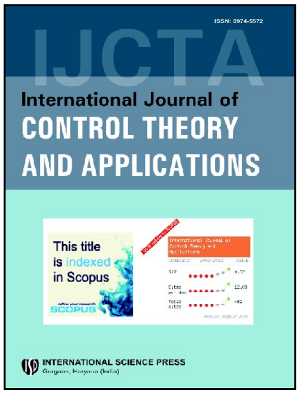



**International Journal of Control Theory and Applications**

ISSN : 0974-5572



Volume 10 • Number 22 • 2017

# A Novel Audio Noise Cancellation using Non-Diagonal Time Frequency Algorithm

**Rajendra Prasad[1*] and Polaiah Bojja[2]**

[1]*Research Scholar, Department of ECE, KL University, Guntur, Andhra Pradesh, India*
*E-mail: rajendra_prasad_ch@srecwarangal.ac.in*
[2]*Professor, Department of ECE, KL University, Guntur, Andhra Pradesh, India, E-mail: paulraj.bojja@gmail.com*

***Abstract:*** Historically, backdrop and hiss noise were generate in audio signals. Audio signals are normally corrupted with backdrop and hiss noise from different audio instruments. Audio de-noising main aim is to decrease the noise holding the un-divulged signals. In this paper a method is proposed to de-noise the said noise form audio signals called non-diagonal block thresholding estimator. In this estimator the block parameters are made sensitive to the signal quality. These block parameters are adjusted based on the signal quality and minimized using SURE estimator of the risk. In this estimator separate spectrogram co-efficient are used for noise removal, by that the error is minimized between the actual signal and improved signal. Audio signal contaminated with adaptive white Gaussian noise is chosen for performance evolution of proposed algorithm. In this paper, we mainly concentrate on compute of Signal to Noise Ratio(SNR) and SSNR (Segmented SNR) to improve the audio quality. The MATLAB simulation shows better results for this proposed algorithm. This algorithm will be applicable for real time implementation of adaptive noise cancellation systems.

***Keywords:*** Novel De-noising, Time Frequency(TF), Block Thresholding(BT), Short Time Fourier Transform(STFT), SNR, SSNR and SURE.

## 1. INTRODUCTION

Audio de-noising is extensively becoming a principal area of research. Because in real world variety of audio enabled equipment's require audio de-noising [1-4]. Noise is an important factor that affects speech recognition performance. The background noise can be significantly suppressed using the proposed adaptive noise cancellation scheme [5-8]. Diagonal time-frequency (TF) audio de-noising structure eliminate the noise by handling all window Fourier or wavelet coefficient individually, with power subtraction[9], empirical Wiener [10], or thresholding operators[11]. These structures construct secluded TF algorithms that are intuited as a "backdrop noise" [12]. Ephraim and Malah presented the backdrop noise is strongly eliminated by the non-diagonal TF estimators that standardize the analysis with aggregating TF coefficients recursively. This analysis has additionally enhanced by rearranging the SNR estimation with parameterized filters that depend on stochastic audio models [13].

Conversely, these parameters have to be tune to the nature of the audio signal. They are empirically fixed in implementation [14]. A novel audio de-noising algorithm through adaptive time-frequency block thresholding(TFBT) using non-diagonal estimator introduced in this paper. Silverman and Cai are introduced in mathematical statistics[15,16] to enhance the asymptotic rot of diagonal thresholding estimations.

For audio TF de-noising, we demonstrate that BT regularizes the estimation and it is effective in reduction of backdrop noise. Block parameters are spontaneously attuned by minimizing risk of Stein estimation[17], and computed systematically from the noisy signal standards. Statistical experimentations reveal that this to state of the art signal de-noising algorithm enhance the SNR and the perceived value.

Section II reviews the TF de-noising algorithms by stressing the contrast among diagonal and non-diagonal techniques. Section III presents proposed de-noising algorithm and calculates the Stein unbiased estimation of resultant risk to alter automatically the block constraints. Section IV presents MATLAB simulation of proposed de-noising algorithm and discussion.

## 2. AUDIO DE-NOISING TOPOLOGIES

The removal of noise in the audio signal uses TF audio-de-noising methods. These methods make use of either wavelet transform/STFT/wavelet packet transform. By which the coefficients are computed for noise removal. These representations declare that TF signal structures can be separated from the noise. We focus on the processing coefficients as against to the possibility of representations. Most commonly statistical experiments are carry out using STFT in audio quality enhancement. Figure 1 shows block diagram of TF audio de-noising.

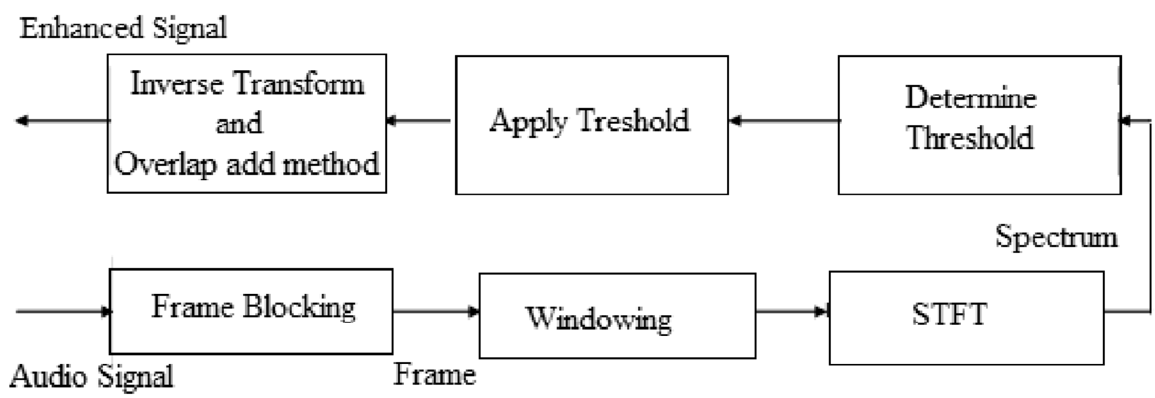



**Figure 1: Block diagram of TF audio-de-noising**

The Audio De-noising techniques are basically divided in to two types depending upon the SNR considered [18, 19]

1. Diagonal Estimator
2. Non-diagonal Estimator

### 2.1. Diagonal Estimator

This estimator determines every attenuation component only from similar noisy coefficient using TF de-noising algorithms. These algorithms produce a musical noise and have a limited performance. Posterior S/N ratio is considered in Diagonal Estimation. SNR of the audio noisy signal is the posterior SNR[20].

The Signal to Noise Ratio $\zeta[p,q]$ are determined from the final SNR in diagonal estimators is defined by

$$\gamma(p,q) = \frac{|S[p,q]|^2}{\sigma^2[p,q]} \tag{1}$$

We can confirm that $\zeta[p,q] = \gamma[p,q] - 1$ is a neutral estimator. The actual Wiener estimator is defined as

$$\alpha[p,q] = \left(1 - \frac{1}{\zeta[p,q]+1}\right)_+ \tag{2}$$

using representation $(Z)_+ = \max(z,0)$

By subtracting the amount of signal distortion and remaining noise power from the noisy signal the coefficients of this empirical Wiener are achieved. The actual Wiener attenuation rule is given as

$$\alpha[p,q] = \left(1 - \lambda\left[\frac{1}{\zeta[p,q]+1}\right]^{\beta_1}\right)^{\beta_2}_+ \tag{3}$$

Where $\beta_1, \beta_2 \geq 0$ and $\lambda \geq 1$ is an over-deduction factor to compensate variation of noise amplitude.

In diagonal estimators, the attenuation factor $\alpha[p,q]$ only depends upon $S[p,q]$ using zero TF regularization. The resultant attenuation coefficients are α[*p*, *q*], *S*[*p*, *q*] shortage of TF reliability. It gives secluded TF coefficients which bring back secluded TF form that are intuited as a buzzing noise. Audio signal de-noising can be implemented with a thresholding in a windowed Fourier frame. It amounts to a simple thresholding of the resultant spectrogram, but it produces a musical noise corresponding to isolated coefficients above threshold. Thresholding estimators decompose noisy signals in a basis or in a frame and set to zero small amplitude coefficients. A diagonal estimator in this basis modifies the amplitude of each coefficient $s_f[m]$ with a factor a[m] and reconstructs

$$\tilde{f} = \sum_{n=1}^{M} \alpha[n] s_f[n] g_n \tag{4}$$

If α[*n*] depends only upon $y_f[n]$ the estimator is said to be diagonal. Soft thresholding process each parameter is threshold freely from its adjoining blocks. BT estimator used to lessen this backdrop noise that considers the way that substantial spectrogram parameters of most audio signals are accumulated together in the TF plane. The calculation of threshold value or attenuation factor from the posterior SNR is called diagonal estimation. The technical term used to characterize the quality of signal is the S/N Ratio. It is expressed as ratio or factor in units of decibels [dB], given by

$$\text{Ratio [dB]} = 20\log\text{(ratio)} \tag{5}$$

Diagonal estimators of the S/N Ratio are processed from the a posteriori S/N Ratio. In diagonal estimators attenuation factor depends on noisy coefficients with zero TF regularization. The resultant attenuated coefficients thus deficit of TF regularity. It produces secluded TF coefficients which bring back secluded TF structures that are seen as a backdrop noise. Some of Diagonal estimation techniques are Ephraim and Malah log-spectrogram amplitude (LSA), Power subtraction (PS), and decision directed SNR estimator.

### 2.2. Non-diagonal Estimator

Non-diagonal Estimation is defined as the calculation of attenuation factor or threshold value from the priori Signal to Noise Ratio[21]. A number of authors proposed to estimate a priori Signal to Noise Ratio using time-recurrence regularization of the posteriori Signal to Noise Ratio to minimize estimation risk as well as the musical noise. In non-diagonal estimator resultant attenuation factors depend on the data values in a whole neighborhood of. In non-diagonal Estimation previous Signal to Noise Ratio is the Signal to Noise Ratio of original sound signal. Non-diagonal estimators depend upon regularization filtering parameters but do better than diagonal estimators. Substantial regularization filters minimize the noise vitality however produce distorted signal. It is advantageous that parameters of the filter are modify based on the audio signal nature. However, parameters are selected empirically. Some of the Non-Diagonal estimator methods are p-point uncertainty model and BT

## 3. PROPOSED METHODOLOGY

### 3.1. De-Noising of Audio

The TF audio quality enhancement strategy figures a STFT of corrupted signal and the resulting parameters are processed to remove the backdrop noise. This proves that noise is separated from these TF signal structures. The sound signal f is debased by a noise that is frequently demonstrated as Gaussian process of zero-mean free of f:

$$s[n] = f[n] + \Psi[n] \;\; m = 0, \;\; 1 \ldots M\text{-}1 \tag{6}$$

Where $\Psi[n]$– Noise

$f$[n] – Original Signal

s[n] – Noisy signal

The original sound signal decomposed over time-recurrence localization indices by using time- recurrence transform. The subsequent coefficients should be composed as

$$S[p,q] = \left\langle y, g_{p,q} \right\rangle = \sum_{m=0}^{M-1} s[m] g_{p,q}^{*}[m] \tag{7}$$

Where, * means the conjugate and q and p are frequency and time limitation lists. These transforms characterize an entire and regularly repetitive signal representation. Assume that these TF molecules characterize a tight frame, which implies that there exists C>0 such that

$$\|s\|^2 = \frac{1}{C} \sum_{p,q} \left| \left( s, g_{p,q} \right) \right|^2 \tag{8}$$

This suggests a straightforward reproduction equation

$$s[m] = \frac{1}{C} \sum_{p,q} S[p,q] g_{p,q}[m] \tag{9}$$

Where C is a constant and it is an excess element is equivalent to one then a tight frame is an orthogonal premise. To offset the noise element the de-noising algorithm adapts time- frequency coefficients by multiply each of coefficient byfactor. TF de-noising calculations be different through the computation of the. The noise coefficient difference should be known or evaluated.

$$\sigma^2[p,q] = E\left\langle \left| \Psi, g_{p,q} \right| \right\rangle \tag{10}$$

## 3.2. Non-Diagonal Estimation Algorithm

To reduce estimation risk and the musical noise, several authors have presented a time- recurrence regularization of the posteriori S/N ratio $\gamma$ [$p$, $q$] to estimate a priori S/N ratio. Resulting attenuation factors thus based on the data values $Y$ [$p$', $q$'] for ($p$', $q$') in an entire neighborhood of and the subsequent estimator is known as non-diagonal and is estimated by,

$$\tilde{f}[m] = \frac{1}{C}\sum_{p,q} \alpha[p,q]S[p,q]g_{p,q}[m] \tag{11}$$

Ephraim and Malah have announced the first order recursive time filtering is used to obtain decision-directed SNR estimator.

$$\zeta[p,q] = \alpha\tilde{\zeta}[p-1,q] + (1-a)\{\gamma[p,q]-1\} \tag{12}$$

Large regularization filters minimize the noise vitality however produce distorted signal. It is advantageous that parameters of the filter are tuned based on audio signal nature. However, parameters are selected empirically. By calculating a single attenuation factor over TF blocks the TFBT estimator regularizes estimate. The adaptive block thresholding selects dimensions by decreasing risk of the estimate. Best block dimensions are determined by reducing this estimated risk.

In BT estimator an upper bound of the hazard is figured by examine independently the variance and bias terms. Here we notice that the upper bound of the prophet hazard with blocks is constantly bigger than the prophet chance without blocks, on the grounds that the previous is gotten from the proportionate minimization yet with less parameter as attenuation factor stay consistent over every block. The signal estimatoris figured from the noisy information s with a steady attenuation factor $a_i$ over every block $B_i$

$$\tilde{f}[m] = \sum_{i=1}^{I} \sum_{(p,q)eB_i} \alpha_i S[p,q]g_{p,q}[m] \tag{13}$$

To see how to process each ai, one relates the Stein estimation hazard, $r = E$ {‖$f$-f^ ‖2} to the frame vitality transformation and given by,

$$\gamma = E\left\{ \left\| f - \tilde{f} \right\| \right\} \leq \sum_{i=1}^{I} \cdot \sum_{(p,q)eB_e} E\left\{ \left| F[p,q] - \tilde{F}[P,Q] \right|^2 \right\} \tag{14}$$

Consider a decision of block dimension and the remaining noise likelihood level ä that one endures the level of threshold ë. Each block width and length, ë is assessed utilizing “Monte Carlo simulation”. Table 1 demonstrates subsequent ë with ä = 0.1%. For a block width W is greater than 1, B# = LXW, have close by ë values.

**Table 1**
**For different block calculated Thresholding level λ and with δ = 0.1%**

| *λ Value* | *W = 1* | *W = 2* | *W = 4* | *W = 8* | *W = 16* |
|---|---|---|---|---|---|
| **L = 2** | 4.7 | 4.7 | 3.5 | 1.8 | 2.0 |
| **L = 4** | 3.5 | 3.5 | 2.5 | 2.0 | 1.8 |
| **L = 8** | 2.5 | 2.5 | 2.0 | 1.8 | 1.5 |

## 4. RESULTS

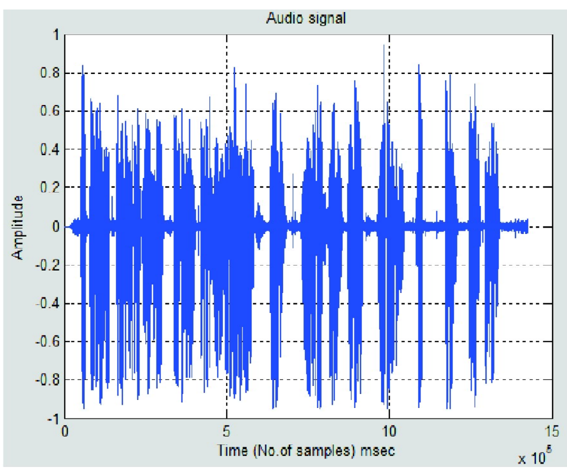
The tests exhibited underneath have been performed on sound signal with a time of 1minute. This signal tainted by Gaussian white noise with different amplitude levels. STFT with Hanning windows were utilized as a part of the implementation. The dimension of windows taken is 50 ms. For each sound, de-noising with "maximum noise removal" were applied, this removes almost all the original noise. The figure 2 shows de-noising of audio signal using non- diagonal thresholding. The target measures are individually S/N Ratio and the segmental SNR characterized as

$$SNR = 10\log_{10}\frac{\sum_{k=1}^{M} f^2[m]}{\sum_{k=1}^{M}(f[m]-s[m])^2} \tag{15}$$

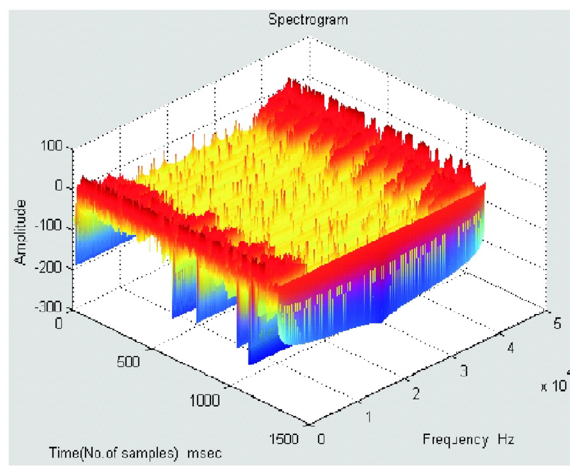
Where $s(m)$ is the noisy signal sequence[22-24]

The improvement gain measures the increase in SNR due to the algorithm, and is defined as

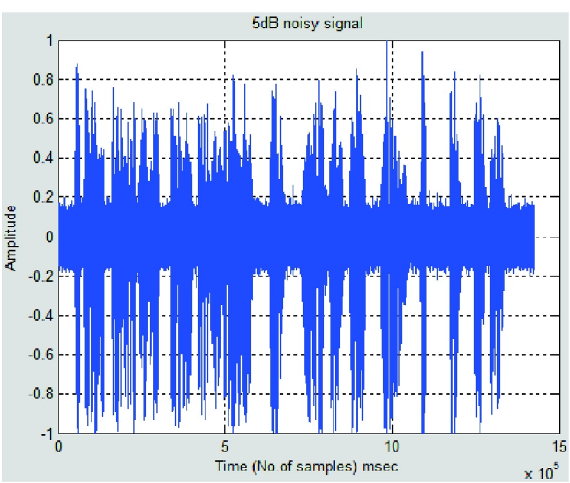
$$\text{Gain} = \text{S/N Ratio (Output)} - \text{S/N Ratio (Input)} \tag{16}$$



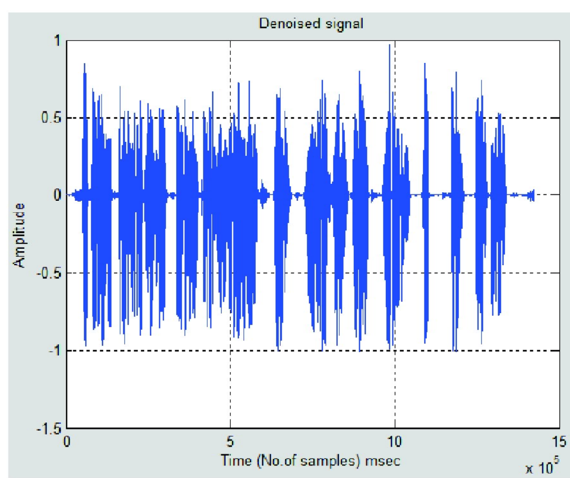

(a) Audio signal




(b) Spectrogram




(c) 5dB noisy signal




(d) De-noised signal

**Figure 2: De-noising of audio signal using non- diagonal thresholding**

Segmental signal to noise ratio defined as

$$SegSNR = \frac{1}{H}\sum_{t=0}^{H-1} T(10\log_{10}\frac{\sum_{n=0}^{S-1} f^2[(\frac{n+ls}{2})]}{\sum_{n=0}^{S-1}(f[\frac{n+ls}{2}]) - \hat{f}[\frac{n+ls}{2}]^2}) \tag{17}$$

Where, T(y) = min. [max. (-10, y), 35]

S = samples in each frame

H = total frames in the signal

Table 2 shows SNR and SSNR values of different noise signals. Signal to noise ratio (SNR) and Segmented SNR (SSNR) values shows actual signal and improved signal are almost equal. These simulation results show that the quality of audio signal is improved.

**Table 2**
**Different Noisy Signals with SNR and SSNR**

| *Audio Signal + SNR* | *Non-Diagonal Estimator* | |
|---|---|---|
| | *SNR((dB)* | *SSNR(dB)* |
| 5.00 | 19.98 | 19.9902 |
| 6.00 | 21.07 | 21.0874 |
| 8.00 | 22.89 | 22.9116 |
| 9.00 | 23.82 | 23.8418 |
| 10.00 | 24.34 | 24.3658 |
| 15.00 | 27.03 | 27.0628 |
| 20.00 | 28.95 | 29.0126 |

## 5. CONCLUSION

In this paper, De-noising of audio signal using non- diagonal thresholding is simulated using MATLAB. By comparing diagonal estimator this Non-diagonal time- frequency estimators introduce less background noise and are more effective and efficient to enhance audio signals. The results are analyzed by computing Signal to noise ratio (SNR) and Segmented SNR (SSNR) of actual signal and improved signal. These simulation results show that the quality of audio signal is improved. Non- diagonal time frequency thresholding is utilized as a part of real time implementation of adaptive noise cancellation systems.

## REFERENCES

[1] E. Bahoura "Wavelet speech enhancement based on timescale adaptation," Speech Communication., vol. 48, no. 12, pp. 1620– 1637, Dec. 2006.

[2] Siddappaji, K.L. Sudha" A New Time-Varying LMS Adaptive Filtering Algorithm in noise Cancellation System" Journal of Basic and Applied Engineering Research Print ISSN: 2350-0077; Online ISSN: 2350-0255; Volume 1, Number 7; October, 2014 pp. 87-92.

[3] Muhammad Amir Shafiq, Saqib Ejaz & Nisar "Hardware Implementation of Adaptive Noise Cancellation over DSP Kit TMS320C6713" International Journal of Signal Processing (SPIJ), Volume (7) : Issue (1) : 2013, pp. 75-86.

[4] TIHAL,BOULANOIR "Denoising sensors signals within mechatronic systems using a Wavelet technique" 978-1-4673-6673-1/15/31:00 2015 IEEE.

[5] Raj Kumar and Agrawal," Hardware Implementation of Adaptive Algorithms for Noise Cancellation" International Journal of Information and Electronics Engineering, Vol. 2, No. 2, March 2012.

[6] Nitish Krishnamurthy • John H.L. Hansen "Environment dependent noise tracking for speech enhancement" Int J Speech Technol DOI 10.1007/s10772-012-9182-0, springer dec-2012.

[7] M. Gorriz, Javier Ramirez, S. Cruces-Alvarez, Carlos G.Puntonent, Elmar W. Lang and Deniz Erdogmus, "A Novel LMS algorithm Applied to Adaptive Noise Cancellation," IEEE Signal Processing Letters, vol. 6, no. 1, pp. 34-37, Jan. 2009.

[8] X. Guan, G. Wu and X. Chen "QX-LMS adaptive FIR filters for system identification" 2nd International Congress on Image and Signal Processing (CISP2009), 2009, pp. 1-5.

[9] S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoustics, Speech, Signal Process., vol. ASSP-27, no. 2, pp. 113–120, Apr. 1979.

[10] M. Li, H. G. McAllister, N. D. Black, and D. T. A. Perez, "Perceptual TF subtraction algorithm for noise reduction in hearing aids," IEEE Trans. Biomed. Eng., vol. 48, no. 9,pp. 979–988, Sep. 2001.

[11] R. J. McAulay and M. L. Malpass, "Speech enhancement using soft decision noise suppression filter," IEEE Trans. Acoust., Speech, SignalProcess., vol. ASSP-28, no. 2, pp. 137–145, Apr. 1980.

[12] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP), 1979, vol. 4, pp. 208–211.

[13] I.Cohen, "Noise spectrum estimation in adverse environments: Improved minima controlled recursive averaging," IEEE Trans. Speech Audio Process., vol. 11, no. 5, pp. 466–475, Sep. 2003.

[14] I. Cohen, "Speech enhancement using a noncausal a priori SNR estimator," IEEE Signal Process. Lett., vol. 11, no. 9, pp. 725–728, Sep. 2004.

[15] T. Cai, "Adaptive wavelet estimation: A BT and oracle inequality approach," Ann. Statist., vol. 27, pp. 898–924, 1999.

[16] Silverman B. W. and T. Cai "Incorporation information on neighboring coefficients into wavelet estimation," Sankhya, vol. 63, pp. 127–148, 2001.

[17] S. Chang and Y. Kwon, "Speech enhancement for non-stationary noise environment by adaptive wavelet packet," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 2002, vol. 1, pp. 561–564.

[18] G. O. Glentis, "Efficient fast recursive least-squares adaptive complex filter using real valued arithmetic," Elsevier Signal Processing, vol. 85, 2005, pp. 1759-1779.

[19] S. Haykin, Adaptive Filter Theory, 4th ed., Prentice Hall, Upper Saddle River, NJ, 2002.

[20] I. Cohen, "Speech enhancement using supergaussian speech models and noncausal a priori SNR estimation," Speech Commun., vol. 47, no. 3, pp. 336–350, Nov. 2005.

[21] O. Capp, "Elimination of the musical noise henomenon with the Ephraim and Malah noise suppressor," IEEE Trans. Speech, Audio Process, vol. 2, pp. 345–349, Apr. 1994.

[22] B.Widrow and E.Walach, Adaptive Inverse Control, Reissue Edition: A Signal Processing Approach, Wiley-IEEE Press, 2007.

[23] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP), 1979, vol. 4, pp. 208–211.

[24] D. Donoho and I. Johnstone, "Idea spatial adaptation via wavelet shrinkage," Biometrika, vol. 81, pp. 425–455, 1994.