AUDIOWATERMARK DETECTION FOR ALL-PASS PIRAT ATTACK: HYBRID BLIND EQUALIZATION/WIENER DECONVOLUTION APPROACH

S. Larbi(1), *M. Jaidane*(1), *N. Moreau*(2)

(1) Laboratoire des Systèmes de Communications
Ecole Nationale d'Ingénieurs de Tunis – ENIT BP 37 – 1002 Tunis – Tunisia
(2) Département Traitement du Signal et des Images
Ecole Nationale Supérieure des Télécommunications – 46 Rue Barrault– 75013 Paris – France
Sonia.Larbi@enit.rnu.tn – Meriem.Jaidane@enit.rnu.tn – moreau@tsi.enst.fr

ABSTRACT

Digital audiowatermarking can be viewed as a communication system (cf. figure 1) where the information v_n is embedded imperceptibly into the digital audio signal x_n through appropriate spectral shaping filter H_j. The resulting watermark t_n should be robust to standard signal manipulations (i.e. compression, A/D-D/A conversion, resampling,...) on one side, and to intentional pirat attacks on the other side. In particular, all-pass filter attacks introduce a nonminimum phase problem in the watermarking detection scheme of figure 1, which strongly degrades the performances of detection techniques based on mean square criteria.
We propose in this paper a solution to the above mentioned problem, combining Wiener deconvolution and blind equalization based on a non quadratic criterion.

1. INTRODUCTION

Considering the watermarking scheme of figure 1 as a communication system, some robustness studies have been achieved concerning MPEG compression and desynchronization [1]. In this context, the all-pass filter attack can be considered as the well known nonminimum phase channel problem in communication systems, since all-pass filtering of the watermarked signal y_n does not affect its audio quality but strongly affects the detection of the embedded watermark v_n. This paper proposes a detection scheme that is robust to all-pass filtering. It is structered as follows : in the second section we focus on the insufficiency of the Wiener deconvolution detection scheme based on mean square error minimization in case of nonminimum phase channel. Since blind linear equalization techniques based on non quadratic criteria [2, 3, 4] are efficient to reduce the effects of nonminimum phase channels in communication systems, we examine in the third section some possible detection schemes based on this approach and we present the performances of the proposed hybrid detection. It is a non causal blind equalization combined with a Wiener deconvolution. The blind equalization is based on a non quadratic criterion. The simulation results of section 4. show that the hybrid scheme is very efficient against all-pass attacks.

2. INSUFFICIENCY OF A WIENER DECONVOLUTION DETECTION SCHEME

2.1. The generic detection scheme

We consider here public key watermarking (i.e. with blind watermark detection) : the unwatermarked original signal x_n is not available at detection. Briefly, the embedded watermark signal t_n is obtained by spectral shaping of the sequence v_n through the filter H_j, which frequency response module matches with the masking threshold over the bloc j of N x_n-samples. The masking threshold is computed by an auditory model and is updated each N samples. For more details about the construction of v_n, the reader may refer to [1]. The detection process of the considered watermarking scheme is achieved by a non causal FIR deconvolution filter described by :

$$\hat{v}_n = \sum_{i=-Q}^{P} g_i \hat{y}_{n-i} = G^T \hat{Y}_n \tag{1}$$

where $G = (g_{-Q}, ..., g_0, ..., g_P)^T$ and $\hat{Y}_n = (\hat{y}_{n+Q}, ..., \hat{y}_n, ..., \hat{y}_{n-P})^T$. The parameter G is optimized in order to minimize $E\left[(v_n - \hat{v}_n)^2\right]$. Since :

$$y_n = h_n * v_n + x_n, \tag{2}$$

(where h_n is the impulse response of the spectral shaping filter H_j), the detection of v_n can be viewed as a Wiener deconvolution filter that conjointly equalizes H_j and reduces the "noise" x_n.

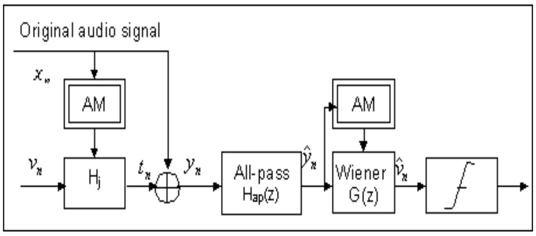

Figure 1: The simplified generic [1] watermarking scheme.

We consider first the case without all-pass attack, i.e. $\hat{y}_n = y_n$. We prove easily (see for example [6]) that the frequency response of the optimal deconvolution filter is

$$G_{opt}(e^{j\omega}) = \frac{\sigma_v^2 A_j(e^{j\omega})}{\sigma_v^2 + |A_j(e^{j\omega})|^2\, \gamma_x(e^{j\omega})} \tag{3}$$

We assume in (3) that v_n is a white noise with power spectral density σ_v^2 and we denote the power spectral density of x_n by $\gamma_x(e^{j\omega})$. The spectral shaping filter H_j is a causal recursive filter with transfer function

$$H_j(z) = \frac{1}{1 - \sum_{i=1}^{M} a_i^{(j)} z^{-i}} = \frac{1}{A_j(z)}. \tag{4}$$

When $\gamma_x(e^{j\omega}) \rightarrow 0$ the optimal filter is the inverse $A_j(z)$ of the spectral shaping filter $H_j(z)$. It is interesting to note that [6] the Wiener filter in (3) can be viewed as a cascade of the inverse filter $A_j(z)$ followed by a non causal Wiener smoothing filter $F(z)$ reducing the filtered noise $X(z)A_j(z)$. Indeed, we have

$$\begin{aligned} G_{opt}(e^{j\omega}) &= A_j(e^{j\omega}) \left[\frac{\sigma_v^2}{\sigma_v^2 + |A_j(e^{j\omega})|^2 \gamma_x(e^{j\omega})}\right] \\ &= A_j(e^{j\omega}) F(e^{j\omega}) \end{aligned} \tag{5}$$

Note that $H_j(z)$ and $A_j(z)$ can be deduced from x_n as well as from y_n or $\hat{y}_n$. Thus, without all-pass filter the optimal Wiener filter can be computed by solving the Wiener-Hopf equations [6] at the detection process.

2.2. The all-pass attack : a nonminimum phase channel estimation problem

We consider now an all-pass attack described by the transfer function $H_{ap}(z)$ (cf. figure 1). Since $|H_{ap}(z)|^2 = 1$, the sound quality of the all-pass filtered signal $\hat{y}_n$ is unaffected. In this case the detection scheme described in the previous section fails completely. For example, with a first order all-pass filter $H_{ap}(z) = \frac{z^{-1}-a}{1-az^{-1}}$ with $a = 0.5$, the BER[1] of the Wiener deconvolution detection scheme grows up to 0.995. We consider in the following the general form of an all-pass filter with real valued impulse response

$$H_{ap}(z) = \prod_{k=1}^{M} \frac{z^{-1} - a_k}{1 - a_k z^{-1}} \prod_{k=1}^{N} \frac{(z^{-1} - b_k^*)(z^{-1} - b_k)}{(1 - b_k z^{-1})(1 - b_k^* z^{-1})} \tag{6}$$

where the a_k's are the real and the b_k's are the complex valued poles of H_{ap}. Then, a causal and stable all-pass filter ($|a_k| < 1$ and $|b_k| < 1$ for all k) is always a nonminimum phase system. Therefore, the detection of the all-pass filtered watermark can be viewed as a blind equalization of a nonminimum phase channel $H_{np}(z)$ (see figure 2).

3. A BLIND EQUALIZATION APPROACH

A well known class of blind equalization algorithms requires the use of non convex cost functions, different from the quadratic ones usually employed in Wiener filtering for example. The pioneering work of Sato [2] was developped by among others Godard [3] and Shalvi and Weinstein [4]. These approaches are based on similar principles and require the computation of suitable defined higher order statistics of the observed signal.

3.1. Reflexion on possible blind detection schemes

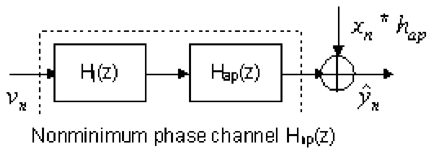

Figure 2: Equivalent nonminimum phase channel H_{np}.

As mentioned in 2.2, the watermark detection becomes a general blind equalization, where the system to be equalized is the nonminimum phase linear time variant channel (figure 2)

$$H_{np}(z) = H_j(z) H_{ap}(z) \tag{7}$$

The system non stationarity is introduced by the updating of the spectral shaping filter H_j each bloc j of N samples, due to the auditory model[2]. The nonminimum phase character of the system (7) is introduced by the all-pass attack. A global blind equalization scheme is possible but it does'nt make use of the a priori knowledge of H_j, which can be extracted from the observed

[1] BER : Bit Error Rate

[2] The used auditory model is the MPEG no. 1 model, which is updated every 512 samples of the analyzed audio signal x_n.

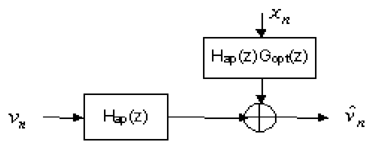

Figure 3: Equivalent, to be equalized channel.

signal $\hat{y}_n$. Indeed, taking advantage from this knowledge at the detection, we rewrite (3) for the global system (7)

$$
\begin{aligned}
G_{opt}^{ap}(e^{j\omega}) &= \frac{\sigma_v^2 H_j^*(e^{j\omega}) H_{ap}^*(e^{j\omega})}{\sigma_v^2 |H_j(e^{j\omega}) H_{ap}(e^{j\omega})|^2 + \gamma_x(e^{j\omega})} \\
&= \frac{\sigma_v^2 H_j^*(e^{j\omega}) H_{ap}^*(e^{j\omega})}{\sigma_v^2 \left|H_j(e^{j\omega})\right|^2 + \gamma_x(e^{j\omega})}
\end{aligned} \tag{8}
$$

due to the fact that $\left|H_{ap}(e^{j\omega})\right|^2 = 1$. It follows that

$$G_{opt}^{ap}(e^{j\omega}) = G_{opt}(e^{j\omega}) H_{ap}^*(e^{j\omega}) \tag{9}$$

Since $H_{ap}^*(e^{j\omega})$ is unknown, the Wiener deconvolution filter in (9) realizes only $G_{opt}(e^{j\omega})$: it is unable to equalize $H_{ap}^*(e^{j\omega})$. So we may see the Wiener filter output signal $\hat{v}_n$ as the noisy output of the all-pass filter excited by v_n (figure 3). Thus, the proposed detection scheme in figure 4 is hybrid : it combines Wiener deconvolution and blind equalization, where the equalizer output $\tilde{v}_n$ is the input of the decision system.

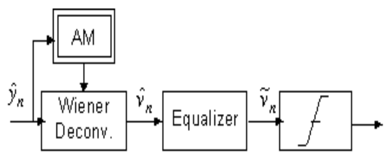

Figure 4: The proposed hybrid detection scheme.

3.2. The Godard blind equalizer [3, 4]

The non quadratic optimization criterion used for the blind equalizer can be viewed as a contrast function where equalizing means maximizing of some distance to gaussianity [5]. In the proposed hybrid detection scheme we maximize the criterion function Φ

$$\Phi = \frac{1}{K(v)} E\left[\left(|\tilde{v}_n|^2 - \alpha\right)^2\right] \tag{10}$$

where $\tilde{v}_n = \sum_{k=-P_{nc}}^{P_c} I_k \hat{v}_{n-k}$, with $I = (I_{-P_{nc}}, \ldots I_{P_c})^T$ the non causal equalizer impulse response and $K(v)$ the kurtosis of v_n and

$$\alpha = \frac{E\left[|v_n|^4\right]}{E\left[|v_n|^2\right]} \tag{11}$$

Assuming real valued signals and $K(v) < 0$, the gradient based algorithm updating the parameters of I is given by [4]

$$I_k(n+1) = I_k(n) - \mu\left\{\left(\tilde{v}_n^2 - \alpha\right) - E\left[\tilde{v}_n^2\right]\right\} \tilde{v}_n \hat{v}_{k-n} \tag{12}$$

where $I_k(n)$ and μ denote respectively the k-th parameter of the equalizer impulse response at iteration n and the gradient algorithm step size. $E\left[\tilde{v}_n^2\right]$ can be estimated by empirical averaging :

$$\langle \tilde{v}^2 \rangle_n = (1-\delta)\langle \tilde{v}^2 \rangle_{n-1} + \delta \tilde{v}_n^2 \tag{13}$$

where δ is the estimation step size.

Remarks

- The non causality of the filters G and I is allowed since we have no real time constraint.
- Since blind equalization algorithms are slow, the detection is made in 3 steps : first the deconvolution filter is applied, then the resulting sequence $\hat{v}_n$ is used as the adaptive equalizer input. Finally and after the convergence of I, the fixed mean equalizer is again applied to $\hat{v}_n$.

4. SIMULATION RESULTS

To demonstrate the performance of the proposed hybrid detection scheme we have considered the following situation : the sequence is binary valued (± 1) and represents a periodic repetition of a signature (private key) composed of 4 possible symbols. v_n is a non gaussian white noise with kurtosis < 0. The unknown all-pass filter is 4th order with zeros $z = (-1.0004; 1.04; 0.95 \pm 0.311j)$. The constellation of figure 5 is the result of the generic detection scheme with only Wiener deconvolution (non causal, P=Q=50, see (1) and (3)) without all-pass attack. The obtained BER was equal zero. Note that '.' means a good detected symbol and '+' a false detected one.
The figure 6 represents the constellation when detecting the all-pass filtered sequence with the previous Wiener filter only. We clearly see that the detection fails completely (BER=1). Now we introduce a non causal adaptive equalizer I of order 20 ($P_c = P_{nc} = 10$) updated with the step size $\mu = 10^{-4}$ cascaded with the Wiener filter. The obtained constellation is shown on

figure 7 (zoomed on). Comparing constellations (6) and (7), it is obvious that the hybrid scheme enhanced the detection quality. Inded, the corresponding BER falls drastically to 0.168.

5. CONCLUSION

In digital audiowatermarking, all-pass filter attacks introduce a problem of nonminimum phase channel inverse identification. This kind of attacksdegrades strongly the performances of the detection techniques based on mean square criteria.

The proposed hybrid blind equalization/Wiener deconvolution scheme, based on a non quadratic criterion, outperforms the Wiener deconvolution detection. In this paper, the performances of the Godard algorithm is analyzed. The simulations show that the hybrid scheme enhanced the detection performances in a situation where only a 4 level-constellation and time invariant all-pass attacks were considered. The performances of the hybrid scheme will be tested for more constrained situations by developping a specific contrast function suited to the probability density function of the watermark.

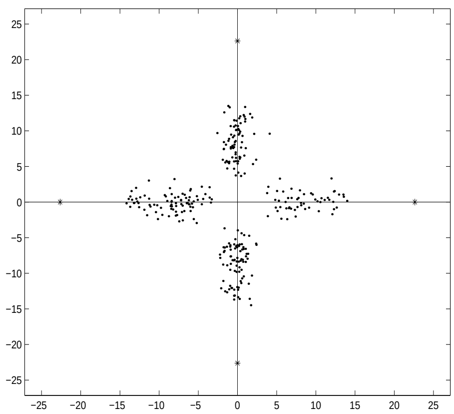

Figure 5: Results of the Wiener deconvolution without all-pass attack (BER=0)

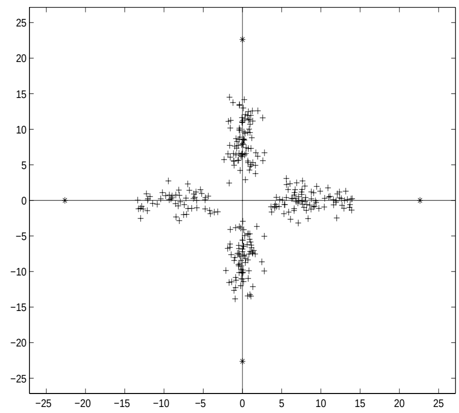

Figure 6: Results of the Wiener deconvolution with all-pass attack (BER=1)

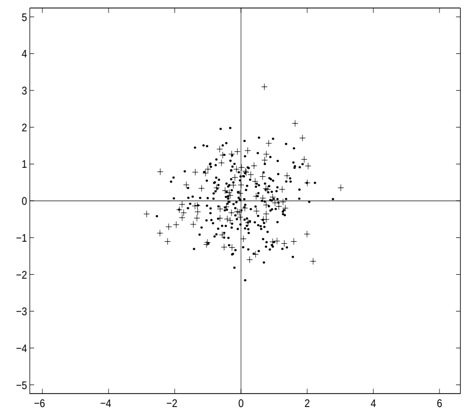

Figure 7: Results of the hybrid detection scheme (BER=0.168)

REFERENCES

[1] L. de C.T. Gomes, E. Gomez, M. Bonnet, N. Moreau, “Resynchronization Methods for Audio Watermarking”, 110th Convention of Audio Engineering Society, New York, September 2001.

[2] Y. Sato, “ A method for self recovering equalization for multilevel amplitude modulation”, IEEE Trans. Commun., vol. COM-23, no. 6, pp.679-682, June 1975.

[3] D. Godard, “Self-recovering equalization and carrier tracking in two-dimensional data communication systems”, IEEE Trans. Commun., vol. COM-28, no. 11, pp.1867-1875, Nov. 1980.

[4] O. Shalvi, E. Weinstein, “New criteria for blind deconvolution of nonminimum phase systems”, IEEE Trans. Information Theory, vol. 36, pp 312-321, March 1990.

[5] P. Comon, “From source separation to blind equalization contrast-based approach”, Int. Conf. on Image and Signal Processing, pp. 20-30, May 2001, Morocco.

[6] M. H. Hayes, “ Statistical Digital Signal Processing and Modeling”, Ed. Wiley, 1996.