SPREAD-SPECTRUM SUBSTITUTION WATERMARKING GAME

Jean-Philippe Boyer, Pierre Duhamel* and Jacques Blanc-Talon*[†]

[*] LSS, 3 rue Joliot-Curie, 91190 Gif sur Yvette - France
{jean-philippe.boyer, pierre.duhamel}@lss.supelec.fr
[†] CTA/GIP, 16 bis av. Prieur de la Côte d'Or, 94114 Arcueil - France
jacques.blanc-talon@etca.fr

ABSTRACT

In the integrity checking context of multimedia contents, a malicious user aims at devising a forged content in order to fool a watermarker by making him use as a genuine content. By considering that the watermark acts as an integrity stamp, the false-alarm probability to recover the watermark signature in a forged content is the criterion of interest. We study and solve a game for this criterion between a watermarker and a falsifier which is allowed to perform a substitution attack, *i.e.* replace the watermarked signal by a non-watermarked content. As for the watermarker, we are concerned with additive spread-spectrum (SS) embedding. Signals are modeled by parallel colored gaussian processes. Due to the intractability of the false-alarm probability, we resort to Chernoff bound as an alternative cost. Our study confirms some common heuristics: the best attacker choice is to substitute the watermarked host signal using a signal which has very close statistics to the original host signal. The best watermarker strategy is to embed the watermark into the weakest frequency power components of the host signal. We finally consider the consequences of these results in terms of frequency embedding domain for an image SS watermarking scheme which has to be robust to compression. This reveals notable differences with informed scalar quantized-based schemes.

1. INTRODUCTION

Zero-bit watermarking (*aka.* one-bit watermarking) addresses the problem to determine whether a tested content contains a given watermark (hypothesis H_1) or not (hypothesis H_0). Regardless to the embedding techniques and detection strategies, this problem has been initially introduced to address copyright protection issues [1]. In this context, the presence of a signature conveys commercial rights information about the tested content. To be reliable, the watermark should have maximal robustness against attacks. More precisely, the watermarker aims at making the miss detection probability $Pr(\text{choose } H_0|H_1 \text{ is in force})$ (denoted $Pr(H_0|H_1)$ for short) as low as possible, keeping the false-alarm probability $Pr(H_1|H_0)$ under an acceptable bound. Reciprocally, an hacker aims at erasing the watermark and hence maximizing the miss detection rate. Both of them are supposed to mildly distorted the media since they are both supposed to be interested in its commercial value. This approach refers to robust watermarking and it can be formalized as a game [2] between two opponents: the watermarker and an attacker.

Several other studies have chosen the overall probability of detection error as performance criterion [2, 3]. Assuming that H_0 and H_1 have equal priors and equal error costs, this criterion reads $\frac{1}{2}Pr(H_0|H_1) + \frac{1}{2}Pr(H_1|H_0)$. The watermarker wants to make it as low as possible whereas the attacker wants to maximize it. Traditionally, both of them are subjected to distortion constraints. Now, one can ask what are the applicative goals of the latter problem? To jointly maximize the sum of both error probabilities, the attacker ideally should perform an attack so that 1) he/she erases the watermark when it is present (in order to maximize the miss detection term $Pr(H_0|H_1)$) and 2) makes the detector use an non-watermarked content as a watermarked one (in order to maximize the false-alarm term $Pr(H_1|H_0)$). The first goal corresponds to robust watermarking. The second one can however be viewed as an integrity checking issue [4] by considering that the watermark acts as an integrity stamp (*i.e.* the presence of the watermark is interpreted as an authenticity evidence). To succeed, an attacker has to devise some forged content (arbitrarily *distinct* from the watermarked one) keeping the watermark detectable. These requirements, which are exactly opposed to the robust watermarking ones, refer to the fragile watermarking problems class[1]. Nevertheless, fragile watermarking has been seldom addressed in a game-theoretic point of view [4]. In particular, what are the best watermarker and attacker choices to respectively optimize the false-alarm cost? In this paper, we investigate this question in considering that the watermarker uses Spread-Spectrum (SS) embedding and the attacker performs a substitution attack, *i.e.* substitute the watermarked content by his own non-watermarked content. By additionally specifying that the system has to be robust to compression in the context of image watermarking, we finally propose an interpretation in terms of frequency embedding domain and a strategy comparisons with scalar quantized-based schemes.

2. CONTEXT AND GAME FORMULATION

Our framework is depicted on Fig. 1. Let $\boldsymbol{x} \in \mathbb{R}^N$ be a N-length random vector to be marked. This host signal could be extracted from an image block or an audio stream. A centered pseudo-random signature $\boldsymbol{w} = \{w_n\}_{1 \le n \le N}$ is produced. A secret key acts as a seed for this pseudo-random pattern. This key is shared at the embedding and detecting sides. $\boldsymbol{w}$ is added to $\boldsymbol{x}$, producing a spread-spectrum watermarked signal $\boldsymbol{y}_{|H_1}$. The latter content is

[1] Note that adopting the overall probability of error is a kind of "mix" approach between robust and fragile watermarking. Although it often provides a clear mathematical framework, the applicative finalities of this model remain less obvious.

1-4244-0367-7/06/$20.00 ©2006 IEEE

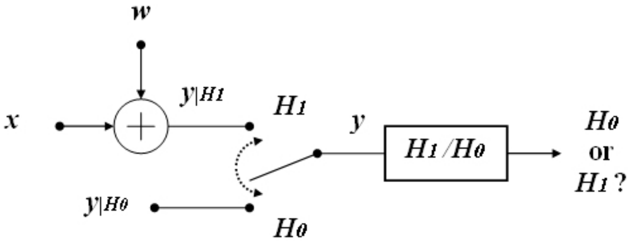

Fig. 1. Spread-spectrum substitution game context.

made publicly available. We consider the situation where a malicious user, a falsifier, would like to tamper this signal for falsification purposes (replace a car number plate by another one for example). Basically, two strategies are possible: modify or substitute the watermarked content. In this paper, we focus our attention on the *substitution* attack, which is also the most studied [5]. That is, the falsifier is allowed to build his own forged non-watermarked signal $\boldsymbol{y}_{|H_0}$ in order to substitute the original content $\boldsymbol{y}_{|H_1}$. To identify illicit contents, a detector has to guess if a tested content $\boldsymbol{y}$ contains watermark $\boldsymbol{w}$ (hypothesis H_1) or not (hypothesis H_0). The detector is not assumed to know the original host signal (blind watermarking). Once a detector structure has been chosen, False-alarm and Miss watermark detection probabilities can be formally defined by $P_F = Pr(H_1|H_0)$ and $P_M = Pr(H_0|H_1)$. As mentioned in the introduction, the forgery has reached its initial goal if the attacker is able to make the detector use the forged content as a genuine watermarked content. Hence, from the attacker point of view, this is equivalent to devise a forged content which maximizes the false-alarm probability P_F. Conversely, the detector does not want to be fool by a non-watermarked content. Thus, he wants to keep the false-alarm probability as low as possible by correctly tuning the embedded watermark. In a game-theoretic formulation, false-alarm probability appears to be the game cost. The associated game is then

$$\min_{Pr(\boldsymbol{w})} \max_{Pr(\boldsymbol{y}_{|H_0})} P_F \qquad (1)$$

where $Pr(\boldsymbol{w})$ and $Pr(\boldsymbol{y}_{|H_0})$ refer to the statistics of the watermark and the forged signal. Additionally, for the system to be reliable in absence of forgery, the watermarker expects to recover the presence of the watermark with a target maximal miss detection rate P_0, *i.e.*

$$P_M \leq P_0. \qquad (2)$$

This last constraint can provide to the watermark a minimal robustness toward possible "innocent" processings (which should not be considered as falsifications). This refers to semi-fragility systems. Besides, it is also classically required that the watermark induces a maximal average embedding distortion D_0 to guarantee the watermark invisibility, *i.e.*

$$D_w \triangleq \frac{1}{N}\mathbb{E}\,\|\boldsymbol{w}\|^2 \leq D_0. \qquad (3)$$

In this framework, note that it makes no sense to assume that the falsifier is subjected to any distortion constraint. The only constraint for the attacker is here to produce any non-watermarked signal.

3. PARALLEL GAUSSIAN SIGNALS AND NEYMAN-PEARSON TEST

In order to make the stated game explicitly solvable, signals are modeled by parallel colored independent gaussian processes such it has been done in [6]. Host and forged signals are respectively modeled as $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_1)$ and $\boldsymbol{y}_{|H_0} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_0)$, where $\boldsymbol{R}_1$ and $\boldsymbol{R}_0$ are symmetrical and positive definite $N \times N$ correlation matrices. Since the detector knows pattern $\boldsymbol{w}$, we thus have $\boldsymbol{y}_{|H_1} \sim \mathcal{N}(\boldsymbol{w}, \boldsymbol{R}_1)$ at the detector side. Note that the falsifier could ideally design a non-centered signal. Nevertheless, our framework assumes that the falsifier is supposed to devise a non-watermarked content. Since the watermark information in $\boldsymbol{y}_{|H_1}$ is conveyed by the mean of the distribution, one way to formalize a substitution attack is to impose the attacker to produce any centered signal. Moreover, most statistical modelings assumed than "natural" signals are generally centered (such as in [7] in the wavelet domain in instance). Hence, the centering assumption implies that the forged signal should be any "natural" signal, which seems to not be too restrictive. We consider watermarks produced by a pseudo-random centered gaussian process[2], that is $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_w)$ where $\boldsymbol{R}_w$ is symmetrical and positive definite $N \times N$ correlation matrices.

Let $\boldsymbol{B}$ be the Karhunen-Loève basis of $\boldsymbol{x}$ which makes $\boldsymbol{R}_1$ diagonal. For stationary gaussian processes, the Karhunen-Loève Transform (KLT) is approximatively equal to the DCT whenever N is large [6]. Each component can be viewed as a frequency component. Hence, due to the good decorrelation property of the DCT, we can assume that the forged signal and the watermark are also decorrelated in $\boldsymbol{B}$, thus have diagonal correlation matrices in $\boldsymbol{B}$. For $1 \leq n \leq N$, we respectively denote $\{\lambda_1(n)\}$, $\{\lambda_0(n)\}$ and $\{\lambda_w(n)\}$ the positive eigenvalues of $\boldsymbol{R}_1$, $\boldsymbol{R}_0$ and $\boldsymbol{R}_w$. Since KLT is unitary, the embedding distortion is preserved in $\boldsymbol{B}$. In the remainder of this paper, we now work in $\boldsymbol{B}$. In particular, we consider that all quantities involved so far correspond to their representation in $\boldsymbol{B}$ without introducing new notations.

In the sequel it is assumed that the attacker knows or is able to know the host signal and watermark distributions. Moreover, we consider that the detector knows or is able to know the forged signal statistics. Then, knowing pattern $\boldsymbol{w}$, the detector has to choose between the two following hypotheses

$$\begin{cases} H_0: \ \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{R}_0) \\ H_1: \ \boldsymbol{y} \sim \mathcal{N}(\boldsymbol{w}, \boldsymbol{R}_1) \end{cases} \qquad (4)$$

for a given tested signal $\boldsymbol{y}$. The optimal procedure is a Neyman-Pearson test [8] based on the likelihood ratio of $\boldsymbol{y}$. This equals

$$\begin{aligned} \Lambda(\boldsymbol{y}) &= \log \frac{Pr(\boldsymbol{y}|H_1)}{Pr(\boldsymbol{y}|H_0)} \qquad (5) \\ &= \frac{1}{2}\left(\log \frac{|\boldsymbol{R}_0|}{|\boldsymbol{R}_1|} - (\boldsymbol{y}-\boldsymbol{w})^T \boldsymbol{R}_1^{-1}(\boldsymbol{y}-\boldsymbol{w}) + \boldsymbol{y}^T \boldsymbol{R}_0^{-1}\boldsymbol{y}\right). \end{aligned}$$

For a given detection threshold τ, the Neyman-Pearson test is performed by the decision rule

$$\Lambda(\boldsymbol{y}) \begin{cases} > \tau \Rightarrow H_1, \\ \leq \tau \Rightarrow H_0. \end{cases} \qquad (6)$$

[2] Noteworthy, one can show that all the stated results in this paper remain valid in also considering watermarks with antipodal form, *i.e.* $w_n = \pm\left(\lambda_w(n)\right)^{\frac{1}{2}}$.

The associated False-alarm and Miss probabilities are then $P_{F|\boldsymbol{w}}(\tau) = Pr(\Lambda(\boldsymbol{y}) > \tau|\boldsymbol{w}, H_0)$ and $P_{M|\boldsymbol{w}}(\tau) = Pr(\Lambda(\boldsymbol{y}) \leq \tau|\boldsymbol{w}, H_1)$. Averaging these quantities over all possible realizations of pattern $\boldsymbol{w}$, we get $P_F(\tau) = \int_{\boldsymbol{w}\in\mathbb{R}^N} Pr(\boldsymbol{w})P_{F|\boldsymbol{w}}(\tau)d\boldsymbol{w}$ and $P_M(\tau) = \int_{\boldsymbol{w}\in\mathbb{R}^N} Pr(\boldsymbol{w})P_{M|\boldsymbol{w}}(\tau)d\boldsymbol{w}$.

The average embedding distortion (3) is expressed by $D_w = \frac{1}{N}\text{Trace}(\boldsymbol{R}_w) = \frac{1}{N}\sum_{n=1}^{N}\lambda_w(n)$. Due to the watermarking context, we additionally assume that the watermark power components remain *locally* smaller than the host signal components, that is $\forall n \in \{1,\dots,N\},\ \lambda_w(n) \ll \lambda_1(n)$. Then, game (1) can be rewritten as

$$\min_{\{\lambda_w(n)\}}\ \max_{\{\lambda_0(n)\}}\ P_F(\tau) \text{ subject to } \begin{cases} P_M(\tau) \leq P_0, \\ \sum_{n=1}^{N}\lambda_w(n) \leq ND_0, \\ \forall n,\ \lambda_w(n) \ll \lambda_1(n). \end{cases} \tag{7}$$

4. CHERNOFF BOUNDING

For game (7) to be solved, expression or estimate of probability $P_{F|\boldsymbol{w}}(\tau)$ is required. In the general case (that is when $\boldsymbol{y}_{|H_1}$ and $\boldsymbol{y}_{|H_0}$ are not assumed to have equal variances and/or equal means), this quantity does not admit closed-form expression. Instead, one possible solution is to work with dissimilarity metrics between rival data distributions (4). In particular, for $s \in [0,1]$, the Chernoff distance

$$D_c(s,\boldsymbol{w}) = -\log\int_{\boldsymbol{y}\in\mathbb{R}^N} Pr(\boldsymbol{y}|\boldsymbol{w},H_0)^{1-s}Pr(\boldsymbol{y}|\boldsymbol{w},H_1)^{s}d\boldsymbol{y} \tag{8}$$

is widely used as a dissimilarity measure between statistical distributions [9]. This distance is relevant since it generally leads to meaningful upper bounds on error probabilities to address hypothesis testing setups [10]. It can be shown that [8]

$$P_{F|\boldsymbol{w}}(\tau) \leq e^{-D_c(s,\boldsymbol{w})-s\tau} \tag{9}$$

where s is determined such as $D'_c(s,\boldsymbol{w}) = -\tau$ holds. Here, $D'_c(s,\boldsymbol{w})$ is the first derivative of $D_c(s,\boldsymbol{w})$ with respect to s. Then, averaging (9) over $\boldsymbol{w}$, we have

$$\begin{aligned} P_F(\tau) &\leq \int_{\boldsymbol{w}\in\mathbb{R}^N} Pr(\boldsymbol{w})e^{-D_c(s,\boldsymbol{w})+sD'_c(s,\boldsymbol{w})}d\boldsymbol{w} \quad (10)\\ &\triangleq d(s). \quad (11)\end{aligned}$$

$d(s)$ acts as a meaningful upper bound which guarantee a prescribed false-alarm rate. Hence, it can also makes sense for the attacker and watermarker in aiming at respectively optimize it. Then, we now choose $d(s)$ as an alternative optimization criterion. Note that this approach holds when addressing strategy setup. On the other hand, a performance analysis should require more accurate estimate techniques (see [8], chap. II.7). In our case, denoting $u(n,s) \triangleq s\lambda_0(n) + (1-s)\lambda_1(n)$, (8) becomes after some algebra

$$D_c(s,\boldsymbol{w}) = \frac{1}{2}\sum_{n=1}^{N} w_n^2\frac{s(1-s)}{u(n,s)} - \log\frac{\lambda_1(n)^{1-s}\lambda_0(n)^s}{u(n,s)} \tag{12}$$

and straightforward computations yield $\log d(s) =$

$$\frac{1}{2}\sum_{n=1}^{N} s\frac{\lambda_0(n)-\lambda_1(n)}{u(n,s)} + \log\frac{\lambda_1(n)\,u(n,s)}{u(n,s)^2 + s^2\lambda_0(n)\lambda_w(n)}. \tag{13}$$

5. ATTACKER STRATEGY

We now return to the game: the attacker aims at maximizing $d(s)$, or equivalently $\log d(s)$, knowing that the watermarker has previously embedded a mark which respects the distortion constraint (3). Furthermore, the detection constraint (2) implies that threshold τ has been set by the detector in such a way that $P_M(\tau) \leq P_0$ holds. Hence, τ is to be tuned after the attacker and the watermarker choices. Hence, it can be considered as fixed during the game[3]. As τ varies, s maps $[0,1]$ according to the relation $D'_c(s) = -\tau$ [8]. Then, P_0 fixes τ, and then s. For fixed $\{\lambda_1(n)\}$, $\{\lambda_w(n)\}$ and s, the attacker strategy part is given solving

$$\max_{\{\lambda_0(n)\}} \log d(s) \tag{14}$$

where the $\{\lambda_0(n)\}$ solely have to remain positive. Evaluating the gradient of $\log d(s)$ with respect to the component $\lambda_0(n), 1 \leq n \leq N$, and setting it to zero are equivalent to solve the N following third order polynomial equation in $\lambda_0(n)$

$$\begin{aligned} 0 \quad = \quad & s^3\lambda_0(n)^3 + s(2-3s)\lambda_1(n)\lambda_0(n)^2 \\ & +\lambda_1(n)\Big((3s-1)(s-1)\lambda_1(n) - s^2\lambda_w(n)\Big)\lambda_0(n) \\ & +(s-1)^2\lambda_1(n)^2\Big(\lambda_w(n)-\lambda_1(n)\Big). \quad (15)\end{aligned}$$

Evaluating the real roots and applying a Taylor expansion with $\frac{\lambda_w(n)}{\lambda_1(n)} \ll 1$, it comes that $\log d(s)$ is possibly maximal for

$$\lambda_0(n) \simeq \lambda_1(n) + (1-2s)\lambda_w(n). \tag{16}$$

Moreover, it can be shown that the second derivative of the criterion keeps a constant sign on a neighborhood of $\lambda_0(n) = \lambda_1(n) + (1-2s)\lambda_w(n)$, which makes it an optimum point. Since $\lambda_w(n) \ll \lambda_1(n)$, (16) shows that the optimal distribution of the forged signal remains very close to the original host signal ones. This confirms a widely observed heuristic in integrity checking which states that the more challenging substitution attack is achieved with the original host content.

6. WATERMARKER STRATEGY

We now address the watermarker strategy. Face to the previous optimal attack (16), the watermarker aims at minimizing $\log d(s)$ with respect to $\{\lambda_w(n)\}$ subjected to the distortion constraint $\sum_{n=1}^{N}\lambda_w(n) \leq ND_0$ and for some fixed s. Taking into account attack (16) and applying a Taylor expansion of $\log d(s)$ with $\frac{\lambda_w(n)}{\lambda_1(n)} \ll 1$, we get

$$\log d(s) \simeq -\frac{s^2}{2}\sum_{n=1}^{N}\frac{\lambda_w(n)}{\lambda_1(n)}. \tag{17}$$

Note that each term of the sum is the watermark-to-host power ratio in the n^{th} component. The initial minimization then becomes equivalent to maximize the sum $\sum_{n=1}^{N}\frac{\lambda_w(n)}{\lambda_1(n)}$. This maximization is solved as follow: let n_0 be the index of the weakest host signal

[3] $P_F(\tau)$ being monotonic decreasing with respect to τ, note that the detector should set τ to a maximal value. And since $P_M(\tau)$ is increasing with respect to τ, this maximal value is the one which verifies $P_M(\tau) \leq P_0$ with equality.

component, *i.e.* $n_0 = \arg\min_n \lambda_1(n)$. Doing so, we straightforwardly derive an upper bound of the sum by

$$\sum_{n=1}^{N} \frac{\lambda_w(n)}{\lambda_1(n)} \leq \frac{1}{\lambda_1(n_0)} \sum_{n=1}^{N} \lambda_w(n) = \frac{1}{\lambda_1(n_0)} N D_0. \qquad (18)$$

It is clear that upper bound (18) is reached in choosing $\lambda_w(n_0) = ND_0$ and $\lambda_w(n) = 0$ for all $n \neq n_0$. Then, the optimal strategy for the watermarker is ideally to spend all the distortion budget on the weakest host power component. Of course, in practice, this strategy cannot be admitted since locally concentrating too much power on a single component would imply some visibility of the watermark, violating the constraint $\lambda_w(n) \ll \lambda_1(n)$ for some component n. In a more realistic context, some Human Visual System [11] will fix a maximal watermark-to-host signal power ratio $R(n)$ in each component n (*i.e.* $\forall n \in \{1, \ldots, N\}$, $\frac{\lambda_w(n)}{\lambda_1(n)} \leq R(n)$) which ensures perceptual transparency. Despite these additional constraints, one can show that the optimal repartition among the components remains unchanged in its principle: fill first the weakest component until the local constraint $\frac{\lambda_w(n_0)}{\lambda_1(n_0)} \leq R(n_0)$ is saturated. Then, fill the second weakest component until the local constraint is saturated, *etc.* Stop when the distortion budget ND_0 has been completely spent.

7. INTERPRETATION AND STRATEGY COMPARISONS WITH QUANTIZED-BASED SCHEMES

In the context of image watermarking, we now propose to further interpret these results in term of embedding frequency domain and to make a comparison with distortion compensated scalar quantization-based scheme (QIM) [3] studied in the semi-fragile integrity checking context [4]. To this end, we specialize the previous framework in considering that signal $\boldsymbol{x}$ which has been initially defined as the host signal is in fact the sum of two signals: the real raw host signal and an additive independent noise which models a compression processing, say a JPEG-like processing. Hence, doing so, we require the system to also be robust toward compression according to constraint (2).

From the attacker point of view, the situation is fundamentally distinct from the forgery strategy with QIM: for SS, Sect.5 has shown that there exists an optimal substitution attack whereas, for QIM, any non-watermarked forged signal provide constant performances [4]. This can be interpreted as a forged signal interference rejection capability. SS embedding thus does not have this feature.

As for the watermarker, when the embedding domain is viewed as the DCT domain, results of Sect.6 point out that the watermark should be inserted in the highest frequencies as they traditionally provides the smallest variances for natural images. However, compression processings cannot be considered as additive in this domain since they traditionally eliminate high frequencies for perceptual reasons. Hence, these components are not expected to be a relevant embedding domain since the watermark no longer survives even if no forgery occurs. Instead, middle frequencies should convey the watermark. This strategy also appears to be distinct from QIM wherein low frequencies have been shown to be the first chosen to embed the mark [4]. This difference can be simply explained: since SS is not interference-rejecting, the most polluting signal is mainly the raw host signal contribution more than the compression contribution. Conversely, in the QIM case, the interference rejection removes the host signal contribution. Thus detection performances are only influenced by the compression noises, which have smaller variances in low frequencies for perceptual reasons. Hence, low frequencies are first involved for QIM.

8. CONCLUSION

Fragile watermarking problem has been specifically addressed through game theory: a falsifier aims at devising a forged content in order to fool a watermarker by making him use as a genuine content. When the watermark acts as an integrity stamp, the false-alarm probability to recover the watermark signature in a forgery content is the criterion of interest. We study and solve a game for this criterion between a watermarker and a falsifier. The attacker is assumed to perform a substitution attack, that is the watermarked signal is replaced by a non-watermarked content. As for the watermarker, we are concerned with additive spread-spectrum embedding. Signals are modeled by parallel colored gaussian processes. Due to the intractability of the false-alarm probability, we resort to Chernoff bound as an alternative cost.

Our study confirms some common heuristics: the best substitution attack uses a signal which has very close statistics to the original host signal. Face to this, the best watermarker strategy is to embed the watermark into the weakest frequency power components of the host signal.

Specifying that a spread-spectrum image watermarking system should also be robust to JPEG-like compression, we have proposed an interpretation in terms of frequency embedding domain and a strategy comparisons with quantized-based schemes. Due to the non-interference rejection capability of spread-spectrum, it has been pointed out that middle frequencies should be chosen first to embed the watermark. This turns out to be distinct from informed quantized-based scheme where interference rejection capability makes the low frequencies the appropriate embedding domain.

9. REFERENCES

[1] J. R. Hernandez and F. Perez-Gonzalez. Statistical analysis of watermarking schemes for copyright protection of images. *Proc. IEEE*, 87:1142–1166, July 1999.

[2] P. Moulin and A. Ivanovic. The zero-rate spread-spectrum watermarking game. *IEEE Trans. on S.P.*, 51(4):1098–1117, April 2003.

[3] J. J. Eggers and B. Girod. Blind watermarking applied to image authentication. In *Proc. IEEE ICASSP*, Salt Lake City, USA, May 2001.

[4] J. P. Boyer, P. Duhamel, and J. Blanc-Talon. Game-theoretic analysis of a semi-fragile watermarking scheme based on scs. In *proc of Int. Conf. Image Processing ICIP*, Genova, Italy, September 2005.

[5] O. Ekici, B. Sankur, B. Coskun, U. Naci, and M. Akcay. Comparative evaluation of semi-fragile watermarking algorithms. *Journal of Electronic Imaging*, January 2004.

[6] P. Moulin and M. K. Mihcak. The parallel-gaussian watermarking game. *IEEE Trans. on I.T.*, 50(2):272–289, February 2004.

[7] S. LoPresto, K. Ramchandran, and M. T. Orchard. Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework. In *Proc. Data Compression Conference 97*, Snowbird, Utah, USA, 1997.

[8] H. L. Van Trees. *Detection, Estimation, and Modulation Theory, vol. 1.* John Wiley and Sons, Inc.

[9] A. Jain, P. Moulin, M. I. Miller, and K. Ramchandran. Information-theoretic bounds on target recognition performance based on degraded image data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(9):1153–1166, September 2002.

[10] A. K. Goteti and P. Moulin. Qim watermarking games. In *Proc. IEEE ICIP '04*, volume 2, pages 717–720, October 2004.

[11] A. B. Watson. Dct quantization matrices visually optimized for individual images. In *Proc. SPIE 1913*, pages 202–216, 1993.