ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 7, July 2013

IMPLEMENTATION OF TIME FREQUENCY BLOCK THRESHOLDING ALGORITHM IN AUDIO NOISE REDUCTION

K.P. OBULESU[1] **P. UDAY KUMAR**[2]
[1]PG Scholar of VLSI System Design in Electronics & Communication Engineering
[2]Assistant Professor in Department of Electronics & Communication Engineering
AVR & SVR College of Engineering & Technology, Nandyal, A.P, India

ABSTRACT

Audio signals are synthetic signals, in which music or speech, are often corrupted by noise during recording and transmission. Speech enhancement is a long standing problem with numerous applications ranging from hearing aids, to coding and automatic recognition of speech signals etc. and assume that the noise is additive and statistically independent of the signal. Audio denoising procedures are designed to attenuate the noise and retain the signal of interest. Reduction of noise from audio signals has two methods, Diagonal & Non Diagonal audio denoising algorithms.

In this paper, Non diagonal method is used in which Block parameters are automatically adjusted to the nature of the audio signal by minimizing a Stein estimator which is calculated analytically from noisy signal values. This Block thresholding method eliminates "musical noise" by grouping Time-frequency coefficients in blocks before being attenuated. This algorithm is robust to variations of signal structures such as short transients and long harmonics. Numerical experiments demonstrate the performance and robustness of this procedure through objective and subjective evaluations. This approach can further be improved by optimizing the SNR estimation with parameterized filters that rely on stochastic audio models and this project is implemented in MATLAB 7.4.0 software.

Index Terms: Audio Denosing, Block Thresholding, Log Spectral Amplitude.

1. INTRODUCTION

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. Applications such as music and speech restoration are numerous. Figure 1 shows the methodology in which Diagonal time-frequency audio denoising algorithms attenuate the noise by processing each window Fourier or wavelet coefficient independently, with empirical wiener power subtraction or thresholding operators[2][3]. This algorithm has limited performance which create isolated time frequency structures that are perceived as a musical noise which is due to lack of time-frequency regularity.

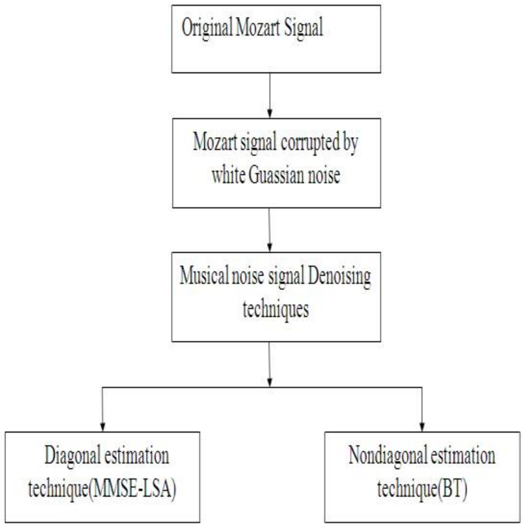

Fig1.Block diagram for Methodology

The musical noise is strongly attenuated with nondiagonal time frequency estimators that regularize the estimation by recursively aggregating time frequency coefficients. For audio time frequency denoising, block thresholding regularizes the estimate and is effective in musical noise reduction. Block parameters are automatically adjusted by minimizing a Stein estimator of the risk, and are calculated analytically from the noisy signal values. But in Non Diagonal audio denoising algorithm (Block Thresholding) through adaptive time frequency block thresholding method, improves the asymptotic decay of diagonal thresholding estimators which regularizes the estimate, and does not create isolated time frequency coefficients which are responsible for musical noise.

2. LITERATURE SURVEY

A new approach to improve the performance of speech enhancement techniques based on wavelet thresholding. First, space adaptation of the threshold is obtained by extending the principle of the level dependent threshold to the Wavelet Packet Transform (WPT)[5].A wavelet function estimation via the approach of block thresholding and ideal adaptation with oracle. Oracle inequalities are derived and serve as guides for the selection of smoothing parameters.

Based on an oracle inequality and motivated by the data compression and localization properties of wavelets, an adaptive wavelet estimator for non parametric regression is proposed and the optimality of the procedure is investigated [2].An asymptotic and numerical properties of a class of block thresholding estimators for wavelet regression.

Consider the effect of block size on global and local adaptivity and the choice of thresholding constant. The optimal rate of convergence for block thresholding with a given block size is derived for both the global and local estimation, the block thresholding as a hypothesis testing problem. The combined results lead naturally to an optimal choice of block size and thresholding constant [3]. Speech is estimated by Wiener filtering in the wavelet packet domain, modified by the signal presence probability. A magnitude decision-directed estimator for the variance of speech, which is closely related to the decision directed estimator of Ephraim and Malah[7].

Consider three different statistical models, two fidelity criteria, and two approaches for the estimation of the variances of the STFT coefficients. The statistical model is either Gaussian, Gamma; the fidelity criteria include minimum mean-squared error (MMSE) of the STFT coefficients and MMSE of the log-spectral amplitude (LSA)[8].Log-Spectral Amplitude estimator, which minimizes the mean-square error of the log-spectra for speech signals under signal presence uncertainty. An estimator for a priori signal-to-noise ratio, and introduce an efficient estimator for the a priori speech absence probability.

A fully Bayesian approach for sparse audio signal regression in an union of two bases, with application to audio denoising. One basis aims at modeling tonal parts and the other at modeling transients. The noisy signal is decomposed as a linear combination of atoms from the two basis, plus a residual part containing the noise[9].Wavelet shrinkage using thresholding is asymptotically optimal in a minimax mean-square error sense over a variety of smoothness spaces. However, for any given signal, the MSE-optimal processing is achieved by the wiener filter, which delivers substantially improved performance [11].

Block thresholding methods have been proposed by Hall and Picard (1995) as a means of obtaining increased adaptivity when estimating a function using wavelet methods. For example, it has been shown that block thresholding reduces mean squared error by rendering the estimator more adaptive to relatively subtle, local changes in curvature, of the type that local bandwidth choice is designed to accommodate in traditional kernel methods. The block thresholding also provides extensive adaptivity to many varieties of aberration, including those of chirp and Doppler type[10].

The standard power spectral subtraction method resulting in superior speech quality and largely reduced musical noise. Threshold value which minimize Bayesian risk. And using entropy, part of the noisy signal into unvoiced signal section and the other signal section to apply each threshold value for each section. Experiment result shows that proposed algorithm is more efficient[12].

3. GAUSSIAN NOISE ANALYSIS

Gaussian noise is properly defined as the noise with a Gaussian amplitude distribution. Gaussian noise is a type of statistical noise in which the amplitude of the noise follows that of a Gaussian distribution whereas additive white Gaussian noise is a linear combination of a Gaussian noise and a white noise

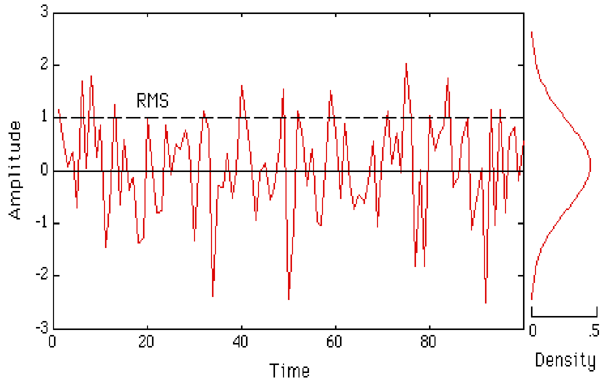

Fig 2. Representation of Gaussian Noise

Gaussian noise is noise that has a random and normal distribution of instantaneous amplitudes over time. The Figure 2 shows a sample of Gaussian noise with the normal (bell) curve drawn to the right to represent the distribution of instantaneous amplitudes.

3.1 CHARACTERISTS OF WHITE NOISE ANALYSIS

Generalized white noise functional (white noise distributions) serve key roles in stochastic analysis. White noise is viewed as a system of idealized elemental random variables. Idealized elemental random variables would be fine if the causality is satisfied, namely for any instant 't' (assumed to be present time) the information contained in the given random system in the past is the same as that contained in idealized elemental random variables, and are constructed up to 't'. If the system is parameterized by continuous time, denoted by 't', and is a generalized stochastic process under mild assumptions. A local block wise hard thresholding procedure with a block size of the order (log n) 2 where n is the sample size. In particular a block size of order log n is optimal in the sense and leads to an estimator which is both globally and locally adaptive. The block size and threshold level play important roles in the performance of a block thresholding estimator. At each resolution level, the procedure, sure block, chooses the block size and threshold empirically by minimizing Stein's Unbiased Risk Estimate (SURE). By empirically selecting both the block size and threshold and allowing them to vary from resolution level to resolution level, the sure block estimator has significant advantages over the more conventional wavelet thresholding estimators with fixed block sizes. Graphical models are a framework for representing and exploiting prior conditional independence structures within distributions using graphs. In the Gaussian case, these models are directly related to the sparsity of the inverse covariance (concentration) matrix and allow for improved covariance estimation with lower computational complexity.

3.2 SIGNAL TO NOISE RATIO The ratio of the amplitude of a desired signal at any point to the amplitude of noise signals at the same point, often expressed in decibels the peak value is usually used for pulse noise, while the root-mean-square (rms) value is used for random noise. The quantity measures, relationship between the strength of an information-carrying signal in an electrical communications system and the random fluctuations in amplitude, phase, and frequency superimposed on the signal and collectively referred to as noise.

4. NOISE ESTIMATION TECHNIQUES

In probability theory and statistics, the variance is used as one of several descriptors of a distribution, and is describes how far values lie from the mean. In particular, the variance is one of the moments of a distribution, and forms the part of systematic approach to distinguishing between probability distributions. However Audio denoising techniques are classified as Diagonal Estimation technique and Non Diagonal Estimation technique

4.1 DIAGONAL ESTIMATION TECHNIQUE

Simple time-frequency denoising algorithm computes each attenuation factor a[l,k] only from the corresponding noisy coefficients y[l,k] and are called diagonal estimators. The algorithm has a limited performance and produce a musical noise. The expression for minimize the quadratic estimation risk.[1]

$$r = E\{\left\| f - \hat{f} \right\|^2\} \le 1/A \sum_{l,k} E\{\left| F[l,k] - \hat{F}[l,k] \right|^2\} \quad \ldots(4.1)$$

STAIN UNBIASED RISK ESTIMATE (SURE)

Stein Unbiased Risk Estimate (SURE)[1] is an unbiased estimator of the mean-squared error of a given estimator, in a deterministic estimation scenario. In other words, and is provides an indication of the accuracy of a given estimator. In deterministic estimation, the true mean-squared error of an estimator generally depends on the value of the unknown parameter, and is cannot be determined completely.

Let θ ∈ R^n be an unknown deterministic parameter and let '*x*' be a measurement vector and is distributed normally with mean θ and covariance $\sigma^2 I$. Suppose *h*(*x*) is an estimator of 'θ' from '*x*'. The importance of SURE is an unbiased estimate of the mean-squared error (or squared error risk) of *h*(*x*).

E(SURE(h))=MSE(h)

Optimal attenuation factor is given[1] by

A. $a[l,k] = 1 - 1/\xi[l,k] + 1$ **…(4.2)**

Where $\xi[l,k] = F^2[l,k]/\sigma^2[l,k]$

Diagonal estimation of the SNR $\xi[l,k]$ are computed from a posteriori SNR defined by

$$\gamma[l,k] = \left| Y[l,k] \right|^2 / \sigma^2[l,k] \ldots(4.3)$$

And an unbiased estimator $\hat{\xi}[l,k] = \gamma[l,k] - 1$

An attenuation factor a[l,k] of the diagonal estimator only depends upon y[l,k] with no time-frequency regularization. The 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, and are perceived as a musical noise.

4.2 NONDIAGONAL ESTIMATION TECHNIQUE

To reduce musical noise as well as the estimation risk, the main task is to estimate the a priori SNR $\xi[l,k]$ with a time-frequency regularization of the a posteriori SNR $\gamma[l,k]$. The resulting attenuation factors depend upon the data values $Y[l',k']$ for (l',k') in a whole neighborhood of (l,k) and the resulting estimator can be written as[1]

$$\hat{f}[n] = 1/A \sum_{l,k} a[l,k] Y[l,k] g_{l,k}[n] \quad \ldots(4.4)$$

The above equation is said to be nondiagonal. The regularization of the SNR estimation reduces musical noise as well as the estimation risk can be written as

$$r = E\{\left\| \hat{f} - f \right\|^2\} \ldots(4.5)$$

MEAN SQUARE ESTIMATOR (MSE)

MSE of an estimator is one of many ways to quantify the difference between an estimator $\hat{f}$ and true value of the error f. The MMSE estimator is defined as the estimator achieving minimal MSE.

MSE= $E[(\hat{f} - f)^2]$

MMSE-LSA audio denoising algorithm robust and effective, and is difficult to tune and adjust for the tradeoff between noise reduction and distortion.

The problem of estimating a priori SNR of a spectral component in analysis frame, a priori SNR should be preestimated in each analysis frame due to non stationary of the audio signal. In priory SNR estimation two methods are considered, the first one is to utilize an audio spectral component variants and second one is based on a decision directed estimation method. In decision directed estimation, an estimation of a priory SNR of estimator is found to be very useful and is combined with either the MMSE.

The noise separation technique has been used recently for the restoration of degraded audio recording because, is a free of the frequently encountered musical noise artifact and is demonstrated of how the aircraft is actually eliminated without bringing distortion to the recorded signal even if the noise is only poorly stationary. A priori SNR estimation is of major consequence in audio enhancement application. A non causal recursive estimator for the a priori SNR based on Gaussian model. Optimal estimators in the STFT domain are often on a Gaussian statistical model. Accordingly, the individual short term spectral components of the audio and noise signals are modeled as statistically independent Gaussian random variables.

A short time spectral amplitude estimator minimizes the mean square error of the spectral magnitude whereas log spectral amplitude estimator minimizes the mean square error of log spectra. The sequence to audio spectral variances is a random process, and is generally correlated with the sequence of audio spectral magnitudes. Causal and non causal estimators for the a priori SNR derived with the model assumption and the estimation of the audio spectral components. Statistical modeling of audio signals in the Short Time Fourier Transform(STFT) domain has recently received

ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 7, July 2013

much attention, but is still a puzzling problem. Hence, consider the model of individual STFT expansion co-efficiency of the audio signal as zero mean statistically independent Gaussian random variables, and is enables to derive useful minimum mean square error estimators for the Short Time Spectral Amplitude(STSA), as well as the Log Spectral Amplitude(LSA) and is underlies the design of many audio enhancement algorithms.

5. BLOCK THRESHOLDING

Thresholding gives amplitude separation. To well separate signal and noise thresholding is used. The purpose of a filter is for frequency separation and frequency signal restoration. For amplitude separation thresholding is used. Thresholding are two types as Soft Thresholding & Hard Thresholding as shown in figure 3.

SOFT THRESHOLDING

In Soft thresholding the coefficients which are within the threshold value are consider as zero and subtract the threshold value from the coefficients which are above the threshold value. Depending upon the changes in the noise threshold value will change.

$$y(n) = x(n) - \in \text{ for } |x(n)| > \in$$

$$y(n) = 0 \text{ for } |x(n)| \leq \in \quad \ldots.(5.1)$$

HARD THRESHOLDING

In Hard thresholding the coefficients which are within the Threshold value are consider as zero and the coefficient which are above the Threshold value remain same and are considered as actual coefficients of the signal. In hard thresholding the threshold value is fixed.

$$y(n) = x(n) \text{ for } |x(n)| > \in \qquad y(n) = 0 \text{ for } |x(n)| < \in \quad \ldots.(5.2)$$

Where y(n) is an output signal of thresholder, x(n) is an input signal of thresholder and ‘ $\in$ ’ is threshold level.

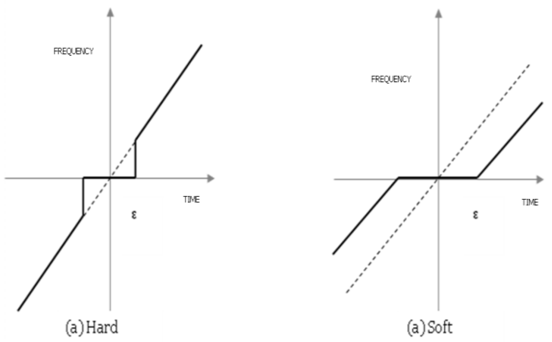

Fig 3.Representation of hard and soft thresholding

The time-frequency plane (l, k) is segmented in ‘i’ blocks ‘B_i’ whose shape may be chosen arbitrarily. The signal estimator ‘ $\hat{f}$ ’ is calculated from the noisy data ‘y’ with a constant attenuation factor ‘a_i’ over each block ‘B_i’.

$$\hat{f}[n] = \sum_{i=1}^{I} \sum_{(l,k)\in B_i} a_i Y[l,k] g_{l,k}[n] \ldots(5.3)$$

$$Y[l,k] = F[l,k] + \in [l,k] \quad \ldots(5.4)$$

Can be minimized by choosing an attenuation factor

$$a_i = 1 - 1/\xi_i + 1 \qquad \ldots(5.5)$$

Where σ_i^2 is variance

$$\xi_i = \ddot{F}_i^2 / \sigma_i^2 \quad \ldots(5.6)$$

is the average a priori SNR in ‘B_i’, and is calculated from

$$\ddot{F}_i^2 = 1/B_i^{\#} \sum_{(l,k)\in B_i} |F[l,k]|^2$$

$$\ddot{\sigma}_i^2 = 1/B_i^{\#} \sum_{(l,k)\in B_i} \sigma^2[l,k] \ldots(5.7)$$

Which are the average signal energy and noise energy in ‘ B_i ’ and ‘ $B_i^{\#}$ ‘ is the number of coefficients $(l,k) \in B_i$.

The block thresholding estimators that estimate the SNR over each ’B_i’ by averaging the noisy signal energy

$$\hat{\xi}_i = \ddot{Y}_i^2 / \sigma^2{}_i - 1 \ldots(5.8)$$

Where ‘ $\hat{\xi}_i$ ’ is an unbiased estimator

$$\ddot{Y}_i^2 = 1/B_i^{\#} \sum_{(l,k)\in B_i} |Y[l,k]|^2 \ldots.(5.9)$$

and attenuation factor

$$a_i = (1 - \lambda / \xi_i + 1) \ldots.(5.10)$$

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

CHOICE OF THRESHOLD LEVEL λ

The parameter ‘λ’ is set depending upon ‘ $B_i^{\#}$ ’ by adjusting the residual noise probability $prob\{\ddot{\in}^2 \rangle \lambda\sigma^2\} = \delta$. as shown in table I where δ is residual noise.

Table I Thresholding level λ calculated for different Block size $B_i^{\#}$ with δ =0.1%[1]

$B_i^{\#}$	4	8	16	32	64	128
λ	4.7	3.5	2.5	2.0	18	1.5

Let $B_i^{\#} = L_i \times W_i$ be a rectangular block size, $L_i \geq 2$ and $W_i \geq 2$ are respectively the block length in time and block width in frequency.

ADAPTIVE BLOCK THRESHOLDING

An adaptive audio block thresholding algorithm adapts all parameters to the time-frequency regularity of the audio signal. The adaptation is performed by minimizing a Stein unbiased risk estimator calculated from the data. A block thresholding segments shown in figure 4 the time-frequency plane disjoint rectangular blocks of length L_i in time and width W_i in frequency. Block size is the choice of block shapes and size among a collection of possibilities. The adaptive block thresholding chooses the sizes by minimizing an estimate of the risk.

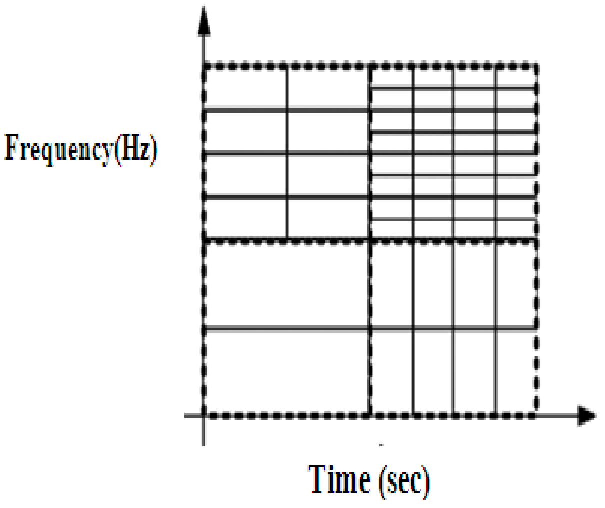

Fig 4. Partition of macro block into blocks of different sizes[1]

The adaptive block thresholding chooses the sizes by minimizing an estimate of the risk.

The Risk estimator

$$R_i = \dddot{\sigma}^2 (B_i^\# + (\lambda^2 B_i^\# - 2\lambda(B_i^\# - 2))/(\dddot{Y}_i^2 / \dddot{\sigma}_i^2) + B_i^\# (\dddot{Y}_i^2 / \dddot{\sigma}_i^2 - 2)) \quad ....(5.11)$$

The size of macro blocks is set to be equal to the maximum block size 8X16 The SNR defined as

$$SNR = 10\log_{10} \sum_{n=0}^{N-1} f^2[n] / \sum_{n=0}^{N-1} (f[n] - \hat{f}[n])^2$$

Table II Percentages of the different block size selected by the block thresholding for Mozart[1]

Mozart signal	W=16	W=8	W=4	W=2	W=1
L=8	25.3	10.4	5.2	4.0	11.5
L=4	10.7	4.2	3.0	1.9	3.6
L=2	3.1	2.5	2.2	3.0	7.3

The adaptive block thresholding groups coefficients in blocks whose sizes LXW are adjusted to minimize the stein risk estimate and it attenuates coefficients in those blocks. In numerical experiments, each macroblock is segmented with 15 possible block sizes with a combination of block length L=8,4,2 and block width w=16,8,4,2,1 as shown in Table II.

ADAPTIVE BLOCKS:

Four threshold selection methods are universal, mini-max, Stein's Unbiased Estimate of Risk (SURE), and Minimum Description Length (MDL) criteria. The application of the threshold to the wavelet coefficients includes global (hard, soft, garrote, and firm), level-dependent, data-dependent, Translation Invariant (TI), and Wavelet Package Transform (WPT) thresholding methods. Time frequency dependent threshold estimation Wavelet denoising involves thresholding in which coefficients below a specified value are set to zero, and is called hard thresholding. Whereas soft-thresholding simply shrinks or scales coefficients below the threshold value [3][4]. A general optimal universal threshold for the Gaussian white noise under a mean squared error criterion [6]. An adaptive block technique is used for large blocks where modules of each coefficient do not have many variations and otherwise refer small ones. The effect of block size on global and local adaptivity is the choice of thresholding constant. The optimal rate of convergence for block thresholding with a given block size is derive both the global and local estimation.

The block size for achieving the global and local adaptivity, consider the choice of thresholding constant for a given block size by treating the block thresholding as a hypothesis testing problem. The combined results lead naturally to an optimal choice of block size and thresholding constraint. An estimator is globally adaptive and can automatically adjust to varying level of overall regularity of the target function. The non causal estimator is capable of discriminating audio onsets and noise irregularities. The non causal estimator yields a higher improvement in the segmental SNR.

Most audio enhancement systems are based on the decomposition of audio and noise in frequency domain using the short time Fourier transform and the modification of the spectral coefficiency with a gain function, and often produce a new randomly fluctuating type of noise, referred to as musical noise. The phenomenon can be explained by noise or signal to noise ratio estimation errors leading to spurious peaks in the processed spectrum. When the enhanced signal is reconstructed in the time domain, the peaks result in short sinusoidal and their frequencies vary from frame to frame. The Absolute Threshold of Hearing (ATH) is the minimum sound level of a pure tone and an average ear with normal hearing can hear with no other sound present. The absolute threshold relates to the sound and can just be heard by the organism. The absolute threshold is not a discrete point, and is therefore classed as the point at which a response is elicited a specified percentage of the time.

ALGORITHM FOR ADAPTIVE TIME FREQUENCY BLOCK THRESHOLDING

1. Determine the SNR of the musical noise signal.
2. Apply Hanning window.
3. Apply half overlapped window
4. Apply STFT.
5. Determine Block size by considering Stein Unbiased Risk Estimator.
6. Apply Block Thresholding.
7. Apply inverse STFT.
8. Obtain denoised signal.

9. Compare the SNR of the musical noise signal and the denoised signal. Fourier transformer converts the signal in time domain to frequency domain. STFT converts the signal from time domain to time-frequency domain and it gives spectral components. In STFT a short time duration of a signal is considered as a window and Fourier transform is applied to that 'T' duration window

LINEAR MINIMUM MEAN SQUARE ESTIMATOR

The MMSE algorithms all rely on an assumption of quasi-stationary and an assumption of uncorrelated spectral components in the signal. The quasi-stationary assumption requires short-time processing. At the same time, the assumption of uncorrelated spectral components can be warranted by assuming the signal to be infinitely long and wide sense stationary [7][8].The temporal power localization within a short frame can be modeled as a non stationary of the signal is not resolved by the short time processing. As a consequence of the stationary and long frame assumptions, the MMSE model the frequency domain signal covariance matrix as a diagonal matrix. Compared to the mentioned frequency domain MMSE method, the known signal subspace methods implicitly avoid the infinite data length assumption, so that the inter frequency correlation caused by the finite length effect is accommodated.

In general, the MMSE model attenuates more in the spectral valleys than the spectral subtraction methods. The signal subspace methods are designed to shape the residual noise power spectrum for a better spectral masking, where the masking threshold is found experimentally. The temporal masking property is not employed by current frequency domain MMSE estimators and the signal subspace models. Both the frequency domain signal covariance matrix and filtering matrix are estimated as complex-valued full matrices, which means the information about inter frequency correlation are not lost and the amplitude and phase spectra are estimated jointly. Specifically, the linear prediction based source filter model to estimate the signal covariance matrix, upon which a time domain or frequency domain LMMSE estimator is built. In the estimation of the signal covariance matrix, the matrix is decomposed into a synthesis filter matrix and an excitation matrix.

The synthesis filter matrix is estimated by a smoothed power spectral subtraction method followed by an autocorrelation Linear Predictive Coding (LPC) method. The excitation matrix is a diagonal matrix with the instantaneous power of the LPC residual as its diagonal elements. The instantaneous power of the LPC residual is estimated by a modified Multi Pulse Linear Predictive Coding (MPLPC) method. Compared to several quasi stationary based estimators, the LMMSE estimator results in a lower spectral distortion to the enhanced audio signal while having higher noise reduction capability. The algorithm applies more attenuation in the valleys between pitch impulses in time domain, while small attenuation is applied around the pitch impulses.

An optimally modified log spectral amplitude estimator, which minimizes the mean-square error of the log-spectra for speech signals under signal presence uncertainty. An estimator for the a priori signal to noise ratio (SNR) introduce an efficient estimator for the a priori speech absence probability.

LOG SPECTRAL AMPLITUDE

The log-spectral amplitude (LSA) estimator, proved very efficient in reducing musical residual noise phenomena. Modification under signal presence uncertainty is obtained by multiplying the spectral gain by the conditional audio presence probability, which is estimated for each frequency bin and each frame. In this paper, we present an optimally modified LSA (OM-LSA) estimator. The spectral gain function is obtained as a weighted geometric mean of the hypothetical gains associated with signal presence and absence. The priori SAP is estimated for each frequency bin and each frame by a soft-decision approach, which exploits the strong correlation of speech presence in neighboring frequency bins of consecutive frames. Objective and subjective evaluation in various environmental conditions show that the proposed modification approach is advantageous, particularly for low input SNRs and nonstationary noise. Excellent noise reduction can be achieved even in the most adverse noise conditions, while avoiding musical residual noise and the attenuation of weak speech components.A frequency domain optimal linear estimator is proposed which incorporates the masking properties of the human auditory system to make the residual noise distortion inaudible. The use of wavelet-threshold multitaper spectra is also proposed for frequency-domain speech enhancement methods as an alternative to the traditional fast Fourier transform (FFT)-based magnitude spectra.

6. RESULTS & CONCLUSIONS

The comparison between Block thresholding and minimum mean square estimation-log spectral amplitude algorithms is shown in table III.

Table III. Different Signals and their SNR

Signal & SSNR(in db)	Block Thresholding	MMSE-LSA
Mozart 2db	13.52	5.78
Mozart 4db	14.90	7.62
Mozart 5db	16.88	10.54
Mozart 10db	17.85	13.50
Mozart 15db	21.76	17.72
Mozart 25db	24.81	21.07

The spectrogram is Mozart signal with 15db noise. Mozart signal is a type of piano signal, whereas Mozart signal with noise containing signal has original Mozart signal and guassian noise. After applying block thresholding and Minimum Mean Square Estimation Log spectral Amplitude algorithm to the noise Mozart signal, noiseless Mozart signals is obtained. With the help of signal to noise ratio values of the signals performance of the both the techniques will be decided whether which technique is better to eliminate the musical noise. For different types of Mozart signals spectrogram will varies. Comparison of both the algorithms of signal shown in figure 5 & 6. Comparison of both the algorithm shows signal to noise ratio values. From the above table BT algorithm have high signal to noise ratio value compare to LSA. So, diagonal technique is more effective than nondiagonal technique to eliminate the noise.

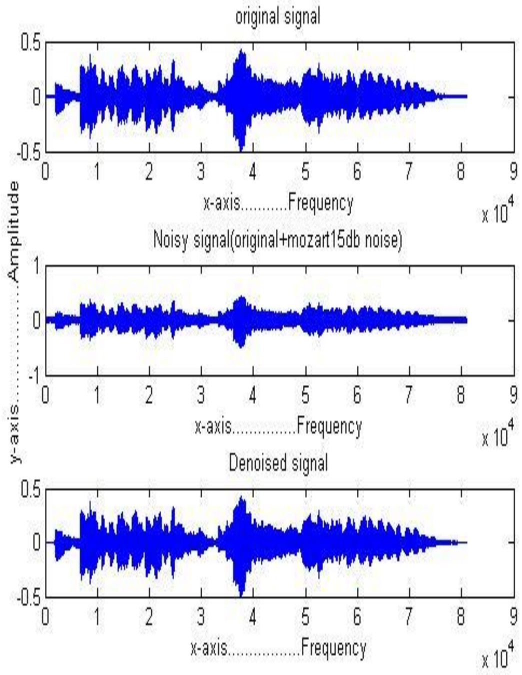

Fig 5. Comparison of different waveforms with Block Thresholding Technique

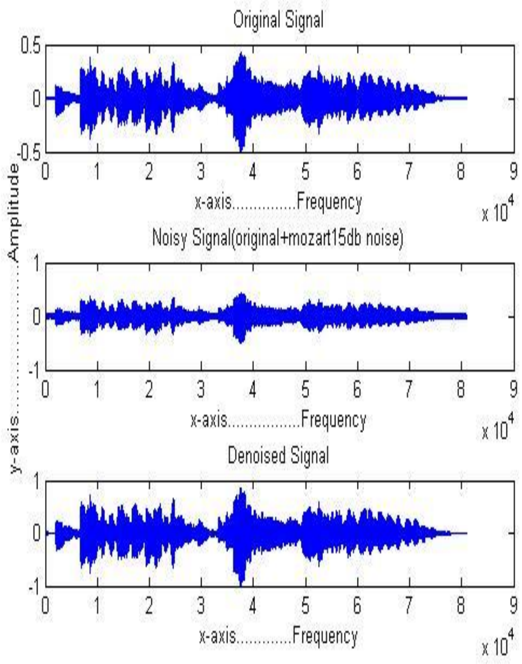

Fig 6.Comparison of different waveforms with Log Spectral Amplitude Method

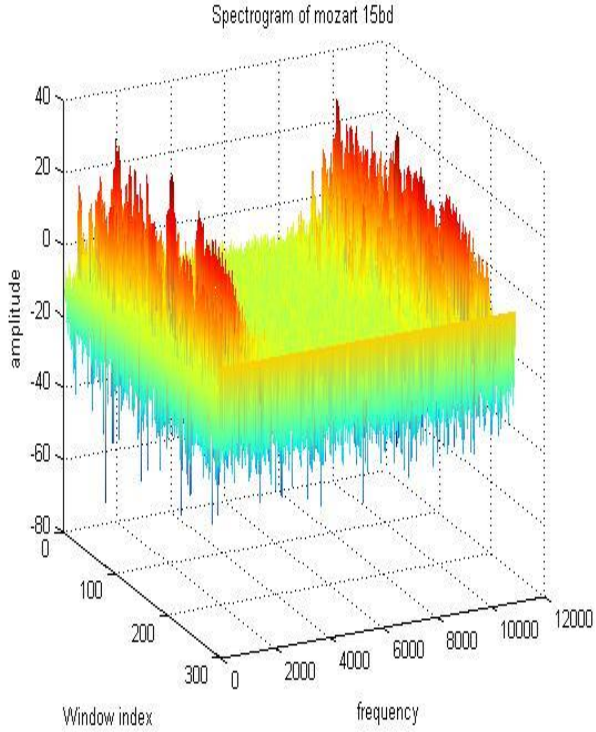

Fig 7. Spectrogram of Mozart 15db

CONCLUSIONS

Diagonal time frequency audio denoising technique attenuates the noise by processing each spectrogram coefficient independently that creates isolated time frequency structures that are perceived as a "musical noise". The algorithm has a limited performance and produce a musical noise. To denoise the musical noise signal more efficiently, an adaptive block thresholding nondiagonal estimation is used, which adjusts all parameters adaptively to signal property by minimizing a Stein estimation of the risk.

The audio signal collected from a noisy environment is predominantly degraded by additive background noise. The Log-Spectral Amplitude (LSA) is the modified technique minimizes MSE of the log-spectra based on Gaussian model for the spectral components of audio signal. The SNR values of BT are greater than the values of LSA. Hence the performance of BT is high. For increasing of Mozart signal with noise the SNR values of both the techniques will be increased. An adaptive audio block thresholding algorithm adapts all parameters to the time-frequency regularity of the audio signal. The adaptation is performed by minimizing a Stein unbiased risk estimator calculated from the data. Optimal estimators in the STFT domain are often on a Gaussian statistical model. Accordingly, the individual short term spectral components of the audio and noise signals are modeled as statistically independent Gaussian random variables.

Hence Nondiagonal time-frequency estimators are more effective than diagonal estimators to remove noise from audio signals because they introduce less musical noise. SNR values are high for nondiagonal technique of block thresholding algorithm, SNR values are low for diagonal technique of MMSE-LSA algorithm.

7.REFERENCES

[1] Guoshen Yu,Stephane Mallet,Emmanuel Bacry "Audio Denoising by Time Frequency Block Thresholding" IEEE Transactions on Signal Processing Volume 56 No.5, May 2008.

[2] Cai .T, "Adaptive wavelet estimation: A block thresholding and oracle inequality approach" vol. 27, 898-924, 1999.

[3] Cai .T, "On block thresholding in wavelet regression: adaptivity, block size, and threshold level" vol 12, 2002.

[4] Cappe .O, "Elimination of the musical noise phenomenon with the Ephariam and Malah noise suppressor" in IEEE trans, 1994.

[5] Bahoura .M and Rouat .J, "A new approach for wavelet speech enhancement" vol.47, pp.173-205, 1993.

[6] Cohen .I, "Speech enhancement using a noncausal a priori SNR estimator" vol.11, No.9, September 1994.

[7] Cohen .I, "Speech enhancement using super guassian speech models and non causal a priori SNR estimation", in 2005.

[8] Cohen.I, "Speech specral modeling and enhancement based on autoregressive conditional heteroscedasticity models", in June 2005.

[9] Godsill .J, "Sparse regression with structured priors: Application to audio denoising", in 2006.

[10] Hall .P, "On the minimax optimality of block thresholded wavelet estimators",vol 9, 33-49, 1999.

[[11] Ghael .P and Sayeed .M, "Improved wavelet denoising via empirical wiener filtering", in 1997.

[12] Kamath .D and Loizou .C, "A multi-band spectral subtraction method for enhancement speech corrupted by colored noise", in 2001.

Authors Profile:

P. Uday Kumar is currently working as an Assistant professor in ECE department. Received the M.Tech degree from JNTU Kakinada. and B.TECH degree from JNTU Hyderabad.

He is an Associate Member of IEI. He reviewed some research papers of International Journals & has many accepted & published papers in International Journals and Conferences. His current research interest includes design of Signal & Image Processing, and Embedded VLSI System design.

K.P.Obulesu is currently a PG scholar of VLSI System Design in ECE Department. He received B.TECH degree from JNTU Anantapur.His Current research interest includes Digital Signal Processing & VLSI System design.