WEIGHTED LEAST-SQUARES BLIND DECONVOLUTION

Simone Fiori and Francesco Piazza

Dept. Electronics and Automatics – University of Ancona
Via Brecce Bianche I, 60131 Ancona Italy
E-mail:simone@eealab.unian.it

ABSTRACT

The aim of this paper is to present a new cost function for blind deconvolution of non-minimum phase systems. The proposed criterion arises as a natural consequence of a fundamental theorem proved by Benveniste, Goursat and Ruget [3], and appears to be the weighted square of the difference among two spectra, thus its minimization leads to a weighted least-squares blind deconvolution technique. In order to assess the new theory some simulation results both on ideal (noiseless) and noisy channels are presented.

1. INTRODUCTION

In digital system deconvolution, the problem of recovering a source sequence $s(t)$ distorted by a physical system (channel), from observations of system's output $x(t)$ only, is dealt. The same problem is termed *blind* [2, 3, 4, 5] when the impulse response $\vec{h}$ of the system is unknown and the source signal is not observable. The channel's linear model writes:

$$x(t) = \vec{h}^T \vec{s}(t) + \nu(t) , \quad (1)$$

where $\vec{s}(t)$ is the input sequence, and $\nu(t)$ is an additive noise whose principal sources are additive channel noise, crosstalk and sampling errors [9]. A linear filter described by its impulse response $\vec{w}$ deconvolves $\vec{h}$ if it reverses the effects produced by $\vec{h}$ on the source signal. Denoting by $\vec{x}(t)$ the observed sequence, the output of the deconvolving system may be written as:

$$z(t) = \vec{w}^T(t)\vec{x}(t) . \quad (2)$$

A schematic of the channel/filter chain is depicted in Figure 1.

When $\vec{h}$ and $s(t)$ are unknown, the filter response $\vec{w}_\star$ such that $z(t)$ equals $s(t)$ except for a finite delay and a scale factor, has to be blindly identified [2, 3, 9]. Moreover, the deconvolution problem is particularly

This research was supported by the Italian MURST.

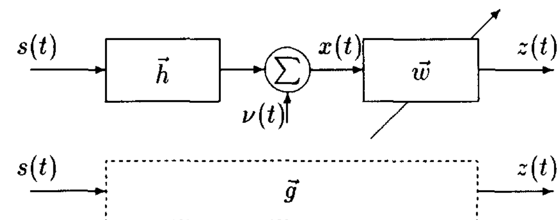

Figure 1: Channel/filter chain and global system.

difficult to solve when $\vec{h}$ represents a non-minimum phase system. In fact, denoting by $H_{\text{nmp}}(q^{-1})$ the system transfer function, by a known circuit theory result it is possible to write:

$$H_{\text{nmp}}(q^{-1}) = H_{\text{mp}}(q^{-1})H_{\text{ap}}(q^{-1}) ,$$

where $H_{\text{mp}}(q^{-1})$ represents the minimum phase part of $H_{\text{nmp}}(q^{-1})$, and $H_{\text{ap}}(q^{-1})$ denotes the remaining all-pass part (here "q^{-1}" denotes the unit-delay operator). By using classical *second-order* methods involving $|H_{\text{nmp}}(q^{-1})|^2$, it is impossible to invert the all-pass part of the system, thus its deconvolution may not be attained; higher-order methods are hence needed. Furthermore, when $\vec{h}$ is non-minimum phase, the inverse $H^{-1}_{\text{nmp}}(q^{-1})$ is unstable and real-time deconvolution is not allowed [5], therefore the inverse has to be approximated by an all-zeros function, that is an FIR deconvolving filter allowing on-line deconvolution with a negligible delay. On the other hand, every time an FIR deconvolving filter is used, an approximation error occurs [5].

Since the pioneering work of Sato (see [3] and references therein), several blind deconvolution methods have been developed and applied to solve communication [2, 3], geophysical measurements [4, 8], and blind image restoration [6] problems. Particularly, much research has been carried off to develop both suitable filtering structures [1, 4, 5, 7] and cost functions [2, 8, 9, 10]. In this paper a simple transversal filter as depicted in Figure 2 is used, while the principal aim is

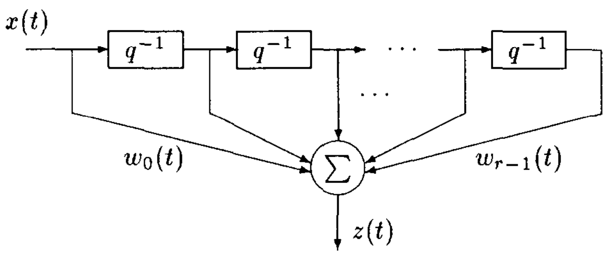

Figure 2: Schematic of a causal FIR filter.

to develop a new cost function. First we recall a fundamental theorem proved by Benveniste, Goursat and Ruget [3], then we use its consequences for defining a Weighted Least-Squares cost function whose gradient-based minimization algorithm may be used for adjusting the filter impulse response. Then we show computer sumulations for assessing our theoretical analysis.

2. WEIGHTED LEAST-SQUARES DECONVOLUTION

By using our notation, a fundamental theorem that the modern system deconvolution theory is based on, can be restated as: *Consider a source sequence s of independent, identically distributed random variables with distribution P_s, P_s being symmetric with finite variance, and a sequence given by $x = \vec{h}^T\vec{s}$. Moreover, consider a system described by its impulse response $\vec{w}$ such that the distribution of the output random variable $z = \vec{w}^T\vec{x}$ is still P_s. Denote with $\vec{g}$ the source-to-output impulse response (see Figure 1), and assume that the distribution P_s is non-Gaussian. Then $\vec{g} = \pm$* identity *except for a possible delay.* (*Benveniste, Goursat and Ruget*, [3].)

Clearly, the main idea underlying this theorem is that $\vec{w}$ deconvolves $\vec{h}$ when, and only when, the distribution P_z of $z(t)$ equals the source distribution. This shows that the following cost function may be considered:

$$C_B(\vec{w}) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \{[P_z(\zeta;\vec{w}) - P_s(\zeta)] \star k(\zeta)\}^2 \, d\zeta \ , \quad (3)$$

where "$\star$" denotes the convolutional product. It deserves to note that given a random process $x(t)$ the probability distribution of $z(t)$ depends upon the current configuration of the response $\vec{w}(t)$, thus we denote P_z as $P_z(z;\vec{w})$. C_B has been designed as a proper measure of distance between the probability density function (pdf) of the source signal s and the pdf of the output of the deconvolving filter. When such a distance vanishes the FIR filter deconvolves its input x. Note that C_B contains a *weighting kernel* $k(z)$, which has been introduced to provide a suitable error low-pass filtering, as will be clarified in the following.

2.1. Algorithm derivation

Expression (3) is not easily tractable, therefore we need to transform it in some way. By using the Parseval theorem and some Fourier transform properties, C_B can be rewritten as ("j" denotes the imaginary unit):

$$C_B(\vec{w}) = \int_{-\infty}^{+\infty} |\Psi_z(\omega;\vec{w}) - \Psi_s(\omega)|^2 |K(\omega)|^2 d\omega \quad (4)$$

$$\Psi(\omega) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} P(\xi) e^{-j\omega\xi} d\xi \ ,$$

$$K(\omega) \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} k(\xi) e^{-j\omega\xi} d\xi \ .$$

The structure of the cost function $C_B(\vec{w})$ is the weighted square of the difference among two spectra, thus we call the present method *Weighted Least-Squares Blind Deconvolution.* The kernel $K(\omega)$ makes the criterion C_B better suited to data, in that it smoothes the effects of rounding and sensor errors and avoids strong variations of $|\Psi_z - \Psi_s|$.

Here it is supposed that functions $\Psi_z(u;\vec{w})$ and $\Psi_s(u)$ admit expansions:

$$\Psi_z(u;\vec{w}) = \sum_{n=0}^{+\infty} \frac{1}{n!} \frac{d^n \Psi_z}{du^n}(0;\vec{w}) u^n \ ,$$

$$\Psi_s(u) = \sum_{n=0}^{+\infty} \frac{1}{n!} \frac{d^n \Psi_s}{du^n}(0) u^n \ .$$

From the above expressions it follows that:

$$\Psi_z(u;\vec{w}) - \Psi_s(u) = \sum_{n=0}^{+\infty} \psi_n(\vec{w}) u^n \ , \quad (5)$$

$$\psi_n(\vec{w}) = \frac{(-j)^n}{n!} \left[\int_{-\infty}^{+\infty} \zeta^n P_z(\zeta;\vec{w}) d\zeta - E[s^n] \right] \ ,$$

where $E[\cdot]$ denotes mathematical expectation. Using expression (5) in equation (4) yields:

$$C_B = \sum_{\ell=0}^{+\infty} \sum_{h=0}^{+\infty} \psi_\ell \psi_h^\star \mathcal{I}_{\ell+h} \ , \quad (6)$$

$$\mathcal{I}_p \stackrel{\text{def}}{=} \int_{-\infty}^{+\infty} \omega^p |K(\omega)|^2 d\omega \ ,$$

where "$\cdot^\star$" denotes complex conjugation. Note that as $|K(\omega)|$ is symmetric, the integral $\mathcal{I}_p$ vanishes for odd values of p.

It is important to note that formula (6) is exactly *equivalent* to the definition (3). Now a suitable approximation of expression (3) can be obtained by limiting in (6) indexes ℓ and h by positive integers L and H. Besides, in practice momenta $E[z^n]$ have to be replaced with their estimates. Formally, as a suitable approximation of $C_B(\vec{w})$ the following function may be considered:

$$\tilde{C}_B(\vec{w}) \stackrel{\text{def}}{=} \sum_{\ell=0}^{L}\sum_{h=0}^{H} \mathcal{G}_{\ell,h}[m_z^{(\ell)}(\vec{w})-m_s^{(\ell)}][m_z^{(h)}(\vec{w})-m_s^{(h)}] , \tag{7}$$

where $m_z^{(n)}(\vec{w})$ denotes an estimate of the momentum $E[z^n]$, $m_s^{(n)}$ is the known momentum of order n of the source signal $s(t)$, and:

$$\mathcal{G}_{\ell,h} \stackrel{\text{def}}{=} (-1)^{\frac{\ell-h}{2}} \frac{1}{\ell!h!}\mathcal{I}_{\ell+h} . \tag{8}$$

In order to estimate $E[z^n]$, simple recursive low-pass filters may be used. They are described by:

$$\Delta m_z^{(n)}(t) = \beta[z^n(t) - m_z^{(n)}(t-1)] \ , n \geq 1 . \tag{9}$$

The smoothing parameter β should belong to $]0,1[$. In order to minimize recursively $\tilde{C}_B$ with respect to the filter weight vector $\vec{w}$, the gradient $\partial\tilde{C}_B/\partial\vec{w}$ is needed. Straightforward calculations show:

$$\begin{aligned}\frac{\partial \tilde{C}_B}{\partial \vec{w}} &= \sum_{\ell=0}^{L}\sum_{h=0}^{H}\mathcal{G}_{\ell,h}\left\{\frac{\partial m_z^{(\ell)}}{\partial z}[m_z^{(h)} - m_s^{(h)}]+\right.\\ &\left. +[m_z^{(\ell)} - m_s^{(\ell)}]\frac{\partial m_z^{(h)}}{\partial z}\right\}\vec{x} .\end{aligned}$$

From (9) it follows that:

$$\frac{\partial m_z^{(n)}}{\partial z} = \begin{cases} \beta\cdot n\cdot z^{n-1} & \text{for} \quad n\geq 1 , \\ 0 & \text{for} \quad n = 0 , \end{cases}$$

thus the gradient steepest descent cost minimization adapting rule:

$$\Delta\vec{w} = -\gamma\frac{\partial\tilde{C}_B}{\partial\vec{w}} , \tag{10}$$

in this case writes:

$$\begin{aligned}\Delta\vec{w} &= -\gamma\beta\sum_{\ell,h=1}^{L,H}\mathcal{G}_{\ell,h}\left\{\ell z^{\ell-1}[m_z^{(h)} - m_s^{(h)}]+\right.\\ &\left. +hz^{h-1}[m_z^{(\ell)} - m_s^{(\ell)}]\right\}\vec{x} .\end{aligned} \tag{11}$$

In the above equation γ is a positive learning stepsize.

2.2. Some implementation details

In practice it is much more easy to assume $H = L$ equal to a positive integer M. From equations (8) and (9) it is easy to verify that for any $(\ell,h)\in\{1,\ldots,M\}^2$ the symmetry property $\mathcal{G}_{\ell,h} = \mathcal{G}_{h,\ell}$ holds true. Hence, defining:

$$\eta \stackrel{\text{def}}{=} 2\gamma\beta , \ m_{zs}^{(n)} \stackrel{\text{def}}{=} m_z^{(n)} - m_s^{(n)} , \ \mathcal{M}_{\ell,h} \stackrel{\text{def}}{=} \ell\cdot\mathcal{G}_{\ell,h} ,$$

equation (11) may be rewritten, equivalently:

$$\Delta\vec{w} = -\eta\vec{x}\sum_{\ell=0}^{M}\sum_{h=0}^{M}\mathcal{M}_{\ell,h}z^{\ell-1}m_{zs}^{(h)} . \tag{12}$$

It deserves to remark that the fundamental result by Benveniste, Goursat and Ruget holds under the hypothesis that the source distribution is a *symmetric* function. Here we suppose for simplicity that this symmetry property implies $p_s(\xi) = p_s(-\xi)$. Hence, from definition of momenta $m_s^{(n)}$, it follows that $m_s^{(n)} = 0$ for n odd.

Finally, consider the following choice for the weighting kernel:

$$k(x) = \exp\left(-\frac{x^2}{2}\right) \ \Rightarrow \ K(\omega) = \exp(-\omega^2) . \tag{13}$$

This kernel gives $\mathcal{I}_{2m} = \sqrt{\pi}\frac{(2m)!}{4^m m!}$.

3. COMPUTER SIMULATIONS

In support of the proposed approach, simulation results obtained with the following data are presented:

- as vector $\vec{h}$ the sampled impulsive response of a typical non-minimum phase telephonic channel [3]: $H_{\text{chan}}(q) = -0.0174q^{-1} + 0.0522q^{-3} + 0.0174q^{-4} - 0.5826q^{-5} - 0.0522q^{-6} + 1.0043q^{-7} + 0.0130q^{-8} + 0.3522q^{-9} + 0.0174q^{-10} + 0.0957q^{-11} + 0.0130q^{-12} + 0.0217q^{-13} + 0.0087q^{-14}$, and as deconvolving structure a transversal filter with 21 weights (as in [3]);
- as learning parameters: $L = H = M = 4$, $\beta = 0.001$ and $\gamma = 4$;
- as source sequence a random signal uniformly distributed within $[-\sqrt{3},\sqrt{3}]$, and for $\vec{w}(0)$ a null vector except for the 10^{th} entry equal to 1.

True momenta $m_s^{(n)}$ were estimated by using formula $m_s^{(n)} = (\sum_{i=1}^{10000} s_i^n)/10000$. As performance index we used the Inter-Symbol Interference (ISI), defined as in [9]:

$$ISI = \frac{\sum_i g_i^2 - g_{\max}^2}{g_{\max}^2} ,$$

where $\vec{g}$ denotes again the convolution between the impulse response of the channel and the impulse response of the equalizing filter, and $g_{\max}$ denotes the entry of $\vec{g}$ having the maximal value.

Figure 3 shows the ISI computed at any epoch (1 epoch = 200 samples) and the convolution product between the channel impulse response and the filter impulse response after 100 epochs, both averaged over 10 realizations of the source sequence for an ideal (noiseless) channel (i.e. $\nu(t) = 0$). Perfect equalization should imply a unique central bar on the convolution graph, but in a real-world context some interference residuals of course must be taken into account. Figure 4 shows instead the averaged ISI and convolution product for a noisy channel, when $\nu(t)$ is a zero-mean Gaussian noise of variance 0.01 that corresponds to a signal-to-noise ratio SNR=20dB.

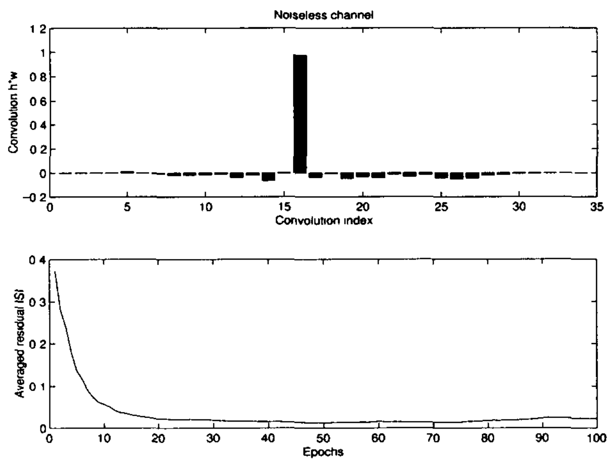

Figure 3: Simulation with an ideal (noiseless) channel.

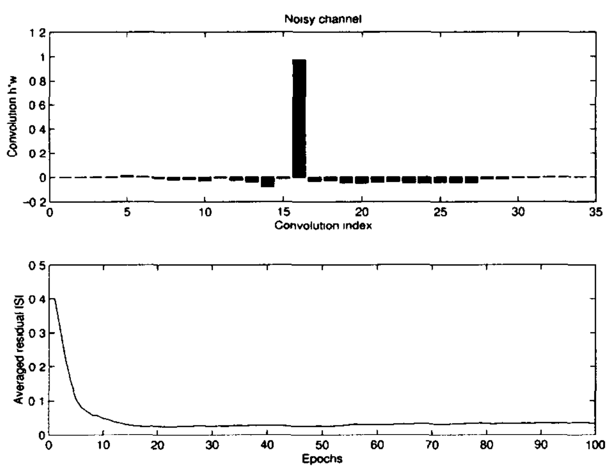

Figure 4: Simulation with a noisy channel.

4. CONCLUSIONS

In this paper a new cost function for blind equalization of non-minimum phase systems has been presented. The new cost function is equal to the weighted square difference among characteristic functions of the source signal and of the equalizing filter output, thus its minimization leads to a weighted least-squares method. Simulation results performed both on an ideal and a noisy channel show the effectiveness of the proposed blind equalization technique and its good robustness against noise.

5. REFERENCES

[1] A. ALKULAIBI AND J.J. SORAGHAN, *Hybrid Higher-Order Cepstrum and Functional Link Network-Based Blind Equaliser*, Signal Processing, Vol. 62, pp. 101 - 109, 1997

[2] S. BELLINI, *Blind Equalization*, Alta Frequenza, Vol. 57, pp. 445 - 450, 1988

[3] A. BENVENISTE, M. GOURSAT AND G. RUGET, *Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer in Data Communication*, IEEE Trans. on Automatic Control, Vol. AC-25, NO. 3, pp. 385 - 399, June 1980

[4] R. GODFREY AND F. ROCCA, *Zero Memory Nonlinear Deconvolution*, Geophisical Prospecting, Vol. 29, pp. 189 - 228, Apr. 1981

[5] S. HAYKIN, (Ed.) *Blind Deconvolution*, Prentice-Hall, 1994

[6] D. KUNDUR AND D. HATZINAKOS, *Blind Image Deconvolution*, IEEE Signal Processing Magazine, Vol. 13 , No. 3, pp. 43 - 64, May 1996

[7] R. PARISI, E.D. DI CLAUDIO, G. ORLANDI AND B.D. RAO, *Fast Adaptive Digital Equalization by Recurrent Neural Networks*, IEEE Trans. on Signal Processing, Vol. 45, No. 11, pp. 2731 - 2739, Nov. 1997

[8] E.H. SATORIUS AND J.J. MULLIGAN, *Minimum Entropy Deconvolution and Blind Equalisation*, Electronics Letters, Vol. 28, No. 16, pp. 1534 - 1535, July 1992

[9] O. SHALVI AND E. WEINSTEIN, *New Criteria for Blind Deconvolution of Nonmimimum Phase Systems (Channels)*, IEEE Trans. on Information Theory, Vol. IT-36, No. 2, March 1990

[10] S. VEMBU, S. VERDÙ, R.A. KENNEDY AND W. SETHARES, *Convex Cost Functions in Blind Equalization*, IEEE Trans. on Signal Processing, Vol. 42, No. 8, pp. 1952 - 1959, Aug. 1994