We are IntechOpen, the world's leading publisher of Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the
TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Speech Enhancement Based on LWT and Artificial Neural Network and Using MMSE Estimate of Spectral Amplitude

Mourad Talbi, Riadh Baazaoui and Med Salim Bouhlel

Abstract

In this chapter, we will detail a new speech enhancement technique based on Lifting Wavelet Transform (LWT) and Artifitial Neural Network (ANN). This technique also uses the $MMSE$ Estimate of Spectral Amplitude. It consists at the first step in applying the LWT to the noisy speech signal in order to obtain two noisy details coefficients, cD_1 and cD_2 and one approximation coefficient, cA_2. After that, cD_1 and cD_2 are denoised by soft thresholding and for their thresholding, we need to use suitable thresholds, $thr_j, 1 \leq j \leq 2$. Those thresholds, $thr_j, 1 \leq j \leq 2$, are determined by using an Artificial Neural Network (ANN). The soft thresholding of those coefficients, cD_1 and cD_2, is performed in order to obtain two denoised coefficients, cDd_1 and cDd_2 . Then the denoising technique based on $MMSE$ Estimate of Spectral Amplitude is applied to the noisy approximation cA_2 in order to obtain a denoised coefficient, cAd_2. Finally, the enhanced speech signal is obtained from the application of the inverse of LWT, LWT^{-1} to cDd_1, cDd_2 and cAd_2. The performance of the proposed speech enhancement technique is justified by the computations of the Signal to Noise Ratio (SNR), Segmental SNR ($SSNR$) and Perceptual Evaluation of Speech Quality ($PESQ$).

Keywords: ANN, Lifting Wavelet Transform, Ideal thresholds, Soft thresholding, MMSE Estimate

1. Introduction

Numerous speech enhancement techniques have been developed in the previous years as speech enhancement is a core target in numerous challenging domains such as speech and speaker recognitions, telecommunications, teleconferencing and hand-free telephony [1]. In such applications, the goal is to recover a speech signal from observations degraded by diverse noises components [2]. The unusual noise components can be of various classes that are frequently present in the environment [3]. Many algorithms and approaches were proposed for resolving the problem of degraded speech signals [4–6]. Furthermore, methods of single or multi-microphones are proposed in order to ameliorate the behaviour of the speech enhancement approaches and also to reduce the acoustic noise components even in very noisy conditions [2]. The most well-known single channel approaches that are extensively known in speech enhancement application is the spectrum subtraction (SS) that requires only one channel signal [7]. It has been embedded in some high-quality mobile phones for the same application. Though, the SS approach is just appropriate for stationary noise environments [2]. Furthermore, it surely introduces music noise problem. In fact, the higher the noise is reduced, the greater the alteration is brought to the speech signal and accordingly the poorer the intelligibility of the enhanced speech is obtained [8, 9]. As a result, ideal enhancement can hardly be attained when the Signal to Noise Ratio (SNR) of the noisy speech is relatively low; below $5\ dB$. However, it has quite good result when the noisy speech SNR is relatively high; above $15\ dB$ [2]. The SS and other speech enhancement methods that are based on SS principal have ameliorated the decision directed (DD) methods in reducing the musical noise components [10–13]. Numerous algorithms that ameliorate the DD methods were suggested in [14]. In [15], a speech enhancement technique based on high order cumulant parameter estimation was proposed. In [16, 17], a subspace technique, which is based on Singular Value Decomposition (SVD) approaches was proposed; the signal is enhanced when the noise subspace is eliminated, and accordingly, the clean speech signal is estimated from the subspace of the noisy speech [2]. Another technique which was widely studied in speech enhancement application, is the adaptive noise cancellation (ANC) approach that was firstly suggested in [18, 19]. Moreover, most important speech enhancement methods employed adaptive approaches for getting the tracking ability of non-stationary noise properties [20, 21]. Numerous adaptive techniques were proposed for speech enhancement application, we can find time domain algorithm, frequency domain adaptive algorithms [22–26] or adaptive spatial filtering methods [27, 28] that frequently employ adaptive SVD methods in order to separate the speech signal space from the noisy one. Another direction of research combines the Blind Source Separation (BSS) methods with adaptive filtering algorithms for enhancing the speech signal and to cancel effectively the acoustic echo components [29–32]. This approach employs at least two microphones configuration in order to update the adaptive filtering algorithms. Also, a multi-microphone speech enhancement technique was proposed for the same aim and ameliorated the existing one-channel and two-channel speech enhancement and noise reduction adaptive algorithms [33, 34]. Numerous research works highlighted the problem of speech enhancement on a simple problem of mixing and unmixing signals with convolutive and instantaneous noisy observations [35–37]. In the last decade, a novel research direction has proven the efficacy of the wavelet domain as a novel effective mean that can ameliorates the speech enhancement approaches, and numerous algorithms and methods have been proposed for the same aim [38, 39]. In this chapter, we propose a novel speech enhancement technique based on Lifting Wavelet Transform (LWT) and Artifitial Neural Network (ANN) and also uses $MMSE$ Estimate of Spectral Amplitude [40]. The presentation of this chapter is given as follows: after the introduction, we will deal with Lifting Wavelet Transform (LWT), in section 2. Section 3 is reserved to describe the proposed speech enhancement technique. Section 4 presents the obtained results and finally we conclude our work in section 5.

2. Lifting wavelet transform (LWT)

The Lifting Wavelet Transform (LWT) becomes a powerful tool for signal analysis due to its effective and faster implementation compared to the Discrete Wavelet Transform (DWT). In the domains of the signal denoising, signal compression and watermarking, the LWT permits to obtain better results compared to the DWT . The LWT permits to saves times and has a better frequency localization feature that overcomes the shortcomings of DWT. The Signal decomposition using LWT needs three steps: splitting, prediction and update.

3. The proposed technique

In this chapter, we propose a new speech enhancement technique based on Lifting Wavelet Transform (LWT) and Artifitial Neural Network (ANN). This technique also uses the $MMSE$ Estimate of Spectral Amplitude. It consists at the first step in applying the LWT to the noisy speech signal in order to obtain two noisy details coefficients, cD_1 and cD_2 and one approximation coefficient, cA_2. After that, cD_1 and cD_2 are denoised by soft thresholding and for their thresholding, we need to use suitable thresholds, $thr_j, 1 \leq j \leq 2$. Those thresholds, $thr_j, 1 \leq j \leq 2$, are determined by using an Artificial Neural Network (ANN). The soft thresholding of those coefficients, cD_1 and cD_2, is performed in order to obtain two denoised coefficients, cDd_1 and cDd_2. Then the denoising technique based on MMSE Estimate of Spectral Amplitude [40] is applied to the noisy approximation cA_2 in order to obtain a denoised coefficient, cAd_2. Finally, the enhanced speech signal is obtained from the application of the inverse of LWT, LWT^{-1} to cDd_1, cDd_2 and cAd_2. For each coefficient, $cD_j, 1 \leq j \leq 2$, the corresponding ideal $thr_j, 1 \leq j \leq 2$ is computed using the following MATLAB function:

```
function [thr] = Compute_Threshold (cc,cb).
R = [];
for(i = 1:length(cb)).
   r = 0;
   for(j = 1:length(cc)).
      r = r + (wthresh(cb(j),'s',abs(cb(i)))-cc(j)).^2;
   end;
      R = [R r];
end;
[guess,ibest] = min(R);
thr = abs(cb(ibest));
```

The inputs of this function are the clean details coefficient, cc and the corresponding noisy coefficient, cb. The output of this function is the ideal threshold, thr. Note that the couple (cc, cb) can be (ccD_1, cD_1) or (ccD_2, cD_2) where ccD_j *and* $cD_j, 1 \leq j \leq 2$ are respectively the clean details coefficient and the corresponding noisy details one at the level j.

In this chapter and as previously mentioned, the ideal threshold at level j is thr_j and is used for soft thresholding of the noisy details coefficient cD_j at that level. In this work, the used ANN is trained by a set of couples (P, T) where P is the input of this neural network and is chosen to be cD_j and T is the Target and is chosen to be the corresponding ideal threshold, thr_j at level j. Consequently, for computing a suitable threshold to be used in soft thresholing of $cD_j, 1 \leq j \leq 2$, we use one ANN so we have two $ANNs$ and two different suitable thresholds. Each of those $ANNs$ is constituted of two layers where the first one is a hidden layer and contains ten neurons having tansigmoid activation function and the second layer is the output one and contains one neurone having purelin activation function (**Figure 1**).

As shown in **Figure 1**, the input of this ANN is the noisy details coefficient at level $1 \leq j \leq 2$, $P = cD_j, 1 \leq j \leq 2$ and the desired output or Target, $T = thr_j, 1 \leq j \leq 2$. The activation functions tansigmoid and pureline are respectively expressed by Eqs. 1 and 2.

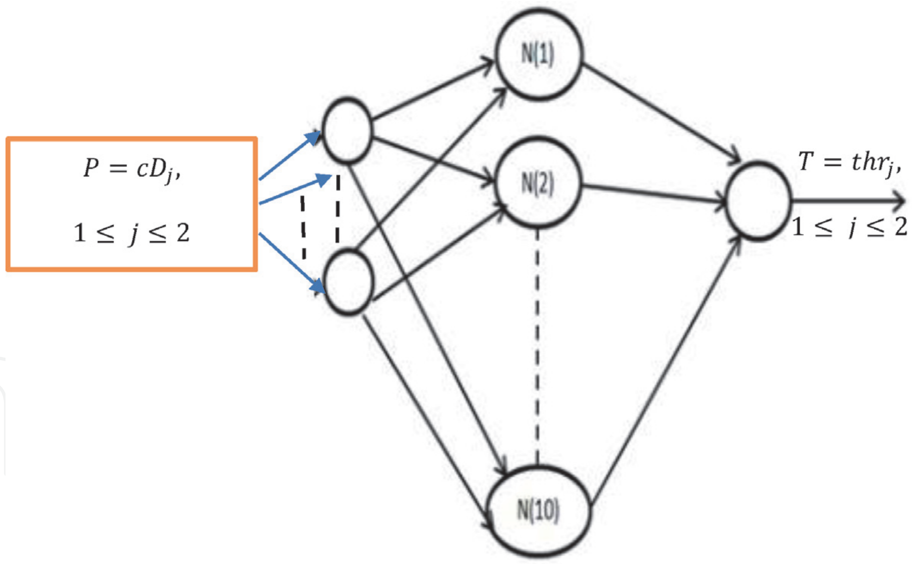

Figure 1.
The architecture of the ANN used in this work.

$$tansig\ (n) = 1/(1 + exp\ (-n)) \tag{1}$$

$$Purelin\ (n) = n \tag{2}$$

Generally, neural networks consist of a minimum of two layers (one hidden layer and another output layer). The input information is connected to the hidden layers through weighted connections where the output data is calculated. The number of hidden layers and the number of neurons in each layer controls the performance of the network. According to [41], there are no guidelines for deciding a manner for selecting the number of neurons along with number of hidden layers for a given problem to give the best performance. And it is still a trial and error design method [41].

For training each *ANN* used in this work we have employed 50 speech signals and 10 others used for testing those networks. Therefore, for training each used ANN, we used 50 couples of Input and Target (P, T). Evidently, the noisy speech signals used for the *ANNs* testing do not belong to the training database. The different parameters used for the training of the used ANNs are the epochs number which is equal to 5000, the momentum, μ or Mu which is equal to 0.1, the gradient minimum which is equal to $1e - 7$. The employed training algorithm is Leverberg-Marquardt.

In summary, the novelty of the proposed technique consists in applying the denoising technique based on MMSE Estimate of Spectral Amplitude [40]. Also, we apply the ANN for computing ideals thresholds to be used for thresholding of the noisy details coefficients obtained from the application of the *LWT* to the noisy speech signal.

4. Results and discussions

For the evaluation of the proposed technique, we have applied it to twenty Arabic speech signals pronounced by a male and female speakers. Those signals are corrupted in artificial manner with additive manner by two types of noise (White Gaussian and Car noises) at different values of *SNRi* before denoising.. The used Arabic speech signals (**Table 1**) are material phonetically balanced and they are sampled at 16 *kHz*.

Also for the evaluation of the proposed speech enhancement approach, we have applied the denoising technique based on *MMSE* Estimate of Spectral Amplitude [40]. This evaluation is performed in terms of Signal to Noise Ratio (*SNR*), the Segmental *SNR* (*SSNR*) and the Perceptual Evaluation of Speech Quality (*PESQ*). In **Tables 2–7** are listed the results obtained from the computations of SNRf (after

Female speaker	Male speaker
Signal1: أحفظ من الأرض	**Signal 1**: لا لن يذيع الخبر
Signal 2: أين المسا فرين	**Signal 2**: أكمل بالإسلام رسالتك
Signal 3: لا لم يستمتع بثمرها	**Signal 3**: سقطت إبرة
Signal 4: سيؤذيهم زماننا	**Signal 4**: من لم ينتفع
Signal 5: كنت قدوة لهم	**Signal 5**: غفل عن ضحكاتها
Signal 6: ازار صائما	**Signal 6**: و لماذا نشف مالهم
Signal 7: كال و غبط الكبش	**Signal 7**: أين زوايانا و قانوننا
Signal 8: هل لذعته بقول	**Signal 8**: صاد الموروث مدلعا
Signal 9: عرف واليا و قائدا	**Signal 9**: نبه آبائكم
Signal 10: خالا بالنا منكما	**Signal 10**: أظهره و قم

Table 1.
The list of the employed Arabic speech sentences.

SNRi (dB)	SNRf (dB)	
	The Denoising approach	
	The proposed speech enhancement technique	The denoising technique based on MMSE Estimate of Spectral Amplitude [40]
−5	**8.3650**	**7.1431**
0	**13.0857**	**11.6110**
5	**16.9010**	**15.5721**
10	**19.8933**	**18.8719**
15	**22.3135**	**21.7972**

Table 2.
Results in term of SNR (signal 4 (female voice) corrupted by Gaussian white noise).

SNRi (dB)	SSNR (dB)	
	The Denoising approach	
	The proposed speech enhancement technique	The denoising technique based on MMSE Estimate of Spectral Amplitude [40]
−5	**−0.0954**	−0.7089
0	**2.5997**	1.7725
5	**4.7373**	3.9719
10	**6.8038**	5.9329
15	**9.5324**	8.7158

Table 3.
Results in term of SSNR (signal 4 (female voice) corrupted by Gaussian white noise).

denosing), of SSNR and PESQ and this for the proposed technique and The denoising technique based on MMSE Estimate of Spectral Amplitude [40].

According to those results listed in **Tables 2–7**, the best results are in Red colour. In terms of *SNRf* (After denoising), and *SSNR*, the best results are those obtained

SNRi (dB)	PESQ	
	The Denoising approach	
	The proposed speech enhancement technique	The denoising technique based on MMSE Estimate of Spectral Amplitude [40]
−5	**1.3225**	**1.3755**
0	**1.5935**	**1.6320**
5	**1.8812**	**1.8977**
10	**2.2016**	**2.2311**
15	**2.5147**	**2.6079**

Table 4.
Results in term of PESQ (signal 4 (female voice) corrupted by Gaussian white noise).

SNRi (dB)	SNRf (dB)	
	The Denoising approach	
	The proposed speech enhancement technique	The denoising technique based on MMSE Estimate of Spectral Amplitude [40]
−5	5.8737	4.2192
0	**9.8414**	8.3451
5	**14.1647**	12.6024
10	**18.5308**	17.4120
15	**22.5102**	21.4578

Table 5.
Results in term of SNR (signal 2 (male voice) corrupted by car noise).

SNRi (dB)	SSNR (dB)	
	The Denoising approach	
	The proposed speech enhancement technique	The denoising technique based on MMSE Estimate of Spectral Amplitude [40]
−5	**0.2145**	−1.1347
0	**2.7478**	1.7861
5	**5.6644**	4.7166
10	**8.8942**	7.8228
15	**11.9663**	10.9850

Table 6.
Results in term of SSNR (signal 8 (male voice) corrupted by car noise).

from the application of the proposed speech enhancement technique. However, in term of PESQ, the denoising technique based on *MMSE* Estimate of Spectral Amplitude [40] is slightly better than the proposed technique.

In **Figures 2–5** are illustrated some examples of speech enhancement using the proposed technique.

These Figures show the efficiency of the proposed speech enhancement technique. In fact, it permits to reduce considerably the noise while conserving the original signal and this especially when the *SNRi* is higher (5, 10 and 15 dB).

In our future work and in order to improve this proposed speech enhancement technique, we will use a Deep Neural Network instead (DNN) instead of a simple ANN and other transforms such as Empirical Mode Decomposition (EMD).

SNRi (dB)	PESQ	
	The Denoising approach	
	The proposed speech enhancement technique	The denoising technique based on MMSE Estimate of Spectral Amplitude [40]
−5	2.2837	2.4021
0	2.5999	2.7163
5	2.8709	3.0184
10	3.1190	3.2461
15	3.3590	3.4789

Table 7.
Results in term of PESQ (signal 8 (male voice) corrupted by car noise).

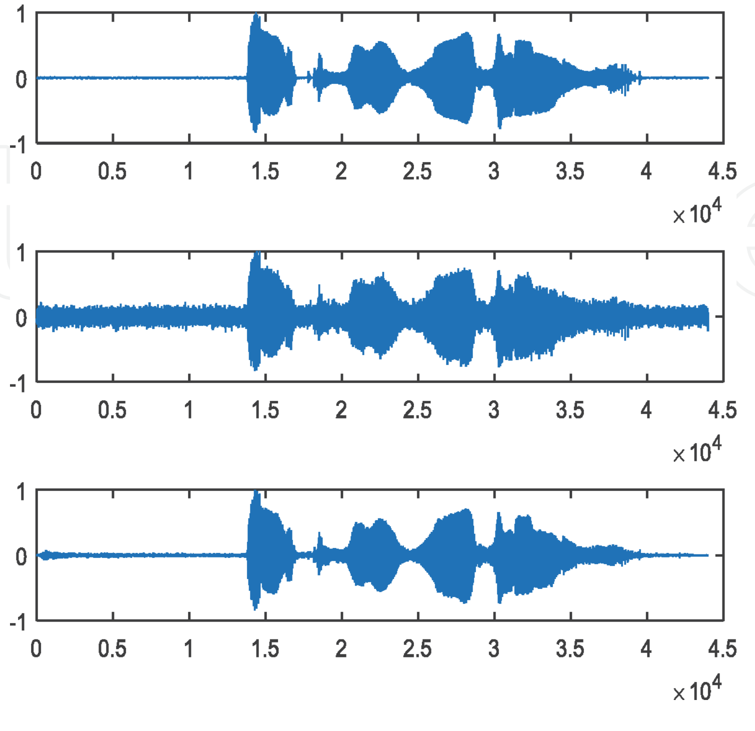

Figure 2.
*An example of speech enhancement applying the proposed technique: Signal 4 (pronounced by a female voice (**Table 1**) corrupted by Gaussian white noise with SNRi = 10dB (before enhancement)). After enhancement we have: SNRf = 19.8933, SSNR = 6.8038 and PESQ = 2.2016.*

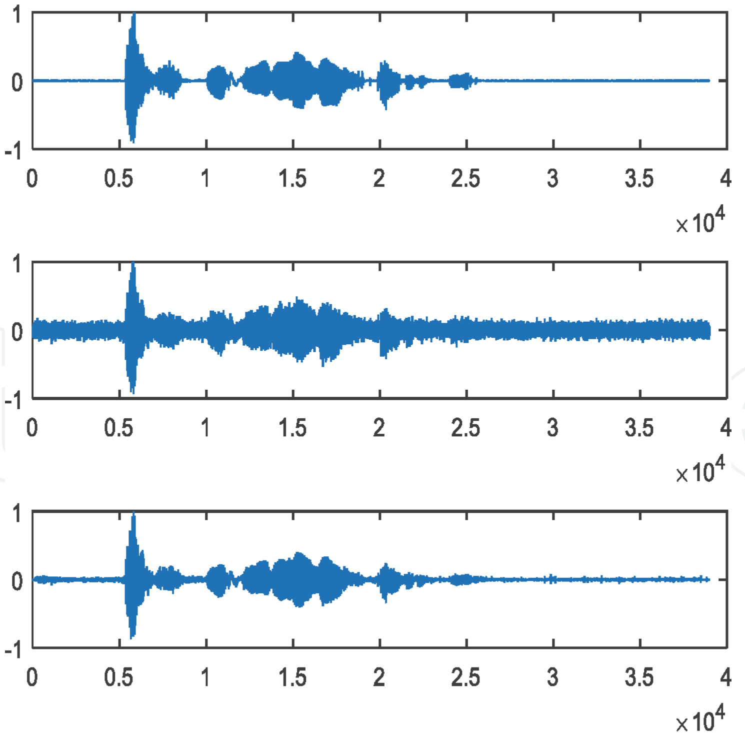

Figure 3.
*An example of speech enhancement applying the proposed technique: Signal 1 (pronounced by a male voice (**Table 1**) corrupted by Gaussian white noise with SNRi = 5dB (before enhancement)). After enhancement we have: SNRf = 13.7710 , SSNR = 0.7135 and PESQ = 2.2350.*

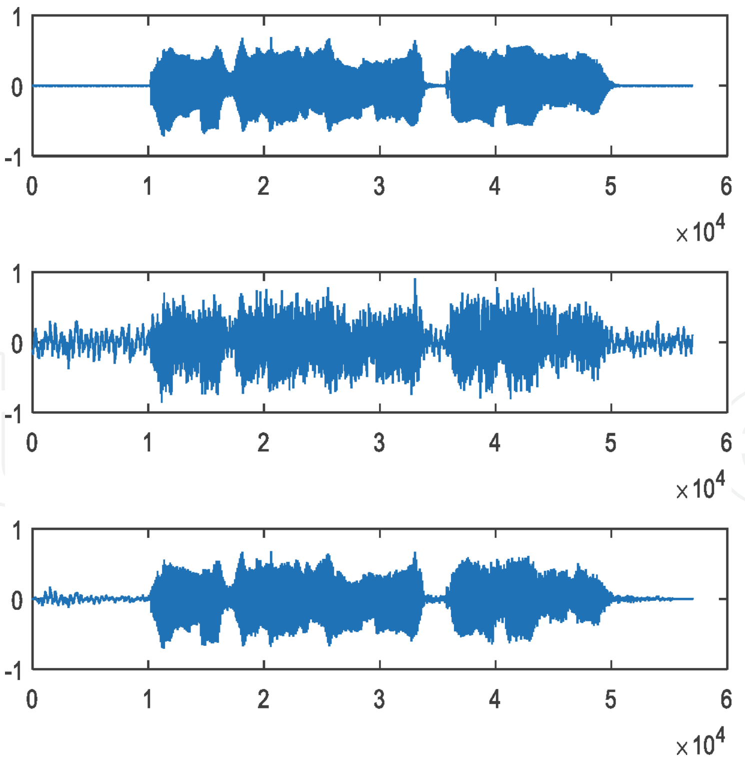

Figure 4.
*An example of speech enhancement applying the proposed technique: Signal 7 (pronounced by a male voice (**Table 1**) corrupted by car noise with SNRi = 5dB (before enhancement)). After enhancement we have: SNRf = 15.1244, SSNR = 8.7594 and PESQ = 3.3304.*

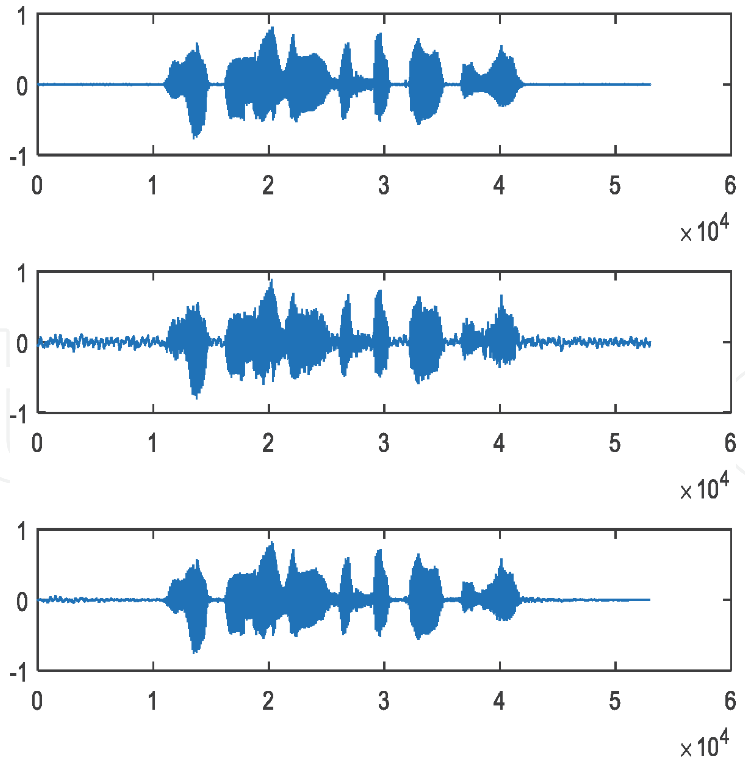

Figure 5.
*An example of speech enhancement applying the proposed technique: Signal 5 (pronounced by a male voice (**Table 1**) corrupted by car noise with SNRi = 10dB (before enhancement)). After enhancement we have: SNRf = 18.8848, SSNR = 6.4497 and PESQ = 3.5469.*

5. Conclusion

In this chapter, we will detail a new speech enhancement technique based on Lifting Wavelet Transform (LWT) and Artifitial Neural Network (ANN). This technique also uses the $MMSE$ Estimate of Spectral Amplitude. It consists at the first step in applying the LWT to the noisy speech signal in order to obtain two noisy details coefficients, cD_1 and cD_2 and one approximation coefficient, cA_2. After that, cD_1 and cD_2 are denoised by soft thresholding and for their thresholding, we need to use suitable thresholds, $thr_j, 1 \leq j \leq 2$. Those thresholds, $thr_j, 1 \leq j \leq 2$, are determined by using an Artificial Neural Network (ANN). The soft thresholding of those coefficients, cD_1 and cD_2, is performed in order to obtain two denoised coefficients, cDd_1 and cDd_2. Then the denoising technique based on $MMSE$ Estimate of Spectral Amplitude is applied to the noisy approximation cA_2 in order to obtain a denoised coefficient, cAd_2. Finally, the enhanced speech signal is obtained from the application of the inverse of LWT, LWT^{-1} to cDd_1, cDd_2 and cAd_2. The performance of the proposed speech enhancement technique is justified by the computations of the Signal to Noise Ratio (SNR), Segmental SNR ($SSNR$) and Perceptual Evaluation of Speech Quality ($PESQ$).

Author details

Mourad Talbi[1]*, Riadh Baazaoui[2] and Med Salim Bouhlel[3]

1 Laboratory of Nanomaterials and Systems for Renewables Energies (LaNSER), Center of Researches and Technologies of Energy of Borj Cedria, Tunis, Tunisia

2 Faculty of Sciences of Tunis, University of Tunis El-Manar, Tunisia

3 Sciences and Technologies of Images and Telecommunications, Higher Institute of Electronics and Telecommunication of Sfax, Sfax University Sfax, Tunisia

*Address all correspondence to: talbi1969@yahoo.fr

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

References

[1] Loizou, P. C. (2007). Speech enhancement: Theory and practice (pp. 589–599). Boca Raton, FL: Taylor and Francis.

[2] Mohamed, D. (2018). An efficient wavelet-based adaptive filtering algorithm for automatic blind speech enhancement. *International Journal of Speech Technology*, 21:355–367.

[3] Djendi, M., Scalart, P., & Gilloire, A. (2013). Analysis of two-sensor forward BSS structure with post-filters in the presence of coherent and incoherent noise. Speech Communication, 55(10), 975–987.

[4] Dixit, S., & Mulge, M. Y. (2014). Review on speech enhancement techniques, International Journal of Computer Science and Mobile Computing, 3(8), 285–290

[5] Bactor, P., & Garg, A. (2012). Different techniques for the enhancement of the intelligibility of a speech signal. International Journal of Engineering Research and Development, 2(2), 57–64.

[6] Scalart, P., Filho, J. (1996). Speech enhancement based on a priori signal to noise estimation. In International Conference on Acoustics, Speech, and Signal Processing. pp. 629–632.

[7] Boll, SF (1979). Suppression of acoustic noise in speech using spectral subtraction. IEEE Transactions on Acoustics, Speech and Signal Processing, 27, 113–120.

[8] Zhang, Y., & Zhao, Y. (2012). Real and imaginary modulation spectral subtraction for speech enhancement. *Journal on Speech Communication*, 55(6), 509–522.

[9] Cappé, O. (1994). Elimination of the musical noise phenomenon with the Ephraïm and Malah noise suppressor. *IEEE Transactions on Speech Audio Processing*, 2(2), 345–349.

[10] Ephraim, Y., & Malah, D. (1984). Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 32(6), 1109–1121.

[11] Ephraim, Y., & Malah, D. (1985). Speech enhancement using a minimum mean-square error log-spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP33(2), 443–445.

[12] Ephraim, Y., LevAri, H., Roberts, W. J. J. (2014). A brief survey of speech enhancement. *IEEE Signal Processing Letters*, 10, 104–106.

[13] Selva Nidhyananthan, S., Shantha Selva Kumari, R., & Arun Prakash, A. (2014). A review on speech enhancement algorithms and why to combine with environment classification. *International Journal of Modern Physics C*, 25(10), 210–225.

[14] Wolfe, P. J., & Godsill, S. J. (2003). Efficient alternatives to the Ephraim and Malah suppression rule for audio signal enhancement. *EURASIP Journal on Applied Signal Processing*, 10, 1043–1051.

[15] Dong, J., Wei, X. P., & Zhang, Q. (2009). Speech enhancement algorithm based on high-order Cumulant parameter estimation. *International Journal of Innovative Computing information and Control*, 5, 2725–2733.

[16] Davila, C. E. (1984). A subspace approach to estimation of autoregressive parameters from noisy measurements. IEEE Transaction on Signal processing, 46, 531–534.

[17] Doclo, S., & Moonen, M. (2002). GSVD-based optimal filtering for signal and multi-microphone speech enhancement. *IEEE Transaction on Signal processing*, 50, 2230–2244.

[18] Widrow, B., Goodlin, R. C. (1975). Adaptive noise cancelling: Principles and applications. *Proceedings of the IEEE*, 63, 1692–1716.

[19] Widrow, B., & Stearns, S. D. (1985). Adaptive signal processing, Upper Saddle River: Prentice-Hall.

[20] Lee, K. A., & Gan, W. S. (2004). Improving convergence of the NLMS algorithm using constrained subband updates. IEEE Signal Processing Letters, 11(9), 736–739.

[21] Weinstein, E., Feder, M., & Oppenheim, A. V. (1993). Multi-channel signal separation by decorrelation. IEEE Transactions on Speech Audio Processing, 1(4), 405–413.

[22] Plapous, C., Marro, C., Scalart, P., Mauuary, L., & Two-Step, A. (2004). Noise reduction technique. In IEEE International Conference on Acoustics, Speech, Signal Processing, Montral, Quebec Canada, 1, pp. 289–292.

[23] Al-Kindi, M. J., & Dunlop, J. (1989). Improved adaptive noise cancellation in the presence of signal leakage on the noise reference channel. Signal Processing, 17(3), 241–250.

[24] Plapous, C., Marro, C., Scalart, P. (2005). Speech enhancement using harmonic regeneration, In IEEE International Conference on Acoustics, Speech, Signal Processing, Philadelphia, PA, USA, 1, pp. 157–160.

[25] Djendi, M., Khemies, F., & Morsli, A. (2015). A Frequency Domain Adaptive Decorrelating Algorithm for Speech Enhancement. In International Conference on Speech and Computer, SPECOM 2015, pp. 51–54.

[26] Djendi, M., Bensafia, S., & Safi, M. (2016). A frequency co-channel adaptive algorithm for speech quality enhancement, In International Conference on Engineering and MIS (ICEMIS).

[27] Tong, R., Bao, G., & Ye, Z. (2015). A higher order subspace algorithm for multichannel speech enhancement. IEEE Signal Processing Letters, 22(11), 2004–2008.

[28] Goldsworthy, R. L. (2014). Two-microphone spatial filtering improves speech reception for cochlear-implant users in reverberant conditions with multiple noise sources. Trends in Hearing, 18, 1–13.

[29] Van Gerven, S., & Van Compernolle, D. (1992). Feed forward and feedback in symmetric adaptive noise canceller: Stability analysis in a simplified case. In European Signal Processing Conference, Brussels, Belgium. pp. 1081–1084.

[30] Djendi, M., Scalart, P., & Gilloire, A. (2006). Noise cancellation using two closely spaced microphones: Experimental study with a specific model and two adaptive algorithms. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 3, pp. 744–747.

[31] Zoulikha, M., & Djendi, M. (2016). A new regularized forward blind source separation algorithm for automatic speech quality enhancement. Applied Acoustics, 112, 192–200

[32] Jin, Y. G., Shin, J. W., & Kim, N. S. (2014). Spectro-temporal filtering for multichannel speech enhancement in short-time Fourier transform domain. IEEE Signal Processing Letters, 21(3), 352–355.

[33] Benesty, J., & Cohen, I. (2017). Multichannel speech enhancement in the STFT domain. In J. Benesty, & I. Cohen (Eds.), Canonical correlation analysis in speech enhancement, Springer briefs in electrical and computer engineering (pp. 79–101). New York: Springer.

[34] Lee, G., & Dae Na, S. (2016). Seong K2, Cho JH3, Nam Kim M4. Wavelet speech enhancement algorithm using exponential semisoft masks filtering. Bioengineered, 7(5), 352–356.

[35] Jutten, C., & Herrault, J. (1991). Blind separation of sources: an adaptive algorithm based on neuromimetic architecture. Signal Processing, 24, 1–10.

[36] Nguyen Thi, H. L., & Jutten, C. (1995). Blind sources separation for convolutive mixtures. Signal Processing, 45, 209–229.

[37] Mansour, A., Jutten, C., & Loubaton, P. (1996). Subspace method for blind separation of sources and for a convolutive mixture model. Signal processing VIII, theories and applications (pp. 2081–2084).

[38] Bouzid, A., & Ellouze, N. (2016). Speech enhancement based on wavelet packet of an improved principal component analysis. Journal Computer Speech and Language, 35, 58–72.

[39] Ghribi, K., Djendi, M., & Berkani, D. (2016). A wavelet-based forward BSS algorithm for acoustic noise reduction and speech enhancement. Applied Acoustics, 105, 55–66.

[40] Timo Gerkmann, and Richard C. Hendriks. Unbiased MMSE-Based Noise Power Estimation With Low Complexity and Low Tracking Delay. IEEE Transactions on Audio, Speech, And Language Processing, Vol. 20, No. 4, May 2012.

[41] http://www.aast.edu/en/openfiles/opencmsfiles/pdf_retreive_cms.php?disp_unit=351/Hatem%20Diab-Thesis.pdf