Audio Signal Enhancement using Non-diagonal Estimator

K. Sreedhar

Abstract— Audio signals are often contaminated by background environment noise and buzzing or humming noise from audio equipments. Audio denoising aims at attenuating the noise while retaining the underlying signals. Removing noise from audio signals requires a nondiagonal processing of time-frequency coefficients to avoid producing "musical noise." A block thresholding estimation procedure is introduced, which adjusts all parameters adaptively to signal property by minimizing a Stein estimation of the risk. Non Diagonal time-frequency audio denoising algorithm attenuates the noise by processing each spectrogram coefficient independently. This Estimator is to minimize the error between clean signal and the enhanced signal. Numerical experiments demonstrate the performance and robustness of this procedure through objective and subjective evaluations.

Index Terms—Audio Denoising, Block Thresholding, Audio signal processing, STFT Transform, Spectrogram, Time-Frequency Audio Denoising, Adaptive Block Thresholding

I. INTRODUCTION

Non Diagonal time-frequency audio denoising algorithms attenuate the noise by processing each window Fourier or wavelet coefficient independently, with thresholding operators. These algorithms create isolated time-frequency structures that are perceived as a "musical noise" is strongly attenuated with nondiagonal time-frequency estimators that regularize the estimation by recursively aggregating time-frequency coefficients. This approach has further been improved by optimizing the SNR estimation with parameterized filters that rely on stochastic audio models. However, these parameters should be adjusted to the nature of the audio signal, which often varies and is unknown. In practice, they are empirically fixed. This paper introduces a new nondiagonal audio denoising algorithm through adaptive time-frequency block thresholding. Block thresholding has been introduced by Cai and Silverman in mathematical statistics to improve the asymptotic decay of diagonal thresholding estimators. For audio time-frequency denoising, block thresholding regularizes the estimate and is thus effective in musical noise reduction. Block parameters are automatically adjusted by minimizing a Stein estimator of the risk, which is calculated analytically from the noisy signal values.

Manuscript received May 30, 2011.

K.Sreedhar, Department of Electronics and Communication Engineering, VITS-SET (N9), Karimnagar, Andhra Pradesh, India, 9908624167., (e-mail: sreedhar_kallem@yahoo.com).

Numerical experiments show that this new adaptive estimator is robust to signal type variations and improves the SNR and the perceived quality with respect to state of the art audio denoising algorithms.

Applications such as music and speech restoration are numerous.

The paper first reviews the proposed approach of time-frequency audio denoising algorithms by emphasizing the difference between diagonal and nondiagonal methods. Section III introduces proposed approach and Section IV computes a Stein unbiased estimate of the resulting risk to adjust automatically the block parameters. Numerical Results and comparisons are presented in Section V, with objective and subjective measures.

II. STATE OF THE PROBLEM

A. Diagonal Estimation

Simple time-frequency denoising algorithms compute each attenuation factor only from the corresponding noisy coefficient and are thus called diagonal estimators. These algorithms have a limited performance and produce a musical noise. In Diagonal Estimation the Posterior SNR is considered. Posterior SNR is the SNR of the Audio Noisy Signal.

Diagonal estimators of the SNR $\xi(l,k)$ are computed from the a posteriori SNR defined by

$$\gamma(l,k)=\frac{|Y[l,k]|^2}{\sigma^2[l,k]}$$

One can verify that $\hat{\xi}[l,k]=\gamma[l,k]-1$ is an unbiased estimator.

The empirical Wiener estimator is defined as

$$a[l,k]=\left(1-\frac{1}{\hat{\xi}[l,k]+1}\right)_+$$

with the notation $(z)_+=\max(z,0)$

Variants of this empirical Wiener are obtained by minimizing a sum of signal distortion and residual noise energy.

The empirical Wiener attenuation rule is given as

$$a[l,k]=\left(1-\lambda\left[\frac{1}{\hat{\xi}[l,k]+1}\right]^{\beta_1}\right)_+^{\beta_2}$$

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

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The attenuation factor $a[l,k]$ of these diagonal estimators only depends upon $Y[l,k]$ with no time-frequency regularization. The resulting attenuated coefficients $a[l,k]$ $Y[l,k]$ thus lack of time-frequency regularity. It produces isolated time-frequency coefficients which restore isolated time-frequency structures that are perceived as a musical noise. A soft thresholding produces a similar phenomenon because each coefficient is also thresholded independently from its neighbors. To remove this musical noise, uses a block thresholding estimator that takes into account the fact that large spectrogram coefficients of most audio sounds are aggregated together in the time-frequency plane.

Nondiagonal time-frequency estimators are more effective than diagonal estimators to remove noise from audio signals because they introduce less musical noise.

III. PROPOSED APPROACH

Methodology of Solution

A. Audio denoising

Time-frequency audio-denoising procedures compute a short-time Fourier transform or a wavelet transform or a wavelet packet transform of the noisy signal, and processes the resulting coefficients to attenuate the noise. These representations reveal the time-frequency signal structures that can be discriminated from the noise. We concentrate on the coefficient processing as opposed to the choice of representations. Numerical experiments are performed with short-time Fourier transforms that are most commonly used in audio processing.

The audio signal *f* is contaminated by a noise that is often modeled as a zero-mean Gaussian process independent of *f*:

$$y[n] = f[n] + \epsilon[n], \quad n=0, 1\ldots N-1$$

A time-frequency transform decomposes the audio signal over time-frequency localization indices. The resulting coefficientsshallbewrittenas

$$Y[l,k] = \langle y, g_{l,k} \rangle = \sum_{n=0}^{N-1} y[n] g^{*}_{l,k}[n]$$

Where, * denotes the conjugate and l & *k* are time and frequency localization indices.

A denoising algorithm modifies time-frequency coefficients by multiplying each of them by an attenuation factor to attenuate the noise component. Time-frequency denoising algorithms differ through the calculation of the attenuation factors *a* [*l*, *k*].The noise coefficient variance is supposed to be known or estimated.

$$\sigma^2 [l, k] = E \{ |\langle \epsilon, g_{l,k} \rangle|^2 \}$$

If the noise is stationary, which is often the case, and then the noise variance does not depend upon time:

$$\sigma^2 [l, k] = \sigma^2 [k]$$

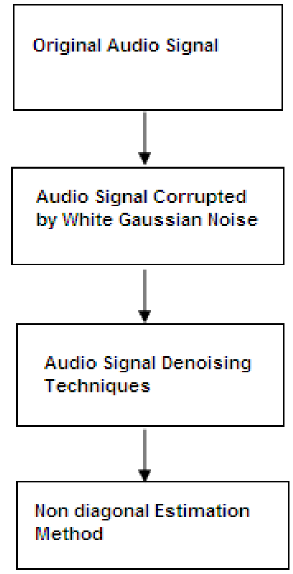


Figure.1. Block Diagram of Denoising Musical Audio noise signal

B. Non diagonal Estimation

To reduce musical noise as well as the estimation risk, several authors have proposed to estimate a priori SNR ξ [*l*, *k*] with a time-frequency regularization of the *posteriori* SNR γ [*l*, *k*]. Resulting attenuation factors *a* [*l*, *k*] thus depend upon the data values Y [*l'*, *k'*] for (l', *k')* in a whole neighborhood of and the resulting estimator is said to be nondiagonal and given by,

$$\hat{f}[n] = (1/A)\sum_{l,k} a[l,k] Y[l,k] g_{l,k}[n]$$

Ephraim and Malah have introduced a *decision-directed SNR* estimator obtained with a first order recursive time filtering:

$$\hat{\xi}[l,k] = \alpha\hat{\xi}[l-1,k] + (1-\alpha)(\gamma[l,k]-1)$$

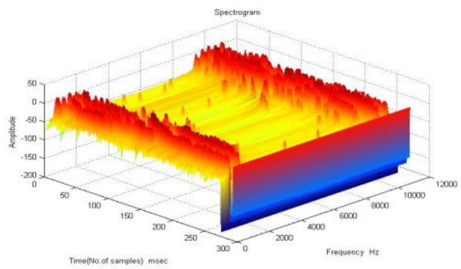
where α € [0, 1] a recursive is filter parameter and is an empirical SNR estimate of based on the previously computed estimate.

In nondiagonal Estimation we consider priori SNR. Pre SNR is the SNR of Audio signal. Nondiagonal estimators clearly outperform diagonal estimators but depend upon regularization filtering parameters.

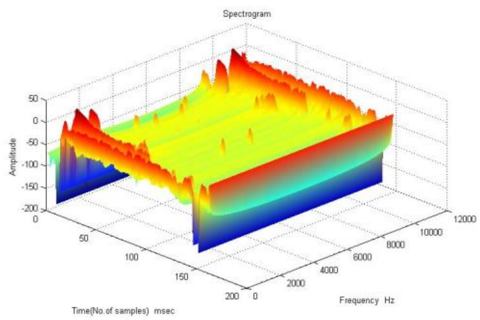
Large regularization filters reduce the noise energy but introduce more signal distortion. It is desirable that filter parameters are adjusted depending upon the nature of audio signals. In practice, however, they are selected empirically.

A time-frequency block thresholding estimator regularizes estimation by calculating a single attenuation factor over time-frequency blocks. The adaptive block thresholding chooses the sizes by minimizing an estimate of the risk. Best block sizes are computed by minimizing this estimated risk.

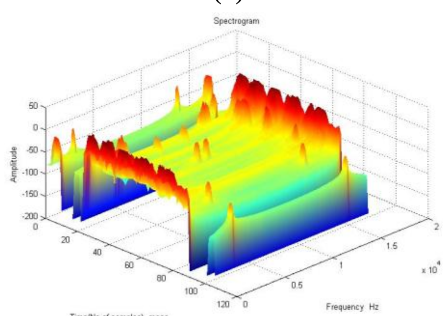
Cohen improved the decision-directed SNR estimator by combining a causal recursive temporal filter with a noncausal compactly supported time-frequency filter to get a first SNR estimation. He then refines this estimation in a Bayesian formulation by computing a new SNR estimation using the MMSE-SP attenuation from the first SNR estimate. This noncausal *a priori* SNR estimator has been combined with attenuation rules derived from Gaussian. Matz and Hlawatsch have also proposed to estimate the SNR with a rectangular time-frequency filter and to use it together with the empirical Wiener estimator.



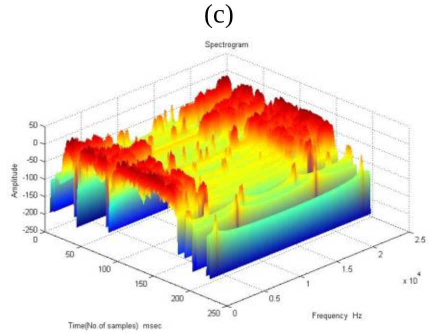
(a)



(b)

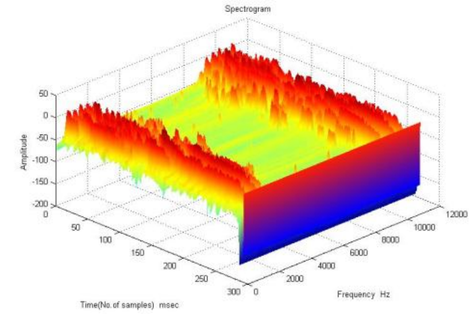


(c)

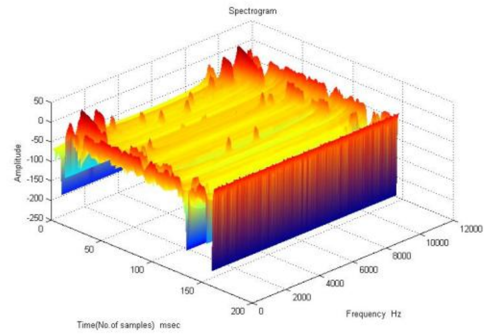


(d)

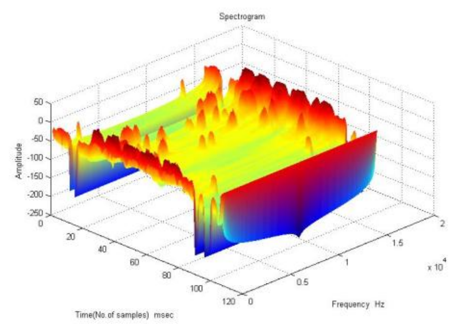
Figure 2. (a), (b), (c), (d): Block Thresholding of denoised 5dB Mozart, Piano, TIMIT and Speech signal



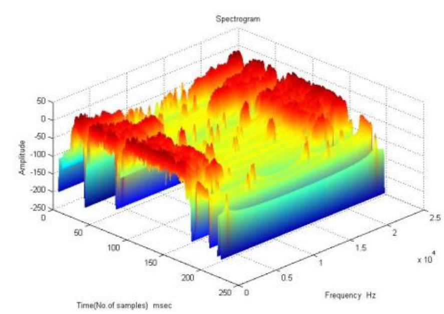
(e)



(f)



(g)



(h)

Figure 3. (e), (f), (g), (h): Block Thresholding of denoised 10dB Mozart, -5dBPiano, 20.63dBTIMIT and 8dBSpeech signal

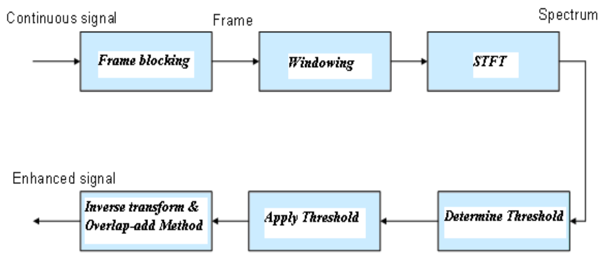


Figure.4. Block diagram of BT Estimator

IV. BLOCK THRESHOLDING ALGORITHM

Block thresholding was introduced in statistics by Cai and Silverman to obtain nearly minimax signal estimators. For audio signal denoising, we describe an adaptive block thresholding nondiagonal estimator that automatically adjusts all parameters.

A. Block Thresholding Estimator Algorithm

A time-frequency block thresholding estimator regularizes estimation by calculating a single attenuation factor over time-frequency blocks. The signal estimator $\hat{f}$ is calculated from the noisy data *y* with a constant attenuation factor a_i over each block B_i

$$\hat{f}[n] = \sum_{i=1}^{I} \sum_{(l,k)\in B_i} a_i Y[l,k] g_{l,k}[n]$$

To understand how to compute each a_i, one relates the Stein estimation risk, r $=E\ \{\|f- \hat{f}\ \|^2\}$ to the frame energy conversion and given by,

$$\gamma = E\left\{\left\| f - \hat{f}\right\|\right\} \leq \frac{1}{A}\sum_{i=1}^{I} \sum_{l,k\in B_i} E\left\{\left|F[l,k] - \hat{F}[L,K]\right|^2\right\}$$

Since $Y[l,k] = F[l,k] + \varepsilon[l,k]$ one can verify that the upper bound is minimized by choosing

$$a_i = 1 - 1/(\xi_i + 1)$$

where $\xi_i = \overline{F_i^2}/\overline{\sigma_i^2}$

is the average *a priori* SNR in B_i. It is calculated from

$$\overline{F_i^2} = \frac{1}{B_i^{\#}} \sum_{(l,k)\in B_i} |F[l,k]|^2 \text{ and } \overline{\sigma_i^2} = \frac{1}{B_i^{\#}} \sum_{(l,k)\in B_i} \sigma^2[l,k]$$

which are the average signal energy and noise energy in B_i and Where $B_i^{\#}$ = no. of coefficients within a block Bi

Cai and Silverman introduced block thresholding estimators that estimate the SNR over each B_i by averaging the noisy signal energy

$$\hat{\xi}_i = \frac{\overline{Y}_i^2}{\overline{\sigma}_i^2} - 1$$

Where

$$\overline{Y}_i^2 = \frac{1}{B_i^{\#}} \sum_{(l,k)\in B_i} \left|Y[l,k]\right|^2$$

The resulting attenuation factor a_i

$$a_i = \left(1 - \frac{\lambda}{\hat{\xi}_i + 1}\right)$$

A block thresholding estimator can thus be interpreted as a nondiagonal estimator derived from averaged SNR estimations over blocks. Each attenuation factor is calculated from all coefficients in each block, which regularizes the time-frequency coefficient estimation.

B. Stein Risk and Choice of λ

An upper bound of the risk of the block thresholding estimator is computed by analyzing separately the bias and variance terms. Observe that the upper bound of the oracle risk with blocks is always larger than that of the oracle risk without blocks, because the former is obtained through the same minimization but with less parameters as attenuation factors remain constant over each block

A time-frequency block thresholding estimator regularizes estimation by calculating a single attenuation factor over time-frequency blocks. The signal estimator $\hat{f}$ is calculated from the noisy data *y* with a constant attenuation factor a_i over each block B_i

$$\hat{f}[n] = \sum_{i=1}^{I} \sum_{(l,k)\in B_i} a_i Y[l,k] g_{l,k}[n]$$

To understand how to compute each a_i, one relates the Stein estimation risk, r $=E\ \{\|f- \hat{f}\ \|^2\}$ to the frame energy conversion and given by,

$$\gamma = E\left\{\left\| f - \hat{f}\right\|\right\} \leq \frac{1}{A}\sum_{i=1}^{I} \sum_{l,k\in B_i} E\left\{\left|F[l,k] - \hat{F}[L,K]\right|^2\right\}$$

Given a choice of block size and the residual noise probability level δ that one tolerates, the thresholding level λ .For each block width and length, λ is estimated using "Monte Carlo simulation" .The below Table shows the resulting λ with δ = 0.1%. Let us remark that for a block width $W > 1$, blocks that contain same number of coefficients, $B^{\#} = LXW$, have close λ values.

λ value	W =16	W = 8	W = 4	W = 2	W = 1
L = 8	1.5	1.8	2.0	2.5	2.5
L = 4	1.8	2.0	2.5	3.5	3.5
L = 2	2.0	2.5	3.5	4.7	4.7

Table.1 Thresholding level λ calculated with different block size B# = L X W and with δ = 0.1%.

The SURE was originally used to estimate the mean of a multivariate normal distribution. We use it to estimate the tuning parameters in the content of speech restoration.

Risk is calculated from the given equation

$$\hat{R}_i = \left\{ \frac{\left[B_i^{\#} + \lambda^2 B_i^{\#^2} - 2\lambda(B_i^{\#} - 2) \right]}{\left[\bar{Y}_i^2 / \bar{\sigma}_i^2 \right]} \Bigg|_{(\bar{Y}_i^2 < \lambda\bar{\sigma}_i^2)} + \left(\frac{\bar{Y}_i^2}{\bar{\sigma}_i^2} + 2B_i^{\#} \right) \Bigg|_{(\bar{Y}_i^2 < \lambda\bar{\sigma}_i^2)} \right\}$$

Where $B_i^{\#}$ = no. of coefficients within a block Bi

λ for Thresholding level

C. Adaptive Block Thresholding

A block thresholding segments the time-frequency plane in disjoint rectangular blocks of length L_i in time and width W in frequency. In the following by "block size" we mean a choice of block shapes and sizes among a collection of possibilities. The adaptive block thresholding chooses the sizes by minimizing an estimate of the risk.

The risk E $\{\|f\text{-}f'\|^2\}$ cannot be calculated since f is unknown, but it can be estimated with a Stein risk estimate. Best block sizes are computed by minimizing this estimated risk.

The Stein estimation risk, r $=E\ \{\|f - \hat{f}\ \|^2\}$ to the frame energy conversion.

If the noise is Gaussian white and the frame is an orthogonal basis then the noise coefficients are uncorrelated with same variance

The adaptive block thresholding groups coefficients in blocks whose sizes are adjusted to minimize the Stein risk estimate and it attenuates coefficients in those blocks. To regularize the adaptive segmentation in blocks, the time-frequency plane is first decomposed in macroblocks $M_{j,}$ j = 1, 2, 3……J. Each macroblock is segmented in blocks B_i of same size which means that $B_i^{\#}$ = p_j is constant over a macroblock $M_{j.}$

The Stein risk estimation over is M_j is (1/A) $\Sigma_{i\in M_j} R_i$

Several such segmentations are possible and we want to choose the one that leads to the smallest risk estimation. The optimal block size and hence p_j is calculated by choosing the block shape that minimizes $\Sigma_{i\in M_j} R_i$. Once the block sizes are computed, coefficients in each are attenuated, where is calculated.

In numerical experiments, each macroblock is segmented with 15 possible block sizes LxW with a combination of block length L = 8, 4, 2 and block width W = 16, 8, 4, 2, 1. The size of macroblocks is set to be equal to the maximum block size 8* 16.

Let $\tilde{f}$ be the block thresholding estimation from the noisy data y. This new attenuation factor is applied on the noisy time-frequency coefficients to reconstruct a second estimator.

$$\hat{f}[n] = (1/A)\sum_{l,k} a[l,k]Y[l,k]g_{l,k}[n]$$

V. ANALYSIS OF RESULTS AND DISCUSSIONS

The experiments presented below have been performed on various types of audio signals: "Piano" is a simple example that contains a single clear clavier stroke; "Mozart" is a musical excerpt that contains relatively quick notes played by a solo oboe; "TIMIT-M" and "TIMIT-F" are, respectively, male and female utterances taken from the TIMIT database . "TIMIT-M" and "TIMIT-F" are sampled at 16 kHz whereas all the other signals are sampled at 11 kHz. They were corrupted by Gaussian white noise of different amplitude. Short-time Fourier transforms with half-overlapping windows were used in the experiments.

These windows are the square root of Hanning windows of size 50 ms for "Piano" and "Mozart" and 20 ms for "TIMIT-M" and "TIMIT-F."1 For each sound, denoising with "partial noise removal" and "maximum noise removal" were applied: the former retains some low-amplitude residual noise; the latter removes almost all the original noise.

Both objective and subjective evaluations have been performed.

The objective measures are respectively the SNR and the segmental SNR defined as

$$SNR = 10\ \log_{10} \frac{\sum_{k=1}^{N} f^2[n]}{\sum_{k=1}^{N} (f\,[n] - y[n])^2}$$

and

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

Where H represents the number of frames in the signal, S is the number of samples per frame that corresponds to 32ms, and T(x) = min [max(x,-10), 35] confines the SNR in each frame to a perceptually meaningful range between 35db and -10db.

Segmental SNR has been shown to have a higher correlation with perceived quality than SNR.

SIGNAL & SNR	NON-DIAGONAL SNR	NON-DIAGONAL SSNR	DIAGONAL SNR	DIAGONAL SSNR
Mozart5dB	15.53	15.5248	13.97	14.4141
Mozart8dB	18.19	18.1867	14.26	14.4869
Mozart9.23dB	19.04	19.0345	14.32	14.5059
Mozart10dB	19.53	19.5243	14.35	14.5150
Mozart15dB	21.98	21.9794	20.24	15.4368
Piano4.75dB	17.11	17.1095	13.76	08.5116
TIMIT10.76dB	18.86	19.0279	17.72	14.3172
Speech5dB	17.52	17.5099	16.69	16.6823
Speech10dB	22.07	22.0618	16.91	16.9012

Table.2 Comparison of Four Types of Noisy Signals with Different Noise Levels.

VI. CONCLUSION

Nondiagonal time-frequency estimators are more effective than diagonal estimators to remove noise from audio signals because they introduce less musical noise. These nondiagonals estimators are derived from a time-frequency SNR estimation performed with parameterized filters applied to time- frequency coefficients. This is mainly used in Music and Speech Restoration.

REFERENCES

1. M. Bahoura and J. Rouat, "Wavelet speech enhancement based on timescale adaptation," *Speech Commun.*, vol. 48, no. 12, pp. 1620–1637, Dec. 2006.
2. 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.
3. 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.
4. T. Cai, "Adaptive wavelet estimation: A block thresholding and oracle inequality approach," *Ann. Statist.*, vol. 27, pp. 898–924, 1999.
5. T. Cai and B. W. Silverman, "Incorporation information on neighboring coefficients into wavelet estimation," *Sankhya*, vol. 63, pp. 127–148, 2001.
6. O. Cappé, "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.
7. S. Chang, Y. Kwon, S. Yang, and I. Kim, "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.
8. I. Cohen, "Speech enhancement using a noncausal *a priori* SNR estimator," *IEEE Signal Process. Lett.*, vol. 11, no. 9, pp. 725–728, Sep. 2004.
9. 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.
10. 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.
11. I. Cohen and B. Berdugo, "Speech enhancement for non-stationary noise environments," *Signal Process.*, vol. 81, no. 11, pp. 2403–2418, Nov. 2001.
12. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean square error short-time spectral amplitude estimator," *IEEE.Trans. Acoust., Speech, Signal Process.*, vol. 32, no. 6, pp. 1109–1121, Dec. 1984.
13. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean square error log-spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-33, no. 2, pp. 443–445, Apr. 1985.
14. R. R. Coifman and D. L. Donoho, "Translation-Invariant De-Noising," in *Lecture Notes in Statistics: Wavelets and Statistics*, A. Antoniadisand G. Oppenheim, Eds. Berlin, Germany: Springer-Verlag, 1995.
15. Y. Ephraim and H. L. V. Trees, "A signal subspace approach for speech enhancement," *IEEE Trans. Speech Signal Process*, vol. 3, no. 4, pp. 251–266, Jul. 1995.
16. G. Matz and F. Hlawatsch, "Minimax robust nonstationary signal estimation based on a p-point uncertainty model," *J. Franklin Inst. (Special Issue on Time-Frequency Signal Analysis and Applications)*, vol. 337, no. 4, pp. 403–419, Jul. 2000.
17. G. Matz, F. Hlawatsch, and A. Raidl, "Signal-adaptive robust timevarying Wiener filters: Best subspace selection and statistical analysis," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Salt Lake City, UT, May 2001, pp. 3945–3948.
18. R. J. McAulay and M. L. Malpass, "Speech enhancement using soft decision noise suppression filter," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-28, no. 2, pp. 137–145, Apr. 1980.
19. S. R. Quackenbush, T. P. Barnwell, and M. A. Clements, *Objective Measures of Speech Quality*. New York: Prentice-Hall, 1988.
20. C. Stein and W. James, "Estimation with quadratic loss," in *Proc. 4th Berkeley Symp. Mathematical Statistics Probability 1*, Berkeley, CA, 1961, pp. 361–379.
21. J. S. Walker and Y.-J. Chen, *Denoising Gabor Transforms*[Online].Available: http://www.uwec.edu/walkerjs/media/DGT.pdf, preprint
22. G. Yu, E. Bacry, and S. Mallat, "Audio signal denoising with complex wavelets and adaptive block attenuation," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Apr. 2007, vol. 3, pp. III-869–III-872.
23. H. Sheikhzadeh and H. R. Abutalebi, "An improved wavelet-based speech enhancement system," *EUROSPEECH*, pp. 1855–1858, 2001.
24. Y. Shao and C. H. Chang, "A generalized time-frequency subtraction method for robust speech enhancement based on wavelet filter bank modeling of human auditory system," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 4, pp. 877–889, Aug. 2007
25. I. J. Kim, S. I. Yang, and Y. Kwon, "Speech enhancement using adaptive wavelet shrinkage," in *Proc. IEEE Int. Symp. Industrial Electronics*, 2001, vol. 1, pp. 501–504.
26. Y. Ephraim, H. Lev-Ari, and W. J. J. Roberts, "A brief survey of speech enhancement," in *The Electronic Handbook*. Boca Raton, FL: CRC Press, 2005.
27. J. S. Garofolo, Getting Started with the DARPA TIMIT CD-ROM: An Acoustic Phonetic Continuous Speech Database. National Institute of Standards and Technology (NIST), Gaithersburgh, MD, 1988.
28. D. Malah, R. V. Cox, and A. J. Accardi, "Tracking speech-presence uncertainty to improve speech enhancement in mon-stationary noise environments," presented at the IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP), Phoenix, AZ, Mar. 1999.
29. Database. National Institute of Standards and Technology (NIST), Gaithersburgh, MD, 1988.
30. D. Malah, R. V. Cox, and A. J. Accardi, "Tracking speech-presence uncertainty to improve speech enhancement in mon-stationary noise environments," presented at the IEEE Int. Conf. Acoust., Speech, Signal Processing (ICASSP), Phoenix, AZ, Mar. 1999.
31. G. Yu, E. Bacry, and S. Mallat, "Audio signal denoising with complex wavelets and adaptive block attenuation," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Apr. 2007, vol. 3, pp. III-869–III-872.
32. J. Yang, "Frequency domain noise suppression approaches in mobile telephone systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 1993, vol. 2, pp. 363–366.
33. P. J. Wolfe and S. J. Godsill, "Simple alternatives to the Ephraim and Malah suppression rule for speech enhancement," in *Proc. IEEE Workshop Statistical Signal Processing*, Aug. 2001, pp. 496–499.
34. J. W. Seok and K. S. Bae, "Speech enhancement with reduction of noise components in thewavelet domain," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, 1997, vol. 2, pp. 1323–1326.
35. Y. Shao and C. H. Chang, "A generalized time-frequency subtraction method for robust speech enhancement based on wavelet filter bank modeling of human auditory system," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 4, pp. 877–889, Aug. 2007.
36. H. Sheikhzadeh and H. R. Abutalebi, "An improved wavelet-based speech enhancement system," *EUROSPEECH*, pp. 1855–1858, 2001.
37. K. V. Sørensen and S. V. Andersen, "Speech enhancement with natural sounding residual noise based on connected time-frequency speech presence regions," *EURASIP J. Appl. Signal Process.*, vol. 18, no. 18, pp. 2954–2964, 2005.
38. M. Johnson, X.Yuan, and Y. Ren, "Speech signal enhancement through adaptive wavelet thresholding," *Speech Commun.*, vol. 49, no. 2, Feb. 2007.
39. N. S. Kim and J. H. Chang, "Spectral enhancement based on global soft decision," *IEEE Signal Process. Lett.*, vol. 7, no. 5, pp. 108–110, May 2000.
40. I. J. Kim, S. I. Yang, and Y. Kwon, "Speech enhancement using adaptive wavelet shrinkage," in *Proc. IEEE Int. Symp. Industrial Electronics*, 2001, vol. 1, pp. 501–504.

41. M. Li, H. G. McAllister, N. D. Black, and D. T. A. Perez, "Perceptual time-frequency subtraction algorithm for noise reduction in hearing aids," *IEEE Trans. Biomed. Eng.*, vol. 48, no. 9, pp. 979–988, Sep. 2001.
42. J. S. Lim and A. V. Oppenheim, "Enhancement and bandwidth compression of noisy speech," *Proc. IEEE*, vol. 67, Dec. 1979.
43. S. F. Lei and Y. K. Tung, "Speech enhancement for nonstationary noises by wavelet packet transform and adaptive noise estimation," in *Proc. Int. Symp. Intelligent Signal Processing Communication Systems*, Dec. 2005, pp. 41–44.
44. C. T. Lu and H. C.Wang, "Enhancement of single channel speech based on masking property andwavelet transform," *Speech Commun.*, vol. 41, no. 2, pp. 409–427(19), Oct. 2003.

AUTHORS PROFILE

K. Sreedhar, received the B.Tech. degree in Electronics and Communication Engineering from JNTUH University, Hyderabad, India in 2005 and M.Tech degree in Communication Systems from JNTUH University, Hyderabad, India in 2009. He attended the International Conference on Technology and Innovation at Chennai. He attended the National Conference at Coimbatore, Tamilnadu, India on INNOVATIVE IN WIRELESS TECHNOLOGY. He is currently working as an Assistant professor in Electronics and Communication Engineering department in VITS-SET (N9) Karimnagar, Andhra Pradesh, India. He has a Life Member ship in ISTE. His areas of interests are Digital Signal Processing and Image Processing.