NON-LINEAR WEIGHTING FUNCTION FOR NON-STATIONARY SIGNAL DENOISING

Farès Abda, David Brie, Radu Ranta

CRAN UMR 7039 CNRS-UHP-INPL, B.P. 239, 54506 Vandœuvre-lès-Nancy, France

ABSTRACT

We propose in this paper a new strategy for non-stationary signals denoising based on designing a time-varying filter adapted to the signal short term spectral characteristics. The basic idea leading us to use a new parametric nonlinear weighting of the measured signal short term spectral amplitude (STSA) is exposed. The overall system consists in combining the estimated STSA and the complex exponential of the noisy phase. The proposed technique results in a significant reduction of the noise for a variety of non-stationary signals including speech signals.

1. INTRODUCTION

The denoising problem consisting in enhancing a signal degraded by uncorrelated additive noise when only the noisy measure is available, is a fundamental task in signal processing. Because of the non-stationarity of the original signal, the short-time Fourier analysis and synthesis concepts have been widely used [1–4]. The noisy signal is decomposed into spectral components by means of the short-time Fourier transform. The advantages of the spectral decomposition results from a good separation of the original signal and noise as well as the decorrelation of spectral components. Therefore, it is possible to treat the frequency bins independently and the estimation problem is simplified since the time varying filtering procedure can be implemented in the frequency domain where multiplicative modifications are applied on the short-time spectrum and the enhanced signal is synthesized from the modified short-time spectrum using the OverLap Add method (OLA). In general, it is significantly easier to estimate the STSA of the original signal than to estimate both amplitude and phase. There are a variety of denoising techniques that capitalize on this aspect by focusing on enhancing only the STSA of the noisy signal. We propose in this article a new strategy for the restoration of the STSA. We introduce a parametric spectral gain formulation based on a continuous non-linear function, namely, the arctangent.

The paper is organized as follows. In section 2 we introduce the necessary concepts of analysis and synthesis using the short-time Fourier transform, with a special attention to the OLA method and the effects of spectral modifications on the reconstructed signal. In section 3, we expose the overall scheme of the denoising system using the nonlinear spectral weighting method. In section 4, we discuss the empirical parameters determination. In section 5, we summarize the paper and draw conclusions.

2. SHORT-TIME FOURIER ANALYSIS AND SYNTHESIS

Let $x(n)$ and $d(n)$ denote the original signal and the statistically independent additive noise respectively. The set of observations expressed using the model $y(n) = x(n) + d(n)$ is analysed using the short-time discrete Fourier transform defined as

$$Y_k(n) = \sum_m w(n-m)y(m)e^{-j2\pi mk/L} \tag{1}$$

where $w(n-m)$ is an appropriate sliding window (e.g. Hamming window) of size N, and $L \geq N$ is the number of analysis frequencies $\omega_k = 2\pi k/L$ for $k = 0, \ldots, L-1$. The denoising procedure may be viewed as the application of a weighting rule, or nonnegative real-valued spectral gain $G_k(n)$, to each bin k of the observed short-time spectrum $Y_k(n)$, in order to form an estimate $\hat{X}_k(n)$ of the original signal short-time spectrum. We recall in the following the OLA method [1] for reconstructing the estimated signal $\hat{x}(n)$ from the modified short-time spectrum $\hat{X}_k(n)$. Expressing a time-varying multiplicative modification to the short-time spectrum of $y(n)$ as

$$\hat{X}_k(n) = Y_k(n)G_k(n), \tag{2}$$

the synthesized signal is obtained by inverse Fourier transform of $\hat{X}_k(n)$, and summing over all the windowed frames

$$\hat{x}(n) = \sum_m \left[\frac{1}{L} \sum_k Y_k(m)G_k(m)e^{j2\pi kn/L} \right]. \tag{3}$$

By defining the time-varying impulse response corresponding to $G_k(n)$ as

$$g_n(m) = \frac{1}{L} \sum_k G_k(m)e^{j2\pi kn/L}, \tag{4}$$

and using the definition of $Y_k(n)$, we obtain from (3)

$$\hat{x}(n) = \sum_l y(l) \left[\sum_m w(m-l)g_m(n-l) \right]. \tag{5}$$

If we let $r = n - l$ or $l = n - r$, the latter relation becomes

$$\hat{x}(n) = \sum_r y(n-r) \left[\sum_m w(m-n+r) g_m(r) \right]. \quad (6)$$

If we define

$$\hat{g}(r-n, r) = \hat{g}(q, r) = \sum_m g_m(r) w(m-q) \quad (7)$$

then (5) becomes

$$\hat{x}(n) = \sum_r y(n-r) \hat{g}(r-n, r). \quad (8)$$

We see from equation (7) that the resulting time-varying filter $\hat{g}(r-n, r)$ is a filtered version of $g_m(n)$ by means of the low-pass filter $w(n)$. However, in our case the spectral time-varying filter $G_k(n)$ is a nonlinearly obtained spectral weighting from $Y_k(n)$. Thus, time-varying modifications applied to the measured short-time spectrum $Y_k(n)$ are only dependent on the latter. Therefore, since $Y_k(n)$ is a slowly time-varying function (by taking into account the band limiting effect of the analysis window), the analysis window will not have any significant effect on the designed time-varying filter.

We do not discuss here the selection of the rate at which $Y_k(n)$ should be sampled in both time and frequency to provide an unaliased representation of $Y_k(n)$. This point is detailed in [1] from which it turns out that a properly sampled short-time transform using a Hamming window requires on the order of four times more information as would be required relative to the original signal. In turn to this redundancy one obtains a very flexible signal representation for which extensive modifications can be made.

3. NONLINEAR WEIGHTING OF THE SHORT-TIME SPECTRAL AMPLITUDE

Most of the denoising techniques operate in the frequency domain by applying a frequency dependent gain function to the spectral components of the noisy signal, in an attempt to heavily attenuate the noise only spectral components, while preserving those corresponding to the original signal. These denoising algorithms commonly consists of three major components: a spectral analysis/synthesis system which was described in section 2, a noise estimation algorithm, and a spectral gain computation. The gain modifies only the Fourier magnitudes of an input frame.

3.1. Minimum statistics noise PSD estimation

Noise estimation usually involves some kind of SNR based signal activity detection which restrict the update of the noise estimate to periods of signal absence. However, these traditional noise estimation approaches are difficult to tune and their application to low SNR signals results often in distorted signals. Due to these reasons, we used a useful noise estimation method, known as the minimum statistics [5–7], which consists in tracking the minima values of a smoothed power estimate of the noisy signal.

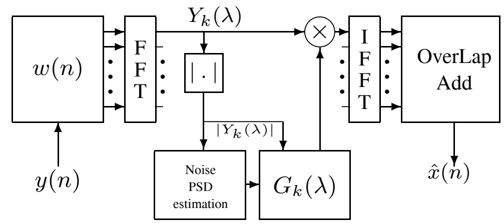


Fig. 1: Block diagram of a single channel signal denoising algorithm.

3.2. The spectral gain calculation

The time-varying filter $G_k(n)$ is specified by a pass region ($G_k(n) \approx 1$) which covers the effective support region of the original signal $x(n)$, a stop region ($G_k(n) \approx 0$) which has to suppress the undesired spectral components representative of the noise and an intermediate region ($0 < G_k(n) < 1$) where both signal and noise are present with values depending on the relative spectral values of the original signal and noise.

Based on this intuitively appealing interpretation of the spectral weighting function $G_k(n)$, and given an estimate of the noise PSD $\hat{\sigma}_k^2(n)$, we propose to design the time-varying filter as

$$G_k^{\alpha,\beta}(n) = \left\{ \frac{1}{\pi} \arctan\left[\beta\left(|Y_k(n)| - \alpha\hat{\sigma}_k(n)\right)\right] + \frac{1}{2} \right\}, \quad (9)$$

where $\alpha, \beta > 0$ are design parameters that have to be chosen in order to adjust at best the shape of $G_k(n)$. The values of α and β control the suppression level of the noise as well as the resulting signal distorsion. This nonlinear gain function is depicted in figure (2). We see that this function takes very low values corresponding to small spectral amplitudes of the measured signal compared to the estimated noise PSD, and takes values close to the unity for large values of the spectral amplitudes. Consequently, this function performs a selective weighting of spectral components based on the noise PSD estimate. However, since the gain function (9) is sensitive to the dynamical range of the spectral amplitudes, the input signal has to be normalised with its maximal value in order to obtain the same variation range for all signals. An other solution to this problem consists in using the modified gain function expressed as

$$\tilde{G}_k^{\alpha,\beta}(n) = \left\{ \frac{1}{\pi} \arctan\left[\beta\left(\sqrt{\gamma_k(n)} - \alpha\right)\right] + \frac{1}{2} \right\}, \quad (10)$$

where

$$\gamma_k(n) = \frac{|Y_k(n)|^2}{\hat{\sigma}_k^2(n)} \quad (11)$$

is the *a posteriori* SNR as defined in [2].

The proposed short-term spectral amplitude estimator is therefore given by

$$|\hat{X}_k^{\alpha,\beta}(n)| = \left\{\frac{1}{\pi}\arctan\left[\beta\left(\sqrt{\gamma_k(n)}-\alpha\right)\right]+\frac{1}{2}\right\}|Y_k(n)| = \tilde{G}_k^{\alpha,\beta}(n)|Y_k(n)|. \quad (12)$$

Finally, by combining this spectral amplitude estimate with the noisy phase of the measured signal, we obtain the short-term spectral estimate as

$$\hat{X}_k^{\alpha,\beta}(n) = \tilde{G}_k^{\alpha,\beta}(n)Y_k(n), \quad (13)$$

and the corresponding synthesized signal $\hat{x}^{\alpha,\beta}(n)$ is obtained using (3).

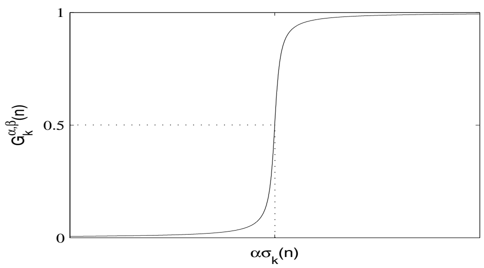


Fig. 2: Weighting function (9) for $\beta = 3$.

4. EMPIRICAL PARAMETER DETERMINATION

We consider in this section the experimental determination of the parameters α and β using the Nelder-Mead algorithm for the minimisation of the MSE between the noise free and the estimated signal calculated as $\frac{1}{M}\sum_{n=1}^{M}\left[x(n)-\hat{x}^{\alpha,\beta}(n)\right]^2$, where M the length of the input signal. We used in a first study four speech signals corrupted by computer-generated white Gaussian noises in order to have different SNRs. The obtained optimal values of α and β with the corresponding output SNRs are given in table (1).

Several simulations have been conducted and we found that reasonable values in the case of speech signals must be chosen such that $2 < \alpha \leq 5$ and $2 < \beta \leq 5$. However, if we choose $\alpha = 2.8$ and $\beta = 3$, the algorithm performs high noise reduction in the most cases. Furthermore, we found that this algorithm, using the proposed values outperforms (in term of SNR improvements) the Wiener [8] and the MMSE-LSA [3, 9] algorithms both using the decision-directed approach for estimating the *a priori* SNR.

Table (2) shows the obtained SNR improvements for four different speech signals using the proposed algorithm, the Wiener and the MMSE-LSA estimators for several input SNRs. The SNR improvement is calculated as

$$\Delta SNR = 10\log_{10}\left[\frac{\sum_{n=1}^{M}\left(x(n)-y(n)\right)^2}{\sum_{n=1}^{M}\left(x(n)-\hat{x}(n)\right)^2}\right] \quad (14)$$

Input SNR		0dB	5dB	10dB	15dB	20dB
Signal 1	α	3.09	2.54	2.25	1.68	0.56
	β	3.18	2.58	1.98	1.45	1.48
Signal 2	α	3.00	2.78	2.55	2.19	1.75
	β	3.26	2.77	2.34	2.36	2.42
Signal 3	α	3.06	2.84	2.62	2.41	2.19
	β	2.93	2.66	2.58	2.76	2.48
Signal 4	α	3.01	2.75	2.43	1.92	1.25
	β	2.77	2.46	2.11	1.97	1.87

Table 1: Optimal values of parameters α and β obtained over 100 simulations for four different speech signals

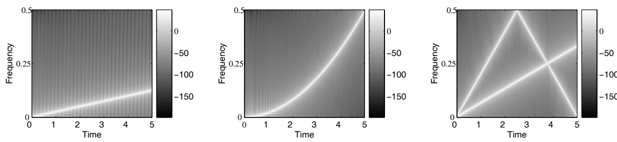
and negative values indicates that there is a signal degradation.

Input SNR		0dB	5dB	10dB	15dB	20dB
Signal 1	A	6.72	4.94	3.01	0.90	-1.87
	W	6.10	3.94	1.78	-0.42	-3.30
	M	6.23	4.24	2.26	0.15	-2.64
Signal 2	A	7.49	5.80	4.04	1.99	-0.55
	W	7.02	4.93	2.96	0.86	-1.74
	M	6.95	5.13	3.29	1.28	-1.28
Signal 3	A	7.69	6.13	4.82	3.56	2.39
	W	7.23	5.40	3.85	2.53	1.20
	M	7.19	5.59	4.16	2.90	1.66
Signal 4	A	7.35	5.53	3.61	1.56	-1.15
	W	6.89	4.80	2.55	0.26	-2.64
	M	6.88	5.03	2.96	0.84	-1.91

Table 2: SNR improvements (in dB) for the proposed algorithm (A), the Wiener (W) and the MMSE-LSA (M) filters for four speech signals and different noise levels.

As we can see from table (2), the proposed method gives the best results. Moreover, the residual noise sounds very similarly to that obtained using the MMSE-LSA algorithm which is known to provide less *musical tones* phenomenon [10].

We also tested the proposed algorithm for denoising simulated *sparse* signals, i.e., a few large coefficients dominate the signal time-frequency representation. In this case, we found it necessary to use a larger value for the parameter α while maintaining the same value for β, and the best results were obtained using the fixed value $\alpha = 5$. Table (3) gives the SNR improvements calculated using (14) for different input SNRs obtained for three test signals, namely, the linear chirp, the quadratic chirp and the two chirps signals corrupted with computer-generated white Gaussian noise. The corresponding noise-free spectrograms are represented in Figure (3). The obtained results confirm the efficiency of the proposed method in denoising such signals.



(a) Linear Chirp (b) Quadratic Chirp (c) Two Chirps

Fig. 3: Sparse test signals.

Input SNR		**0dB**	**5dB**	**10dB**	**15dB**	**20dB**
Linear Chirp	A	16.09	14.92	13.07	10.00	5.88
	W	13.64	13.00	11.82	9.30	5.60
	M	11.62	11.23	10.44	8.46	5.22
Quadratic Chirp	A	15.08	13.96	11.83	8.13	3.74
	W	13.35	12.17	10.78	7.57	3.57
	M	11.50	10.66	9.65	7.01	3.34
Two Chirps	A	12.24	11.97	11.22	9.51	6.47
	W	11.64	11.21	10.39	8.76	6.14
	M	10.30	10.00	9.38	8.08	5.75

Table 3: SNR improvements (in dB) for the proposed algorithm (A), the Wiener (W) and the MMSE-LSA (M) filters for three sparse test signals and different noise levels.

5. CONCLUSION

We proposed a simple and intuitively appealing method for non-stationary signal denoising consisting in a non-linear weighting of the STSA coefficients. We believe that the full potential of the proposed approach is not yet fully exploited, although very encouraging results were obtained. A future work will concern the development of regularization-based parameters estimation from only the noisy measure by taking into account *a priori* knowledge about the original signal. The same non-linear function can be applied to develop a shrinkage function in a wavelets denoising context [11]. In that respect, the derivation of a SureShrink-type threshold for the wavelet based denoising approach is also thinkable since the corresponding shrinking function have a bounded weak derivative [12]. These issues are now being investigated.

6. REFERENCES

[1] J. B. Allen and L. R. Rabiner, "A unified approach to short-time fourier analysis and synthesis," *Proceedings of the IEEE*, vol. 65, no. 11, pp. 1558–1564, November 1977.

[2] R. J. McAulay and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 2, pp. 137–145, April 1980.

[3] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 32, no. 6, pp. 1109–1121, December 1984.

[4] P. J. Wolfe and S. J. Godsill, "Simple alternatives to the Ephraim and Malah suppression rule for speech enhancement," *in Proc. 11th IEEE Workshop on Statistical Signal Processing*, vol. II, pp. 496–499, 2001.

[5] Douglas B. Paul, "The spectral envelope estimation vocoder," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 29, no. 4, pp. 786–794, August 1981.

[6] R. Martin, "Spectral subtraction based on minimum statistics," *in Proc. Eur. Signal Processing Conf.*, pp. 1182–1185, 1994.

[7] R. Martin, "Noise power spectral density estimation based on optimal smoothing and minimum statistics," *IEEE Transactions on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, July 2001.

[8] J. S. Lim and A. V. Oppeheim, "Enhancement and bandwidth compression of noisy speech," *Proceedings of the IEEE*, vol. 67, no. 12, pp. 1586–1604, December 1979.

[9] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 443–445, April 1985.

[10] O. Cappé, "Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor," *IEEE transactions on Speech and Audio Processing*, vol. 2, pp. 345–349, April 1994.

[11] F. Abda, "Débruitage de signaux non-stationnaires : approches temps-fréquence et temps-échelle," Tech. Rep., Master degree, Université Henri Poincaré, Nancy 1, October 2005.

[12] J. Bigot A. Antoniadis and T. Sapatinas, "Wavelets estimators in nonparametric regression : A comparative simulation study," *Journal of Statistical Software*, vol. 6, pp. 1–83, 2001.