A NEW INVERSE FILTER CRITERION FOR BLIND DECONVOLUTION OF SPIKY SIGNALS USING GAUSSIAN MIXTURES

I. Santamaría-Caballero[1], *C. Pantaleón-Prieto*[1], *F. Díaz-de-María*[2] *and A. Artés-Rodríguez*[2]

[1]Dpto. Ing. Comunicaciones, ETSII Telecom-UC, Avda. Los Castros, 39005 Santander, Spain
[2]DSSR, ETSI Telecomunicación - UPM ,Ciudad Universitaria, 28040 Madrid, Spain
Phone: 34-42-201552, Fax: 34-42-201488, e-mail: labgts@elmhp4.dicom.unican.es

ABSTRACT

This paper presents a new Bussgang-type technique for blind deconvolution of spiky signals. Based on a Gaussian mixture model for the spiky signal, the method obtains a deconvolution filter and a zero-memory nonlinearity to estimate the signal. A new updating procedure for the mixture parameters (and, therefore, for the nonlinear estimator) is included in the algorithm: it allows to apply the algorithm without any prior knowledge about the signal and noise. A simulation example illustrates the performance of the proposed method.

1. INTRODUCTION

The subject of this paper is the restoration of a spiky signal $\{x_i\}$ distorted by a linear time-invariant system $\{h_i\}$ (possibly nonminimum-phase) and corrupted by additive noise $\{n_i\}$, given only the output data sequence $z_i = x_i * h_i + n_i$, and some statistical knowledge of the input signal (i.e., blind deconvolution). This problem has a wide variety of applications in digital signal processing like geophysical exploration modeling (seismic deconvolution), ultrasonic analysis or biomedical engineering.

It is well known that the conventional linear prediction methods based on second-order statistics are incapable to solve the problem when the system is nonminimum-phase. Consequently, many techniques for blind deconvolution have been proposed in the literature.

An approach is maximum likelihood deconvolution [1]: it assumes a Bernoulli-Gaussian model for the input and, based on this model, obtains the maximum likelihood estimates of the nonzero positions and amplitudes of the sparse signal, the model parameters and the filter coefficients. However, the maximization of the global likelihood function is rather difficult and computationally expensive.

Recently, other algorithms that utilize the higher-order statistics (cumulants) of the observations have been proposed in blind deconvolution and equalization problems [2,3]. These methods guarantee global convergence to the desired solution but they are also computationally expensive. Besides, to obtain reliable estimates of the higher-order statistics, long data lengths are needed.

From a different point of view, Wiggins [4,5] introduced minimum-entropy deconvolution (MED) in seismic deconvolution, seeking an inverse filter that maximizes the kurtosis of the deconvolved data (or, equivalently, minimizing the entropy or randomness at the output of the filter). Related with the MED-type algorithms, Godfrey and Rocca proposed the zero memory non-linear deconvolution [6] (also called Bussgang method [2]). This technique makes use of rough estimates of the input data obtained from the observations by means of a matched nonlinear function (like a soft threshold), which depends on the probability distribution function (pdf) of the input data. In fact, as it is pointed out in [7], the MED-type algorithms are also Bussgang methods: in this case the deconvolution filter is the solution of a nonlinear set of equations which can be solved using an iterative procedure similar to the one proposed in [6]. The type of nonlinearity controls the final solution and the direction of convergence, thus establishing the differences among these methods. In general, these nonlinearities depend on the pdf of the input signal and the convolutional noise; consequently, the application of these techniques requires a careful selection of the estimator parameters.

In this paper, we propose a new Bussgang-type algorithm for blind deconvolution of spiky signals which includes a procedure for updating the estimator parameters; therefore it can be applied without any prior knowledge about the signal and noise.

2. A NEW INVERSE FILTER CRITERION

The proposed method assumes that the distribution of the spiky signal can be approximated with a mixture of a narrow (subscript n) and a broad (subscript b) zero-mean Gaussians. The narrow Gaussian models the smaller peaks, whereas the broad one models the true peaks (or reflectors):

$$p(x) = \frac{\pi_n}{\sqrt{2\pi}\sigma_n} e^{-\frac{x^2}{2\sigma_n^2}} + \frac{\pi_b}{\sqrt{2\pi}\sigma_b} e^{-\frac{x^2}{2\sigma_b^2}} \quad (1)$$

where π_n and π_b are the mixing proportions of the two Gaussians and are therefore constrained to sum 1.

As long as σ_n^2 is small, the Gaussian mixture approximates the distribution of a spiky signal; moreover, when σ_n^2 tends to zero, (1) becomes a Bernoulli-Gaussian distribution [1] which is a model widely used in seismic deconvolution cases.

Let $\{y_i\}$ be the output of a deconvolution filter $\{f_i\}$ of length $L+1$; considering a sequence of observations $\{z_i\}$ of length M we have

$$y_i = \sum_{j=0}^{L} f_j z_{i-j}, \qquad i = 0, \cdots, N-1 \quad (2)$$

being $N = M + L$.

As it is shown in [5], convolution always increases the Gaussian character of the pdf, therefore the deconvolution filter should remove this effect by making its output to fit (1) again. For doing so, we propose to obtain an inverse filter which maximizes at its output a measure of the relative entropy between mixture model (1) and a Gaussian distribution with the same variance, i.e.,

$$J = \sum_i \log \sum_j \pi_j p_j(y_i) - \sum_i \log f(y_i) \quad (3)$$

where $p_j(y_i)$ is the probability density of y_i under Gaussian j in the mixture, and $f(y_i)$ is a Gaussian pdf with variance $\sigma^2 = \pi_n \sigma_n^2 + \pi_b \sigma_b^2$. This objective function is used to drive the pdf of the inverse filter's output away from the initial Gaussian distribution $f(y)$ towards the mixture model $p(y)$.

The maximization of (3) with respect to the filter coefficients gives

$$\frac{\partial J}{\partial f_m} = \sum_i \left\{ \frac{y_i}{\sigma^2 \sum_j \frac{r_j(y_i)}{\sigma_j^2}} - y_i \right\} \frac{\partial y_i}{\partial f_m} = 0 \quad (4)$$

where the factors $r_j(y_i)$ are the posterior probabilities of a particular sample y_i being generated by a particular Gaussian j, and they are given by

$$r_j(y_i) = \frac{\pi_j p_j(y_i)}{\sum_k \pi_k p_k(y_i)}. \quad (5)$$

Now, taking into account that $\partial y_i / \partial f_m = z_{i-m}$ and substituting (2) for the rightmost y_i term in (4), we obtain

$$\sum_l f_l \sum_i z_{i-l} z_{i-m} = \sum_i \frac{y_i}{\sigma^2 \sum_j \frac{r_j(y_i)}{\sigma_j^2}} z_{i-m}. \quad (6)$$

The set of equations (6) can be written in matrix notation as

$$\mathbf{R}_{zz}\mathbf{f} = \mathbf{g} \quad (7)$$

where $\mathbf{R}_{zz}$ is the Toeplitz autocorrelation matrix of the observations and $\mathbf{g}$ is the crosscorrelation vector between the observations and a nonlinear estimation of the input signal, which is given by

$$\hat{x} = g(y) = \frac{y}{\sigma^2 \sum_j \frac{r_j(y)}{\sigma_j^2}} \quad (8)$$

The set of equations (6) can be solved using the iterative procedure proposed in [6]: starting from an initial inverse filter an estimate of the input signal is obtained using (8); this new estimate is croscorrelated with the observations and a new inverse filter is obtained solving (7). In each iteration the energy of the estimated signal must be normalized to a fixed value. This procedure is iterated until convergence is obtained.

3. NONLINEARITY OPTIMIZATION

A complete application of the proposed method requires a careful selection of the parameters defining the nonlinearity. Conventional Bussgang approaches use a fixed nonlinearity: for instance, from three parameters that define the nonlinear estimator in [6], two are fixed in advance (independent of iteration) and the other changes to make the algorithm data-dependent.

To avoid this *a priori* selection, in this paper we propose a method, based on model (1), that updates the mixture parameters (proportions and variances of the Gaussians), and therefore the nonlinear function $g(y)$, in each step of the deconvolution process.

The mixture parameters can be grouped in vector $\theta = (\pi_n, \sigma_n, \pi_b, \sigma_b)$. A maximum likelihood estimate of θ can be obtained applying the Expectation-Maximization (EM) algorithm [8]. To develop this idea, let us start by defining the observed incomplete data as the estimate of the input signal obtained from (8) after iteration k: $\{\hat{x}_k\}$. On the other hand, the unobserved data are given by $\chi = (\mathbf{d}_1, \mathbf{d}_2)$, where $\mathbf{d}_j$, $j = 1, 2$; is a set of Bernoulli random variables selecting the Gaussian associated to each sample, i.e.,

$$d_{i,j} = \begin{cases} 1; & \text{if} \quad x_i \in G_j \\ 0; & \text{if} \quad x_i \notin G_j \end{cases} \quad (9)$$

Using this particular choice for the complete data: $(\{\hat{x}_k\}, \mathbf{d}_1, \mathbf{d}_2)$, and denoting the current estimate of θ after k iterations of the EM algorithm as θ_k; it is easy to see that the E-step of the next iteration is given by

$$\mathrm{E}\left[d_{i,j} | \{\hat{x}_k\}, \theta_k\right] = r_j(\hat{x}_{i,k}) \quad (10)$$

where E[·] denotes expectation; then, the E-step is equivalent to recompute the posterior probabilities for the estimated signal. Once $r_j(\hat{x}_{i,k})$ is known, in the M-step we maximize J with respect to θ; taking the derivative and equating it to zero gives

$$\sigma^2_{j,k+1} = \frac{\sum_i \hat{x}^2_{i,k} r_j(\hat{x}_{i,k})}{\sum_i r_j(\hat{x}_{i,k})}, \quad (11)$$

$$\pi_{j,k+1} = \frac{1}{N}\sum_i r_j(\hat{x}_{i,k}). \quad (12)$$

It is known that the convergence rate of the EM algorithm may be slow; to avoid this problem, we propose the following modification: after each new $\{\hat{x}_k\}$ is obtained, only one iteration of the EM algorithm is carried out to obtain a new estimate of θ.

On the other hand, note that each new estimate of θ changes the cost function; therefore, to avoid stability problems, it is important to force a slow change in the parameters of the mixture. For this reason, we choose the following updating procedure:

$$\sigma^2_{j,k+1} = \gamma\sigma^2_{j,k} + (1-\gamma)\frac{\sum_i \hat{x}^2_{i,k} r_j(\hat{x}_{i,k})}{\sum_i r_j(\hat{x}_{i,k})}, \quad (13)$$

$$\pi_{j,k+1} = \gamma\pi_{j,k} + (1-\gamma)\frac{1}{N}\sum_i r_j(\hat{x}_{i,k}), \quad (14)$$

γ being a constant near to 1.

In fact, note that from four parameters in θ only two must be estimated in each iteration, since $\pi_b = 1-\pi_n$, and one of the variances is fixed by the following constraint

$$\sigma^2_{x,k} = \pi_n\sigma^2_n + \pi_b\sigma^2_b \quad (15)$$

where $\sigma^2_{x,k}$ is the variance of the estimate $\hat{x}$ at iteration k.

This updating procedure allows to start the algorithm with a very soft nonlinearity and, progressively, to increase the nonlinear character of the estimator as iteration proceeds; thus giving to the proposed algorithm a greater flexibility than other Bussgang approaches. Specifically, if we are looking for a fully sparse signal, the final nonlinearity can be used as a detector; this is important in high noise situations since in these cases the deconvolution filter is unable to remove completely the forward distorting filter and to eliminate the noise at the same time. The partial removal of the forward filter can be corrected after the nonlinearity. The price we paid is that this updating procedure makes the method slower than other Bussgang approaches which use a fixed and, generally, more aggressive nonlinearity.

4. SIMULATION RESULTS

In this example we evaluate the performance of our algorithm using synthetic signals according to the Bernoulli-Gaussian model [1], for which the signal follows a Gaussian distribution with variance σ^2_x with probability λ, and its value is zero with probability $1-\lambda$. This model can be viewed as a Gaussian mixture if we choose $\pi_n = 1-\lambda$ and $\sigma^2_n = 0$. Registers of five hundred samples were generated according to the above model (with $\lambda = 0.1$ and $\sigma^2_x = 10$), and then convolved with a third-order nonminimum-phase ARMA system taken from [9], with transfer function given by

$$H(z) = \frac{1 + 0.1z^{-1} - 3.2725z^{-2} + 1.41125z^{-3}}{1 - 1.9z^{-1} + 1.1525z^{-2} - 0.1625z^{-3}}. \quad (16)$$

Finally, a zero-mean Gaussian noise was added to the result to produce a SNR=20 dB; Fig. 1 shows the observations $\{z_i\}$.

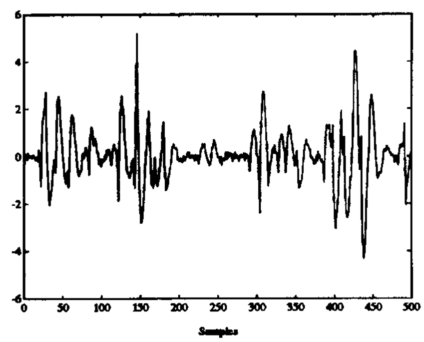

Figure 1: Observations

A 25-tap deconvolution filter was used, with its central tap initialized to $f_0 = 1$. To apply the proposed method we initialize the mixture parameters with the following values: $\pi_n = \pi_b = 0.5$, $\sigma^2_n = \sigma^2_z/2$, where σ^2_z is the variance of the observations, and σ^2_b is selected according to (15). Figs. 2(a) and 2(b) show the estimated sparse sequence obtained with the proposed method and with Godfrey's method, respectively. The proposed method obtains a very accurate estimate while Godfrey's method tends to underestimate the small reflectors. Both methods fail to resolve closely spaced peaks (sometimes two close peaks are merged). It should be noted that although the deconvolved signals (i.e., the output of the deconvolution filter) obtained by both methods can be rather similar, the estimated sparse sequences are very different since they are obtained by using two different nonlinear estimators.

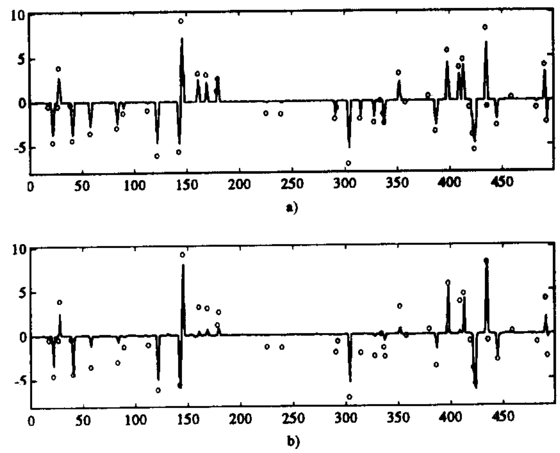

Figure 2: Sparse sequence estimated with the proposed method a) and with Godfrey's method b). Circles depict true spikes. The delay between the input signal and the estimate has been removed.

Finally, Fig. 3 shows a comparison between the nonlinear estimator used in the proposed method and the one used in Godfrey's method. Our method starts with a function almost linear; the updating procedure for the mixture parameter modifies the estimator and after convergence (20 iterations) it yields a reasonable nonlinear mapping. On the other hand, the fixed nonlinear estimator used in Godfrey's method is much more aggressive with the smaller peaks: this improves the convergence rate but leads to worse estimates.

5. CONCLUSIONS

This paper has presented a new Bussgang-type algorithm for blind deconvolution of spiky signals. An improvement and novelty in comparison with other Bussgang approaches is the use of an updating procedure for the parameters defining the nonlinear estimator (based on a Gaussian mixture model for the input signal). It has been shown that this technique achieves better estimates than Godfrey's method (mainly in high noise situations), but with a higher computational cost.

REFERENCES

[1] J. M. Mendel, *Maximum Likelihood Deconvolution*, New York: Springer-Verlag,1990.

[2] S. Haykin, *Blind Deconvolution*, Englewood Cliffs: Prentice-Hall, 1994.

[3] C. L. Nikias, A. P. Petropulu, *Higher-Order Spectra Analysis*, Englewood Cliffs: Prentice-Hall, 1993.

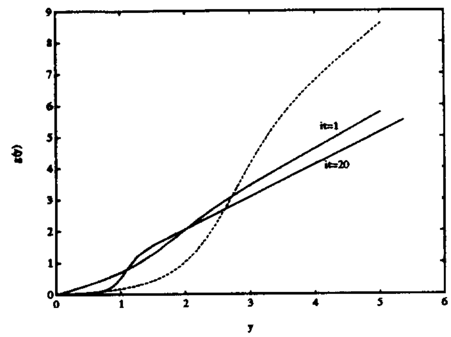

Figure 3: Zero-memory nonlinearity for the proposed method (solid line) in iterations 1 and 20 (final result), and for Godfrey's method (dashed line). Input and output are normalized by the standard deviation.

[4] R. A. Wiggins, "Minimum Entropy Deconvolution", *Geoexploration*, vol. 16, pp. 21-35, 1978.

[5] D. Donoho, "On Minimum Entropy Deconvolution", In *Applied Time Series Analysis II*, ed. D. Findley, pp. 556-608, New York: Academic Press, 1981.

[6] R. Godfrey, F. Rocca, "Zero Memory Nonlinear Deconvolution", *Geophysical Prospecting*, vol. 29, pp. 189-228, 1981.

[7] A. T. Walden, Non-Gaussian Reflectivity, Entropy and Deconvolution", *Geophysics*, vol. 50, no. 12, pp. 2862-2888, Dec. 1985.

[8] A. P. Dempster, N. M. Laird, D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm", *J. Royal Statist. Soc.*, vol. B-39, no. 1, pp. 1-37, Jan. 1977.

[9] C.-Y. Chi, M.-C. Wu, "Inverse Filter Criteria for Blind Deconvolution and equalization using two cumulants", *Signal Processing*, vol. 43, no. 1, pp. 55-63, April 1995.