# **COMPARATIVE PERFORMANCE ANALYSIS OF CHANNEL** NORMALIZATION TECHNIQUES

# Rupa Patel<sup>1</sup>, Aparna Gurjar<sup>2</sup>

<sup>1</sup>Asst. Professor, Department of Computer Application, Ramdeobaba College of Engineering and Management, Nagpur, Maharashtra, India

<sup>2</sup>Asst. Professor, Department of Computer Application, Ramdeobaba College of Engineering and Management, Nagpur, Maharashtra, India

# Abstract

A major part of the interaction between humans takes place via speech communication. The speech signal carries both useful and unwanted information. Processing of such signals involve enhancing the useful information. The intelligibility of speech signals is significantly reduced due to the presence of unwanted information such as noise. Channel normalization algorithms suppress such additive noise introduced in the speech signals by transmission channel or by recording environment conditions. Enhancing the quality and intelligibility of speech signals improve the performance of speech systems such as Automatic speech recognition (ASR), voice communication and hearing aids to name the few. Based on the experimental results the comparative analysis of channel normalization techniques have been presented in this paper to find out the most suitable algorithm for enhancing the speech signals.

\_\_\_\_\_\*\*\*\_\_\_\_\_

Keywords: Cepstral Mean Normalization, Spectral Subtraction, Weiner filter, Signal to Noise Ratio

# **1. INTRODUCTION**

Automatic Speech Recognition (ASR) is a technology that transforms human speech to a symbolic representation. Recognizer performs the transformation with the goal that it can handle spontaneous speech from any speaker in any environment.

The speech waveform produced by a speaker is transmitted over some channel before it reaches the recording device, and the channel disturbs the original speech signal. Most of the channel normalization techniques either deal with channel transfer characteristics or additive noise. Study reveals that different channel normalization techniques have been developed to minimize the effects of channel noises in general on speech recognition systems.

This paper focuses on three channel normalization techniques Cepstral Mean Normalization, Spectral Subtraction and Weiner filtering for reduction of noise in speech signals. The techniques are summarized in table 1.

Table -1:	Channel	Normalization	Techniques
-----------	---------	---------------	------------

a 1 14					
Cepstral Mean	<ul> <li>Noise compensation technique.</li> </ul>				
Normalization	• Also refereed as Channel Mean				
(CIMIN)	Normalization.				
	• Reduce distortion caused by				
	transmission channel.				
	• Applied to cepstral data.				
	• Considered as a high pass filtering				
	process [1].				
	• Based on the observation that a linear				
	channel distortion becomes a				

	constant offset in the cepstral domain
	[2].
	• Cepstral mean value is calculated across the whole speech utterance (combination of cepstral vectors), this calculated cepstral mean is subtracted from each frame single cepstral vector. [3].
Spectral	• Voice activity detector first detects
Subtraction (SS)	whether the frame is noisy or speech
	frame. Noise spectrum is obtained for
	is obtained by subtracting noise
	spectrum from corrupted speech
	spectrum so that signal-to-noise ratio
	(SNR) is improved [4, 5, 6, 7, 8].
Weiner Filter	• Adaptive filter [3].
Weiner Filter	<ul><li>Adaptive filter [3].</li><li>Useful for additive noise reduction</li></ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that SNRpost=SNRprior +1[10].</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that SNRpost=SNRprior +1[10].</li> <li>Noise present in the signal is reduced by accomparison with an actimate of</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that SNRpost=SNRprior +1[10].</li> <li>Noise present in the signal is reduced by comparison with an estimate of the desired clean signal with that of</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that SNRpost=SNRprior +1[10].</li> <li>Noise present in the signal is reduced by comparison with an estimate of the desired clean signal with that of noisy speech signals. [11,12]. This</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that SNRpost=SNRprior +1[10].</li> <li>Noise present in the signal is reduced by comparison with an estimate of the desired clean signal with that of noisy speech signals. [11,12]. This estimate is obtained by minimizing</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that SNRpost=SNRprior +1[10].</li> <li>Noise present in the signal is reduced by comparison with an estimate of the desired clean signal with that of noisy speech signals. [11,12]. This estimate is obtained by minimizing the Mean Square Error (MSE)</li> </ul>
Weiner Filter	<ul> <li>Adaptive filter [3].</li> <li>Useful for additive noise reduction and signal restoration. [3].</li> <li>Based on tracking a priori SNR using Decision-Directed method, proposed by Scalart et al 96 [9].</li> <li>In this method it is assumed that SNRpost=SNRprior +1[10].</li> <li>Noise present in the signal is reduced by comparison with an estimate of the desired clean signal with that of noisy speech signals. [11,12]. This estimate is obtained by minimizing the Mean Square Error (MSE) between the desired signal and the</li> </ul>

#### 2. EXPERIMENTAL RESULTS

The setup to analyze the various algorithms to determine the suitability of the techniques in speech systems has been implemented in MATLAB. Results of seven samples have been discussed here. The seven speech signals are represented as case 1 to case 7 respectively. Testing data is with additive noise.

# **Preprocessing Step**

The speech signals are analyzed in short time segments, referred as analysis frames [13]. Processing of only voiced speech signals is important. The input signals must be first classified as voiced or unvoiced. Voice activity detection (VAD) algorithm detects the presence of human speech in the signal In VAD algorithms, the parameters used for speech detection are based for voiced /unvoiced classification zero-crossing rate and short time energy [14].

Zero-crossing analysis is a simple kind of voice timedomain analysis. Considering audio data as discrete signals Zero crossing is said to occur if successive samples have different algebraic signs. The rate at which zero crossings occur is the measure of the frequency content of a signal. Zero-crossing rate is the number of times that the sample changes the symbols. Zero-crossing rate is a measure of number of times in a given time interval/frame that the amplitude of the speech signals passes through a value of zero. The zero-crossing rate is one of the useful parameter for estimating whether speech is voiced or unvoiced [15].

Energy provides a representation that reflects the amplitude variations. The amplitude of the speech signal varies with time. The amplitude of unvoiced speech segments is much lower than the amplitude of voiced segments. The energy of the speech signal provides a representation that reflects these amplitude variations [16].

To classify input speech signals as voiced or unvoiced signals Voice activity detection algorithm is applied to each of the case samples. The results of pre-processing step are recorded in table 2.

Table -2:	Results	of `	Voice	Activ	ity	Detector

Signals	Energy	ZCR
Case 1	0.60	0.07
Case 2	0.36	0.10
Case 3	0.68	0.07
Case 4	1.02	0.07
Case 5	0.36	0.10
Case 6	0.47	0.08
Case 7	0.43	0.07

All the seven speech signals are classified as voiced speech waveform.

The energy plot and zero crossing plot for case1 is shown in figure below.

For Case 1:



Almost similar energy plot and zero crossing plot for the remaining cases have been obtained.

#### 2.1 Cepstral Mean Normalization

All Cepstral features were mean normalized and normalization scheme were performed on the full utterance. Signal to noise ratio (SNR), which is the measure of signal strength relative to background noise of original voiced speech segment is computed. CMN is applied to it and again SNR is computed. Results are tabulated in table 3 and also in comparative chart 1.

Signal	SNR(db)		
	Original	CMN	Improvement
Case 1	2.4565	33.6335	31.177
Case 2	0.7504	25.1226	24.3722
Case 3	7.1191	35.2533	28.1342
Case 4	6.7261	45.1343	38.4069
Case 5	9.4415	37.9693	28.5278
Case 6	1.2681	27.7092	26.4411
Case 7	3.4327	23.2372	19.8045

 Table -3: Results of CMN



Chart -1: Comparison of SNR before & after CMN

# 2.2 Spectral Subtraction

SNR of original noisy voiced speech segment is computed. Spectral subtraction is applied to it and SNR is again computed. Results are tabulated in table 4 and depicted in chart 2. Spectral subtraction estimates the clean speech spectrum by subtracting the estimated additive noise spectrum from the noisy speech spectrum.

Table -4: Results of Spectral Subtraction

	SNR(db)		
Signal	Original	Spectral	Improvement
		Subtraction	
Case 1	2.4565	14.5997	12.1432
Case 2	0.7504	6.7826	6.0322
Case 3	7.1191	21.6013	14.4822
Case 4	6.7261	24.5745	17.8484
Case 5	9.4415	31.2225	21.781
Case 6	1.268157	5.8731	4.5921
Case 7	3.4327	11.7850	8.3523





Improvement in SNR values indicate that the additive noise have been reduced from the speech signal

### 2.3 Weiner Filer

Weiner filtering is the most basic approach used for reducing the noise from the signal. Signal to noise ratio, is estimated before and after filtering signal. Table 5 and chart 3 shows the results of Weiner filtering.

Table -5: Results of Weiner					
Signal	SNR (dB)				
	Original	Weiner	Improvement		
Case 1	2.4565	24.8676	22.4111		
Case 2	0.7504	6.0803	5.3299		
Case 3	7.1191	23.6743	16.5552		
Case 4	6.7261	23.6149	16.8888		
Case 5	9.4415	23.2383	14.3616		
Case 6	1.268157	12.1107	10.8425		
Case 7	3.4327	16.6735	13.4208		



Chart -3: Comparison of SNR before & after Weiner

From the computed SNR values and comparative chart one can observe that there is an improvement in SNR values which indicates that the noise from the signal have been reduced after the implementation of channel normalization techniques.

# **3. ANALYSIS**

Table 3 and chart 1 summarize the results obtained by using cepstral mean normalization technique. The improvement in SNR values is in the range of 20 db and 30 db which shows that CMN algorithms significantly enhance the signal. The evaluation is repeated with Spectral subtraction and weiner filter. The SNR improvement in case of spectral subtraction is approximately in between 4db to 21db as can be observed from table 4 and chart 2. While in case of weiner filter it lies in the range of 10db to 22 db as can be found from table 5 and chart 3.

Comparative analysis of these techniques is depicted in chart 4.



Chart -4: Improvements in SNR for CMN, Spectral Subtraction and Weiner

From chart 4 one can observe that the results obtained for Spectral subtraction and Weiner are pretty much similar and findings shows that CMN gives better result.

Cepstral mean normalization enhances the original signal approximately by 31%, spectral subtraction by 13% and weiner by 15%. These clearly indicate that Cepstral Mean Normalization technique is useful in reducing the impact of noise as compared to two other mentioned algorithms.

# 4. CONCLUSION

The effectiveness of different normalization techniques has been evaluated and results obtained have been summarized. From the results it is clear that the channel normalization technique cepstral mean normalization reduce distortion and proved to be effective

# REFERENCES

[1]. Stern, Richard M., Bhiksha Raj, and Pedro J. Moreno. "Compensation for environmental degradation in automatic speech recognition." Robust Speech Recognition for Unknown Communication Channels. 1997.

[2]. Garner, Philip N. "Cepstral normalization and the signal to noise ratio spectrum in automatic speech recognition." Speech Communication 53.8 (2011): 991-1001.

[3]. http://recognize-speech.com/preprocessing/10preprocessing.

[4]. N.S. Mahlanyane and D.J. Mashao, "Channel Normalization for GSM Speech Recognition", SATNAC conference paper, UCT, 2003.

[5]. J.S. Lim, A.V. Oppenheim "Enhancement and bandwidth compression of noisy speech ",Proc IEEE, vol 67, pp 1586-1604, December 1979.

[6]. S. Boll, "Suppression of Acoustic Noise in Speech using Spectral Subtraction," IEEE Trans. Acoust., Speech, Signal Process., vol.27, pp. 113-120, Apr. 1979.

[7]. http://sound.eti.pg.gda.pl/denoise/noise.html

[8]. Kim, Chanwoo. Signal processing for robust speech recognition motivated by auditory processing. Diss. Johns Hopkins University, 2010.

[9]. Scalart, Pascal. "Speech enhancement based on a priori signal to noise estimation." Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on. Vol. 2. IEEE, 1996. [10]. www.mathworks.com

[11]. Md. Jahangir Alam, Md. Faqrul Alam Chowdhury, Md. Fasiul Alam, "Comparative Study of a Priori Signal -To- Noise Ratio(SNR) Estimation Approaches for Speech Enhancement", Journal of electrical and electronics engineering, 2009,vol 1

[12]. Rupa Patel, Urmila Shrawankar, Dr. V.M.Thakare ,"Hiding speaker Characteristics for Security", ICCCNT'12, 26th -28th July 2012, Coimbatore

[13]. L. R. Rabiner, R.W. Schafer ," Digital processing of speech signals", Prentice-Hall, Inc., Reprint 2009

[14]. Wei Li et.al, "Voice-Based Recognition System for Non-Semantics Information by Language and Gender", Third International Symposium on Electronic Commerce and Security,2010

[15].http://en.wikipedia.org/wiki/Audio\_normalization#Lou dness\_normalization

[16]. Bachu, R. G., et al. "Separation of Voiced and Unvoiced using Zero crossing rate and Energy of the Speech Signal." American Society for Engineering Education (ASEE) Zone Conference Proceedings. 2008.