SPREAD SPECTRUM WATERMARKING FOR REAL IMAGES: IS EVERYTHING SO HOPELESS?

O. Koval[§], *S.Voloshynovskiy*[§], *F. Pérez-González*[†], *F. Deguillaume*[§], *and T. Pun*[§]

[§] University of Geneva - CUI, 24 rue General Dufour, CH 1211, Geneva 4, Switzerland
[†] Signal Theory and Communications Department, University of Vigo, E-36200 Vigo, Spain

ABSTRACT

In this paper we perform the capacity analysis of known-host-statitistics watermarking methods based on spread spectrum (SS) under Additive Whight Gaussian noise (AWGN) attack. The reason of our research is based on the paradox that being non-effective in theory SS-based practical watermarking systems outperform known-host-state methods when a commonly accepted benchmarking strategy (Stirmark benchmark) is used. We show that the gap in capacity of SS-based techniques with respect to quantization-based techniques at high WNR regime could be significantly reduced, if the embedding scenario is designed using a proper stochastic model of the host image at the encoder. We show how the practical solution of watermark power allocation problem relates to the optimal one. In particular, we model the host image as an Autoregressive Process of the first order (AR(1)) and show the results of power allocation based on the water-pouring principle. It is pointed out that watermark spectrum in a real system that is properly shaped using Contrast Sensitivity Function (CSF) well-approximates the water-pouring solution. Finally, we perform several tests to analyze the modified SS capacity facing AWGN attack. Experimental results show that host interference in this case is significantly weakened, which leads to a noticeable capacity improvement.

1. INTRODUCTION

Capacity analysis of data-hiding methods is a complex and challenging research task. The main complexity is coming from significant diversity of attacking strategies that potentially can be applied to decrease the rate of reliable communications. For a while, no general solution to this problem has been proposed for the case when the whole diversity of possible attacks (for instance, from Stirmark 3.1 benchmark [11]).

Since it was established by Cox et al. [4] that data hiding problem could be regarded as communications with side information available at the encoder and due to the striking result of Costa [2] about zero host interference for such communications protocols for the case of i.i.d Gaussian host, watermark and channel, it became popular to perform the comparison of embedding methods by their capacity-approaching abilities for the ideal AWGN channel [5].

Mainly, two classes of embedding strategies are of particular interest [10]. The methods of the first class usually called "spread-spectrum" or known-host-statistics methods, do not utilize any knowledge about host state but perform embedding based on its statistics to satisfy the necessary embedding distortion constraint. The methods of the second class, regarded to as as known-host-state methods, are developed to approach the Costa's result and are based mostly on the quantization operation. It was shown that known-host-state methods have significantly higher embedding rate than known-host-statistics methods due to the host interference cancellation.

This conclusion is evidently correct for the case of i.i.d. signals but it is questionable in the case of real images. The reason for doubts is twofold. Firstly, the assumption about i.i.d. Gaussian host is not valid in the case of real images neither in the coordinate [6], nor in the transform [8] domains. Therefore some improvement of theoretical performance is expected when proper stochastic image model is applied. The second reason is coming from the analysis of practical watermarking systems using more extended benchmarking. The results of this analysis are showing that the spread spectrum based technique [13, 14] outperfoms all the competitors in Stirmark 3.1.

Motivated by this lack of coincidence, we formulate the main goal of the paper as a theoretical justification for the fundamental performance limits of spread spectrum watermarking in the case of real images.

The paper is organized as follows. Section 2 presents the theoretical model of known-host-statistics data hiding for real images. The results of benchmarking of the developed embedding method is given in Section 3. Finally, Section 4 concludes the paper.

Notation. We use capital letters to denote scalar random variables X, bold capital letters to denote vector random variables $\mathbf{X}$, corresponding small letters x and $\mathbf{x}$ to denote the realizations of scalar and vector random variables, respectively. The power spectrum of $\mathbf{x}$ is denoted $S_{XX}(\omega)$. The superscript N is used to designate length-N vectors $\mathbf{x} = x^N = [x_1, x_2, ..., x_N]^T$ with ith element x_i. We use $X \sim p_X(x)$ or simply $X \sim p(x)$ to indicate that a random variable X is distributed according to $p_X(x)$. The variance of X is denoted σ_X^2. $\mathbf{I}_N$ denotes the $N \times N$ identity matrix.

2. PROBLEM FORMULATION

Communications of i.i.d. signals : The diagram of watermarking system as a communications with side information available at the encoder is presented in Fig.1. It is assumed that the watermark $\mathbf{w} = [w_1, ..., w_N]$ and host image $\mathbf{x} = [x_1, ..., x_N]$ are i.i.d. Gaussian random variables, i.e. $\mathbf{W} \sim \mathcal{N}(\mathbf{0}, \sigma_W^2 \mathbf{I}_N)$ and $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \sigma_X^2 \mathbf{I}_N)$ and they undergo independent AWGN $\mathbf{N} \sim \mathcal{N}(\mathbf{0}, \sigma_N^2 \mathbf{I}_N)$ attack. Here N=$M_1 M_2$, where $M_1 \times M_2$ is the size of the host image. Let also $S_{WW}(\omega)$, $S_{XX}(\omega)$ and $S_{NN}(\omega)$ denote corresponding power spectral densities of $\mathbf{W}$, $\mathbf{X}$ and $\mathbf{N}$.

When the switch is closed (Fig. 1), the host image is

Further author information: (Send correspondence to S.Voloshynovskiy): E-mail: svolos@cui.unige.ch, http://sip.unige.ch

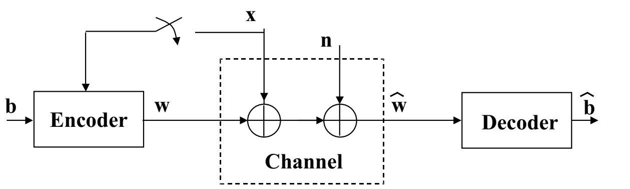
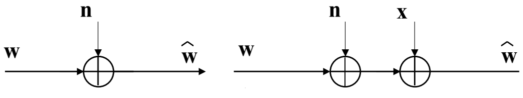


Figure 1: Communications with side information available at the encoder.



(a) (b)

Figure 2: Equivalent communication channels when side information is available (a) and unavailable (b) at the encoder.

available at the encoder and it is possible to approach capacity of the ideal AWGN channel (Fig. 2,a) based on Costa's scheme:

$$C=\frac{1}{2}\log_2\left(1+WNR\right), \tag{1}$$

where Watermark-to-Noise Ratio $WNR = 10\log_{10}\left(\frac{\sigma_W^2}{\sigma_N^2}\right)$. In case of the open switch, side information is not presented at the encoder and system performance is sacrificed from host interference (Fig. 2,b). The corresponding capacity formula is determined as:

$$C=\frac{1}{2}\log_2\left(1+\frac{\sigma_W^2}{\sigma_N^2+\sigma_X^2}\right). \tag{2}$$

It is evident that direct application of (1) and (2) to real images demonstrates obvious performance advantages of known-host-state methods under AWGN attack.

Data hiding for real images : Digital watermarking in still images is based on the specific properties of this type of media. As it was mentioned in the introductory part of the paper, the stochastic image model plays the crucial role. Several models have been proposed in the literature for real images both in the coordinate and in the transform domain [7, 8, 15]. To capture local correlation of natural images we use 1-D autoregressive process of the first order. The advantage of this model is its simplicity and tractability both in the coordinate and in the transform domain. Additionally, it provides a good fit to the power spectral density of real images.

In this case the power spectral density of the host signal is determined by [7]:

$$S_{XX}(\omega)=\sigma_X^2\frac{1-\rho^2}{1+\rho^2-2\rho\cos(\omega)}, \tag{3}$$

where $-1\le\rho\le 1$ is a correlation coefficient and $-\pi\le\omega\le\pi$. Power spectral density of AR(1) process is shown in Fig. 3.

It is clear (Fig. 3) that a white Gaussian watermark in (2) is not any more optimal because of non-effective allocation of watermark energy.

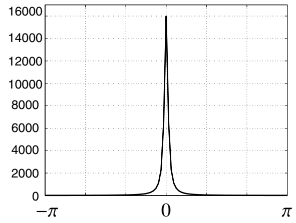


Figure 3: Power spectrum of AR(1) process; σ_X^2=410, ρ=0.95.

Therefore, the problem of the optimal watermark power allocation for such correlated data and AWGN channel in the frequency domain could be formulated based on the water-pouring principle [3]: it is necessary to maximize the mutual information between channel input and output subject to the following power constraint:

$$\max_{\frac{1}{2\pi}\int_\Omega S_{WW}(\omega)d\omega=P_{emb}} \mathbf{I}(\mathbf{W};\hat{\mathbf{W}}) \tag{4}$$

and for $S_{WW}(\omega)\ge 0$. Using methods of Lagrange multipliers and taking into account that the noise power spectrum is constant, $S_{NN}(\omega)$=σ_N^2, this maximization problem could be formulated in the following form:

$$\begin{aligned} &max_{S_{WW}(\omega)}J(S_{WW}(\omega)) = \\ &=\frac{1}{2\pi}\int_\Omega\frac{1}{2}\log_2\left(1+\frac{S_{WW}(\omega)}{S_{XX}(\omega)+\sigma_N^2}\right)- \\ &-\lambda\left(\frac{1}{2\pi}\int_\Omega S_{WW}(\omega)d\omega-P_{emb}\right). \end{aligned} \tag{5}$$

The solution to this problem is given by:

$$S_{WW}^{opt}(\omega)=\begin{cases} \Theta-S_{XX}(\omega)-\sigma_N^2, \text{ if} \\ \quad S_{XX}(\omega)+\sigma_N^2<\Theta, \\ 0, \quad \text{otherwise}, \end{cases} \tag{6}$$

where the constant Θ is selected to satisfy the power constraint in (4). Thus, the obtained result determines those frequency channels where watermark energy should be allocated: no energy will be distributed to low frequency components containing most of the host signal power. The internal boundaries of the frequency range [-π; $-\omega'$]$\cup$[ω'; π] of the watermark power spectrum (Fig. 4) can be obtained from the solution of the following equation:

$$\begin{aligned} &\frac{\sigma^2(1-\rho^2)}{1+\rho^2-2\rho cos(\omega')}\left(1-\frac{\omega'}{\pi}\right)-\sigma^2+ \\ &+\frac{2\sigma^2}{\pi}arctg\left(\frac{1+\rho}{1-\rho}tg\left(\frac{\omega'}{2}\right)\right)=P_{emb}. \end{aligned} \tag{7}$$

We will refer to the embedding according to (6) and (7) as to *optimized spread spectrum* (*OSS*).

Having solved eq. (7) with respect to ω' one obtains

$$S_{WW}^{opt}(\omega)=S_{XX}(\omega')-S_{XX}(\omega), \tag{8}$$

and the resulting capacity of OSS is:

$$C=\frac{1}{\pi}\int_{\omega'}^{\pi}\frac{1}{2}\log_2\left(\frac{S_{XX}(\omega')+\sigma_N^2}{S_{XX}(\omega)+\sigma_N^2}\right)d\omega, \tag{9}$$

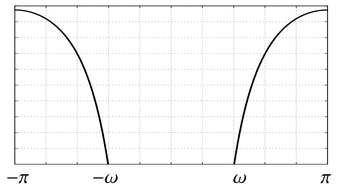


Figure 4: Optimal solution of the watermark power allocation problem.

where $S_{XX}(\omega')$ is a value of the power spectrum of the host image on frequency ω=ω'.

While being optimal from a pure communications point of view the obtained solution can be not perfect for practical digital image watermarking systems due to the specific additional requirements. One of them is perceptual invisibility, meaning that the embedding of information should be performed in a manner adapted to the properties of HVS. In general, three factors, luminance sensitivity (determines the visibility of the noise depending on the background brightness level), frequency sensitivity (characterize the apprehensibility of HVS to the distortion on different frequencies) and texture sensitivity (reflects sensitivity of HVS to the noise in flat, edge and texture regions) should be taken into account.

In the general case, watermark visibility can be determined by the product of the above mentioned maskings [14]:

$$M = M_l \cdot M_f \cdot M_t, \tag{10}$$

where M_l is a luminance masking, M_f is a frequency masking and M_t is a texture masking.

Although all three maskings are important to achieve the optimal system performance, we concentrate in this paper only on the influence of M_f, assuming M_l=M_t=1.

Frequency masking of HVS could be modeled by a *Contrast Sensivity Function* (*CSF*) that for the case of isotropic angular sensivity approximation is determined by [7]:

$$CSF(f) = A\left[\alpha + \frac{f}{f_0}\right] exp\left[-\left(\frac{f}{f_0}\right)^{\phi}\right], \tag{11}$$

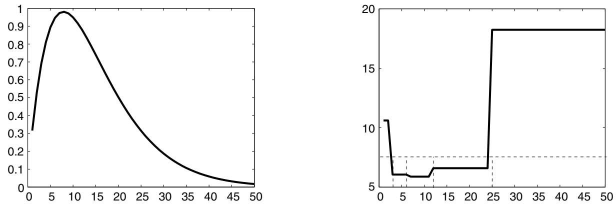
where $f = \sqrt{f_1^2 + f_2^2}$, f_1,f_2 are spatial sampling frequencies in two dimensions in cycles per degree (cpd), A=2.6, $f_0 = 8.772$ and ϕ=1. If normal conditions of viewing are assumed (image resolution is equal to 300 dpi and distance to the image is 0.5 m), f∈[0, 50] cpd.

Embedding according to the CSF means in particular that watermark spectrum should be bounded in each frequency component by the inverse of this function.

If hiding is performed in the wavelet transform domain, one of the possible solutions [14] is to approximate weights $M_f(V_{i,j})$ of each subband $V_{i,j}$, where i, $i \in [1,N]$ corresponds to the level of decomposition and j, $j \in [1,3]$ denotes subband spatial orientation, by:

$$M_f(V_{i,j}) = min_{f \in V_{i,j}} CSF^{-1}(f). \tag{12}$$

For the case of one-dimensional signals and positive frequencies, CSF and its inverse (with corresponding frequency splitting by wavelet transform) are represented in Fig. 5.



(a) (b)

Figure 5: CSF: (a) CSF and (b) approximated inverse CSF.

Assuming an i.i.d. Gaussian watermark, after frequency shaping that depends on global fidelity constraint P_{emb}, it will have properties approximating the optimal communications solution. The only differences consist in (a) embedding of a small portion of the energy in the low frequency part of power spectrum and (b) in the non-smooth character of practical solution. Even in such approximate form, the practical solution approaches the optimal one much closer than the commonly accepted white watermark.

3. EXPERIMENTAL RESULTS

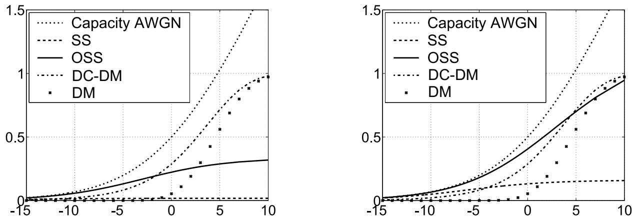
In this section we investigate the information-theoretical performance limits of OSS watermarking and compare them with the state-of-the-art techniques: Dither Modulation and Distortion Compensation DM (DC-DM) [1]. The classical SS is also included to illustrate performance improvement. Capacity of the zero interference AWGN channel is added to show performance gap to the virtually host interference free communications.

Benchmarking set up : In our experiments we assume that global power embedding constraint should guarantee stego image quality of 38 dB in Peak Signal-to-Noise Ratio (PSNR), PSNR=$10\log_{10}\frac{255^2}{\sigma_W^2}$. It corresponds to P_{emb}≈10. While efficiency of known-host-state methods does not directly depend on distribution of thehost, for the SS and the OSS watermarking we review two different regimes characterized in terms of Watermark-to-Image Ratio (*WIR*), WIR = $10\log_{10}\left(\frac{\sigma_W^2}{\sigma_X^2}\right)$=$10\log_{10}\left(\frac{P_{emb}}{\sigma_X^2}\right)$, WIR_1=-6dB and WIR_2=-16 dB. Therefore, the variance of the host image is equal σ_1^2≈ 40 and σ_2^2≈ 410, respectively. The correlation coefficient ρ=0.95 is used in both cases [9, 12]. The range of possible AWGN powers is selected to have *WNR* within the following interval WNR∈[-15 dB; 10 dB]. It should be also noted that we assume Gaussian watermark for the SS and OSS watermarking, while for the case of the DM and DC-DM binary watermark is used. This assumption does not influence the performance of the quantization-based methods because for the target WNR regime it does not depends on the cardinality of watermark alphabet.

Due to the fact that no masking is applied in all watermarking methods which are included in our benchmarking, the OSS embedding is performed without bounding the power spectrum of the watermark by approximated inverse of CSF (Fig. 5, b).

Evaluation of (8) is performed numerically using the following simplifications: spectral range $[\omega'; \pi]$ is divided on $10^3(\pi\text{-}\omega')$ non-overlapping intervals covering Ω. It is assumed that power spectral density of the host image is flat on each interval and its value is determined by left interval boundary.

The results of performed experiments are presented in Fig. 6.



(a)

(b)

Figure 6: Capasity analysis: khown-host-state versus khown-host-statistics watermarking of real images: (a) WIR=-16 dB and (b) WIR=-6 dB.

The obtained experimental results show how significantly SS watermarking performance can be improved using model based embedding. As a reference we use the crossing point of the capacity plots of the SS and the DC-DM and the OSS and DC-DM. In the the case of WIR=-16 dB this point is moved from -10 dB to -2 dB and from -4 db to 4 dB in case of WIR=-6 dB.

Therefore, modified embedding in the OSS allows to enchance the performance of the classical SS on 8 dB in each case.

4. CONCLUSIONS

In this paper we present one of several possible improvements of classical spread spectrum watermarking under AWGN attack. Motivated by the superior practical performance of the state-of-the-art watermarking technology that is based on SS principle, our approach consists in modifying the watermark power spectrum depending on the power spectrum of the host. In this situation, when the host image is modeled as the AR(1) process, we show the solution of the optimal watermark power allocation problem. The developed data-hiding method called optimized SS watermarking illustrates a significant performance enhancement over classical known-host-statistics methods.

5. ACKNOWLEDGMENTS

This paper was partially supported by the Swiss SNF grant No 21-064837.01, SNF Professorship grant No PP002-68653/1, the CTI-CRYMEDA-SA and Interactive Multimodal Information Management (IM2) projects. The authors are thankful to Jose Vila, Emre Topak and Yuriy Rytsar for many helpful and interesting discussions. Special thanks are to J. Eggers (Siemens former University of Erlangen-Nürnberg) for many interesting discussions about the SCS and corresponding capacity results of performance of quantization-based schemes.

REFERENCES

[1] B. Chen and G. W. Wornell. Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. *IEEE Trans. on Information Theory*, 47:1423–1443, May 2001.

[2] M. Costa. Writing on dirty paper. *IEEE Trans. on Information Theory*, 29(3):439–441, May 1983.

[3] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley and Sons, New York, 1991.

[4] I. Cox, J. Killian, T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for images, audio and video. In *Proceedings of the IEEE Int. Conf. on Image Processing ICIP-96*, pages 243–246, Lausanne, Switzerland, 1996.

[5] J. Eggers, J. Su, and B. Girod. Performance of a practical blind watermarking scheme. In *Proc. SPIE Security and Watermarking of Multimedia Contents 01*, San Jose, CA, January 2001.

[6] M.H. Hayes. *Statistical Digital Signal Processing and Modeling Programming*. Wiley, New York, 1996.

[7] A. L. Jain. *Fundamentals of Digital Image Processing*. Prentice-Hall, 1989.

[8] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. on PAMI*, 11:674–693, 1989.

[9] P. Moulin and M. K. Mihcak. A framework for evaluating the data-hiding capacity of image sources. *IEEE Trans. on Image Processing*, 11(9):1029–1042, September 2002.

[10] F. Perez-Gonzlez, F. Balado, and J. R. Hernndez. Performance analysis of existing and new methods for data hiding with known-host information in additive channels. *IEEE Trans. on Signal Processing, Special Issue on Signal Processing for Data Hiding in Digital Media and Secure Content Delivery*, 51(4), April 2003.

[11] F. A. P. Petitcolas. Stirmark 3.1. 1999. http:// www.cl.cam.ac.uk/ fapp2/watermarking/stirmark31/.

[12] J. K. Su, J. J. Eggers, and B. Girod. Analysis of digital watermarks subjected to optimum linear filtering and additive noise. *Signal Processing, Special Issue on Information Theoretic Issues in Digital Watermarking*, 81(6):1141–1175, 2001.

[13] S. Voloshynovskiy, F. Deguillaume, O. Koval, and T. Pun. Robust watermarking with channel state estimation, part I: Theoretical analysis. *Signal Processing*, page Accepted for publication, 2004.

[14] S. Voloshynovskiy, F. Deguillaume, O. Koval, and T. Pun. Robust watermarking with channel state estimation, part II: Applied robust watermarking. *Signal Processing*, page Accepted for publication, 2004.

[15] S. Voloshynovskiy, A. Herrigel, N. Baumgaertner, and T. Pun. A stochastic approach to content adaptive digital image watermarking. In *Third International Workshop on Information Hiding*, volume 1768, pages 212–236, Dresden, Germany, Sept. 29 - Oct. 1st 1999.