provide higher resolution to model the testing environment than the entire Speech signal data base. For Speech Recognition the parameters belonging to different groups are estimated individually, and therefore the overall estimation obtained accurately [1].

Although we need to conduct several stochastic matching procedures instead of once, partitioning high dimensional is favourable in applications with limited resources of online operation. Recognition results indicate that ESSEM achieves better performance with a Speech enhancement.

I had implemented the Speech enhancement and speech recognition by improving PSNR value with an database formed by 34 different environments in the first phase. I believe the same approach can be extended to more environments for different Automatic Speech Recognition tasks [1].

Recommendations for further study can be as follows

1) In many online speech recognition issues are also critical to the Speech enhancement and speech recognition system performance [1].

2) In the experiments section, an enhanced environment configuration can facilitate both accuracy and efficiency in operating conditions [1].

3) To cover very different testing conditions, we may need to include more environments in the training set or use a more complex online mapping function [1].

REFERENCES

[1] Yu Tsao and chin-Hui Lee, “An Ensemble speaker and speaking environment modeling Approach for robust speech recognition” IEEE transactions on audio, speech and language processing, vol. 17,No.5, JULY. 2009,

[2] Wiley, A course in digital signal processing.

[3] Dag stanneby and William walker “digital signal processing and applications “. 2nd edition, 2004.

BIOGRAPHIES

G. VENKAT RAO Received the B.Tech in Electronics & Communications Engineering from JNTUH University in 2006 and M.Tech in JNTUH University in 2011.His area of interest is EMBEDDED SYSTEMS. Presently he is working as a Assistant Professor, Department of Electronics and Communication Engineering, Lakireddy Bali Reddy College of Engineering, Andhra Pradesh.