ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

Speech Enhancement with Reduction of Noise Components in the Wavelet Transforms

Ravindra D. Dhage, Pravinkumar R. Badadapure

M.E Scholar, Department of Electronics & Telecommunication Engineering, JSPM's Imperial College of Engineering and Research, Pune, M.HIndia

Professor, Department of Electronics & Telecommunication Engineering, JSPM's Imperial College of Engineering and Research, Pune, M.H, India

ABSTRACT: A new speech enhancement method based on time and scale adaptation of wavelet thresholds, time dependency is introduced by approximating the Teager Energy of the wavelet coefficients. During the past decade, Wavelet Transforms (WT) have been applied to various research areas, applications include signal and image de-noising, compression detection and pattern recognition. De-noising methods based on the wavelet thresholding have not been successfully applied to speech enhancement. Wavelet transform combined with other signal processing tools has been proposed, this technique does not require an explicit estimation of the noise level or of the a priori knowledge of the SNR as is usually needed in most of the popular enhancement methods. Proposed method is evaluated on speech recorded in real conditions (plane, car, exhibition hall, restaurant, airport and train station) and artificially added noise, show that this method to improves the speech enhancement rates for low SNRs. Comparison in terms of Signal-to-Noise Ratio (SNR) is reported for time adaptation and time-scale adaptation thresholding of the wavelet coefficients thresholding.

I. INTRODUCTION

New speech based applications such as automatic speech translation, internet search tools, multimedia teaching, training and multimodal computer interactions are under development in many public and private research laboratories. These systems strong limitation is the inadequacy of processing corrupted or noisy speech.

Many approaches have been studied to increase the robustness of speech processing systems. Speech or speaker recognizers based on a pattern recognition process that separates analysis from recognition. Speech enhancement already has a strong potential. It can also be used in coding and with various apparatus such as audio prostheses and drawback of these enhancement techniques is the necessity to estimate the noise. This can be a strong limitation when recording with non-stationary noise, to improve speech enhancement in non-stationary noise propose to control the gain of the update of the estimated noise spectrum during speech presence in a modified Minimum Mean-Square Error Log-Spectral Amplitude (MMSE-LSA) estimator.

This is new techniques that do not require any a priori knowledge of the noise and that can complement the contemporary de-noising systems by taking into account speech characteristics. Speech enhancement system would probably combine both approaches 1) cleaning of the noise when it can be estimated and 2) enhancement by taking into consideration the speech structure when it is not possible to know anything regarding the noise.

We propose a new procedure based on a time-scale threshold of wavelet packet coefficients without requirement or knowledge of the noise level. The thresholds are modulated with a nonlinear mask that reflects the spatial and time dominance evolution of speech on noise. Their applications include signal and image de-noising. Wavelet shrinkage is a simple de-noising technique based on a thresholding of the wavelet coefficients to define the limit between the wavelet coefficients of the noise and those of the target signal, it is not always possible to separate the components corresponding to the target signal from those of noise by a simple thresholding. Unvoiced segments are comparable to

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

those of noise, thresholding uniformly to all wavelet coefficients not only suppresses additional noise but also some speech components like unvoiced ones the perceptive quality of the filtered speech is greatly affected.

The discriminative threshold in various sub-bands is time and spatially adapted in relation with the speech components when speech dominates the noise. The proposed techniques are tested on noisy speech recorded in real environments and with artificial noise. The first evaluation is based on a comparison of Signal-to-Noise Ratio increases. It is observed that the proposed Time and Scale Adaptation (TSA) of the wavelet coefficients yields a greater increase of SNR for very noisy speech (-10 dB to 10 dB). EMF is preferred to the Times Scale Adaptation that uses a continuous derivative thresholding function.

II. NOISE REDUCTION

We present the most popular de-noising methods based on the wavelet transform complete representation of the signal and they proposed an iterative algorithm to remove noise. This method based on the spatial correlation between the wavelet coefficients over adjacent scales. Speech enhancement is closely related to speech restoration to the original speech signal often leads to speech enhancement. Some important differences between enhancement and restoration an ideal speech signal is degraded and the objective is to make the processed speech signal as close as possible to the original. An objective of speech enhancement is to make the processed signal sound better than the unprocessed signal. Signals are often corrupted by noise which is unwanted signal one of the most common sources of noise is background noise. Theses Channel noise which affects both analog and digital transmission quantization noise which results from over compression of speech signals, reverberation noise or delayed version of noise are also present in some situations to recover a signal s(t) from noisy data x(t) with a Gaussian white noise bi.

$$xi = si + bi \qquad i = 1, \ldots, N \qquad (1)$$

$$T_s(\lambda, \omega_k) = \begin{cases} sgn(\omega_k)(|\omega_k| - \lambda) & if\,|\omega_k| > \lambda \\ 0 & if\,|\omega_k| \leq \lambda \end{cases} \qquad (2)$$

Where ω_k represents the wavelet coefficients, they proposed a universal threshold λ for the WT,

$$\lambda = \sigma\sqrt{2Log(N)} \qquad (3)$$

With σ = MAD/0.6745, where N is the length of x and σ is the noise level absolute value of the wavelet's coefficients estimated on the first scale. In the Wavelet Packets Transform (WPT) case the threshold is defined as,

$$\lambda = \sigma\sqrt{2Log(NLog_2N)} \qquad (4)$$

If the wavelet transform has been extensively combined with other methods to improve the speech quality of corrupted speech signal. Speech enhancement our method to enhance speech by spatially and time adapting the thresholds in the context of Wavelet Packet Transforms (WPT).

III. APPLICATION TO SPEECH ENHANCEMENT

3.1 Wavelet thresholding

Speech enhancement algorithm to prevent the speech quality deterioration during the thresholding process the unvoiced regions are classified first and then thresholding is applied in different ways, the problem is not completely resolved but this approach constitutes an interesting step to avoid the speech degradation.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

3.2 Wiener filtering in the wavelet domain

The wavelet transform based Wiener filtering is a special application. This idea arises from the fact that wavelet transforms tend to de-correlate data. A multi-microphone system is proposed for speech enhancement and Wiener filtering performances in the wavelet domain are better than those obtained in the Fourier domain and another version that combines Wiener and coherence in the wavelet domain is also proposed [3].

3.3 Wavelet filter bank

The frequency bands of the cochlear filters are not uniformly distributed. Several transformations (scales) are proposed to take into account the perceptive aspect of hearing, wavelet transform is used as a bank of filters (not uniformly distributed) to improve performance of the speech enhancement method based on the spectral subtraction [4]. Speech enhancement method based on the coherence function is proposed. The wavelet transform is also used as a filter bank.

IV. WAVELET ENHANCEMENT METHOD

4.1 Wavelet packet analysis

Wavelet thresholding techniques have not be successfully applied to speech enhancement and difficulties are related to the speech signal complexity and to the nature of the noise improve the wavelet thresholding performance. For a given level j the WPT decomposes the noisy signal x (n) into 2j sub-bands corresponding to wavelet coefficient sets wj,

$$\omega_{k,m}^{j} = WP\{X(n), j\} \qquad n = 1, \dots . N \qquad (5)$$

By increasing the value of j the bandwidth of all sub-bands decreases which improves the scale adaptation of the discriminatory threshold the noise is considerably reduced but the quality of the reconstructed speech is affected, the resulting wavelet coefficients $w_{k,m}^{4}$ of each sub-band k, each sub-band k by smoothing the corresponding TEO coefficients and normalizing

$$t_{k,m}^{4} = [w_{k,m}^{4}]^{2} - w_{k,m-1}^{4} w_{k,m+1}^{4} \qquad (6)$$

$$\mathrm{M}_{k,m}^{4} = \frac{t_{k,l}^{4} * h_{k}(m)}{\max(|t_{k,l}^{4} * h_{k}(m)|)} \qquad (7)$$

4.2 Time modulation:

For each wavelet sub-band k the corresponding threshold λ_k should be time adapted only for speech like frames and kept unchanged for noisy. The cleaning of noise will be maximal when noise is dominant in the wavelet's sub-band for the speech frame. To distinguish these frames we define a parameter S_k^4 named offset that estimates. It is given by the abscissa of the maximum of the amplitude distribution H of the corresponding mask $M_{k.m}^4$ and is estimated over the analyzed frame,

$$S_k^4 = \mathrm{abscissa}[H(M_{k,m}^4)] \qquad (8)$$

V. SIGNAL TO NOISE

Speech enhancement process aim to improve the quality speech intelligibility in order to reduce the listener fatigue depending on specific application. The wavelet transform plays an important role in signal analysis and widely used in

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

many applications such as signal detection and De-noising. A signal x (n) is created by combining a clean speech sentence s (n) and a noise b (n),

$$x(n) = s(n) + b(n)$$

$$\mathrm{SNR}_U = 10\log\frac{\sum_{n=1}^{N} s(n)^2}{\sum_{n=1}^{N} b(n)^2} \tag{9}$$

Where, N is the length of the sentence expressed in number of samples

VI. MEL-SCALE MULTI-RATE FILTER BANK

The speech signal was enhanced by using a bank of wavelet filters uniformly distributed in the frequency domain. These proposed method for speech recognition and speaker identification, these applications use respectively 21 sub-bands and 24 sub-bands, the impact of the filter bank, we test three other filter banks (MEL1, MEL2, and MEL3) based on Daubechies wavelets. MEL1 and MEL2 are based respectively on 21 and 24 filters. The last filter bank MEL3 is obtained by dividing the low frequencies of MEL2 and comprises 32 sub-bands.

VII. PERFORMANCE ANALYSIS

Wavelet thresholding method has been initially proposed to remove additive white noise. In this section we use speech signals also corrupted by narrow band noises that are not white experiments show that our thresholding approach is also efficient for that kind of noise and evaluated using speech corrupted with white noise, car noises and speech recorded in real environment. The performance of this method for clean speech corrupted by additional noise at various SNR. The size of the analysis frame has been set equal to the length of the speech file one estimate of the noise has been used for the Ephraim and Malah algorithm (EMF).

7.1 Time adaptation of the threshold (TA)

Time adapted thresholding technique to white wide band noise and to narrow band noise.

Table No.1 SNR tests for white noise corrupted speech

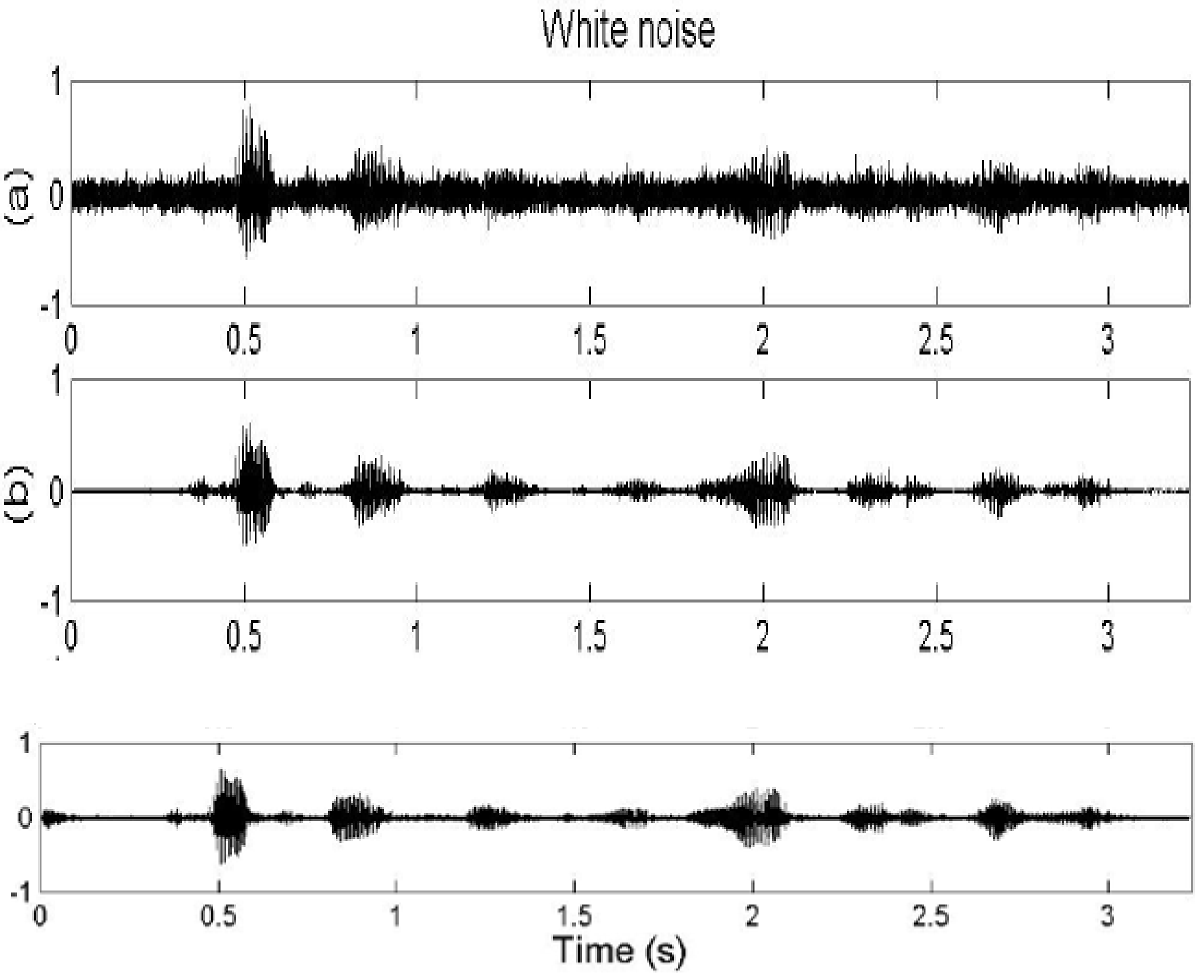
SNR(dB)	TA(dB)	EMF(dB)
-10	0.99	-0.56
-5	3.23	2.66
0	6.36	4.96
5	9.36	7.48
10	12.17	9.80

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015



(c)

Fig.1 a) Speech corrupted with white noise (SNR=0dB) enhancement results using b) Time Adaptation filtering c) EMF

VIII. CONCLUSION

This proposed method is one of the first successful applications of the wavelet thresholding method for speech enhancement. When the increase in SNR is used as criteria it has been observed that for artificial white noise TA is superior to TSA which is also better than EMF. With the Time Adapted threshold (TA) performance is better when not using the MEL scale same remark is valid for the Time–Scale Adaptation technique (TSA). The car noise is very low frequency and the TSAMEL uses a bank of wavelets with an important resolution in low frequencies (32 sub-bands). It is observed that the EMF is better for SNR 5 dB. It is observed that TA is usually sufficient for wide band noise while it is absolutely inefficient with narrow band noises (car and airplane).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 6, June 2015

REFERENCES

[1] J. Sika and V. Davidek, “Multi-channel noise reduction using wavelet filter bank,” in EuroSpeech’97, 1997.
[2] Bahoura, M., Rouat, J., January 2001b. Wavelet speech enhancement using the Teager energy operator. IEEE Signal Processing Letters 8, 10–12.
[3]Pan, Q., Zhang, L., Dai, G., Zhang, H., December 1999. Two denoising methods by wavelet transform. IEEE Trans. Signal Processing 47, 3401–3406.
[4] Sarikaya, R., Pellom, B., Hansen, J., June 1998. Wavelet packet transform features with application to speaker identification. In: NORSIG-98 IEEE Norsic Signal Processing Symposium. Vigso, Denmark, pp. 81–84.
[5] Seok, J., Bae, K., April 1997. Speech enhancement with reduction of noise components in the wavelet domain. In: ICASSP 97. Munich, Germany, pp. 1223–1326.
[6] Sika, J., Davidek, V., Spetember 1997. Multi-channel noise reduction using wavelet filter bank. In: EuroSpeech’97. Rhodes, Greece, pp. 2595 – 2598.
[7] Vidakovic, B., Lozoya, C., September 1998. On time-dependant wavelet denoising. IEEE Trans. Signal Processing 46, 2549–2554.
[8] Xu, Y., Weaver, J., Healy, D., Lu, J., November 1994. Wavelet transform domain filters: A spatially selective noise filtration technique. IEEE Trans Image Processing 3, 747–758.
[9] Bahoura, M., Rouat, J., January 2001b. Wavelet speech enhancement using the Teager energy operator. IEEE Signal Processing Letters 8, 10–12.
[10] Bahoura, M., Rouat, J., September 3-7 2001a. New approach for wavelet speech enhancement. In: Eurospeech 2001. Aalborg, Denmark, pp. 1937– 1940.
[11] Chen, S. H., Wang, J. F., 2004. Speech enhancement using perceptual wavelet packet decomposition and teager energy operator. J. VLSI Signal Process. Syst. 36 (2-3), 125–139.
[12] Cohen, I., September 3-7 2001. Enhancement of speech using bark-scaled wavelet packet decomposition. In: Eurospeech 2001. Aalborg, Denmark, pp. 1933–1936.