ACOUSTICS 2012

# Alleviation of uniform ambient noise in underwater acoustic communication receivers using spectral subtraction

S. Kamal, S. Menonkattil Hariharan and S. P.R. Pillai

Department of Electronics, Cochin University of Science & Technology, 682022 Ernakulam, India
supriyadoe@gmail.com

The acoustic medium of the ocean is exceedingly turbid with a multitude of simultaneously active noise sources like water currents, shipping noise, drilling, construction, off-shore oil rigs, marine life, shore waves and other hydrodynamic activities. These inherent mechanisms produce spectrally overlapping acoustic signals that radiate into the depths of the ocean, and the superposition of these signals from across a large area of the sea cumulatively constitutes a spatially homogeneous ambient noise field. It is often difficult to utilize such a cluttered and exacerbated acoustic channel for communication or other observational purposes. In this paper a method based on spectral subtraction, for alleviating the acoustic ambient noise with underwater acoustic receivers is presented. Spectral subtraction is a technique used to retrieve the power spectrum of a non- stationary signal of interest in a uniform noise field by subtracting an estimated average of noise spectrum from the originally observed signal spectrum. The noise spectrum is estimated and updated during each iteration, while the signal of interest is absent or silent as well as at intentionally put blanking periods in communication signals.

# 1 Introduction

In the realms of the ocean, the absence of light in few meters of depth, makes hearing more important than sight. This results in marine life as well as manmade systems utilizing acoustic energy for observation and communication in ocean depths. During the last few decades, the acoustic ambient noise level of the ocean medium has been increased multifold in low frequencies and mid-frequencies mostly due to the surface and sub-surface human activities. The anthropogenic noise, emanated from shipping, pile driving and construction, sonar and related research equipment, underwater explosive blasting, offshore oil exploration and production are becoming pervasive, contributing severely to the existing background noise level of the ocean. [1]. The quality of the acoustic channel is cumulatively deteriorated severely by these unwanted signals. The rise in the ambient noise level has thus made it increasingly difficult to make acoustic observations, measurements and communication through the ocean medium.

Despite all the good engineering efforts made to minimize the effects, a minimum level of background noise will continue to remain in every receiver system. The acoustic ambience of the receiver has a primary influence on these remnants. The noise margin of a receiver system is defined by the amount by which a signal exceeds the minimum amount for proper operation. The sensitivity of the acoustic receiver system, is limited by the inherent noise source, both internal and external, that always contribute to the observations and also constraints the minimum value of detectable signals. The adjacency of the immense noise field originated from unwanted background sources lifts the receiver noise floor up, so that most of the interesting signals obliterate or are buried within the noise. While there is no signal activity in the channel, the receiver noise floor will be limited by the magnitude of ambient noise and in the best case, limited by the receiver electronics. The total sound field is, by the principle of superposition, a linear combination of pressure fields due to individual sources. The signals picked up by the hydrophones in the receiver frontend are a complex mixture of such independent sources, present in the acoustic scene. In underwater acoustic receivers, the noise floor is often determined by the level of ambient noise rather than the system electronics. Because of this complex manifolds it is often found much difficulty in filtering the noise from the observations. A generic hierarchical overview of the scenario is depicted in Figure 1.

Although several techniques are being developed for curing the ill effects of the noise in underwater acoustic receivers, multi-sensor beamforming is the mostly adopted technique that relies on the assumption that individual sources are spatially separated. But the isotropic nature of the noise field, created by surface winds and waves as well as shipping activities, which have uniformly distributed noise power makes spatial filtering based on hydrophone directivity rather difficult.

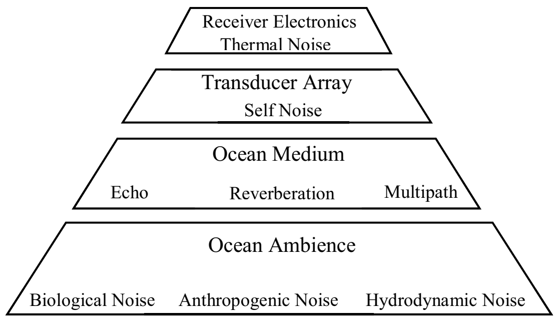



Figure 1. Hierarchy of acoustic reception and noise sources

For effective listening, the noise margin of the acoustic receivers should be increased to an acceptable level. In this paper a method based on Spectral Subtraction is proposed to alleviate the uniform ambient noise in acoustic receivers. The method of spectral subtraction is one of the earliest and widely accepted noise reduction techniques for speech enhancement that has been proposed by S.F. Boll, [2, 3,4] in which the noise interference is subtracted away from the measurements by means of spectral estimates. The signal and the noise processes are assumed to be uncorrelated ergodic processes and mix up in additive mode. In such a scenario, the original signal spectrum can be extracted directly by subtracting an estimate of the noise spectral magnitude obtained during no signal activity in the signal stream. The final signal is then computed by of the previously extracted signal spectrum. The method is rather simple and computationally efficient and can be used for online signal enhancement applications.

# 2 Underwater Ambient Noise

According to the nature of the sources and their mixing mechanism in the channel, underwater noise can be classified broadly into additive and non-additive. The additive noise is often created by linear mixing of multiple sources or linear combination of the vibrations induced separately by two or more simple harmonic excitations on a single source. Most of the underwater noise like tonal noise, periodical noise, pulse noise and broadband noise are additive in nature [5]. However, in underwater ambience, due to the non-linear characteristics of the water medium, addition is not the only process governs the mixing process of noise with the signal of interest. As echo and reverberations are pervasive in the ocean medium, multiplication and convolution are common means of mixing signals together. Homomorphic transforms can be used to inverse the nonlinear mixing process to a linear one [6]. That is, the problem is converted to the same structure as a linear system. The noise field generated by ship submarine or a boring engine is a periodical noise. Continuous noise can be further classified as periodic, such as the sound from rotating machinery or pumps, or aperiodic, such as the sound of a ship breaking ice [1]. The acoustic ambience includes many kinds of broadband noise like cavitation, thermal, surface wind, rain, quantization, and random noise like white noise. The approximated model of the scenario is depicted in figure 2. The statistical independence of the individual sources guaranties uncorrelatedness. The basic assumptions needed for the application of spectral subtraction can be satisfied by considering the noise to be stationary as well as converting the non-additive noise sources into additive, using homomorphic transforms.

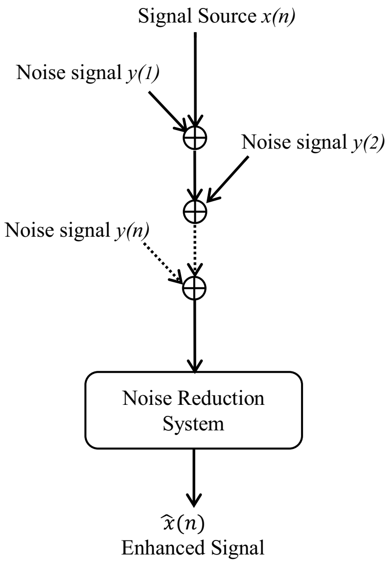



Figure 2. The mixing process of ambient noise

# 3 The Stationarity of Noise

One of the basic assumptions made by spectral subtraction algorithm is the stationarity of noise field, i.e. the noise spectrum remains stationary for minimum time. As described, the acoustic signals received by the hydrophones are a complex mixture or a superposition of all the sources acting together in the medium. The stationarity of a signal can be described mathematically as follows.

Generally for most of the acoustic sources the intensity is distributed nonuniformly both in time and frequency [5]. The distribution can be represented as spectral density,

$$i = \Delta I / \Delta f \tag{1}$$

where $i$ shows the intensity of the signal in the frequency range of $\Delta f = 1$ Hz and the total intensity $I$ contained in the range is,

$$I = \int_{f_1}^{f_2} i \, df \tag{2}$$

where $f_1$ and $f_2$ determine the lower and upper frequencies. The instantaneous spectral density $i(t)$ is a time dependent quantity for almost all random noise sources. The value $i$ in Eq. (2) is averaged over a period $\tau$, $i = \langle i(t) \rangle_\tau$. If $i$ is approximately invariant for each frequency, independent of the time while the averaging is being performed, the noise can be considered stationary. These calculations are performed during each overlapping window frames (refer Section 4), to ensure that the noise spectrum is stationary. If spectral shifts are detected, the noise template is updated accordingly.

# 4 Detecting Signals in Noise

While trying to attend to one voice in a cocktail party, or trying to listen to an enemy submarine lurking around in the ocean depths, the task is to extract useful information from the background noise. A source alone or many sources simultaneously contribute to a signal of interest while other sources present in the background produce the noise [5]. The dividing line between the useful signal and noise is essentially determined by the specific application of the receiver. The presence of a valid signal is to be determined out of the information yielded by the hydrophone measurements alone. The detection of the signal activity in the presence of noise can be regarded as a subjective judgment. Several methods for detecting signal activity in the background noise have been proposed in literature [7, 8, 9]. The functional parameters of the signal detection systems are often described by detection and statistical decision theory.

## 4.1 Signal Detection

With a single hydrophone receiver frontend, as there is no auxiliary channel available for noise estimation, the noise spectrum should be estimated during the no signal activity frames of the observation. Hence, an effective signal activity detection method is needed to isolate noise only frames. Once a noise frame is isolated the spectrum is estimated and the spectral template to be subtracted is updated. However, when the receiver is used for communication purposes rather than passive observations, the occurrence of signals can be largely predictable. In digital communication modems the detection is rather easy as it relies on defined protocols. The signal can be expected to arrive in predictive intervals between transceivers. Intentional blanking periods can be incorporated between packet bursts. During these silent intervals, the noise spectrum can be estimated without much hassle.

In order to detect the signal activity in noise, it needs to be confirmed that the signal stream captured by the hydrophone contains only noise, or a combination of signal and noise. For that, the observation $y(m)$ can be termed as

$$y(m) = b(m)x(m) + n(m) \tag{3}$$

where $b(m)$ is the signal flag, a $high$ represents a valid signal activity and a *low* indicates the absence of a signal, i.e. noise alone, $x(m)$ is the target signal and $n(m)$ is the noise. In digital communication frontends, where the pattern of the signal may be known *a priori*, a correlator or a matched filter can be used to detect the signal [10]. The impulse response $h(m)$ of the matched filter to detect $x(m)$, is taken as the time-reversed replica of $x(m)$. i.e.,

$$\mathrm{h(m) = x(N-1-m)} \qquad 0 \le \mathrm{m} \le \mathrm{N}-1 \qquad (4)$$

where N shows the size of $x(m)$. The output of the filter is determined by,

$$z(m) = \sum_{m=0}^{N-1} h(m-k)y(m) \qquad (5)$$

The output of the filter is then compared against a threshold to obtain $b(m)$, the signal flag sequence which is used for isolating noise only frames. The same sequence can also be used for auto squelching the communication receiver outputs. Another approach is to measure the average energy contained in the segmented frames of the signal sequence, and mark noise only frames by thresholding and gating. Apart from this, zero-crossing rate and linear prediction coefficients are widely used features to determine signal and noise frames. Figure 3(a) and 3 (b) shows the output of the signal activity detection stage in time domain and spectrogram. The noise only frames are used to estimate the noise spectral template to be subtracted.

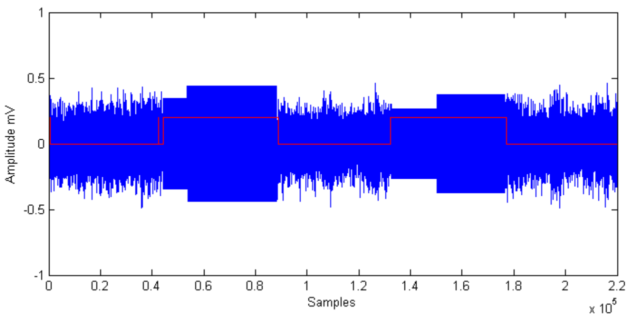



Figure 3(a). The peaks of the red line show signal activity. The test signal is a sine wave tone with step amplitude variation

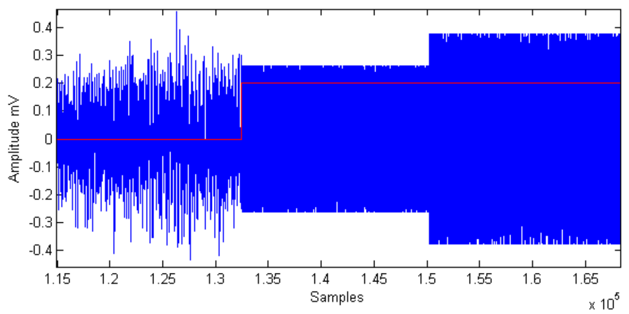


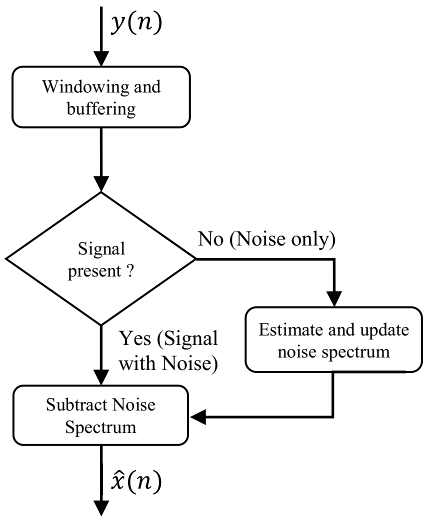

Figure 3(b). A close look on the noise to signal transition

# 5 Spectral Subtraction

Spectral subtraction is a widely used method of noise removal by subtracting an estimated average noise spectrum from the noisy signal spectrum and restoring the power or magnitude spectrum of a signal observed in the additive noise. Several prior assumptions about the nature of the target signal and the noise have been made in order to perform spectral subtraction. The primary assumption is that the spectrum of the input signal corrupted by noise can be described as a linear sum of the signal spectrum and the noise spectrum. The other main assumptions are:

- Target signals and noise signals are statistically independent, hence uncorrelated.
- The gathered signal stream contains noise only frames where the target signals are absent.
- The noise spectrum is relatively stationary than the target signal spectrum.
- The noise parameters are more or less stationary between the consecutive updates.

The basic operational flow chart of the spectral subtraction method is presented in figure:




The ambient noise $L(\mathrm{n})$ is additive to the target signal $x(n)$, the noisy signal $y(n)$ can be written as,

$$y(n) = x(n) + L(n), \qquad for\ 0 \le n \le N-1 \qquad (6)$$

where $n$ is the time index and $N$ is the total number of samples under consideration. The objective of target signal enhancement is to find the enhanced target signal $\hat{x}(n)$ from the observed $y(n)$, with the assumption that $L(n)$ is uncorrelated with $x(n)$.

## 5.1 Algorithm

Initially, the input signal $y(n)$ is segmented into $K$ segments of uniform length. The time-domain signals can be transformed to the frequency-domain as,

$$Y_k(\omega) = X_k(\omega) + L_k(\omega), \qquad for\ 0 \le k \le K-1 \qquad (7)$$

where $k$ is the segment index, $\mathrm{Y_k(\omega)}$, $\mathrm{X_k(\omega)}$ and $\mathrm{L_k(\omega)}$ are the short-time DFT magnitudes taken of $y(n)$, $x(n)$, and $L(n)$ respectively. Spectral subtraction is given as,

$$|X_k(\omega)|^a = |Y_k(\omega)|^a - |L_k(\omega)|^a \qquad (8)$$

the DFT magnitudes are raised to a power $a$ where, $a = 1$ corresponds to magnitude spectral subtraction, $a = 2$ corresponds to power spectrum subtraction. If an estimate of the self-noise spectrum $\hat{L}_k$ can be obtained, then an approximation of interested signal $\hat{X}_k$ can be obtained from

$$\hat{X}_k = Y_K(\omega) - \hat{L}_k(\omega) \quad (9)$$

The noise magnitude spectrum $\hat{L}_k(\omega)$ can be estimated as a running average of those signal blocks determined to be primarily noise alone. The average noise magnitude spectrum is then subtracted from the magnitude spectrum of the incoming signal; negative differences are set to zero. The modified magnitude spectrum is further supplemented with the phase spectrum for generating modified spectrum and the output signal is recovered via an inverse FFT.

# 6 Implementation

After the incoming signal stream is buffered to match the speed of processing, the hydrophone observations $y(n)$, are segmented into frames of $N$ samples. The magnitude and phase spectrum of the signal stream are then estimated within small windows using FFT. The noise magnitude spectrum is calculated from the noise only frames isolated by the signal activity detector. The noise spectral template is then subtracted from the magnitude spectrum of the incoming signal $y(n)$. The negative differences are replaced by zeros to preserve timing. The obtained signal magnitude spectrum is then recombined with the original phase spectrum, and the output signal is retrieved using inverse FFT. Figure 4 shows the detailed block diagram of the spectral subtraction system.

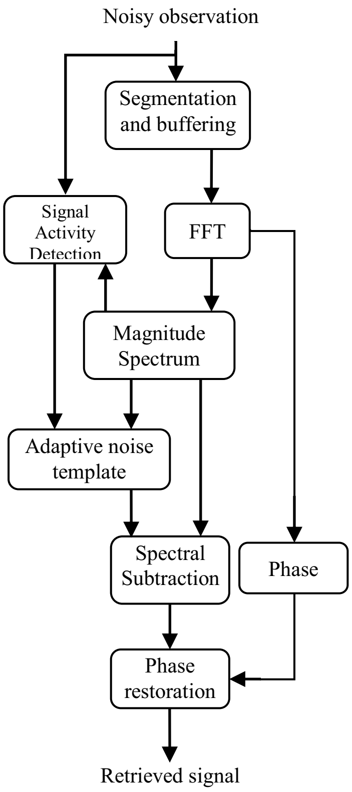



Figure 4. Block diagram of spectral subtraction process

## 6.1 Threading scheme

The entire process of spectral subtraction is divided into several concurrent threads. While the main thread progresses with the noise subtraction; the auxiliary threads can go for sampling the data, look for large spectral shifts and updating the spectral template of the noise to be subtracted. Each thread shares its output with other threads as they are running in a common process space. Threads are brought to run or sleep state according to the availability of data. The threading scheme presented in Figure 5, helps to efficiently utilize the processing power, improve the realtime performance of the system and generate adaptive noise templates online.

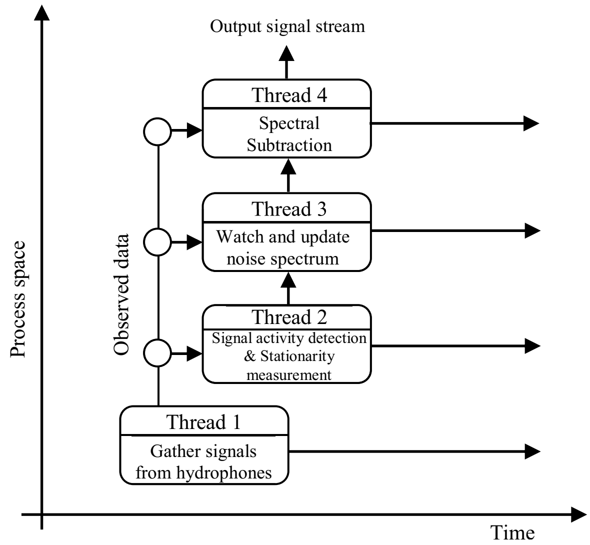



Figure 5. The threading scheme of spectral subtraction process

# 7 Results and discussions

The specific objective of the attempt was to improve the tangibility of signals of interest that are additively mixed in stationary ambient acoustic signals received by hydrophone arrays. With the intense levels of noise background, it is often difficult to extract any vital signals from the received signal. The proposed method improves the SNR considerably by incorporating spectral subtraction based noise reduction technique. The method achieves the goal by subtracting an estimate of the noise spectrum from the observation spectrum. Several test runs have been conducted with different noise conditions. The first target was to extract an FSK modem signal corrupted primarily by rain noise. The time domain and spectral plots are shown in Figure 6(a) and 6(b) respectively. From the spectrogram of the observation, it is clear that the modem signals are too faint.

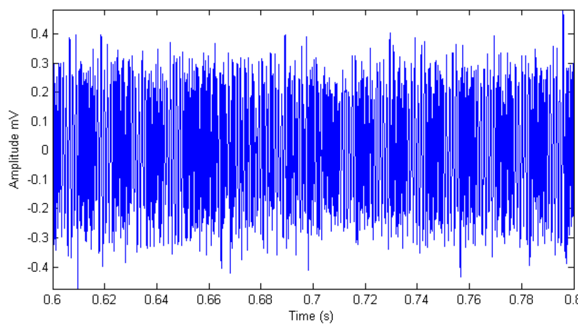



Figure 6(a) An FSK modem signal corrupted by rain and other ambient sources

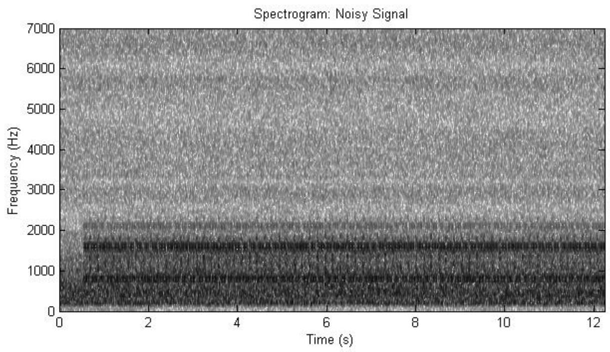



Figure 6(b) The spectrogram of the observed signal

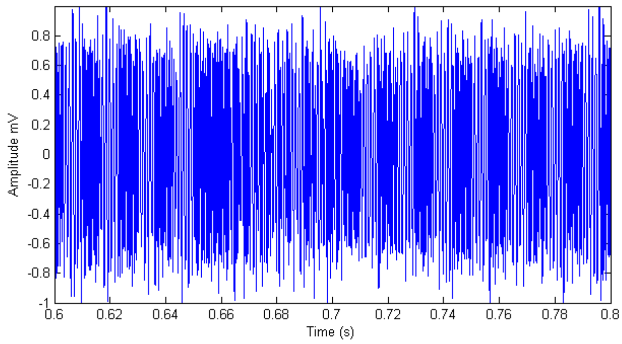



Figure 6(c) The retrieved signal after spectral subtraction

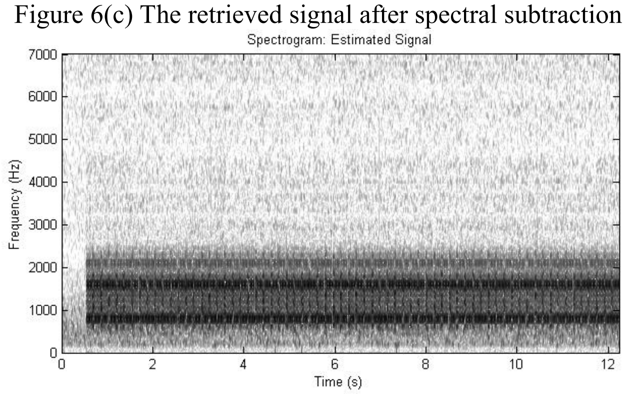



Figure 6(d) Spectrogram of the retrieved signal. The background noise is visibly reduced

Figures 6(c) and 6(d) shows the estimated signal using spectral subtraction. In the spectrogram, the modem signal is clearly visible while the background noise is considerably reduced. The algorithm performed reasonably well on a series of synthesized signal noise mixtures.

## 8 Conclusions

The proposed system employed spectral subtraction algorithm as an active noise reduction method for alleviating ambient noise in underwater acoustic receiver. The observations are assumed to be mixed additively with stationary ambient noise. The spectral template of the noise is estimated during the absence of vital signals in the signal stream using a signal activity detector. The estimated noise template is subtracted from the observation spectrum to obtain the target signal. The method performed consistently over the testing process, giving significant reduction in the noise interference and helped in pulling up the salient signals obliterated by the noise ambience. The synthesized ocean acoustic signal mixtures were utilized for the experiments. Hardware based real world observations from underwater test facilities can be evaluated with the system in future. On the basis of the empirical results obtained, the system can be further expanded for online noise reduction, as it is practically usable with underwater acoustic receivers.

## Acknowledgments

The authors gratefully acknowledge Naval Research Board, New Delhi for the financial assistance and the Department of Electronics, Cochin University of Science and Technology for extending all the facilities for carrying out this work.

## References

[1] John Hildebrand, “Sources of anthropogenic sound in the marine environment”, Int. Policy Workshop on Sound and Marine Mammals, London, 28-30 September 2004.

[2] S.F. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” *IEEE Trans. Acoust., Speech, Signal Process*. vol.27, pp. 113-120, Apr. 1979.

[3] Weiss, M., Aschkenasy, E., Parsons, T., 1974. “Study and the development of the INTEL technique for improving speech intelligibility” *Technical Report NSC-FR/4023, Nicolet Scientific Corporation.*

[4] Ekaterina Verteletskaya, Boris Simak, “Noise Reduction Based on Modified Spectral Subtraction Method”. IAENG International Journal of Computer Science, 38:1, IJCS.38.1.10.

[5] Kinsler, Lawrence et al, “Fundamentals of Acoustics”, 4th ed., Wiley & Sons, 1999

[6] Steven W. Smith, Digital Signal Processing A Practical Guide Newnes Press, London (2003)

[7] K.Sakhnov, E.Verteletskaya, B. Šimák, “Dynamical Energy-Based Speech/Silence Detector for Speech Enhancement Applications”. World Congress of Engeneering 2009 Proceedings. Hong Kong:, 2009, pp.801-806.

[8] V. Prasad, R. Sangwan et al., “Comparison of voice activity detection algorithms for VoIP”, proc. of the Seventh International Symposium on Computers and Communications, Taormina, Italy, 2002, pp. 530-532.

[9] Minkoff J. Signal Processing Fundamentals and Applications for Communications and Sensing Systems. Norwood, MA, USA: Artech House; 2002.

[10] Saeed Vaseghi. Advanced Digital Signal Processing and Noise Reduction. John Wiley & Sons, West Sussex, U.K., third edition, 2006.