Proceedings of the first biennial UMaT International Conference on Mining & Mineral Processing, "Expanding the Frontiers of Mining Technology", Tarkwa, Ghana, 4th – 7th August, 2010.

Limitations of the Spectral Subtraction Method on Noise Reduction

Olufeagba, B. J. *, Oyetunji, S. A. ** and Alimi, I. A. ***

Contact e-mail: lobojay12@yahoo.com

Abstract

Noise reduction has become very practicable with the development of sophisticated special purpose microcomputers and signal processors. As a result, many applications, such as mobile communication, speech recognition, military voice communications, broadcast link enhancement and restoration of musical recordings, very efficient noise reduction techniques are needed.

A useful noise reduction algorithm is the spectral subtraction method which under some restrictions can extract an improved facsimile of the original signal by computing the spectrum of the noise and subtracting it from the combined corrupted signal spectrum.

In this paper, the predicted performance of noise reduction using the spectral subtraction technique was investigated. During the study, signals were corrupted with channel noise derived from various sources and at different signal to noise ratios. The technique was applied to reduce the noise and found to be effective although the improvement of signal to noise ratio was found to be limited to 15 dB.

*, ** Federal University of Technology Akure, Ondo State, Nigeria

*** Federal Radio Corporation of Nigeria (FRCN), Akure, Ondo State, Nigeria

1. Introduction

The spectral subtraction method is a frequency domain technique which operates under the assumption that the spectrum of a corrupted input signal can be expressed as the sum of the speech spectrum and the noise spectrum (Malca and Wulich 1996, Dominic, 2002). It then reduces stationary noise added to speech by subtracting the spectral noise calculated during non-speech activity.

The most important requirement is to be able to evaluate the spectrum of two signals in near real-time in order to exploit the apparently simple relationship that akin to stripping the noise from the signal. The other requirement is finding a suitable sample from which to determine the properties of the noise. This process places one of the limits on the extent to which the two signals can be separated especially since the noise estimate will always be over or under the value of the true noise level no matter the speech/pause detecting method adopted.

2. Elements Of The Spectral Subtraction Method

Spectral Subtraction (SS) is a technique used to reduce the effect of noise acoustically added to speech signals. The application here relies on the fact that the spectrum of the input signal can be expressed as the sum of the speech spectrum and the noise spectrum (Halevy, 2002). The noise sources are assumed to be white. To implement SS, the noise spectrum is estimated from regions that can be referred to as "noise-only" and subtracting the result from the noisy speech signal.

If it can be assumed that the noise properties remains relatively invariant during speech activity, the block diagram of figure 1 outlines a possible procedure process for removing noise from noisy signal (Boll, 1979). The evaluation of the spectra employs the fast Fourier transform computations to estimate them as follows.

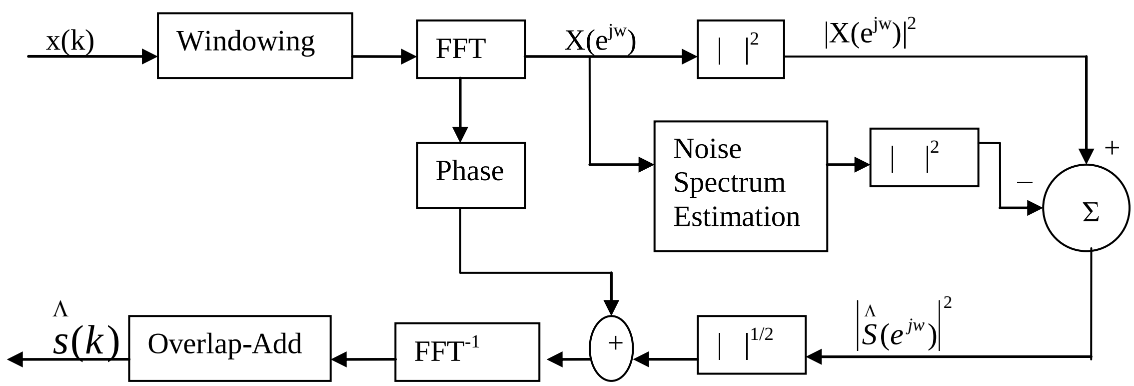

Fig. 1: Block Diagram of Spectral Subtraction Technique

The noisy speech signal is windowed, such that the sums of these windowed sequences add up to the original (Ykhlef, et al, 2002). Let *s(k)* and *n(k)* be represented by a windowed speech signal and noise signal respectively (Baher, 1990).. The sum of the two denoted by *x(k)* can be expressed as follows:

$$x(k) = s(k) + n(k). \qquad (2.1)$$

Taking the Fourier Transform of both sides gives

$$X(e^{j\omega}) = S(e^{j\omega}) + N(e^{j\omega}) \qquad (2.2)$$

where :

$$x(k) \leftrightarrow X(e^{j\omega})$$

$$X(e^{j\omega}) = \sum_{k=0}^{L-1} x(k)e^{-j\omega k}$$

$$x(k) = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(e^{j\omega})e^{j\omega k}d\omega. \tag{2.3}$$

If the magnitude of $N(e^{j\omega})$ is found and replaced by its average value $\mu(e^{j\omega})$ where the latter is the signal taken during the regions estimated as "noise-only" (George, (2005)) the phase $\theta_N(e^{j\omega})$ of $N(e^{j\omega})$ can be replaced by the phase $\theta_X(e^{j\omega})$ of $X(e^{j\omega})$. The spectral subtraction estimator $\hat{S}(e^{j\omega})$ then becomes (Evans, et al, 2002):

$$\left|\hat{S}(e^{jw})\right|^2 \approx \left[\left|X(e^{jw})\right|^2 - \left|\mu(e^{jw})\right|^2\right]e^{j\theta_x(e^{jw})} \tag{2.4}$$

The resulting error from this estimator is:

$$\varepsilon(e^{j\omega}) = \hat{S}(e^{j\omega}) - S(e^{j\omega}) = N(e^{j\omega}) - \mu(e^{j\omega})e^{j\theta_x}. \tag{2.5}$$

$$\hat{S}(e^{jw}) = \left|\hat{S}(e^{jw})\right| \angle X(e^{jw}) \tag{2.6}$$

$$\hat{s}(k) = F^{-1}\left\{\hat{S}(e^{jw})\right\} \tag{2.7}$$

The subtraction process can cause some anomalies such as those arising when the spectrum becomes negative. This is usually treated by half-wave rectification, which may introduce "musical" tones artifacts in the processed signal, and can be expressed by equation *(2.8)*.

$$\left|\hat{S}(e^{jw})\right| = \begin{cases} \left|\hat{S}(e^{jw})\right| & if \left|\hat{S}(e^{jw})\right| \rangle 0 \\ 0 & elsewhere \end{cases} \tag{2.8}$$

After subtraction of noise spectral and half-wave rectification (where necessary), the inverse transform is taken for each window and overlap added to reconstitute the output speech sequence (Berouti, et al., (1979).

3. Experimental Analysis And Results

During the course of the study, a noisy broadcast signal was digitized and captured onto a computer and the data was then imported into the MATLAB Software (environment).The captured signal was studied to establish important characteristics including the spectral features of both the signal and the channel noise. The method adopted relies on information about the properties of the noise which was captured during pauses in speech. Having thus determined the statistical properties of the channel noise, a spectral subtraction technique was adopted.

Studies of the channel noise indicated that it was very close to synthesized noise obtained from MATLAB and white noise obtained from web sources

To test the procedure, a clean speech signal was added to a noise signal then captured using the sound card of an MMX enabled PC. The results were then imported into the MATLAB software for study (MATLAB, 2002).

Their statistical properties were studied and are as shown in table 1.

Table 1: Statistical Properties of the Clean Signal

Signal Type	Means	Standard Deviations(σ)
Clean Signal	-2.695e-005	0.155
Noise Signal (VAI=0)	-5.175e-005	0.121

Parts of the imported signals were subjected to spectral subtraction in order to assess its effectiveness for speech enhancement. It was observed that the results conform reasonable well with theoretical results as the statistical properties (mean and standard deviation) of the restored signal were similar to that of the clean test signal as displayed in table 2. Figure 2 shows plots of noise signal (2a), clean signal (2b), composite signal (2c) and the restored signal (2d).

The Autocorrelation function of an uncorrupted signal (3a) and restored signal (3b) were determined and are shown in Figure 3. It was observed that they conform to theoretical results.

Table 2: Statistical Properties of the Analyzed Clean Signal

Signal Type	Means	Standard Deviations(σ)
Part of Clean Signal	6.5331e-005	0.146
Part of Noise Signal	-7.9750e-004	0.122
Composite Signal	-7.4235e-004	0.189
Restored Signal	5.7518e-005	0.146

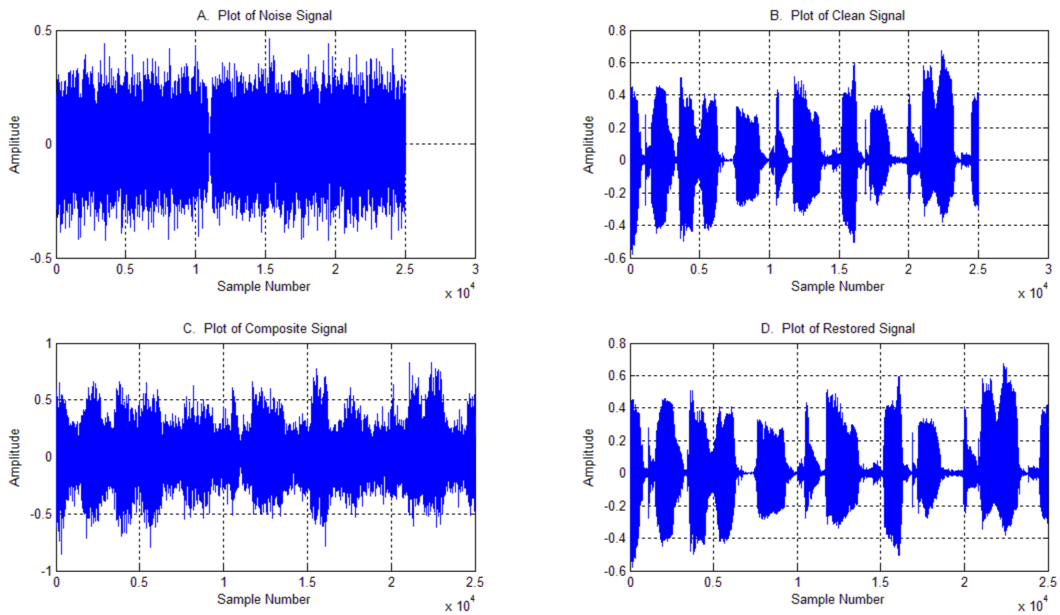

Fig. 2: Plots of Clean Signal and Noise Signal

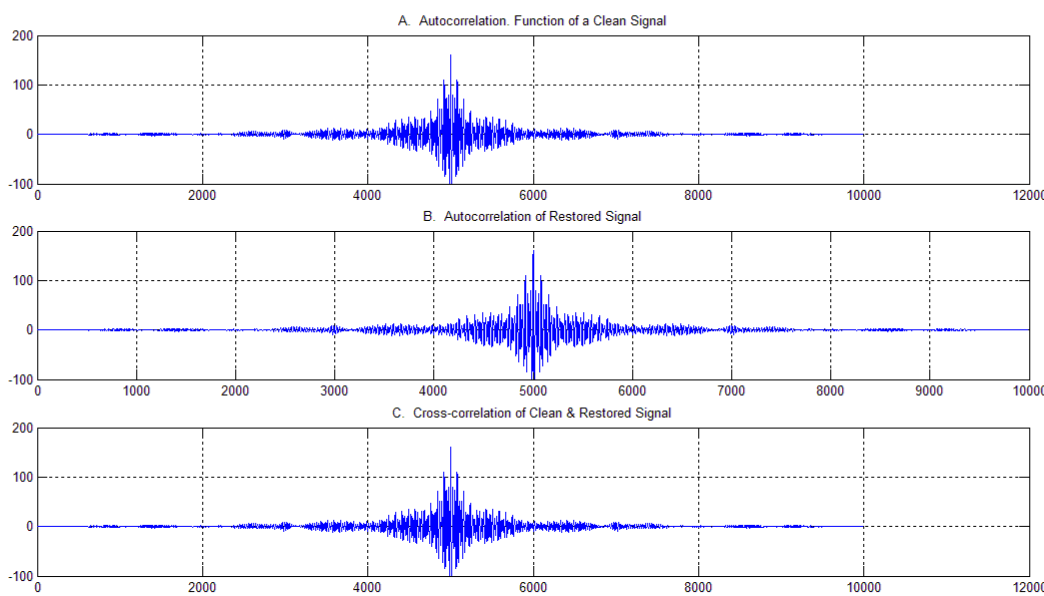

Fig. 3: Autocorrelation and Cross-Correlation of Clean Signal and Noise Signal

The noise signal was added to the clean signal and the amplitude of the clean signal was varied in other to simulate different signal to noise ratio conditions and spectra of signals were determined. It was observed that as the signals magnitudes were decreased the noise became more prominent. The results are as shown in table 3 with the corresponding plots (5a, 5c, 5e, 5g, 6a, 6c, 6e, 6g) with the spectra (5b, 5d, 5f, 5h, 6b, 6d, 6f, 6h) are shown in figure 5 and 6 respectively.

Table 3: Varying levels of the Clean Signal

Signal Type	Mean	Standard Deviation (σ)	Variance (σ^2)	Signal to Noise Ratio $(SNR)_{i/p}$ dB	Signal to Noise Ratio $(SNR)_{o/p}$ dB
Noise when VAI=0	-7.9750e-004	0.1216	0.0148	-	-
Signal (x)	6.5331e-005	0.1456	0.0212	1.5649	15.8109
x/4	1.6333e-005	0.0364	0.0013	-10.4763	3.7697
x/16	4.0832e-006	0.0091	8.2805e-005	-22.5175	-8.2715
x/32	2.0416e-006	0.0045	2.0701e-005	-28.5381	-14.2921

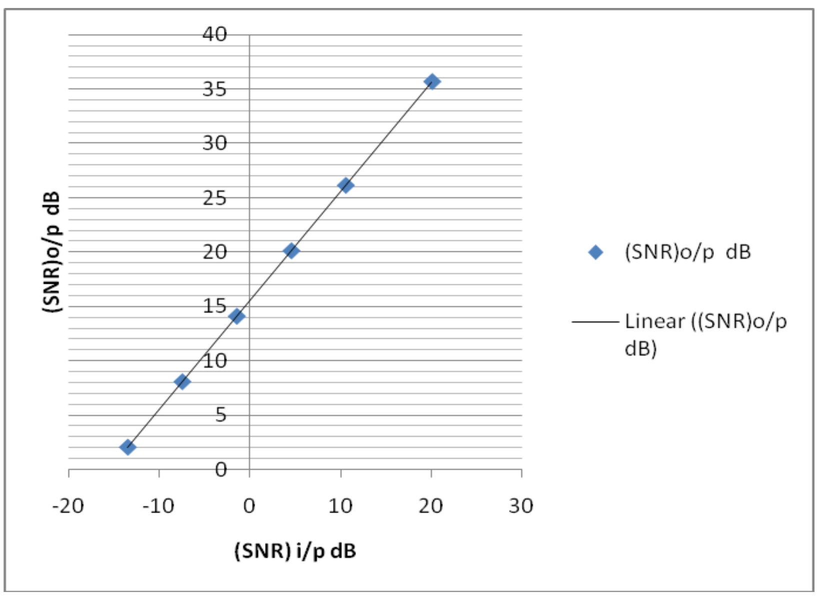

Fig. 4: Plots of Signal to Noise Ratio of Clean Signals

Table 3 implies that if a good noise signal is available, the output of the SST can give an improved signal. The input and output signal to noise ratio in table 3 can be represented graphically by figure 4. The relationship between the signal to noise ratios can be seen to be quite linear with the intersection on the Signal to Noise ratio axis corresponding to an improvement of almost 15dB. The same result is evident from the table where the difference between the ratios is close to 15 dB.

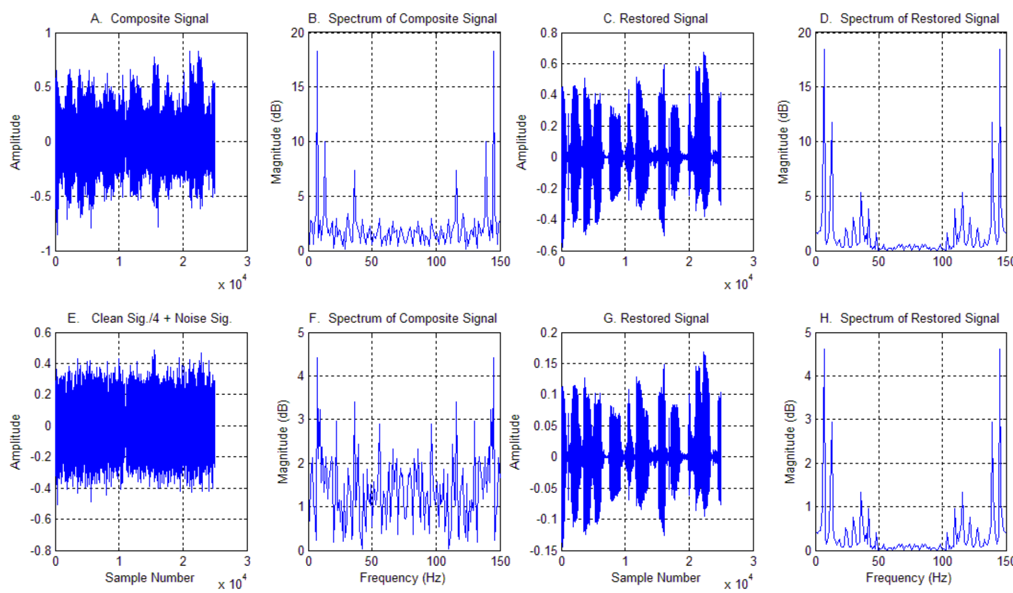

Fig. 5: Plots and Spectra of different levels of Composite Signals

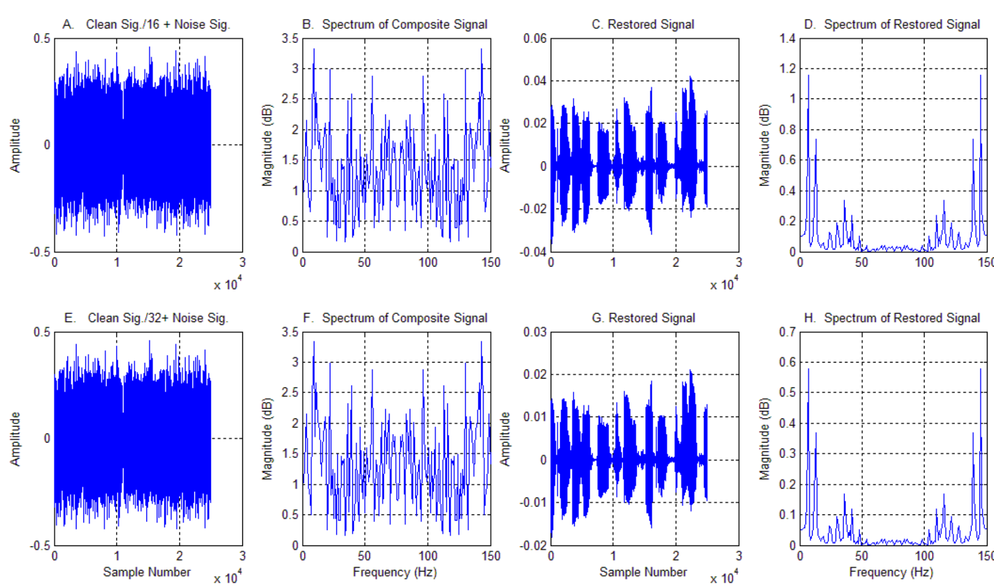

Fig. 6: Plots and Spectra of different levels of Composite Signals

4. Discussion And Conclusions

The performance of the spectral subtraction method is limited by the necessity to estimate the noise characteristics obtained during speech pauses. The accuracy of noise estimation depends on the performance of a voice activity detector (VAD) which is best automatically controlled..
The resulting estimation errors with the respect to the PSD of the noisy signal and the noise PSD result in the estimated magnitude spectrum consisting of a succession of randomly spaced spectral peaks (short duration narrow band energy). The most significant outcome of this work is the rather interesting result that the method can yield improvements of up to 30 dB which can be very useful in many applications. The problems at very low signal levels were not studied here but are known to introduce extraneous musical tones which are still to be fully dealt with. In many applications the improvement realized can certainly make a very marked difference in the quality of sound obtained.

5. References

Baher, H. (1990). *Analog and Digital Signal Processing.* 1st ed. New York, John Wiley & Sons, pp. 435.

Berouti, M. Schwartz, R. and Makhoul, J. (1979). *Enhancement of Speech Corrupted by Acoustic Noise. Proc ICASSP*, pp. 208-211.

Boll, S. F. (1979). Suppression of Acoustic Noise in Speech using Spectral Subtraction", *IEEE Trans ASSP.27(2),* pp. 113-120.

Dominic, K. C. (2002). *Speech Enhancement: Concepts and Methodology.* Technical paper No. 1. Columbia, University of Missouri, pp. 7-9.

Evans, W. D., Mason, J. S., Liu W. M. and Fauve B.(2002) *On the Fundamental Limitations of Spectral Subtraction*: An Assessment by Automatic Speech Recognition. School of Engineering, University of Wales Swansea, pp1

George A. S. (2005). *Using Spectral Subtraction to Enhance Speech and Increase Performance in Automatic Speech Recognition.* Technical report. University of Maryland, Institute of Systems Research & Electrical & Computer Eng. Department, pp. 4-6.

Halevy, A. (2002). *Speech Enhancement: Reduction of Additive Noise in the Digital Processing of Speech (Unpublished M.Sc Thesis)* University of Maryland, Department of Mathematics, pp. 2.

Malca Y. and Wulich D. (1996). *Improved Spectral Subtraction for Speech Enhancement*. Technical report No. 1. University of the Negev, Department of Electrical & Computer Engineering, pp. 1-2.

MATLAB (2002) *Using MATLAB*, MathWork Inc, USA.

Ykhlef, F. Guessoum, A. and Berkani, D. (2002). *Speech Enhancement Based on a Combination of Spectral Subtraction and a Minimum Mean-Square Error Short-Time Log-Spectral Amplitude Estimator in Wavelet Domain*. Technical report. University of Blida, Department of Electronics, Faculty of Engineering Sciences, pp1-2.