ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Evaluation of Speech Quality Using CDR Technique to Remove Reverberation

Shruti Gangappa Bidari[1], Chandrashekarappa K [2]

PG Student [DEC], Dept. of ECE, Bangalore Institute of Technology, Bengaluru, Karnataka, India[1]

Associate Professor, Dept. of ECE, Bangalore Institute of Technology, Bengaluru, Karnataka, India[2]

ABSTRACT: In this paper, we discuss about CDR Technique to remove background noise, interferences, echo signals. First Signal with noise and reverberation are described. Noise need to be separated using CDR technique. CDR estimators are combined into one frame as post filter to remove reverberation and evaluated by processing reverberant speech ASR recognition accuracy as well as various signal quality measures.CDR estimation is done using geometric analysis as partial removal of error from the speech. As final step spectral subtraction is carried out to remove remaining reverberation speech. The output of the result shows that the unbiased estimator has a practical benefit over present estimators, and proposed DOA-independent estimator used for vanishing reverberation. This all process is carried using special tools known as MATLAB.

KEYWORDS: Dereverberation; Reverberation suppression; diffuse noise suppression; Spatial Coherence.

I.INTRODUCTION

Speech or voice is basic way for humans to convey the information and thoughts. Speech has different frequency for different peoples. Frequency usually ranges from 3Hz to 4 KHz depending on their character. Speech consist of many background noise Clear speech cannot be processed because of noise. The purpose of removing noise is to improve the quality of the voice. Noise must be shrike to using speech enhancement algorithm.

The main objective is to improve the quality by removing noise. Techniques consist of CDR methods. Existing model Signal to reverberant ration (SRR).SRR model are complex and omnidirectional signal outperformer.

Proposed model Coherent to Diffuse power ratio (CDR) is present signal based quality enhancement and automatic speech recognition method. Coherence spectral is identifying relation between two signals. Proposed CDR has advantage over existing estimators. CDR can also be explained as ratio between direct and diffused signal component. The method is carried out in many different stages remove the noise. Partially noise removed with post filtering. Still completed noise is not removed. Depending upon the requirement ratio can be computed frequency dependent or frequency independent.

Initially first sound which is received is through free field. Sound is reflected through wall, speakers and enters the microphone. Microphone signals are combined by squaring magnitude and using phase of one microphone signal. Like that different microphone signal are combine and feed into preprocessing.

Using STFT domain spatial magnitude averaging is reduces the variation of spectral in microphone or post filter. Improve the system performance. Preprocessor are used signal enhancement. Its utilize here with the purpose behind reducing the variety in the exchange function which are caused by destructive and constructive disturbances and calculate the post filter gain and short time estimation of special coherence which have been evaluated direct signal or reverberation coherence, were direct coherence is derived and estimated from TDOA. Spectral magnitude subtraction is calculated for post filter. Desired signal is dependent on TDOA. Reverberation signal is made assumption and both signals are feed back to CDR estimator. CDR estimator signal is feed into post filter gain and preprocessing signal is also post filter gain then magnitude subtraction is done. As final output clear voice is received.

A. Scope and motivation

The work displayed in this paper is persuaded by the quickly developing business sector of discourse interchanges frameworks.. The primary advantage of sans hands phones is that they enable the client to walk around unreservedly without wearing a headset or microphone, and subsequently give a characteristic method for correspondence. Voice-controlled frameworks are, for instance utilized as a part of a working room where they permit specialists and medical attendants to unreservedly move around the patient. The principle advantage of listening device applications is to expand the listening to limit empowering a portable amplifier client to cooperate better with other individuals.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

B. Reverberation in Enclosed Spaces

Reverberation a focal topic of this exposition is instinctively portrayed by the idea of reflections. The desired source produces wave front, which proliferate outward from the source. The wave fronts reflect off the dividers of the room and superimpose at the mouthpiece. In Fig.1 this is shown with a case of an immediate way and a single reflection. Because of contrasts in the lengths of the engendering ways to the microphone and in the measure of sound vitality consumed by the dividers, every wave front touches base at the receiver with an alternate adequacy and stage. The term Reverberation assigns the presence of deferred and constricted duplicates of the source signal in the received signal.

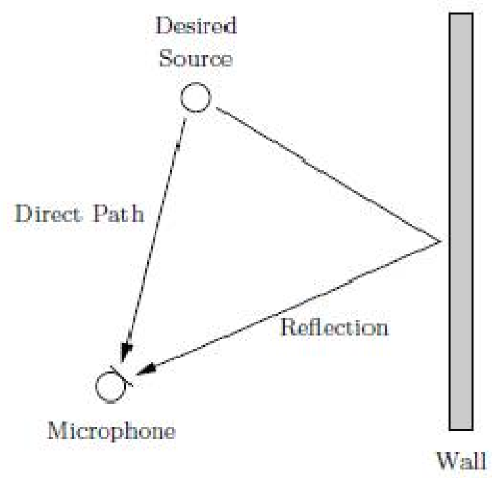

Fig.1. Illustration of the direct path and a single reflection from the desired source to the microphone.

Reverberation can be defined as the process of multi-way propagation of an acoustic sign from its source to the amplifier. The received signal flag for the most part comprises of an direct sound, reflections that arrive not long after the direct sound (generally called early Reverberation), and reflections that touch base after the early Reverberation (regularly called late Reverberation). The blend of the immediate sound and early Reverberation is now and again alluded to as the early stable segment.

Direct Sound: The primary sound that is received through free-field, i.e., without reflection, is the called direct sound. In that the source is not in viewable pathway of the path there is no direct sound. The different between the underlying excitation of the source and its perception is distance on the separation and the speed of the sound.

Early Reverberation: Later the sounds which were reflected off one or more surfaces (dividers, floor, furniture, and so on.) will be received. These reflected sounds are isolated in both time and direction from the direct sound. The reflected sounds shape a sound segment which is generally called early reverberation. Early reverberation will shift as the source moves inside the space, and gives us data about the extent of the space and the position of the source in the space. Reverberation before are additionally causes an otherworldly bending called coloration.

Late Reverberation: Late reverberation results from reflections which land with bigger deferrals after the entry of the direct sound. They are seen either as isolated echoes, or as reverberation, and hinder discourse understand ability.

II. PROBLEM FORMULATION

The recording of an echo signal or noisy speech signal by multi-microphones with a spacing d, located in the same horizontal plane .The signal $m_i(t)$ of the i-th microphone is composed of a desired signal component $s_i(t)$and undesired signal $n_i(t)$ i.e.

$$m_i(t) = s_i(t) + n_i(t) \tag{1}$$

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Short time Fourier transform representation is corresponding uppercase letters, $M_i(l,f)$, $S_i(l,f)$ and $N_i(l,f)$ respectively, where l is discrete-time frame index and f is continuous frequency

Using STFT domain ,the short –time auto and cross power spectra between two signals $x(t)$ and $y(t)$ are given by

$$\varphi_{xy}(l,f) = \varepsilon\{X(l,f)Y^*(l,f)\} \tag{2}$$

Whereε is the expectation operator.

The time-and frequency-dependent signal to noise ratio(SNR) of the microphone signals can be defined as

$$SNR(l,f)=\varphi_s(l,f)/\varphi_n(l,f) \tag{3}$$

It is assumed that signal and noise are mutually orthogonal, such that

$$\varphi_m(l,f) = \varphi_s(l,f) + \varphi_n(l,f) \tag{4}$$

The complex spatial coherence of mixed sound field can then be written as a function of the SNR and the signal and noise coherence functions:

$$\Gamma m(l,f) = \frac{SNR(l,f)\,\Gamma s(f) + \Gamma n(f)}{SNR(l,f)+1} \tag{5}$$

Equation (5) can be rewritten as parametric line equation in the complex plane, highlighting that Γm lies on straight line connecting $\Gamma n$and Γs:

$$\Gamma m(l,f) = \Gamma s(f) + \frac{1}{CDR(l,f)+1} + \left(\Gamma n(f) - \Gamma s(f)\right) \tag{6}$$

Note that the line parameter $D(l,f) =[CDR(l,f) +1]^{-1}$ is equivalent to diffuseness

A. Desired Signal

The desired signal component is modelled as a plane wave with the direction of arrival (DOA) $\theta = 0^{o}$corresponds to broad side direction i.e.

$$\Gamma s(f)= e^{jkd\sin(\theta)} \tag{7}$$

Where wave number $k=2\pi f/c$,the time difference of arrival (TDOA) $\Delta t = d \sin(\theta)/c$, c is speed of sound

B. Reverberation as Isotropic Sound Field

In array signal processing, environmental noise is often modelled by the superposition of an infinite number of uncorrelated, spatially distributed noise sources. The most common assumption for the spatial distribution is a sphere centred on the receiver, which corresponds to diffuse. The spatial coherence function between multi sensors in a diffuse noise field

$$\Gamma diffuse(f) = \frac{\sin(2\pi f d/c)}{2\pi f d/c} \tag{8}$$

III. PROPOSED METHODOLOGY

The project flow and block diagram is clearly shown in Fig.1. Proposed method work can be divided into two parts, Reverberation of input audio and application of resulting is De-reverberation.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

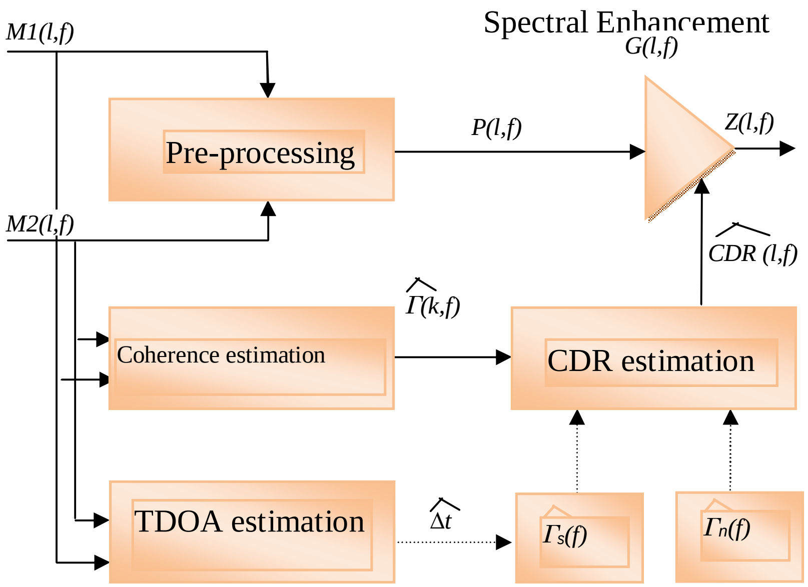

Fig.2. Block diagram of CDR technique to remove reverberation.

The proposed structure of reverberation or noise suppression system is based on CDR gauges. To start with the microphone signals are joined by averaging the squared extent and utilizing the stage from one of the microphone signals:

$$P(l,f) = \frac{\sqrt{|M1(l,f)|^2 + |M1(l,f)|^2}}{2} e^{jargM1(l,f)}$$

Spatial magnitude averaging in the STFT space is being used to decrease the difference of estimator for the calculation of microphone postfilters, yet likewise been utilized as a preprocessor for signal enhancement. It is utilized here with the purpose behind reducing the varieties in the exchange function which are caused by destructive and constructive disturbance of early reflection component by direct path. For the calculation of the coherence based post filter gain $G(l,f)$,short –time estimates $\Gamma_x(l,f)$ of the spatial coherence are initially obtained by spectra which have been evaluated direct signal or reverberation coherence, where the direct coherence is derived and estimated from TDOA. A post filter addition is then computed using spectral magnitude subtraction:

$$G(l,f) = \max\{Gmin, 1 - \sqrt{\frac{\mu}{CDR\widehat{(l,f)} + 1}}$$

With the over subtraction phase μ and the increase floor G_{min}The yield sign is calculated by applying the postfilter and addition to the pre-processed signal $P(l,f)$, i.e.,$Z(l,f)=G(l,f)P(l,f)$, and changed once more into the time space .subsequent to the pre-processor output does not have any spatial separating impact, the postfilter increase can be specifically connected to the pre-processor output ,and does not require a rectification to spatial filtering as it would be the case for beam former as pre-processor.

Note that, while utilizing a DOA-free CDR estimator, the proposed signal improvement system is totally independent of DOA of the target signal.

A. Short time Fourier transform

The brief short time Fourier transform change (STFT), or on the other hand transient Fourier change, is a Fourier-related change used to decide the sinusoidal recurrence and stage substance of nearby segments of a sign as it changes after some time. By and by, the strategy for registering STFTs is to partition a more extended time signal into shorter fragments of equivalent length and after that figure the Fourier change independently on each shorter portion. This uncovers the Fourier range on each shorter portion. One then normally plots the changing spectra as an element of time.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

B. Fast Fourier transform

A fast Fourier transform (FFT) algorithm computes the discrete Fourier transform (DFT) of a sequence, or its inverse. Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa.

Flow chart:

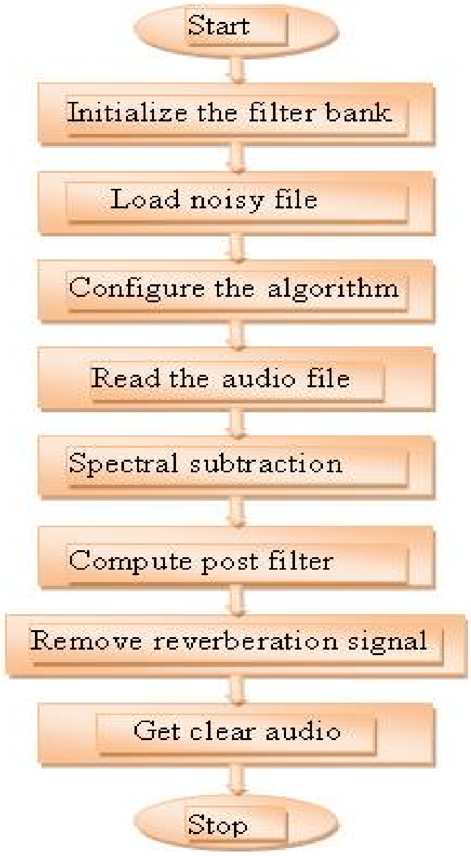

Fig 3: Flow Chart for Dereverberation

First for ours proposing system, input is original signal, which contains information, some noise, and also surrounding noise. Then after it enters its spectrum, here it estimates the distortion present in the original signal. After that estimate It applied to the Post filter that also estimate the accuracy in the distortion percentage, and can be remove the reverberation signal.

The CDR technique is given by the following flowchart.

Step1: Initialize the filter bank those are FFT size, Frame shift, Filter length.

Step2: Load the noisy file using in.mat format.

Step3: Configure the algorithm using sampling rate, speed of sound, Time difference of arrival etc.

Step4: Read the input audio .wav file and then estimating the noise is carried out by taking the Short time Fourier transform.Step5: After reading file spectral subtraction take place.

Step6: Compute the postfilter.

Step7 : The output signal is computed by applying postfilter to pre-processing signal will get speech enhancement .

IV. RESULTS

The proposed algorithms for CDR technique for de-reverberation are programmed in MATLAB environment. Fig.4 shows the Re-sampling signal after creating reverberant audio wave file; this is the input noisy signal. Fig.5(a) shows the

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

noise frame index before evaluating the CDR technique, and Fig.5(b) shows the clean signal after applying the filter gain.

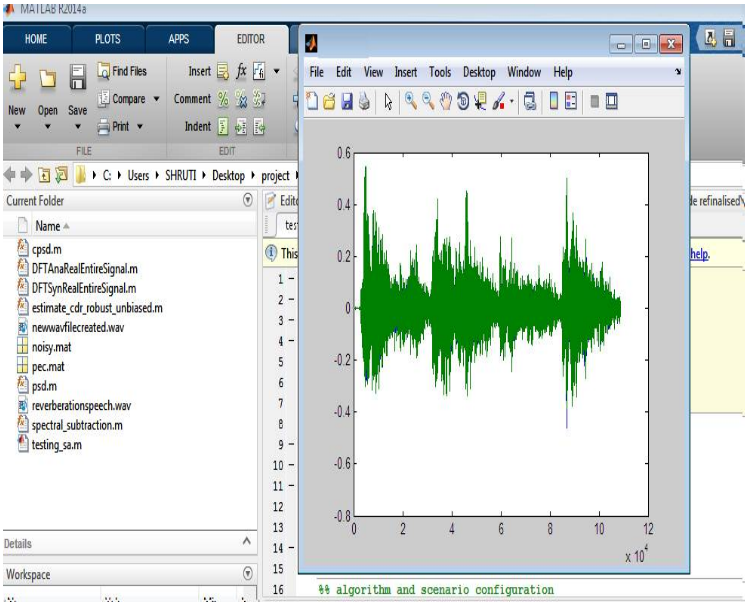

Fig.4 .Re-sampling input audio signal

The following graph shows that the Re-sampling input audio signal .this is the time and frequency dependent. In this graph first we load the noise signal, using this noise signal we create the audio file and after resample that audio file will get above distortion signal.

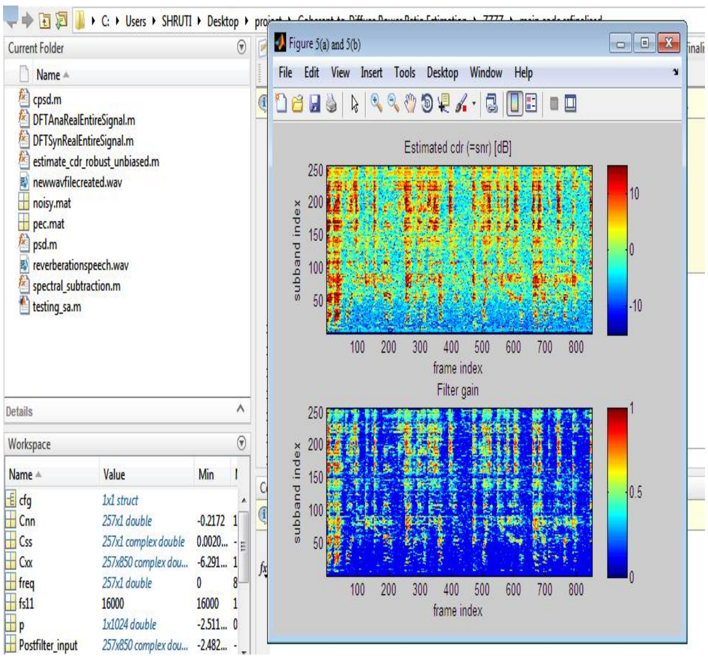

Fig.5(a) The noise frame index before evaluating the CDR technique, and Fig.5(b) shows the clean signal after applying the filter gain.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

The above graph shows that the a) the noise frame index before evaluating CDR technique and b) the clean frame index signal after applying the filter gain. In noise frame index CDR technique it consist more noise before evaluating and there signal to noise ratio is also more and then fig. 5b. has clean signal after evaluating CDR its apply to the technique we get the clean signal.

V. CONCLUSION

The result of CDR estimation techniques and their application to dereverberation have been investigated. In any case, the known CDR estimators were observed to be either one-sided or not sufficiently strong for handy application to signal enhancement. It has been demonstrated that few variations of impartial estimators can be inferred which enhance power towards model errors, and that learning of either the sign DOA or noise coherence is sufficient for estimation of the CDR. Utilizing the enhanced estimators for dereverberation has been appeared to prompt enhanced dereverberation execution. Utilizing the DOA-free estimator, the proposed signal upgrade plan constitutes a totally remove reverberation framework which requires no information or estimation of the sign DOA

ACKNOWLEDGMENT

I acknowledge with the support rendered by Dr.A.G.Nataraj, Principal of BIT, Bangalore. With immense pleasure I express my gratitude to Dr.K.V.Prasad, Head of the Department. Dept of ECE for his motivation for the completion of project, I would like to thank my guide Chandrashekarappa K. Associate professor. Dept of ECE, for guiding me in creating a ground for this project work and being an inspiration to come out successfully with this work. Finally I am grateful to my family and friends for their support during the course of my studies.

REFERENCES

[1] R. Le Bouquin-Jeannes, A. Azirani, and G. Faucon, “Enhancement ofspeech degraded by coherent and incoherent noise using a cross-spectralestimator,” IEEE Trans. Speech and Audio Process., vol. 5, no. 5, pp.484–487, Sep. 1997.

[2] S. Goetze, A. Warzybok, I. Kodrasi, J. O. Jungmann, B. CauchiJ. Rennies, E. A. P. Habets, A. Mertins, T. Gerkmann, S. Doclo, andB. Kollmeier, “A study on speech quality and speech intelligibility measure for quality assessment of single-channel dereverberation algorithms,”in Proc. IWAENC, 2014.

[3] The website [Online]. Available :http://www.wikipedia.com.

[4] MATLAB products help by Mathworks

[5] C.D. Yoo, “Speech enhancement based on the generalized dual excitation model withadaptive analysis window,” in Proc. of the IEEE International Conference on Acoustics,Speech, and Signal Processing (ICASSP’95), 1995, vol. 1, pp. 832–835.

BIOGRAPHY

1. SHRUTI GANGAPPA BIDARI pursuing a Master’s Degree in Digital Electronics and Communication from Bangalore Institute of Technology, Bengaluru.
2.
3. CHANDRASHEKARAPPA is an Associate Professor, Department of Electronics and communication engineering, Bangalore Institute of Technology, Bengaluru.