Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Signal Processing: *Image Communication* 18 (2003) 283–296
SIGNAL PROCESSING: IMAGE COMMUNICATION
www.elsevier.com/locate/image

Practical watermarking scheme based on wide spread spectrum and game theory

Stéphane Pateux*, Gaëtan Le Guelvouit

IRISA-INRIA, Campus de Beaulieu, 35042 Rennes Cedex, France

Abstract

In this paper, we consider the implementation of robust watermarking scheme for non-i.i.d. Gaussian signals and distortion based on perceptual metrics. We consider this problem as a communication problem and formulated it as a game between an attacker and an embedder in order to establish its theoretical performance. We first show that known parallel Gaussian channels technique does not lead to valid practical implementation, and then propose a new scheme based on Wide Spread Spectrum and Side Information. Theoretical performances of this scheme are established and shown to be very close to the upper bound on capacity defined by Parallel Gaussian channels. Practical implementation of this scheme is then presented and influence of the different parameters on performance is discussed. Finally, experimental results for image watermarking are presented and validate the proposed scheme.

Keywords: Watermarking; Side information; Game theory; Performance evaluation; Spread transform

1. Introduction

A lot of effort has been dedicated over the last years for designing practical watermarking systems. The approaches were often viewing the media content as noise from the watermark detection perspective, hence regarding watermarking as a form of wide spread spectrum communication (WSS) with various forms of distortion measures (MSE or weighted MSE) and of channel characterizations [13,16,28]. The authors in [6] suggest to take into account the perceptual properties of the content and to embed in perceptually significant frequency components. Other approaches based on WSS and exploiting the perceptual sensitivity of the host data can also be found in [24,26,30]. However these techniques are based on empirical assumptions.

Attacks have often been modeled as the addition of White Gaussian noise (AWGN) [21,27], and more recently as linear filtering plus white or colored additive noise [19,29]. It has been shown in [1,7] that expressing the problem of watermarking as a problem of communication with side information leads to optimal performances. Costa [5] has indeed shown that in the context of attacks modeled by AWGN, the capacity is not dependent on the cover signal. However the solution proposed, known as the *Ideal Costa Scheme* (ICS), requires very large codebooks, hence is not realistic. Different approaches have then been proposed to reach performances of Costa's scheme using structured codebooks; Scalar Costa Scheme

*Corresponding author.
E-mail address: stephane.pateux@irisa.fr (S. Pateux).

0923-5965/03/$ - see front matter
doi:10.1016/S0923-5965(02)00145-5

(SCS) [10], syndrome based coding [2] or more recently trellis with multiple paths [18]. Dithered quantization techniques [1,25] may also be seen as techniques exploiting side information. Most of these schemes are defined for i.i.d. signals which is not a valid assumption for usual considered signals. Techniques based on parallel-Gaussian channels [4,20] have then been proposed to deal with such non-i.i.d. Gaussian signals. However practical implementation of Parallel-Gaussian necessitates to know the original signal [17] and does not lead to a valid implementation (see discussion in Section 2).

This paper deals with the robust data hiding problem, assuming a blind (the extraction system has no knowledge of the host signal) and symmetric (same private key for embedding and extraction) system. In this paper we consider this problem as a communication problem: one seeks the maximum hiding capacity (or rate of reliable transmission) over any hiding and attack strategies. The rate obviously depends on the perceptual distortion levels considered admissible and on the watermark channel (or attack scenarios) characterization. We especially present a technique based on Wide Spread Spectrum and Side Information facing Scaling and Additive White Gaussian Noise (SAWGN) optimized by considering Game Theory formalism in order to define performance limits. Practical implementation as well as efficiency of the proposed technique are further presented.

This paper is organized as follows. In Section 2, general consideration about watermarking of non-i.i.d. signals is first presented as well as limitation of the previously proposed techniques. In Section 3, we then present an optimized watermarking scheme based on Wide Spread Spectrum and Side Information and discuss about its practical implementation. In Section 4, experimental results are shown for image watermarking. Finally Section 5 concludes this work.

2. Watermarking of non-i.i.d. signals

Most of the techniques proposed in watermarking are assuming i.i.d. signals, however this assumption is rarely valid. For example, when performing embedding in a transform domain, coefficients are generally not i.i.d. (e.g. for images, low frequency coefficients have higher energy than high frequency coefficients). In [29], author then showed that in order to resist to filtering attacks, power spectrum of the watermark should be proportional to the power spectrum of the host signal, what they called the PSC condition. In [19,14], optimizations of watermarking techniques based on wide spread spectrum have been proposed for non-i.i.d. Gaussian signals considering Scaling and Additive White Gaussian Noise. While exploiting statistical properties of the host signal, those techniques are still not optimal since they do not exploit the realization of the host signal.

In [20,4], theoretical analysis of watermarking for non-i.i.d. signals have been carried. Capacity bounds have been derived by considering a game between an attacker and the embedder. First for i.i.d. Gaussian signals, capacity can be expressed as

$$C = \frac{1}{2}\log_2\left[1 + \frac{D_1}{D_2 - D_1}\left(1 - \frac{D_2}{\sigma_X^2}\right)\right], \qquad (1)$$

where D_1, D_2 corresponds respectively to the embedding distortion and to the attack distortion;[1] σ_X^2 corresponds to the variance of the host signal X. Optimal strategies for the embedder and the attacker take the following forms:

$$\begin{aligned} Y &= \gamma_1(X + W),\\ Y' &= \gamma_2(Y + \delta) \end{aligned} \qquad (2)$$

with

$$\begin{aligned} \gamma_1 &= \frac{\sigma_X^2 - D_1}{\sigma_X^2},\\ \gamma_2 &= \frac{\sigma_X^2 - D_2}{\sigma_X^2 - D_1},\\ \sigma_\delta^2 &= (D_2 - D_1)\frac{\sigma_X^2 - D_1}{\sigma_X^2 - D_2}, \end{aligned} \qquad (3)$$

where X corresponds to the host signal, W to the watermark (that is defined taking into account

[1] In [4], capacity formulation differs since author considered the attacker using a measure of distortion between the attacked signal and the watermarked one—noted as type X constraints in [20]. As discussed in [20], distortion between the attacked signal and the original one—type S constraints, is more suited.