IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)
e-ISSN: 2278-1676,p-ISSN: 2320-3331, Volume 10, Issue 3 Ver. III (May – Jun. 2015), PP 45-48
www.iosrjournals.org

Speech Enhancement using Signal Subspace Algorithm

Ajay Kaliraman[1], Summi Khurana[2], Dr. Hardeep Saini[3]
[1,2 and 3]*Punjab Technical University, Jalandhar Punjab INDIA*

Abstract: *In speech communication, quality and intelligibility of speech is of utmost importance for ease and accuracy of information exchange. The speech processing systems used to communicate or store speech are usually designed for a noise free environment but in a real-world environment, the presence of background interference in the form of additive background noise and channel noise drastically degrades the performance of these systems, causing inaccurate information exchange and listener fatigue. Speech enhancement algorithms attempt to improve the performance of communication systems when their input or output signals are corrupted by noise. Speech Enhancement in general has three major objectives: (a) To improve the perceptual aspects such as quality and intelligibility of the processed speech i.e. to make it sound better or clearer to the human listener; (b) to improve the robustness of the speech coders which tend to be severely affected by presence of noise; and (c) to increase the accuracy of speech recognition systems operating in less than ideal locations.*

I. Introduction

Several techniques have been proposed for this purpose. The most commonly used methods for Speech Enhancement are spectral subtraction method, Iterative Wiener filtering, Kalman filtering, Linear Predictive coding (LPC) analysis, Signal Subspace method. The performances of these techniques depend on the quality and intelligibility of the processed speech signal. The improvement in the speech signal-to-noise ratio (SNR) is the target of most techniques.

The basic method for speech enhancement is Spectral Subtraction approach .It is very simple method and easy to implement. The conventional power spectral subtraction method substantially reduces the noise levels in the noisy speech. But it introduces an annoying distortion in the speech signal called musical noise. With the passage of time

Spectral Subtraction has undergone many modifications. Evaluation of spectral subtractive algorithms revealed that these algorithms improve speech quality and not affect much more on intelligibility of speech signal. So the other method called Signal Subspace method [6] is proposed that allows better and more suppression of the noise. The aim of this method is to improve the quality, while minimising any loss in intelligibility.

II. Related Works

Yong Xu, Jun Du, Li-Rong Dai, and Chin-Hui Lee; (2014) "An Experimental Study on Speech Enhancement Based on Deep Neural Networks," Here we presents a regression-based speech enhancement framework using deep neural networks (DNNs) with a multiple-layer deep architecture. In the DNN learning process a large training set ensures a powerful modeling capability to estimate the complicated nonlinear mapping from observed noisy speech to desired clean signals.

Shenoy, R.R. ; Seelamantula, C.S.(2014) Frequency domain linear prediction based on temporal analysis. This states that Frequency-domain linear prediction (FDLP) is widely used in speech coding for modeling envelopes of transients signals, such as voiced and unvoiced stops, plosives, etc.

Ante Juki´c1, Toon van Waterschoot2, Timo Gerkmann1, Simon Doclo1(2014) Speech Dereverbersation with Multi-Channels Linear Prediction And Sparse Priors For The Desired SIGNAL. This states that the quality of recorded speech signals can be substantially affected by room reverberation. In this paper we focus on a blind method for speech dereverberation based on the multi-channel linear prediction model in the short-time Fourier domain, where the parameters of the model are estimated using a maximum-likelihood procedure.

Kumar, N. ; Van Segbroeck, M. ; Audhkhasi, K. ; Drotar, P. ;Narayanan, S.S. (2014) Fusion of diverse de-noising systems for robust automatic speech recognition. This states that we present a framework for combining different de-noising front-ends for robust speech enhancement for recognition in noisy conditions.

Peddinti, V. ; Hermansky, H.(2013) Filter-bank optimization for Frequency Domain This states that The sub-band Frequency Domain Linear Prediction (FDLP) technique estimates autoregressive models of Hilbert envelopes of sub band signals, from segments of discrete cosine transform (DCT) of a speech signal, using windows. Shapes of the windows and their positions on the cosine transform of the signal determine implied filtering of the signal.

Amro, I. Higher Compression Rates for Code Excited Linear Prediction Coding Using Lossless Compression (2013).This states that In this paper, we exploit the Hamming Correction Code Compressor (HCDC) Code Excited Linear Prediction frame's Parameter. These parameters includes Linear Prediction Coefficients, Gain and Excitation Bits; which is DCT residual for the signal frame, consist of 40 coefficients, each is quantized using 4 bits. For the signals used in experiments; the total bits in frame were 261 bits with transmission rate 0f 5.22 Kbps.

Meng Guo ; Elmedyb, T.B. ; Jensen, S.H. ; Jensen, J.(2010) Analysis of adaptive feedback and echo cancelation algorithms in a general multiple-microphone and single-loudspeaker system. This states that In this paper, we analyze a general multiple-microphone and single-loudspeaker system, where an adaptive algorithm is used to cancel acoustic feedback/echo and a beam former processes the feedback/echo canceled signals.

Wen Jin Xin Liu ; Scordilis, M.S. ; Lu Han (2009) Speech Enhancement Using Harmonic Emphasisand Adaptive Comb Filtering. This states that An enhancement method for single-channel speech degraded by additive noise is proposed. A spectral weighting function is derived by constrained optimization to suppress noise in the frequency domain. Two design parameters are included in the suppression gain, namely, the frequency-dependent noise-flooring parameter (FDNFP) and the gain factor. The FDNFP

Zhimin Xiang and Yuantao GuAdaptive; (2013) " Speech Enhancement Using Sparse Prior Information," In recent years, sparse representation is adopted to improve the quality of noise corrupted speech. However, the representation of noise is also found to be sparse in some special cases, which degrades the performance of sparsity based speech enhancement. An adaptive speech enhancement algorithm using sparse prior information is planned.

III. Signal Subspace Algorithm

Subspace methods, also known as high-resolution methods or super-resolution methods, generate frequency component estimates for a signal based on an eigen analysis or eigen decomposition of the correlation matrix. Examples are the multiple signal classification (MUSIC) method or the eigenvector (EV) method. These methods are best suited for line spectra — that is, spectra of sinusoidal signals — and are effective in the detection of sinusoids buried in noise, especially when the signal to noise ratios are low.

Vector spaces may be formed from subsets of other vectors spaces. These are called subspaces. A subspace of a vector space V is a subset H of V that has three properties:

a. The zero vector of V is in H.
b. For each u and v are in H, u _ v is in H. (In this case we say H is closed under vector addition.)
c. For each u in H and each scalar c, cu is in H. i.e. H is closed under scalar multiplication.

Let s(k) represent the clean-speech samples and let n(k) be the zero-mean, additive white noise distortion that is assumed to be uncorrelated with the clean speech. The observed noisy speech x(k) is then given by

$$x(k) = s(k) + n(k)$$

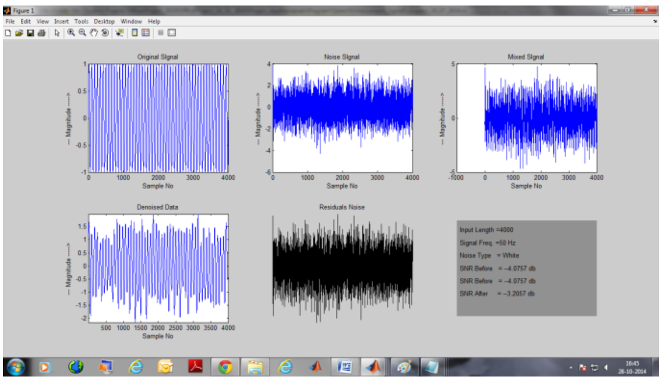
Further, let Rx, Rs, and Rn be ($q \times q$) (with $q > p$) true autocorrelation matrices of x(k), s(k), and n(k), respectively. Due to the assumption of uncorrelated speech and noise, it is clear that

$$Rx = \bar{}Rs + \bar{}Rn.$$

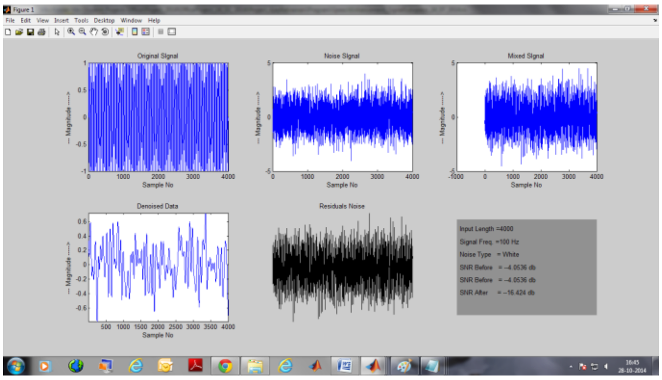
Regardless of the specific optimisation criterion, speech
enhancement is now obtained by

(1 restricting the enhanced speech to occupy solely the signal subspace by nulling its components in the noise subspace,

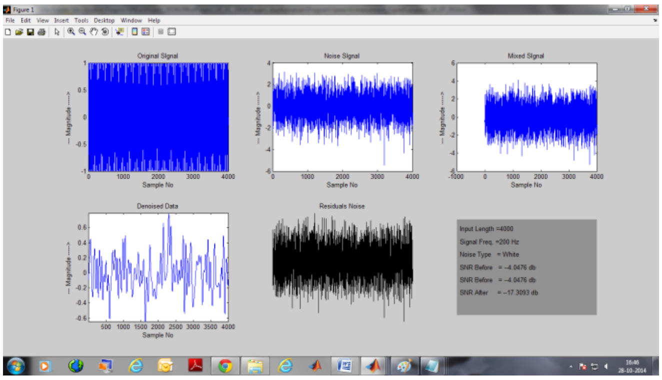
(2) changing (i.e., lowering) the eigenvalues that correspond to the signal subspace.



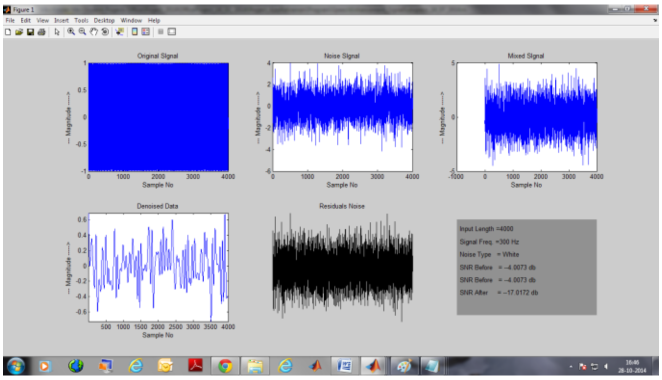
Input Signal at 50 Hz Freq.



Input Signal at 100 Hz Freq.



Input Signal at 200 Hz Freq.



Input Signal at 300 Hz Freq.

IV. Results and Conclusion

S. No.	Input Signal Freq.	Type of Input Signal	Gaussian Noise	MSE	SNR
1.	50 Hz	Sine Wave	0.25*Random()	24.15	78.24
			0.50*Random()	23.25	75.32
			0.75*Random()	21.56	72.54
			1.0*Random()	20.84	70.81

S. No.	Input Signal Freq.	Type of Input Signal	Gaussian Noise	MSE	SNR
1.	100 Hz	Sine Wave	0.25*Random()	12.25	51.26
			0.50*Random()	13.54	52.24
			0.75*Random()	11.65	51.06
			1.0*Random()	10.58	50.41

S. No.	Input Signal Freq.	Type of Input Signal	Gaussian Noise	MSE	SNR
1.	200 Hz	Sine Wave	0.25*Random()	11.21	48.21
			0.50*Random()	10.53	47.25
			0.75*Random()	11.81	46.98
			1.0*Random()	10.67	45.32

V. Results and Conclusion

The result table shows fair improvements in peak signal to noise ratio in the speech signal. The algorithm has been tested on different frequency signals and with different levels of noise in the signal. At each level, the psnr is in improved as compared to the noisy signal. This shows that the noise is removed to a great extent from the noisy speech signal.

References

[1]. Anis Ben Aicha", *IEEE* "Adaptive Noise Smoothing for Speech Enhancement Filters**",** vol. 27, no. 2, pp. 113-120, Apr. 1979..enhancing speech corrupted by coloured noise", in *Proc. ICASSP2002*, vol. 4, pp. IV-4164, May 2002.

[2]. Ante Juki´c1, Toon van Waterschoot2, Timo Gerkmann1, Simon Doclo1(2014), "Speech Dereverbersation with Multi-Channels Linear Prediction And Sparse Priors For The Desired SIGNAL", Hands-free Speech Communication and Microphone Arrays (HSCMA), 2014 4th Joint Workshop pp.23-26.

[3]. Amro, I.(2013), "Higher Compression Rates for Code Excited Linear Prediction Coding Using Lossless Compression", DOI: 10.1109/ICCE-Berlin.2013.6698057 IEEE Third International Conference.

[4]. A. Rezayee and S. Gazor, "An adaptive KLT approach for speech enhancement", *IEEE* Transactions on Speech and Audio Processing, vol. 9, no. 2, pp. 87-95, Feb. 2001.

[5]. C. L. Philipos, "Speech enhancement: theory and practice", Taylor and Francis, 2007.

[6]. C. D. Sigg, T. Dikk and J. M. Buhmann, "Speech enhancement with sparse coding in learned dictionaries", in *Proc.* ICASSP2010, pp. 4758-4761, March 2010.

[7]. D. Giacobello, M. G. Christensen, M. N. Murthi, S. H. Jensen and M. Moonen, "Retrieving sparse patterns using a compressed sensing framework: applications to speech coding based on sparse linear prediction", IEEE Signal Processing Letters, vol. 17, no. 1, pp. 103-106, Jan. 2010.

[8]. D. Giacobello, M. G. Christensen, M. N. Murthi, S. H. Jensen and M. Moonen, "Sparse linear prediction and its applications to speech processing", IEEE Transaction on Acoustic, Speech and Signal Processing, vol. 20, no. 5, pp. 1644-1657, July 2012.

[9]. H. Yongjun, H. Jiqing, D. Shiwen, Z. Tieran and Z. Guibin, "A solution to residual noise in speech denoising with sparse representation", in Proc. ICASSP2012, pp. 4653-4656, March 2012.

[10]. H. Feng, L. Tan and W. B. Kleijn, "Transform-domain wiener filter for speech periodicity enhancement",in *Proc.* ICASSP2012, pp. 4577-4580, March 2012.

Author Profile
S pursuing

The author [1] is pursuing his M.Tech. in ECE from, Punjab Technical University, Punjab India. His field of interest is in signal processing and application system designing.