WAVELET NOISE REDUCTION: APPLICATION TO SPEECH ENHANCEMENT

Mohammed BAHOURA and Jean ROUAT
ERMETIS, DSA, Université du Québec à Chicoutimi, Chicoutimi, Québec, Canada, G7H 2B

1. Introduction

During the past decade, the Wavelet Transforms (WT) have been applied to various research areas. Their applications include signal and image denoising, compression, detection, and pattern recognition.

To our knowledge, denoising methods based on the wavelet thresholding have not been successfully applied to speech enhancement. However, wavelet transform combined with other signal processing tools has been proposed.

In this paper, we pass in review the principal wavelet denoising methods published in the literature. Then, we expose the speech enhancement techniques using wavelet transform. Next, we point out the difficulties encountered in the wavelet thresholding methods and finely we propose some solutions.

2. Noise reduction by wavelets

2.1 Principe

Two basic approaches are proposed to remove noise by wavelet transform. The first is based on the singularity information analysis [1], whereas the second is based on the thresholding of the wavelet coefficients [2].

Mallat et al. [1] proved that modulus maxima of the wavelet coefficients give the complete representation of the signal and they proposed an iterative algorithm to remove noise. In the singularity analysis context, Xu et al. [3] developed a noise filtration method based on the spatial correlation between the wavelet coefficients over adjacent scales. An improved version is proposed by Pan et al. [4]. The thresholding method will be described in the next section.

2.2 Wavelet shrinkage

Donoho [2] proposed a powerful approach for noise reduction. It is based on the thresholding of the wavelet coefficients:

- Transform the noisy signal y into wavelet coefficients w,
- Apply a soft or hard threshold λ at each scale,
- Transform back the resulting coefficients, and get the estimated signal.

Donoho and Johnstone [5,6] define the soft threshold by

$$T_\lambda(w) = \begin{cases} sgn(w)(w-\lambda) & \text{if } |w| \geq \lambda \\ 0 & \text{if } |w| < \lambda \end{cases} \quad (1)$$

The authors proposed a universal threshold λ for the WT

$$\lambda = \sigma\sqrt{2\log(N)} \quad (2)$$

with $\sigma = MAD/0.6745$, where N is the length of y and σ is the noise level. *MAD* represent the absolute median estimated on the first scale. In the Wavelet Packets Transform (WPT) case, the threshold becomes:

$$\lambda = \sigma\sqrt{2\log(N\log_2 N)} \quad (3)$$

Johnstone and Silverman [7] studied the correlated noise situation and proposed a "level-dependent" threshold

$$\lambda_j = \sigma_j\sqrt{2\log(N_j)} \quad (4)$$

with $\sigma_j = MAD_j/0.6745$, and N_j is the number of samples in scale j.

The discriminatory threshold is also defined using other criterion like Minimax and SURE (Stain's Unbiased Risk Estimate) [6,7].

The time-adaptation of the threshold that takes into consideration the time behaviour of the noisy signal constitutes an interesting approach [10].

3. Application to speech enhancement

The wavelet transforms are successfully applied to improve the performance of speech enhancement methods. Unfortunately, the wavelet thresholding technique can not be applied directly because the simple threshold can not discriminate very efficiently the speech components from the noise ones. In this section we summarize the principal speech enhancement technique's using wavelet.

3.1 Wavelet thresholding

An algorithm based on the wavelet thresholding is proposed for speech enhancement algorithm [11]. 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.

3.2 Wiener filtering in the wavelet domain

The wavelet transform based Wiener filtering is a special application of the Wiener filtering. This idea arises from the fact that wavelet transforms tend to decorrelate data.

A multi-microphone system is proposed for speech enhancement [12]. The Wiener filtering performances in the wavelet domain are better than those obtained in the Fourier domain. Another version that combines Wiener and coherence in the wavelet domain is also proposed [13].

3.3 Wavelet filter bank

Most of the speech enhancement systems are conceived around filter banks. This tendency is justified by the behavior of the cochlea, which operates like a filter bank. In addition, it is recognized that 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 (Mel, Bark, etc...). The 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 [14].

A modified version of the speech enhancement method based on the coherence function is proposed [15]. The wavelet transform is also used as a filter bank.

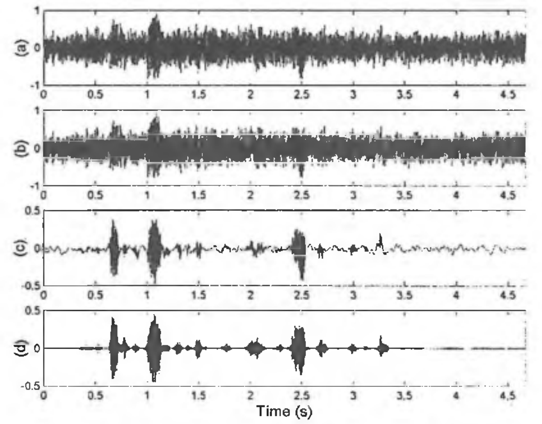

Fig. 1. Speech enhancement results: a) noisy speech recorded in a aircraft plane, b) enhancement using WT with an universal threshold, c) enhancement using WT with level-dependent threshold, and d) enhancement using WPT with level-dependent threshold.

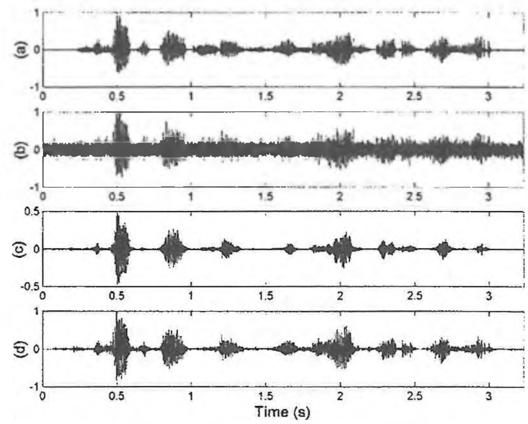

Fig. 2. Speech enhancement results: a) clean signal, b) noisy version (SNR=0dB), c) enhancement using WPT with level-dependent threshold, d) enhancement using WPT with level-dependent and *time-adapted* threshold.

4. Proposed approach

The wavelet thresholding technique is a simple method to remove noise from corrupted signal. As pointed out previously, it was not successfully applied in speech enhancement. These difficulties are simultaneously associated to the speech signal complexity and to the nature of the noise.

To improve the wavelet thresholding enhancement, we then propose:

- The use of the WPT instead of the WT,
- To extend the concept of the level-dependent threshold (Equation4) to the WPT,
- The use of *time-adapted* threshold based on the speech waveform energy.

5. Results

The proposed approach is tested using speech sound corrupted by white noise and speech recorded in real environments (in a sawmill, in a car, and aircraft cockpit).

The speech signal of Fig. 1-a is corrupted by a narrow-band noise. The standard threshold (the same value for all scales) of the WT (Fig. 1-b) is inefficient to remove noise (the same result is obtained by the standard threshold of the WPT). However, the filtered results using the level-dependent thresholding of the WPT are better when obtained by using the WT (respectively represented on Fig. 1-d and 1-c).

The *time-adapting* effect is carried out using speech from the TIMIT database (Fig. 2-a) that is corrupted by white noise (Fig. 2-b). It can be seen that the enhanced speech using the *time-adapted* threshold becomes very close to the original clean speech (Fig. 2-d).

5. Conclusion

The proposed method constitutes a successful application of the wavelet thresholding for speech enhancement. The level-dependent threshold using WPT permits to remove various environmental noises (narrow or large frequency-band). Whereas, the *time-adapting* of the threshold avoids the speech degradation quality during the thresholding process.

References

[1] S. Mallat and W.L. Hwang: Singularity Detection and Processing with Wavelets, IEEE Trans Inform Theory, 38 617-643, 1992.

[2] D.L. Donoho. Nonlinear wavelet methods for recovering signals, images, and densities from indirect and noisy data. Proceedings of Synposia in Applied Mathematics. 47 :173-205, 1993.

[3] Y. Xu, J.B. Weaver, D.M Healy, and J. Lu. Wavelet transform domain filters: A spatially selective noise filtration technique. IEEE Trans. Image Processing, 3(6) :747–758, Nov. 1994.

[4] Q. Pan, L. Zhang, G. Dai, and H. Zhang, Two denoising methods by wavelet transform, IEEE Trans. Signal Processing, vol. 47, no. 12, pp. 3401–3406, December 1999.

[5] D.L. Donoho and I.M. Johnstone. Ideal spatial adaptation by wavelet shrinkage. Biometrika, 81(3) :425–455, 1994.

[6] D.L. Donoho. De-noising by soft-thresholding. IEEE Trans. Inform. Theory, 41(3) :613–627, May 1995.

[7] I.M. Johnstone and B.W. Silverman. Wavelet threshold estimators for data with correlated noise. J. Roy. Statist. Soc. B, 59 :319–351, 1997.

[8] D.L. Donoho and I.M. Johnstone. Adapting to unknow smoothness via wavelet shrinkage. J. Amer. Stat. Assoc., 1200–1224, 1995.

[9] X.P. Zhang and M.T. Desai. Adaptative denoising based on SURE risk. IEEE Signal Processing Letters, 5(10) :265–267, October 1998.

[10] B. Vidakovic and C.B. Lozoya. On time-dependant wavelet denoising. IEEE Trans. Signal Processing, 46(9) :2549–2554, September 1998.

[11] J.W. Seok and K.S. Bae. Speech enhancement with reduction of noise components in the wavelet domain. In ICASSP 1997, pages 1223–1326, Munich, Germany, April 1997.

[12] D. Mahmoudi. A microphone array for speech enhancement using multiresolution wavelet transform. In Proc. Of Eurospeech'97, pages 339–342, Rhodes, Greece, Sept. 1997.

[13] D. Mahmoudi and A. Drygajlo. Combined wiener and coherence filtering in wavelet domain for microphone array speech enhancement. In ICASSP98, pages 385 –388, 1998.

[14] T. Gulzow, A. Engelsberg, and U. Heute. Comparison of a discrete wavelet transformation and nonuniform polyphase filterbank applied to spectral-subtraction speech enhancement. Signal Processing, 64 :5–19, 1998.

[15] J. Sika and V. Davidek. Multi-channel noise reduction using wavelet filter bank. In EuroSpeech'97, pages 2591-2594, Rhodes, Greece, Sept. 1997.