BLIND DECONVOLUTION BY MODIFIED BUSSGANG ALGORITHM

Simone Fiori, Aurelio Uncini, and Francesco Piazza

Dept. Electronics and Automatics – University of Ancona (Italy)
E-mail:simone@eealab.unian.it

ABSTRACT

The 'Bussgang' is one of the most known blind deconvolution algorithms. It requires the prior knowledge of the source statistics as well as the deconvolution noise characteristics. In this paper we present a first attempt for making the algorithm 'more blind' by replacing the original Bayesian estimator with a flexible parametric function whose parameters adapt through time. To assess the effectiveness of the proposed method, computer simulations are also presented and discussed.

1. INTRODUCTION

Blind deconvolution [1, 2, 3, 10, 13, 16] concerns the problem of recovering a source signal $s(t)$ distorted by a linear channel with impulse response $\vec{h}$, from observations of the channel output $x(t)$, without knowledge about $\vec{h}$ nor the statistics and temporal characteristics of the source. In the vector notation [2, 9] the linear model writes:

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

where $\vec{s}(t)$ is a vector containing the input samples:

$$s(t),\ s(t-1),\ s(t-2),\ \ldots,\ s(t-\ell+1) ,$$

with ℓ being the number of entries in $\vec{h}$, and $N(t)$ being an additive noise that originates by many simultaneous effects [15].

A transversal filter described by its impulse response $\vec{w}$ is a channel equalizer if $\vec{w}$ cancels the effects of $\vec{h}$ on the source signal. Denoting by $\vec{x}(t)$ the vector containing samples:

$$x(t),\ x(t-1),\ x(t-2),\ \ldots,\ x(t-m+1) ,$$

where m is the number of tap-weights in $\vec{w}$, the output of the filter writes [9]:

$$z(t) = \vec{w}^T(t)\vec{x}(t) , \qquad (2)$$

Since $\vec{h}$ and $s(t)$ are unknown, the equalizer $\vec{w}_\star$ such that $z(t) \sim s(t)$ has to be *blindly* found usually by means of

---

an iterative algorithm [1, 12]. When $\vec{h}$ represents a non-minimum phase system, its inversion cannot be performed by means of an FIR filter, therefore every time an FIR equalizer is used an approximation error occurs [2, 12]. In formulas we get:

$$z(t) = cs(t-\delta) + n(t) , \qquad (3)$$

where $n(t)$ is the so-called *deconvolution noise*, c is an amplitude factor and δ is a finite delay. A suitable representation of $n(t)$ is a Gaussian random process [12] with variance denoted here with σ^2 (called 'deconvolution noise power'). It is worth noticing that the same model for $z(t)$ takes into account the error due to the fact that during the whole adaptation phase $\vec{w} \neq \vec{w}_\star$.

One of the most known blind equalization algorithm is the 'Bussgang' by Prof. Bellini [1, 12], based on a memoryless Bayesian estimation $\hat{s} = g(z)$ of s by the knowledge of z and a LMS-style adjustment of $\vec{w}$ with the quantity $\hat{s} - z$ as error, under the hypothesis of IID (Independent Identically Distributed) source sequences. In the original Bellini's theory, optimal $g(\cdot)$ depends on the statistics of the deconvolution noise and of the source sequence, thus models of them are required [1]. He considered different cases [13], and recently Destro-Filho *et al.* developed a special algorithm suited for binary sources [6]. Under the hypothesis that $n(t)$ is Gaussian and the source sequence has a uniform distribution [1, 13], $g(z)$ depends on σ^2, thus a problem of that algorithm is to estimate the deconvolution noise power in the best way.

We propose a self-tuning procedure that allows to automatically determine optimal parameters of a flexible approximated estimator $g(z)$, in connection with a different error-minimization algorithm based on the Gradient Steepest Descent technique. Such a self-tuning behavior allows to overcome the problem of finding a suitable value of σ^2.

2. MODIFIED BUSSGANG ALGORITHM

In [1] an error criterion like:

$$U(\vec{w}) \stackrel{\text{def}}{=} \frac{1}{2} E_{\vec{x}}[(g(z) - z)^2 | \vec{w}] \qquad (4)$$

is proposed, where $E_{\vec{x}}[f|\vec{w}]$ denotes mathematical expectation with respect to $\vec{x}$ under the hypothesis $\vec{w}$, hereafter simply denoted by $E[f]$. The function $g(z)$ provides an estimate of the source signal based on the Bayesian technique, so the deconvolving filter $\vec{w}_\star$ minimizes $U(\vec{w})$. As a method to find iteratively the optimal filter given observations of the channel output $x(t)$, an LMS-style algorithm was used [12]. By interpreting the difference $g(z) - z$ as an 'error', the structure of the algorithm proposed by Bellini is:

$$\Delta \vec{w} = \mu [g(z) - z]\vec{x} , \qquad (5)$$

where μ is a positive learning stepsize.

It deserves to note that there seem to be no theoretical reasons to use the LMS-style procedure rather than others. So, the minimization of the same cost function U can be performed also by means of a stochastic Gradient Steepest Descent (GSD) algorithm described by $\Delta \vec{w} \propto -\frac{\partial U}{\partial \vec{w}}$. In the present context this rule assumes the following expression:

$$\Delta \vec{w} = -\eta E[(g'(z) - 1)(g(z) - z)\vec{x}] , \qquad (6)$$

where η is a positive stepsize and $g'(z)$ denotes the derivative of the function $g(z)$ with respect to z. By comparing equations (5) and (6) one gathers they coincide if in (5) the variable stepsize $\mu(z) \overset{\text{def}}{=} -\eta[g'(z) - 1]$ is used, and the operator $E[\cdot]$ is dropped down.

The Bellini's expression for $g(z)$ is dependent upon the deconvolution noise power σ^2 [1, 12]. The choice of a suitable estimate for this parameter is quite difficult; moreover, an optimal value for σ^2 probably does not exist since it should be changed through time accordingly with the adaptation progress, as already outlined in [1, 12, 13]. Despite this, for a wide range of the noise power a suitable approximation of $g(z)$ seems to be [12] the bilateral sigmoid:

$$g(z) \overset{\text{def}}{=} a \tanh(bz) , \qquad (7)$$

with a and b being properly chosen parameters.

In [12] a pair of values for a and b is obtained by fitting the expression (7) with the actual Bellini's function. Anyway, it is clear that as an optimal constant value for σ^2 cannot be found, a suitable pair of *constant* parameters a and b cannot be fixed, too. In order to get rid of this drawback, we propose to adapt through time their values by means of a GSD algorithm applied to U (thought of as a function of a, b and z). In formulas we get:

$$\Delta a = -\alpha \frac{\partial U}{\partial a} = -\alpha E\left[(g - z)\frac{g}{a}\right] , \qquad (8)$$

$$\Delta b = -\beta \frac{\partial U}{\partial b} = -\beta E\left[(g - z)(a^2 - g^2)\frac{z}{a}\right] , \qquad (9)$$

where α and β are constant positive learning stepsizes.

It deserves to note that owing to the structure of U due to the expression (7), the problem of minimizing U is now

ill-posed, because a simple way to minimize U is to vanish $||\vec{w}||$. To prevent such a behavior, it is possible to embed a simple constraint on the norm of $\vec{w}$, that is $\vec{w}^T \vec{w} - \kappa^2 = 0$, where κ^2 is an arbitrarily chosen non-null constant that provides an amplification of the filter output with a factor $|\kappa|$. This condition can be taken into account by defining a new criterion J as:

$$J \overset{\text{def}}{=} U + \lambda(\vec{w}^T \vec{w} - \kappa^2) , \qquad (10)$$

where λ is a *Lagrange multiplier*. Using again the GSD algorithm $\Delta \vec{w} = -\eta \frac{\partial J}{\partial \vec{w}}$ to search for the minimum of J, we can replace the unconstrained rule (6) with:

$$\Delta \vec{w} = -\eta E\left[\left(ab - \frac{b}{a}g^2 - 1\right)(g - z)\left(\vec{x} - \frac{z}{\kappa^2}\vec{w}\right)\right] . \qquad (11)$$

Equations (11), (8) and (9) give a new gradient-based blind equalization method with the flexible estimator (7).

3. THEORETICAL ANALYSIS: HOW TO DETERMINE OPTIMAL PARAMETERS WHEN SOURCE STATISTICS ARE KNOWN

Suppose that the source statistics are known. Let it be:

$$\gamma(z) \overset{\text{def}}{=} [1 - g'(z)][g(z) - z] , \qquad (12)$$

the adapting equation (6) recasts into:

$$\Delta \vec{w} = +\eta E[\gamma(z)\vec{x}] , \qquad (13)$$

thus stationarity occurs when the relationship:

$$E[\gamma(z)\vec{x}] = \vec{0} \qquad (14)$$

holds true. If $g(\cdot)$ is an odd function (like the $\tanh$ one), then $\gamma(z)$ is odd too, and it can be expanded as:

$$\gamma(z) = \sum_{\ell=0}^{+\infty} \gamma_\ell z^{2\ell+1} .$$

Recalling that (ideally) at convergence $z(t) = cs(t)$, with c being an arbitrary scaling parameter, and that:

$$x(t - i) = \sum_{k=-\infty}^{\infty} h_k s(t - k - i) ,$$

supposing that filters $\vec{h}$ and $\vec{w}$ have infinitely many elements [2], the convergence condition writes:

$$\forall i : E\left[\sum_{k=-\infty}^{+\infty} h_k s(t - k - i) \sum_{\ell=0}^{+\infty} \gamma_\ell (cs)^{2\ell+1}\right] = 0 , \Rightarrow$$

$$\forall i : \sum_{k=-\infty}^{+\infty} h_k \sum_{\ell=0}^{+\infty} \gamma_\ell c^{2\ell+1} E[s(t - k - i)s^{2\ell+1}(t)] = 0 .$$

III-2

Since by hypothesis $s(t)$ is an IID sequence, the property $E\left[s(t-k-i)s^{2\ell+1}(t)\right]=0$ for $k+i\neq 0$ holds true, then we remain with:

$$\forall\, i:\ h_{-i}\sum_{\ell=0}^{+\infty}\gamma_\ell E\left[s^{2\ell+2}\right]=h_{-i}E\left[s\gamma(s)\right]=0\ .$$

More explicitly, assuming for simplicity $E[s^2]=1$, the convergence condition reads:

$$E\left[sg'(s)g(s)\right]-E\left[sg(s)\right]-E\left[sg'(s)\right]=1\ .\qquad(15)$$

This condition establishes a relationship between a and b. Furthermore, by using the fundamental stability results proven by Benveniste, Goursat and Ruget in [2], we found that: *For sub-Gaussian sources and for channels with impulse responses and inverse impulse responses endowed with finite energy, the algorithm (6) finds the appropriate equalizer if $g(\cdot)$ satisfy* [7]:

$$3g''(x)\left[1-g'(x)\right]+g'''(x)\left[1-g(x)\right]<0\text{ for }x>0\ .\ (16)$$

Then, conditions (15) and (16) allow finding a and b.

4. COMPUTER SIMULATION RESULTS

In support of the new deconvolution theory, as an experimental case we present simulations performed as follows:

- as vector $\vec{h}$ we take the sampled impulse response of a typical non-minimum phase telephonic channel with $\ell=14$ used in [2]; the bar-graph of $\vec{h}$ is shown in Figure 1;

- as source signal a sub-Gaussian random process uniformly distributed within $[-\sqrt{3},\sqrt{3}]$ has been taken, like that described in [1] to develop Bellini's theory;

- as deconvolving structure a transversal filter with $m=21$ taps as in [2] is used;

- the algorithm starts with null weights except for the 10th one equal to 1, $a(0)=b(0)=1$, and runs with $\eta=0.08$, $\alpha=0.08$ and $\beta=0.08$; as filter output amplitude gain we chose $\kappa=2$.

The actual deconvolution accuracy degree is measured by means of the residual ISI defined as in [15]:

$$ISI=\frac{\|\vec{\nu}\|^2-\nu_{\max}^2}{\nu_{\max}^2}\ ,\qquad(17)$$

where $\vec{\nu}$ denotes the convolution between $\vec{w}$ and $\vec{h}$, and $\nu_{\max}$ is the component of $\vec{\nu}$ having the maximal absolute value.

The Figure 2 shows the convolution $\vec{\nu}$ between the channel impulse response and the filter impulse response learnt
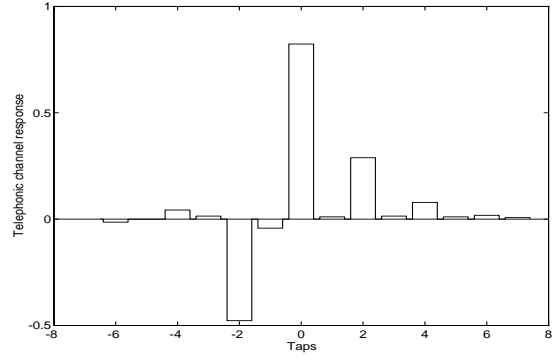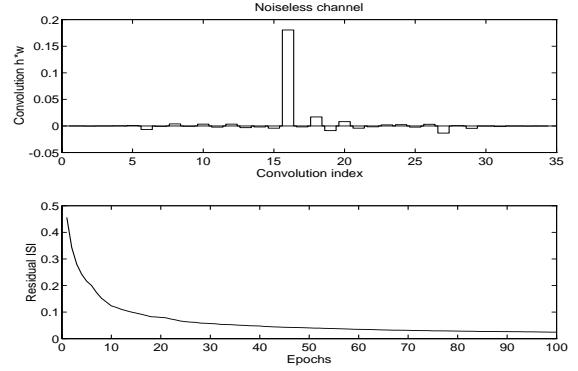

Figure 1: Sampled telephonic channel response $\vec{h}$.

Figure 2: Noiseless case. Convolution between $\vec{h}$ and the learnt $\vec{w}$. Averaged interference residual after 100 epochs.

after 100 epochs (1 epoch corresponds to 200 input samples); it also shows the ISI computed at the end of any epochs, both averaged over 20 realizations of the source sequence in the noiseless case (i.e. $N(t)=0$).

Figure 3 refers instead to a noisy channel where $N(t)$ is a zero-mean AWGN of variance 0.01, that means having a signal-to-noise ration equal to 20dB. In both cases the algorithm seems to perform well, and the second simulation shows it is rather insensitive with respect to Gaussian additive noise.

5. CONCLUSION

The aim of this paper was to propose a first report on the improvement to the Bellini's 'Bussgang' algorithm that relies on using a flexible approximation of the Bayesian estimator required in the original theory. Computer simulation results show the effectiveness of the proposed approach both with noiseless and noisy channels.

We believe the approximation capabilities of the nonlinear function $g(\cdot)$ as well as the performances of the algo-
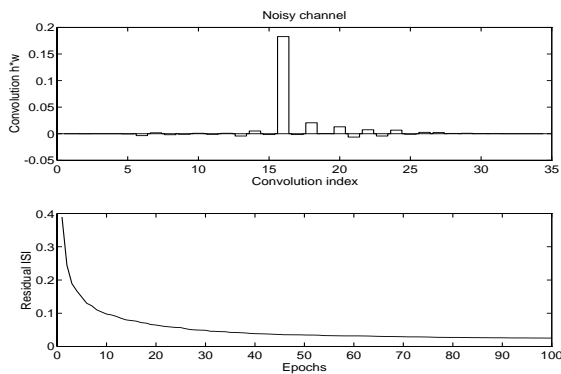
Figure 3: Noisy case. Convolution between $\vec{h}$ and the learnt $\vec{w}$. Averaged interference residual after 100 epochs.

rithm could be enhanced by the use of more flexible functions like:

$$g(x) = a \tanh\left[\sum_{i=0}^{n} c_i x^i\right] \ ,$$

or by the use of functional-link neural units endowed with adjustable activation functions, already proven to be effective in Blind Source Separation by Fiori *et al.* in [8]. Moreover some efforts should lead to a deeper comprehension of the important theoretical aspects of the method, like convergence and stability properties. Furthermore, extension to blind deconvolution of linear complex channels and of non-linear channels [4, 5, 11, 14] are currently under investigation.

6. REFERENCES

[1] S. BELLINI, *Bussgang Techniques for Blind Equalization*, in IEEE Global Telecommunication Conf. Rec., pp. 1634 – 1640, Dec. 1986

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

[3] J.A. CADZOW, *Blind Deconvolution via Cumulant Extrema*, IEEE Signal Processing Magazine, pp. 24 – 42, May 1996

[4] S. CHEN, S. MCLAUGHLIN, AND B. MULGREW, *Complex-Valued Radial Basis Function Network, Part. I: Network Architectures and Learning Algorithms*, Signal processing, Vol. 35, pp. 19 – 31, 1994

[5] S. CHEN, S. MCLAUGHLIN, AND B. MULGREW, *Complex-Valued Radial Basis Function Network, Part. II: Applications to Digital Communications Channel Equalisation*, Signal processing, Vol. 36, pp. 175 – 188, 1994

[6] J.B. DESTRO FILHO, G. FAVIER, AND J.M. TRAVASSON ROMANO, *New Bussgang Methods for Blind Equalization*, Proc. of International Conference on Acoustics, Speech and Signal Processing (IEEE-ICASSP), pp. 2269 – 2272, 1997

[7] S. FIORI, *Neural Artificial Unsupervised Learning Models for Blind Signal Processing*, M.Sc. Dissertation, Dept. of Electronics and Automatics, University of Ancona, Ancona (Italy), July 1996

[8] S. FIORI, P. BALDASSARRI, AND F. PIAZZA, *An Efficient Architecture for Independent Component Analysis*, These proceedings

[9] S. FIORI AND F. PIAZZA, *BLADE: A New On-Line Blind Equalization Method Based on the Burelian Distortion Measure*, Proc. of International Symposium on Circuits and Systems (IEEE-ISCAS), Vol. IV, pp. 441 – 444, 1998

[10] D.N. GODARD, *Self Recovering Data Communication System*, IEEE Trans. on Communications, Vol. COM-28, No. 11, Nov. 1980

[11] S. HE AND Z. HE, *Blind Equalization of Nonlinear Communication Channels Using Recurrent Wavelet Neural Networks*, Proc. of International Conference on Acoustics, Speech and Signal Processing (IEEE-ICASSP), Vol. 4, pp. 3305 – 3309, 1997

[12] S. HAYKIN, *Adaptive Filtering*, Ed. Prentice-Hall, 1991 (Chap. 20: Blind Deconvolution)

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

[14] J.P. SAINI AND M.C. SRIVASTAVA, *A New Polynomial Structure for Channel Equalization and ACI Suppressioni in 64-QAM Reception*, Proc. of International Symposium on Circuits and Systems (IEEE-ISCAS), Vol. IV, pp. 457 – 460, 1998

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

[16] O. TANRIKULU, B. BUYURMAN, A.G. CONSTANTINIDES, AND J.A. CHAMBERS, *Soft Constrained Satisfaction Blind Channel Equalization Algorithms*, Int. Journal of Adaptive Control and Signal Processing, Vol. 12, pp. 117 – 134, 1998