Asian Journal of Information Technology 12 (5): 154-159, 2013
ISSN: 1682-3915

A Noise Reduction Approach for Speech Signal Based on Stein's Unbiased Risk Estimate

[1]S. Karthikeyan, [2]P. Ganesh Kumar and [1, 3]K. Saran
[1, 3]Department of ECE, PSNA College of Engineering and Technology, Dindigul-624622, Tamilnadu, India
[2]Department of ECE, KLN College of Engineering and Technology, Sivagangai Dt-630611, Tamilnadu, India

Abstract: The proposed research in the denoising of speech signal is based on the concepts of Wavelet Thresholding by imposing quantum parameters. The idea of signal denoising is to preserve the signal features while reducing the noise level. Various denoising approaches exist in which wavelet-based point wise thresholding approaches are extensively adopted in many application fields like speech processing applications, medical applications, etc. For signal denoising based on wavelet thresholding, there are two decisive aspects, namely, the use of a proper thresholding function and the estimate of the noise standard deviation. Both greatly manipulate the quality of the denoised signal. In this study, a simple wavelet-based denoising approach is performed for real time speech signal which uses the modified linear expansion of thresholds based on Stein's Unbiased Risk Estimation (SURE) and the noise standard deviation estimation depending on the number of vanishing moments of the wavelet transform. Investigational results demonstrate that higher Signal to Noise Ratio (SNR) with lower mean square error.

Key words: Wavelet thresholding, denoising, hilbert transform, mean square error, orthonormal, SNR

INTRODUCTION

Applied scientists and engineers who work with data obtained from the real world know that signals do not exist without noise. Under ultimate conditions, this noise may decrease to such insignificant levels while the signal increases to such significant levels that for all practical purposes denoising are not essential. Unfortunately, the noise corrupting the signal, more often than not must be removed in order to recover the signal and proceed with further data analysis. The noise removal may take place in the original signal (time-space) domain or in a transform domain.

If the later, should it be the time-frequency domain via the Fourier transform or the time-scale domain via the Wavelet transform. Enthusiastic supporters have described the development of wavelet transforms as revolutionizing modern signal and image processing over the past two decades.

For a review of available software libraries and an introduction to some of the wavelet literature, refer to the survey by Benazza-Benyahia and Pesquest (2005). Meanwhile, some waveforms such as the mean frequency waveform and the spectral width waveform are very useful to evaluate the MSE and standard deviation showed by Eldar *et al.* (2004, 2005) and Eldar and Merhav (2004). The spectrogram indices are all extracted from the maximum frequency waveform of the speech signal spectrogram which can be calculated using the Short-Time Fourier Transform (STFT) (Donoho and Johnstone, 1994). Their precision, together with the precision of the mean frequency waveform and the spectral width waveform which is estimated for continuous speech signals is directly influenced by the estimation resolution of the maximum frequency waveform (Stein, 1981). Any extra frequency component in the speech signal coming from noise may reduce the estimation resolution which harms further processing. Therefore, it is a preliminary and important step to denoise the continuous speech signal, especially when the SNR is low (<10 dB).

Because of the non-stationary characteristic of the speech signal, the traditional spectrogram enhancement methods such as the adaptive filtering method (Donoho and Johnstone, 1994, 1995) which denoises the signal by simply optimizing the Mean Square Error (MSE) are not sufficient. Once a good denoising method is used, the MSE would be minimized with no additional frequency components induced algorithm is used.

Corresponding Author: S. Karthikeyan, Department of ECE, PSNA College of Engineering and Technology, Dindigul-624622, Tamilnadu, India

Minimax estimation is a cautious approach aiming at finding optimal estimators in the worst case situation compatible with the available information. Minimax problems have been the focus of many works in statistics. More recently, a renewal of interest has been observed for minimax signal estimation problems and among many interesting new results, connections between these problems and linear matrix inequalities have been shown by Eldar *et al*. (2004, 2005) and Eldar and Merhav (2004) for some kind of minimax mean square estimation problems. In the meantime, non-linear estimation has gained popularity in signal processing problems.

For example, wavelet regression methods using thresholding operators have been developed and they have been shown to be optimal from an asymptotic minimax viewpoint for certain classes of regular signals (Iusem, 2003). A similar approach was used by Combettes (1996, 2003) and Combettes and Pesquet (2004) to develop more sophisticated multivariate estimates for multi component image denoising. Like the Fourier Transforms (FT), a Wavelet Transform (WT) represents a signal in another domain (Strange and Nguyen, 1996) a time-frequency domain. However, wavelet transforms are more general than the Fourier transform. Unlike the Fourier transform, wavelet transforms may describe localized signals more efficiently. For example, a wavelet transform may describe a function into different frequency components and then study each component with a resolution matched to its scale.

In this study, a different approach for optimizing nonlinear estimators relies on the use of Stein's Unbiased Risk Estimator (SURE) in problems involving additive Gaussian noise. The SURE Method was used to build adaptive thresholding estimators. The Stein's Unbiased Risk Estimate (SURE) principle was also applied by Oppenheim *et al.* (2001) for signal restoration problems formulated as constrained convex optimization problems.

MATERIALS AND METHODS

Wavelet thresholding: Assume the observed signal $y = [y_1, y_2, ..., y_N]^T$ given by:

$$y_i = x_i + z_i \quad i = 1,..., N \tag{1}$$

Where:

x_i = Samples of noise-free signal x

z_i = Gaussian white noise with independent identical distribution (i.i.d.), i.e., z_i N(0, σ^2) shown in Fig. 1

From noisy signal y, researchers want to find an approximation $\hat{x}$ to the original signal x that minimizes the Mean Squared Error (MSE) as follows:

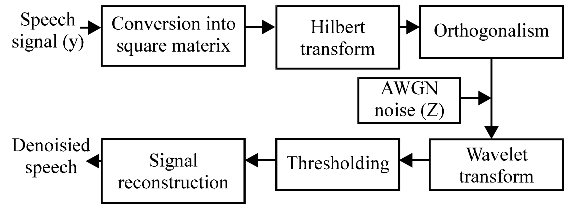

Fig. 1: Wavelet denoising process

$$R(\hat{x}, x) = \frac{1}{N}\left\|\hat{x} - x\right\|^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{x}_i - x_i)^2 \tag{2}$$

where, $x = [x_1, x_2, x_3, ..., x_N]^T$ and $\hat{x} = [\hat{x}_1, \hat{x}_2, ..., \hat{x}_N]^T$. Here, use mean instead of the mathematical expectation because the optimal solution is desired for each individual noisy signal. Let W be an orthonormal wavelet transform. After applying w to Eq. 4 then Eq. 4 can:

$$w = s + b \tag{3}$$

with $w = [w_1, w_2, ..., w_N]^T = W_y$, $s = [s_1, s_2, ..., s_N]^T = W_x$ and $b = [b_1, b_2, ..., b_N]^T = W_z$. Since, W is an orthonormal transform, the noise remains Gaussian with same statistics in the orthonormal wavelet domain, i.e., $b_j \sim N(0,\sigma^2)$. Let $\theta(\cdot)$ be a wavelet thresholding function. Then, the estimate can be expressed as follows:

$$\hat{x} = w^{-1}(\theta(W_y)) \tag{4}$$

Thus, the risk function given in Eq. 2 can be expressed as:

$$R(\hat{x}, x) = \frac{1}{N}(\hat{x}_i - x_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(\hat{s}_i - s_i)^2 \tag{5}$$

where, $\hat{s}_i = \theta(\omega_i)$. Many thresholding functions have been presented. The most well-known thresholding functions are the hard thresholding function and the soft thresholding function.

Soft and hard thresholding: Signal denoising using the (Discrete Wavelet Transform) DWT consists of three successive procedures, specifically, signal decomposition, thresholding of the DWT coefficients and signal reconstruction. Firstly, researchers carry out the wavelet analysis of a noisy signal up to a chosen level N. Secondly, researchers perform thresholding of the detail coefficients from level 1 to N. Lastly and researchers synthesize the signal using the altered detail coefficients from level 1 to N and approximation coefficients of level N. However, it is generally impossible to remove all the noise without corrupting the signal. As for thresholding, researchers can settle either a level-dependent threshold vector of length N or a global threshold of a constant value for all levels. The threshold estimate a for denoising with an orthonormal basis is given by:

$$\delta = \sigma\sqrt{2\log L} \tag{6}$$

Where the noise is Gaussian with standard deviation σ of the DWT coefficients and L is the number of samples or pixels of the processed signal. From a different point of view, thresholding can be soft or hard. Hard thresholding zeroes out all the signal values lesser than δ. Soft thresholding does the same thing and apart form that subtracts the values larger than δ. On the contrary to hard thresholding, soft thresholding causes no discontinuities in the resulting signal.

Threshold determination: A small threshold may yield a result close to the input, yet it is noisy. On the other hand a large threshold produces a signal with a large number of zero coefficients. This leads to a smooth signal. Providing too much attention to smoothness, however, destroys details and in image processing may cause blur and artifacts. To investigate the effect of threshold selection, researchers performed wavelet denoising using hard and soft thresholds on four signals popular in wavelet literature, hard and soft thresholding with threshold (Fig. 2 and 3) are defined as follows. The hard thresholding operator is defined as:

$$\theta_h(w) = \begin{cases} w & \text{if } |w| \geq \delta \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

On the other hand the soft thresholding operator is defined as:

$$\theta_s(w) = \begin{cases} \text{sign}(w)(|w| - \delta) & \text{if } |w| \geq \delta \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The main idea of soft thresholding is that noise is removed by shrinking or killing coefficients that are insignificant relative to some threshold value.

Threshold selection rules: According to the basic noise model, four threshold selection procedures are implemented. Each rule corresponds to a tptr option in the command, thr = thselect (y, tptr) which returns the threshold value:

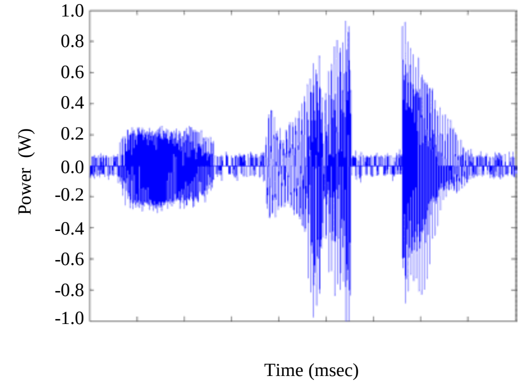

Fig. 2: Result of Hard thresholding

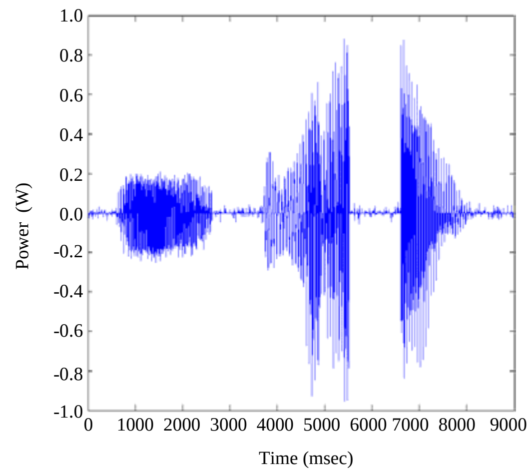

Fig. 3: Result of soft thresholding

- Option tptr = 'rigrsure' uses for the soft threshold estimator a threshold selection rule based on Stein's Unbiased Estimate of Risk (quadratic loss function). You get a measure of the risk for a particular threshold value t. reducing the risks in t gives a selection of the threshold value
- Option tptr = 'sqtwolog' uses a fixed form threshold yielding minimax performance multiplied by a small factor proportional to log(length(s))
- Option tptr = 'heursure' is a mixture of the two previous options. As a result, the SURE estimate is very noisy if the signal-to-noise ratio is very small. The fixed form threshold is used in such situations
- Option tptr = 'minimaxi' uses a fixed threshold chosen to produce minimax performance for mean square error against an ideal procedure. The minimax principle is used in statistics to design estimators. Since, the de-noised signal can be incorporated to the estimator of the unknown regression function, the minimax estimator is the option that recognizes

the minimum, over a given set of functions of the maximum mean square error. y is a signal which is to be denoised

Signal transformation: In signal processing, wavelets are used for denoising, detecting trends, breakdown points, discontinuities in higher derivatives and self-similarity in signals. At first, researchers focus on discontinuity detection. For the speech signal Fig. 4 where demonstrates the use of the db4 wavelet for impulse detection, i.e., detection of a discontinuity in frequency. The impulse is generated artificially for the purpose. The db4 wavelet is chosen because of its good performance in this case. The decomposition runs up to level 3 which is enough to make the discontinuity apparent. This study describes signal denoising with the application on the speech signal. Input is a continuous speech signal given through microphone.

The signal is obtained as a column vector. This column vector is converted into a square matrix. Now Hilbert transform is performed on this matrix so that the numerical values of the signal can be obtained. FFT is performed on the signal so that the spectral values of the signal can be obtained. As the concept of QSP is to be satisfied, now the spectral matrix is being converted into orthogonal matrix using Gram-Schmidt orthogonalization procedure. In the orthogonal matrix, white noise with zero mean and unit standard deviation added to the signal and then researchers take wavelet transform to noisy input speech signal. The wavelet transform consist of three steps:

- Calculate the wavelet transform of the noisy signal
- Modify the wavelet coefficients according to some thresholding rule
- Calculate the inverse transform using the modified wavelet coefficients

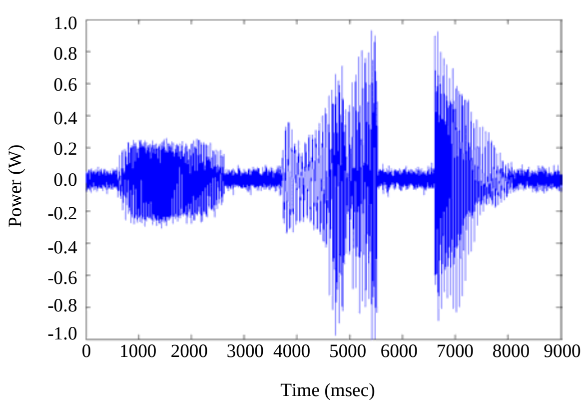

Fig. 4: Noise interference input signal

Stein's unbiased risk estimator: To recover a function of unknown smoothness from noisy, sampled data, researchers set up a procedure, sure shrink which suppresses noise by thresholding the empirical wavelet coefficients. The thresholding is adaptive threshold level is assigned to each dyadic resolution level by the principle of minimizing the stein unbiased estimate of risk (sure) for threshold estimates. The computational attempt of the overall procedure is order N log (N) as a function of the sample size N. Sure shrink is smoothness-adaptive: if the unknown function contains jumps, the reconstruction (essentially) does also; if the unknown function has a smooth piece, the reconstruction is (essentially) as smooth as the mother wavelet will allow.

The procedure is in a sense optimally smoothness-adaptive: it is near-minimax simultaneously over a whole interval of the Besov scale the size of this interval depends on the choice of mother wavelet. Traditional smoothing methods: kernels, splines and orthogonal series estimates even with optimal choices of the smoothing parameter would be unable to performing a near-Minimax way over many spaces in the Besov scale. The proposed SURE-based shrinkage function which is defined as:

$$\theta_{sure}(w_i) = \sum_{k=1}^{K} a_k \varphi_k(w_i) \tag{9}$$

where, $\alpha_k(k\in([1,\ K])$ are the unknown parameters. However, the shrinkage function θ Sure(.) does not consider the intrascale correlations between wavelet coefficients which can be used to further improve the performance of the noise reduction algorithm. Moreover, the computation time can be reduced by decreasing the number of parameters ak. In the following study, researchers therefore propose a simple and effective SURE-based intrascale shrinkage method which exploits these correlations and decreases the number of unknown parameters (Fig. 5).

Researchers use the averaged magnitude of the neighbouring coefficients to quantify the intrascale correlations between wavelet coefficients which can be expressed by:

$$c_i = \frac{1}{M} \sum_{j=i-M,\, j\neq i}^{i+M} |w_j| \tag{10}$$

where, M is the number of neighboring coefficients. In practice, it is sufficient to set M = 2. In the following study, the impact of M on the performance of the proposed method is discussed. In order to use the intrascale correlation coefficient ci, researchers introduce the following shrinkage function:

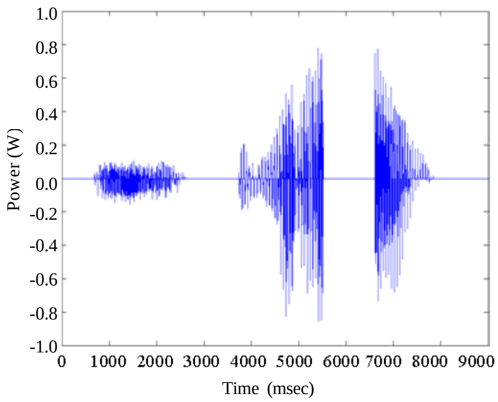

Fig. 5: Result of sure thresholding

$$\theta_{new}(w_i) = w_i + a\mu(c_i)\mu(w_i)w_i \tag{11}$$

where, a is the unknown parameter and:

$$\mu(x) = \exp\left(\frac{-x^2}{12\sigma_n^{\ 2}}\right) \tag{12}$$

This shrinkage function can be regarded as a modified version of θ Sure(.) by decreasing the number of parameters ak and using the intrascale correlations between wavelet coefficients. The function $\mu(x)$ is a simplified version of $\varphi(x)$. The Eq. 7 is the proposed intrascale shrinkage function which contains two unknown parameters and σ_n. In Eq. 12, a robust estimate of σ_n is given by:

$$\sigma_n^{\ 2} = \frac{MAD}{0.6745}\sqrt{\frac{2p}{2p+1}} \tag{13}$$

Where:

MAD = The median absolute deviation of the finest wavelet coefficients

p = The number of vanishing moments of the wavelet transform

For fixed σ_n, researchers can obtain the optimal estimate of a by minimizing the Mean Squared Error (MSE), i.e.:

$$a = \arg\min \frac{1}{N}\sum_{i=1}^{\hat{N}} \left|\theta_{new}(w_i) - u_i\right|^2 \tag{14}$$

RESULTS AND DISCUSSION

In the simulations, researchers use the standard input speech signal. The noisy signals are created by adding Gaussian white noise with different noise levels to the signals. The proposed method is evaluated by comparing

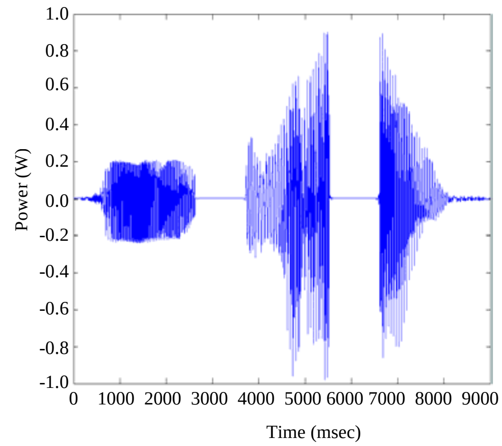

Fig. 6: Input speech signal

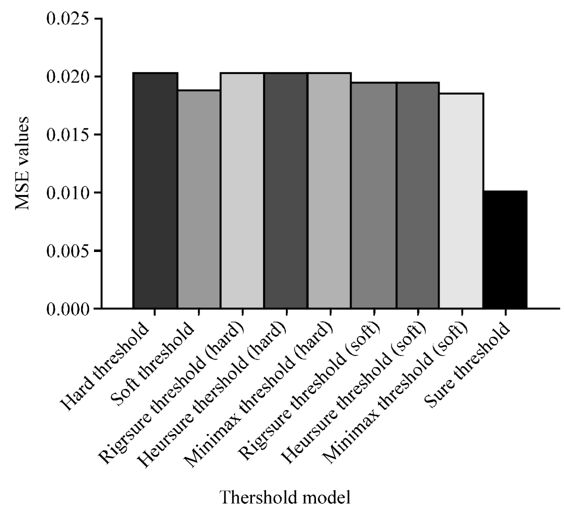

Fig. 7: Comparision of MSE values for various threshold models

it with some most commonly used shrinkage functions, namely, Heursure Shrinkage Function (HSF), Rigrsure (RSURESF) and Minimax (MSF).

For the sake of evaluating the performance of these methods, the Mean Square Error (MSE) is used as the quantitative criterion. In all comparisons, researchers use the stationary Daubechies wavelet with four vanishing moments over three decomposition levels which means p = 4. Figure 6 shows the visual quality of the various algorithms for speech signal. As can be observed, the proposed method exhibits less distortion than the others.

Table 1 displays the MSE results generated by all algorithms for the test signals. The method clearly outperforms HSF, RSURESF and MSF in terms of MSE and the average values are nearly 0.02, 0.02 and 0.01, respectively (Fig. 7).

Table 1: MSE value for different threshold models

Method of thresholding	MSE value
Hard thresholding	0.020366
Soft thresholding	0.018797
Rigrsure thresholding (hard)	0.020384
Heursure thresholding (hard)	0.020384
Minimax thresholding (hard)	0.020304
Rigrsure thresholding (soft)	0.019581
Heursure thresholding (soft)	0.019581
Minimax thresholding (soft)	0.018601
Sure thresholding	0.010027

CONCLUSION

In this study, researchers improved the performance of Sure thresholding function by decreasing the number of unknown parameters and using the new estimate of noise standard deviation. The new approach does not need any prior statistical modelization of the wavelet coefficients. The experimental results demonstrate the efficiency of the new approach which can obtain lower MSE than other denoising thresholding approaches for noisy signals.

REFERENCES

Benazza-Benyahia, A. and J.C. Pesquet, 2005. Building robust Wavelet estimators for multicomponent images using Stein's principle. IEEE Trans. Image Process., 14: 1814-1830.

Combettes, P.L. and J.C. Pesquet, 2004. Wavelet constrained image restoration. Int. J. Wavelets Multiresolution Inform. Process., 2: 371-389.

Combettes, P.L., 1996. The convex feasibility problem in image recovery. Adv. Imag. Electron. Phys., 95: 155-270.

Combettes, P.L., 2003. A block-iterative surrogate constraint splitting method for quadratic signal recovery. IEEE Trans. Signal Process., 51: 1771-1782.

Donoho, D.L. and I.M. Johnstone, 1994. Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81: 425-455.

Donoho, D.L. and I.M. Johnstone, 1995. Adapting to unknown smoothness via wavelet shrinkage. J. Am. Statistical Assoc., 90: 1200-1224.

Eldar, Y.C. and N. Merhav, 2004. A competitive minimax approach to robust estimation of random parameters. IEEE Trans. Signal Process., 52: 1931-1946.

Eldar, Y.C., A. Ben-Tal and A. Nemirovski, 2004. Linear minimax regret estimation of deterministic parameters with bounded data uncertainties. IEEE Trans. Signal Process., 52: 2177-2188.

Eldar, Y.C., A. Ben-Tal and A. Nemirovski, 2005. Robust mean squared error estimation in the presence of model uncertainties. IEEE Trans. Signal Process., 53: 168-181.

Iusem, A.N., 2003. On the convergence properties of the projected gradient method for convex optimization. Mat. Aplied Comput., 22: 37-52.

Oppenheim, G., J. Poggi, M. Misiti and Y. Misiti, 2001. Wavelet Toolbox. The Math Works, Inc., Natick, Massachusetts.

Stein, C., 1981. Estimation of the mean of a multivariate Normal distribution. Ann. Stat., 9: 1135-1151.

Strang, G. and T. Nguyen, 1996. Wavelets and Filter Banks. 2nd Edn., Wellesley-Cambridge Press, USA., ISBN: 9780961408879, Pages: 490.