Automatic Speech Recognition Incorporating Modulation Domain Enhancement

Ms. Jyoti Londhe [1]

[1,2]*Dept. Of Electronics and Telecommunication Engineering, TPCT's college of engineering , Osmanabad, Maharashtra, India*

Introduction

A clean and clear speech signal is linked to the amount of sound in speech development. There are many ways to make a speech signal without a clamor signal. In this paper we are going to study emotion recognition by using KNN filter(Wiener filter), FFT, and melcepts methods.In many places Phones and cell phones get noisy air Domains like cars, airports, roads, trains, stations. Therefore, we attempt to eliminate the clamor signal by using a spectral reduction method. The main purpose of this paper is for real-time device to reduce or decrease background sound with a visual speech signal, this is called speech development. Variety of languages of speech are present, in that background noise lowering speech. Applications like mobile communication can be learned a lot in recent years, speech improvement is required. The purpose of this speech development is to reverse the noise from a noisy speech, such as the speech phase or accessibility. It is often difficult to remove the background sound without interrupting the speech, therefore, the exit of the speech enhancement system is not allowed between speech contradiction and noise reduction.

There may be other techniques such as Wiener filtering, wavelet-based, dynamic filtering and optical output remain a useful method. In order to reduce the spectral, we must measure the clamor spectrum and reduce it from the clamorous acoustic spectrum. Completely this approach, there are the following three scenarios to consider: sound adds speech signal and sound is not related to a single channel in the market. During this paper, we attempted to reduce the audio spectrum in order to improve distorted speech by using spectral output. We have described the method tested in the actual speech data frame in the MATLAB area. The signals we receive from Real speech signals are a website used for various tests. Then we suggest how to reduce the noise between the average noise level and the noise spectrum.

In general, only one medium system is set up based on a variety of speech data and unwanted screaming that, it works in difficult situations where no previous clamor intelligence is available. Genres often assume that sound is stable whenever the speech is alert. They usually allow for disturbed sound during speech operations but in reality, when the sound is not moving, the performance of the speech signal is greatly reduced.

WAYS TO REDUCE NOISE

How to Remove the Spectral

Many additional variations of the various conclusions are designed to improve speech. The one we are using is thought to be the end of the phantom release. This type works within the scope of the spread and creates the hope that the help cycle is said to be due to the phantom additon sound and clamor phantom. The action is specified within the image below and contains two main components.

Convolutional Denoising Autoencoder (CDAE)

Convolutional Denoising Autoencoder (CDAE) Promotes the same function of default encoders by adding convolutional encoding and decoding layers. CDAE has a 2D layout in the dialog and adjusts the input to the 2D alignment structure using the embedded space and the desktop as shown in the picture. Thus, in the study we proposed, So from this method we can propose that the CDAE is best suited for speech development which is As shown in Eq.

(3), on any map included,

$hi = f(Wi * \alpha + ai)$ (3)

where, the process of conversion and the value of bias is Measuring the background of the spectrum.

Reduce the spectrum noise from the sound speech spectrum

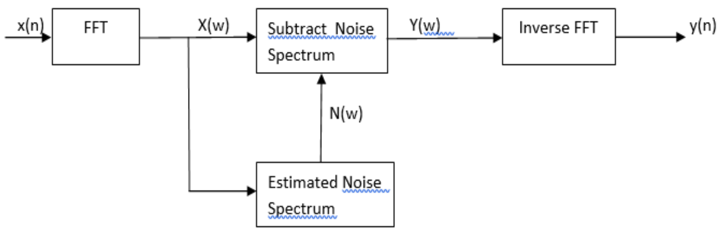

Figure 1: How to Remove the Spectral

Assuming that x (n) can be sound and the modified speech signal is considered to be x (w) is a sound signal spectrum will be, N (w) the spectrum of rated intervals and the range of speech processed by Y (w) and the pure speech signal the first is y (n). therefore, the processed spectrum of the unpolluted speech signal will be provided as follows:

Y (w) = X (w) - N (w)

Next Drawing Speech Block Impression By Spectral Removal.

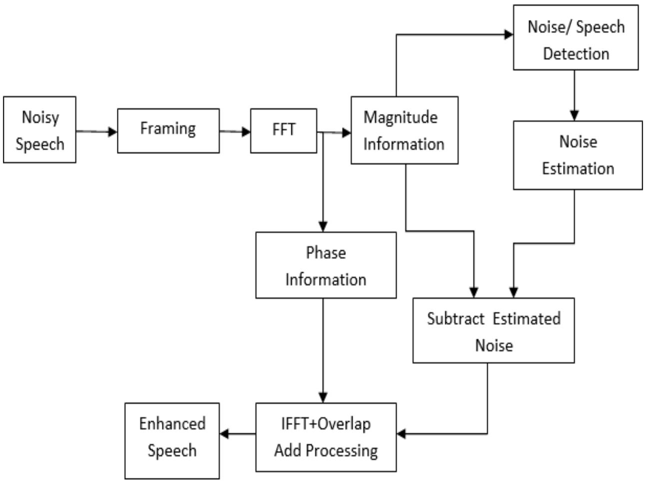

Figure 2: Speech development block diagram

Here, create a continuous output signal. Each frame is 50% overwhelmed. FFT (Finite Fourier transform) is used in each frame.

FFTs are often useful in determining a signal hidden in a noisy time zone in parts of frequency. The functions Y = FFT (x) and y = IFFT (X) suggest the alternating rotation given the vectors length N by

$$X(k) = \sum_{j=1}^{N} x(j)\omega_N^{(j-1)(k-1)}$$

$$x(j) = (1/N)\sum_{k=1}^{N} X(k)\omega_N^{-(j-1)(k-1)}$$

Here,

The proposed method of noise reduction is provided for visual removal that relied on the reduction of the background sound level and the development of the speech signal. Evaluate the volume of sound enhanced by reduced visual acuity and the type of sound reduction recommended.

Sound spectrum measurement:

$$\omega_N = e^{(-2\pi i)/N}$$

The spectrum sound cannot be calculated in advance, but it will be about the time for some instance speech is not present within the sound speech. When the people are speaking, one should definitely pause to take a deep breath. We can take these gaps within speech to balance the background.

We can calculate the average size which is calculated for any frame in the last few seconds.

Removing Noise Spectrum:

After removing, all data of spectral signal which is appear unlikely in positive values. Some chances are available for removing unwanted parts. Fourier conversions, using section segments directly from the Fourier conversion unit, and additional additions are made to reconstruct the speech rate within the time zone.

The basic idea is to reduce the noise from the input signal:

Y (w) = X (w) - N (w)

Suggest workflow and simulation results

FLOWCHART

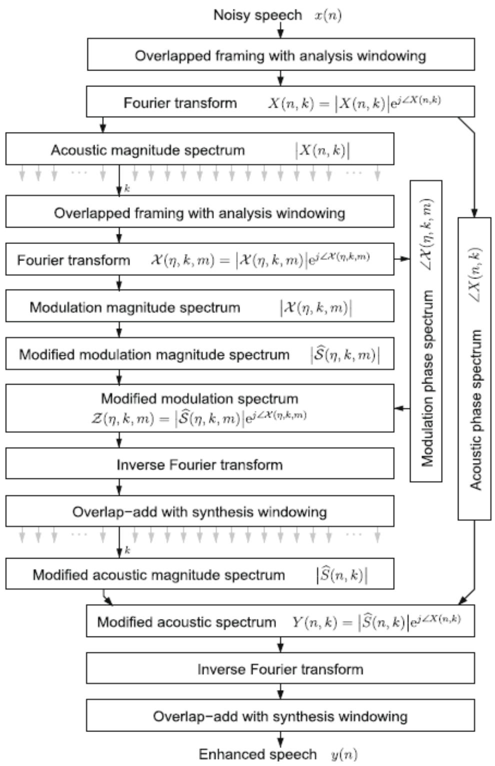

Fig. 3 Flow Chart

The Overlap add-on method is used to interrupt and keep the signals into smaller segments to make the process easier. The scattering method depends on: (1) signal decode in basic parts, (2) the procedure of every part, (3) reassembling the cleared parts in an additional speech. In order to execute domain of frequency process to reduce visibility, it's important to divide the constant wisdom signal speech into compact fragments which is called by frames. After completing procedure, the frames are then reassembled.

The flowchart above shows the proposed method, which involves collecting audio speech information and passing through a window to extract existing art objects within the Clamorous signal and use the FFT algorithm that detects the phase and magnitude of the audio signal, during which the Technology focuses on the size of the clamor to produce a different refined speech signal. By using the type of visual reduction the noise is measured and reduced by the value of the required magnitude. Then use the IFFT algorithm and the overlap adds the process to wish for a refined expression in the time zone.

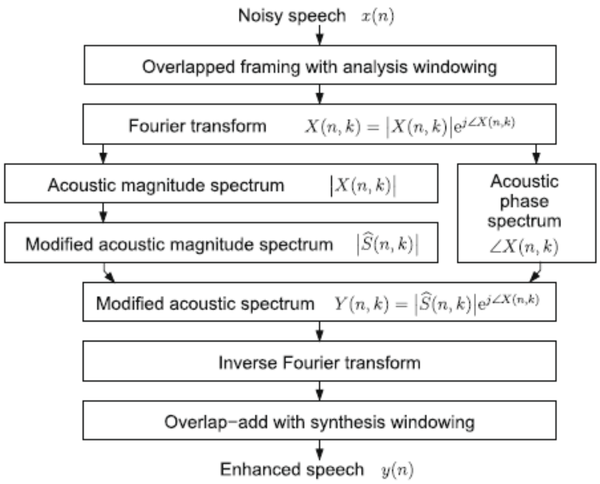

Fig. 4 AMS Framework

Feature extraction techniques for speech recognition Feature extraction (ShahamShabani&YaserNorouzi, 2016) is the most essential part of speech recognition as it distinguishes one speech from other. The extracted features should meet the following criteria:

It should occur frequently and naturally in speech

It should not be receptive to mimicry

It should show less variation from one speaking environment

to another Easy to measure extracted speech feature

It should be balanced over time

RESULTS THAT WE GOT FROM SPEECH SIGANAL AND COMPARED WITH BACKGROUND NOISE IN DB FOR DIFFERENT SPEECH SIGNA

Table –

S = Reduced Noise
F(T) = Trained Signals
mfc = Melcepst Value
TS=Test Sample
W = Winner Class(Classify KNN)
IE = Identified Emotion

Input sig (1) – Identified emotion = **Happy**

Sr. no	Parameters	Minimum Value	Maximus Value
1	S	-0.1817	0.1755
2	F(T)	-0.1245	-0.1245
3	mfc	-9.8171	4.9611
4	TS	-7.1795	3.9534
5	W	1	1
6	IE	Happy	Happy

Table No-(1)

Input sig (2) – Identified emotion = **Happy**

Sr. no	Parameters	Minimum Value	Maximus Value
1	S	-0.9906	1.0093
2	F(T)	0.3048	0.3048
3	mfc	-2.0390	16.0743
4	TS	-1.1279	13.4064
5	W	1	1
6	IE	Happy	Happy

Table No- (2)

Input sig (3) – Identified emotion = **Happy**

Sr. no	Parameters	Minimum Value	Maximus Value
1	S	-0.4381	0.5596
2	F(T)	-0.2610	-0.2610
3	mfc	-12.2700	3.6541
4	TS	-4.3800	0.5596
5	W	1	1
6	IE	Happy	Happy

Table No-(3)

Input sig (4) – Identified emotion = **Sad**

Sr. no	Parameters	Minimum Value	Maximus Value
1	S	-0.1080	0.1333
2	F(T)	0.0217	0.0217
3	mfc	-8.2300	6.8399
4	TS	-7.5872	6.3223
5	W	2	2
6	IE	Sad	Sad

Table No-(4)

Input sig (5) – Identified emotion = **Sad**

Sr. no	Parameters	Minimum Value	Maximus Value
1	S	-0.3765	0.3768
2	F(T)	-0.7768	1.6781
3	mfc	-3.3124	5.2504
4	TS	-0.7768	1.6781
5	W	2	2
6	IE	Sad	Sad

Table No-(5)

Input sig (6) – Identified emotion = **Sad**

Sr. no	Parameters	Minimum Value	Maximus Value
1	S	-0.2380	0.2939
2	F(T)	-0.0324	-0.0324
3	mfc	-2.1412	3.9279
4	TS	-1.2917	1.4274
5	W	2	2
6	IE	Sad	Sad

Table No-(6)

From Reduced noise level for male, It will be found the amplitude stages of male and female sound getting input and output different. This indicates that it has a significant effect on sound speech. It promotes the clamor reduction phase. This type of spectral reduction creates a signal of speech but also of shouting. Noise reduction was not properly guaranteed and can be guaranteed if separated by input and output.

The Clamor Stage of the input and output speech signal and amplitude reduction updates are different from using the table. Table 1 shows the background stage division for the male and female speech signal site.

Emotions Recognition Accuracy Table :

Sr.No	Sample signal	Emotion Recognition	Yes / No Accuracy	Accuracy in Percentage (%)
1	jyoti.mp3	Happy	yes	100%
2	Baba.mp3	Happy	yes	100%
3	jyoti.mp3	Happy	yes	100%

Table No (7)-Recognized happy emotion

Sr.No	Sample signal	Emotion Recognition	Yes / No Accuracy	Accuracy in Percentage (%)
1	umesh.mp3	Sad	yes	66.66%
2	jyoti.mp3	Sad	yes	66.66%
3	Baba.mp3	Happy	No	66.66%

Table No (8)- Recognized Sad emotion

In this paper we are going to study emotion recognition by using KNN filter(Wiener filter), FFT, and melcepts methods. Above table shows the accuracy of our project output. Table No(7) showing output accuracy of sample input signals for happy emotion recognition. And Table No(8) showing output accuracy of sample input signals for sad emotions. We used Matlab data base to store the training recorded signals, and input signal by creating various folders and Files. We gave three input signal to recognize and calculate accuracy of each input signals same emotion recognition results for all those input three saple signals then the accuracy should be 100%. If we don't get each time same output results for three input signals then accuracy of our project should be less that 100%.

Formula to calculate project accuracy :-

Sample input signals in Output (yes/no) × 100 = Accuracy of project %

For Table No-(7) we gave three input sample signals and every time we get correct emotions recognition output. For table no (7) we can calculate Accuracy of project by using above metion formula.

$$\frac{3}{3} \times 100 = 100\ \%$$

For Table No-(8) we gave three input sample signals and for two time we get correct emotions but at 3rd input signal we not got correct output recognition output for table no-(8) We can calculate Accuracy of project by using above mention formula.

$$\frac{2}{3} \times 100 = 66.66\ \%$$

CONCLUSIONS

In this presentation paper we successfully recognize emotion of speech like sad, Happy, Angry, etc. by using wiener filter, KNN filter, FFT, and it is helpful to develop a better spectral reduction algorithm and recognize emotions of a speech by using methods like melcepts, MFCC, Spectral subtraction. It has recently been observed in the fictional results that the suggested type significantly reduces low noise compared to the algorithm of many ways to reduce visibility. This type of speech causes the signal to be visible but is accompanied by a negative screaming. The sound is not very low and it is very low compared to input and output. These forces will also be adjusted and expanded to accommodate static noise. From this type of domain design we have reached 70% and that we can build this system for embedded organizations that are subject to speech processing or communication purpose. For better comparisons, we showed results and therefore spectrograms of clear, sound and prepared speech. Delete the site-prepared speech sentence.

REFERENCES

[1] S. Kamath, and P. Loizou, A multi-band spectral subtraction method for enhancing speech corrupted by colored noise, Proceedings of ICASSP-2002, Orlando, FL, May 2002. R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[2] S. F. Boll, Suppression of acoustic noise in speech using spectral subtraction, IEEE Trans. on Acoust. Speech & Signal Processing, Vol. ASSP-27 , April 1979, 113- 120.

[3] M. Berouti, R. Schwartz, and J. Makhoul, Enhancement of speech corrupted by acoustic noise, Proc. IEEE ICASSP , Washington DC, April 1979, 208- 211.

[4] H. Sameti, H. Sheikhzadeh, Li Deng, R. L. Brennan, HMM-Based Strategies for Improvement of Speech Signals Embedded in Non-stationary Noise, IEEE Transactions on Speech and Audio Processing, Vol. 6, No. 5, September 1998. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 04 Issue: 03 | Mar -2017 www.irjet.net p-ISSN: 2395-0072 © 2017, IRJET | Impact Factor value: 5.181 | ISO 9001:2008 Certified Journal | Page 2488

[5] Yasser Ghanbari, Mohammad Reza Karami-Mollaei, Behnam Ameritard "Improved Multi-Band Spectral Subtraction Method For Speech Enhancement" IEEE International Conference On Signal And Image Processing, august 2004.

[6] Pinki, Sahil Gupta "Speech Enhancement using Spectral Subtraction-type Algorithms" IEEE International Journal of Engineering, Oct 2015.

[7] S. Kamath, A multi-band spectral subtraction method for speech enhancement, MSc Thesis in Electrical Eng., University of Texas at Dallas, December 2001.

[8] Ganga Prasad, Surendar "A Review of Different Approaches of Spectral Subtraction Algorithms for Speech Enhancement" Current Research in Engineering, Science and Technology (CREST) Journals, Vol 01 | Issue 02 | April 2013 | 57-64.

[9] Lalchhandami and Rajat Gupta "Different Approaches of Spectral Subtraction Method for Speech Enhancement" International Journal of Mathematical Sciences, Technology and Humanities 95 (2013) 1056 – 1062 ISSN 2249 5460 .

[10]Ekaterina Verteletskaya, Boris Simak "Enhanced spectral subtraction method for noise reduction with minimal speech distortion" IWSSIP 2010 - 17th International Conference on Systems, Signals and Image Processing .

[11] D. Deepa, A. Shanmugam "Spectral Subtraction Method of Speech Enhancement using Adaptive Estimation of Noise with PDE method as a preprocessing technique" ICTACT Journal Of Communication Technology, March 2010, Issue: 01.

[12]M. Berouti, R. Schwartz, & J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," Proc. ICASSP, pp. 208-211, 1979.

[13] Prince Priya Malla, Amrit Mukherjee, Sidheswar Routary, G Palai "Design and analysis of direction of arrival using hybrid expectation-maximization and MUSIC for wireless communication" IJLEO International Journal for Light and Electron Optics, Vol. 170,October 2018.

[14] Pavan Paikrao, Amit Mukherjee, Deepak kumar Jain "Smart emotion recognition framework: A secured IOVT perspective", IEEE Consumer Electronics Society, March 2021.

[15] Amrit Mukhrerjee, Pratik Goswani, Lixia Yang "DAI based wireless sensor network for multimedia applications" Multimedia Tools & Applications, May 2021, Vol.80 Issue 11, p16619-16633.15p.

[16] Sweta Alpana, Sagarika Choudhary, Amrit Mukhrjee, Amlan Datta "Analysis of different error correcting methods for high data rates cognitive radio applications based on energy detection technique" , Journal of Advanced Research in Dynamic and Control Systems, Volume:9 | Issue:4.

[17] Pavan D Paikrao, Sanjay L. Nalbalwar, 'Analysis Modification synthesis based Opti-mized Modulation Spectral Subtraction for speech enhancement ',International jour-nal of Circuits Systems and Signal Processing, Vol.11, pg 343-352,2017.