ON THE TRADEOFF BETWEEN SECURITY AND ROBUSTNESS OF THE TRELLIS CODED QUANTIZATION SCHEME

Sofiane BRACI, Rémy BOYER and Claude DELPHA

Laboratiore des Signaux et Systèmes (L2S)
CNRS, Université Paris-Sud XI (UPS), SUPELEC
{sofiane.braci,claude.delpha,remy.boyer}@lss.supelec.fr

ABSTRACT

The steganographic security in the Cachin's work is defined as the statistical invisibility between the host signal and its marked version. At contrary, the robustness to an attack is not a prime goal. In robust watermarking, this is exactly the inverse. The Scalar Costa Scheme (SCS) is a typical example of this fact. Indeed, this scheme is robust to Additive White Gaussian Noise (AWGN) attack but is drastically insecure since its probability density function for Gaussian host signal is severely discontinuous. An improved scheme has been proposed by Guillon et al. which increases the security to the detriment of the robustness. In this paper, we propose a new watermarking scheme, based on the combination of the Spread Transform (ST) and the Trellis Coded Quantization (TCQ) which is secure and robust to AWGN attack.

Index Terms— Security, Robustness.

1. INTRODUCTION

In the context of robust watermarking, and towards the half of 90s, Spread Spectrum watermarking [4] appears and widely interest the scientific community during several years. Its advantage is the robustness against Additive White Gaussian Noise (AWGN) attack at the price of a low capacity. In 1999, *Chen and Wornel* [1] have introduced a new class of watermarking scheme called *Quantization Index modulation* (QIM), where the host signal is considered as the side information of a Costa's scheme. By leading on the results of [7], the capacity of informed watermarking schemes is optimal since the side information is not considered as a nuisance signal. A practical and efficient implementation of the Costa's ideas is the Scalar Costa Scheme (SCS) proposed by Eggers et al. [5] which is quiet similar to the Distortion Compensated QIM (DCQIM) watermarking [1].

In the context of the steganography, Cachin in [10] has defined the notion of secure scheme by the closeness of the Probability Density Function (PDF) of the host and marked signals. The distance criterion is the Kullback-Leibler Distance (KLD) or also called relative entropy. A steganographic analysis of the SCS [6] shows that this scheme is not secure according to the Cachin's criterion [10]. Indeed, the SCS introduces many artifacts in the PDF of the watermarked signal. So, a simple inspection of the statistics of the watermarked signal gives away the presence of the watermark. In reference [9], the authors propose a modification of the SCS which leads to considerably improve the steganographical security of the scheme. But, as a price to pay, the proposed scheme imposes many constraints in terms of robustness. The aim of this work is to find a quantization-based watermarking scheme which is steganographicaly secure and robust.

The authors would like to thank the ESTIVALE project from ANR (Agence Nationale de la Recherche) for funding.

2. QUANTIZATION-BASED WATERMARKING SCHEMES AND SECURITY ANALYSIS

2.1. The Scalar Costa Scheme

Eggers et al. in [5] have introduced a sub-optimal scheme based on the Costa's ideas [7]. The authors propose to construct a codebook from the reconstruction points of a scalar quantizer. This approach is called *Scalar Costa Scheme* and is robust to AWGN attack for optimal value of Costa's factor α . However, it has been shown in reference [6] that the regular partitioning of the scalar quantizer generates many artifacts in the Probability Density Function (PDF) of the marked signal. Consequently, in a steganographic point of view, the security of this scheme is low since the relative entropy between the marked and the host signals is high (*cf.* Fig. 1-a). To avoid this problem, Guillon et al. [9] have proposed a new scheme to improve the security of the SCS. The main idea is to insert the message in the host signal with a uniform PDF and with $\alpha = 0.5$. By doing this, it can be shown on Fig. 1-b (and see in [9] for theoretical considerations) that the PDF of the marked signal is close to the one of the host signal, which in particular, decreases the relative entropy. Toward this end, this scheme uses a compressor step before embedding the watermark to equalize with a non-linear function the histogram of the host signal. After encoding, an inverse compression step is applied to the watermarked signal with uniform PDF in order to reconstruct the signal with the original PDF. Unfortunately, the gain in security leads to several constraints, since we have to choose $\alpha = 0.5$, it is impossible to choose the optimal value of α which allows a good robustness facing an AWGN attack. In addition, as figured on Fig. 2, the Guillon et al. scheme is less robust than the SCS. This is due to the compression and decompression which increases the distortions. In the next section, we recall the interest to use a structured codebook based on the Trellis Coded Quantization (TCQ).

2.2. The Trellis Coded Quantization (TCQ) scheme

The TCQ [8] is a trellis-based quantization scheme associated with a structured codebook. This approach is based on the Trellis Coded Modulation introduced in reference [3]. The TCQ allows to reduce the complexity cost of the watermarking system and allows to decrease the watermark distortions. More precisely, in the watermarking-based TCQ, the paths in the trellis are forced by the

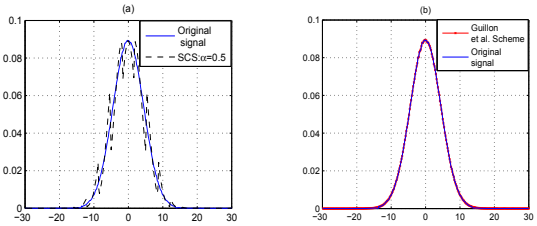

Fig. 1. Probability Density Function (PDF) of the host signal and of the watermarked signal for the (a) SCS, (b) Guillon et al. scheme.

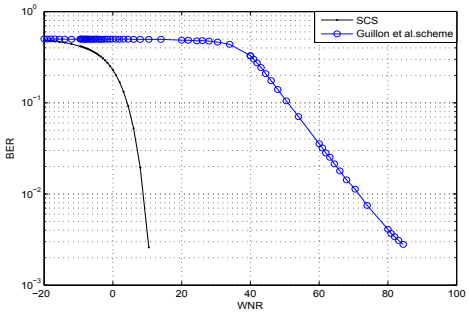

Fig. 2. Bit Error Rate (BER) Vs. Watermark to Noise Ratio (WNR) for the SCS and for the improved SCS of Guillon et al. with DWR = 13 dB

.

values of the watermark and the samples of host signal are quantized with the codebook corresponding to the trellis path with a rate of one bit per sample. This approach is called the TCQ path selection (TCQ-PS) and can be described in the following manner. Consider the trellis defined by the transition function :

$$\begin{aligned} E \times \{0,1\} &\rightarrow E \\ t : (e_i, m[i]) &\mapsto e_{i+1} \end{aligned} \tag{1}$$

where $E = \{0, 1, ..., 2^{r-1}\}$ represents the set of all possible states of the trellis (in the sequel we take r=9). The distortion due to the watermark embedding depends on the previous states of the trellis and of the input symbol :

$$\begin{aligned} E \times \{0,1\} &\rightarrow [-\frac{\Delta}{2}, \frac{\Delta}{2}] \\ o : (e_i, m[i]) &\mapsto d[i] \end{aligned} \tag{2}$$

then, the sub-codebooks can be written as :

$$\mathcal{U}_m[i] = \{k\Delta + o(s_i, m[i]), k \in \mathcal{Z}\} \tag{3}$$

where Δ is the path of scalar quantization associated to the trellis, s_i represents original signal sample and $m[i]$ is the information bit. The closest codeword $u^\star \in \mathcal{U}_m$ to the host signal sample s is determined by the $Viterbi$ algorithm [2] :

$$u^\star = \arg \min_{u \in \mathcal{U}_m} \sum_{i=1}^{N} (s[i] - u[i])^2. \tag{4}$$

At the decoder, the received signal is requantized, with the *Viterbi* algorithm, to find the best path of the trellis. The transition set allows to recover the inserted watermark.

2.3. Security analysis of the SCS and TCQ schemes

In the Cachin's work [10], the concept of ε-secure scheme is introduced. This can be described according to the following argumentation. Let s be the host signal, w the watermark and x the marked signal, then a ε-secure [10] scheme has to satisfy

$$\mathrm{KLD} = \int_{-\infty}^{+\infty} P_S(z) \ln \frac{P_S(z)}{P_X(z)} dz \leq \varepsilon \tag{5}$$

where KLD is the KL distance or the relative entropy, P_S is the PDF of the host signal, and P_X is the PDF of the watermarked signal. So, on Fig. 3, we have reported the KLD as a function of the Document to Watermark Ratio (DWR) for the SCS with optimal α, for the TCQ and the lowest KLD for Gaussian signals. This lower bound is defined as the relative entropy between the distribution of the host signal, assumed to be Gaussian of variance σ_s^2 and the distribution of marked signal, also assumed to be Gaussian of variance $\sigma_x^2 = \sigma_s^2 + \sigma_w^2$ where σ_w^2 is the variance of the watermark signal w. More specifically, it is straightforward to see that this bound is given by

$$\mathrm{KLD}_{\mathrm{theo}}(\mathrm{DWR}) = \frac{1}{2}\left[\frac{1}{1+10^{\frac{\mathrm{DWR}}{10}}} - \ln\left(1+10^{-\frac{\mathrm{DWR}}{10}}\right)\right] \tag{6}$$

where the DWR is given by $\mathrm{DWR} = 10\log_{10}\frac{\sigma_s^2}{\sigma_w^2}$. As the derivative of $\mathrm{KLD}_{\mathrm{theo}}$ with respect to the DWR is

$$\frac{\partial \mathrm{KLD}_{\mathrm{theo}}}{\partial \mathrm{DWR}} = -\frac{\ln 10}{20} \cdot \frac{2 \cdot 10^{\mathrm{DWR}} + 1}{1+10^{\frac{\mathrm{DWR}}{10}}} < 0. \tag{7}$$

Then, the $\mathrm{KLD}_{\mathrm{theo}}(\mathrm{DWR})$ is a strictly decreasing function. It is highly desirable for a watermarking scheme to be as close as possible to this theoretical lower bound. According to Fig. 3, we can note that the KLD for the TCQ is much lower than the one for the SCS, because as we have seen before this scheme generates artifacts in the PDF of the watermarked signal.

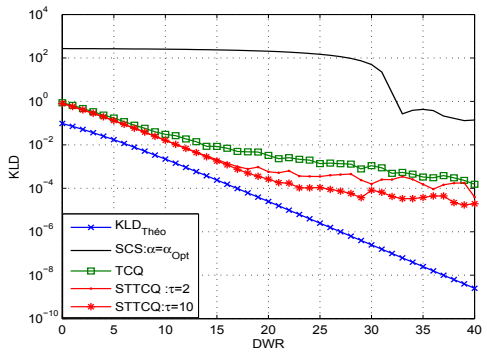

Fig. 3. Kullback-Leibler Distance (KLD) Vs. Document to Watermark Ratio (DWR).

3. ANALYSIS OF THE TRADEOFF BETWEEN ROBUSTNESS AND SECURITY

A desirable property of robust watermarking is the robustness (small BER) to AWGN attack. It is well known that Spread-Transform (ST) based quantization scheme is an efficient way to improve the robustness to AWGN attack [5]. So, we briefly recall this technique.

3.1. Spread Transform

Chen and Wornel in paper [1] have introduced an efficient watermarking scheme which allows to spread the message on several host signal samples. The global process is given on Fig. 4. The spreading

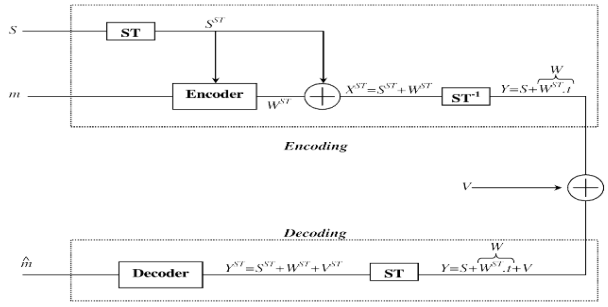

Fig. 4. Scheme of Spread Transform combined with a side information watermarking system.

of the host signal, denoted by $\underline{s} = [s_0, ..., s_{M-1}]$, is given by

$$s_l^{ST} = \sum_{i=\tau l}^{\tau l+\tau-1} s_i t_i \tag{8}$$

where $\tau \in \mathcal{N}^\star$ is the spreading factor. The above operation is in fact a projection along direction $\underline{t}$, where $\underline{t}$ is a unitary vector. Next, the inverse transformation is applied to the watermarked signal according to $x^{ST} = s^{ST} + w^{ST}$ where w^{ST} is the watermark in the transform domain and we have

$$x = s + \underbrace{w^{ST} \cdot \underline{t}}_{w}.$$

The attack is modelled as an AWGN during the transmission across the channel and finally, the decoder receives $y = s + w^{ST} \cdot \underline{t} + v$. Before decoding the message, we make transformation to the received signal y, which allows to have $y^{ST} = s^{ST} + w^{ST} + v^{ST}$, in order to extract the inserted message m which is spreading on the host signal.

Eggers [5] explains the good robustness of watermakring systems which use the ST with the following expression :

$$\mathrm{WNR}_\tau = \mathrm{WNR}_1 + 10\log_{10}\tau \tag{9}$$

where WNR_τ and WNR_1 are, respectively, the watermark to noise ratios after the spreading operation on τ samples and without transformation ($\tau = 1$). This expression is useful for the following reason. Let $\mathrm{WNR} = [-3, 5]$ dB be the desired working range (common case), then for a spreading factor of $\tau = 10$, the BER of the watermarking system combined with the ST is in fact the BER associated with higher WNRs belonging to the range $[7, 15]$ dB. So, using the ST pre-processing allows to multiply the WNR by a factor τ. The "dual" expression of (9) is

$$\mathrm{DWR} = \mathrm{DWR}_\tau + 10\log_{10}\tau \tag{10}$$

where DWR (*resp.* DWR_τ) is the Document to Watermark Ratio (*resp.* in the transformed domain). According to expression (10), we can note that the ST attenuates the watermark before embedded it. As the DWR is increased, this leads to a smaller KLD according to expression (6). So, a potential gain in security can be obtained.

3.2. Spread Transform Trellis Coded Quantization

We have seen that the ST scheme allows to increase the WNR. On the other hand, the TCQ scheme has a good robustness for high WNRs. So, it is natural to combine the two systems. This new quantization-based watermarking scheme is called the STTCQ. Fig. 5 shows the variation of the BER as function of the WNR . We can note that the STTCQ is more robust against AWGN attack than the SCS and the TCQ, the robustness of the STTCQ increases when the spreading factor τ increases. However, we can not increase the spreading factor indefinitely, because its value is constrained by the payload.

Regarding the security, we can remark that if the STTCQ schemes is

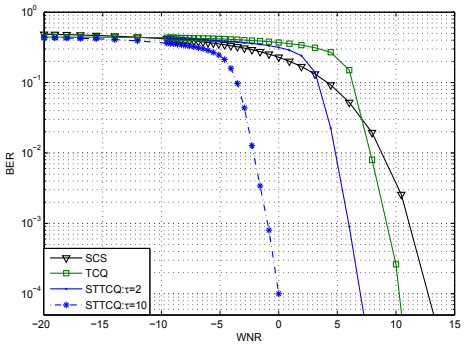

Fig. 5. Bit Error Rate (BER) as function of watermark to noise ratio (WNR) for SCS, TCQ and STTCQ watermarking for DWR = 13 dB.

used in the same range of DWR as the TCQ, the security level will be similar (see Fig.3). Fig. 6 shows that the compromise Robustness-Invisibility of the STTCQ is the best and this tradeoff is better for large spreading factor τ, we note also that Guillon et al. has not a good compromise Robustness-Invisbility which is due to its low level of the robustness.

Fig. 7 shows that for a DWR of 35 dB the watermark is imperceptible for all previous systems (SCS, TCQ, and STTCQ ($\tau = 10$)). In Fig. 7-f, the image watermarked with the STTCQ is attacked with an AWGN attack such as WNR = 2 dB, we can extract the message with zero errors. However, the message is not decoded correctly if we carry out the same experience with the other systems as it is shown in table 1. Fig. 8 shows a comparison between the effect of the SCS, the TCQ and the STTCQ watermarking on the statistics of a real image, we note that SCS is not secure in comparison with the TCQ and the STTCQ, also, in table 1 we can see the difference between the security level of the SCS and the two other watermarking systems.

	SCS	TCQ	STTCQ
BER	0.1730	0.3392	0
KLD	38.9903	0.0059	0.0058

Table 1. The bit error rate for : SCS, TCQ and STTCQ when the message is inserted in real image with size 480×640 and WNR = 2 dB.

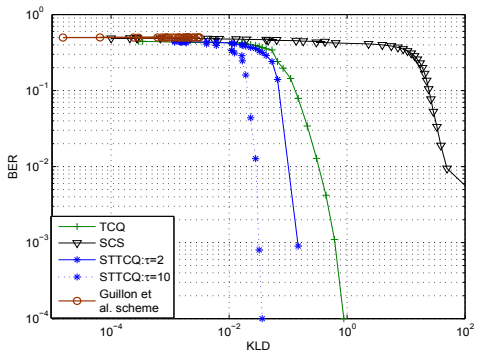

Fig. 6. BER Vs. KLD for the SCS, Guillon et al. scheme, TCQ and STTCQ watermarking schemes. Such as WNR in [-20,12] dB and DWR in [0,40] dB.

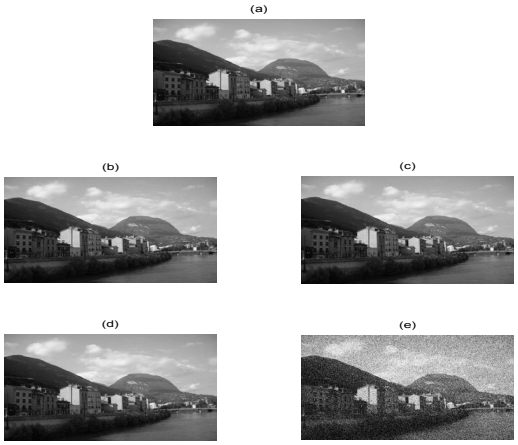

Fig. 7. Real image (figure (a)) with size 480×640 watermarked by : (b) SCS, (c) TCQ and (d) STTCQ ($\tau = 10$) system, such as the Document to watermark ratio (DWR) $= 35$ dB. Figure (e) represents a watermarked images with STTCQ and attacked with an AWGN attack such as WNR $= 2$ dB.

4. CONCLUSIONS

Steganography and robust watermarking are often animated by contradictory goals since for steganography, the prime goal is to have a high level of security, defined as the closeness of the PDF of the host and marked signals, often to the detriment of the robustness to an attack. For robust watermarking, this is precisely the inverse. So, it is interesting to design a watermarking scheme which is secure and robust. Toward this end, we have proposed a new quantization-based watermarking scheme based on the Trellis Coded Quantization on a spread transform domain which allows a good tradeoff between security and robustness with respect to other standard quantization-based watermarking schemes.

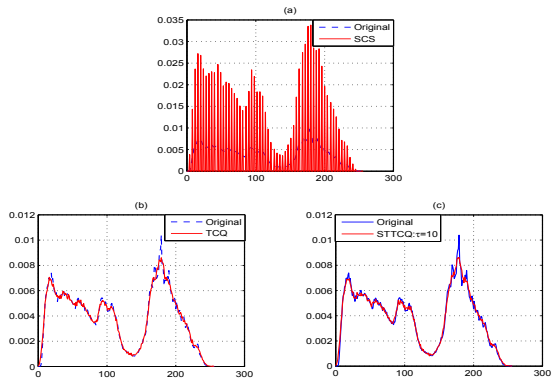

Fig. 8. Probability Density Function (PDF) of the real image with size 480×640 and of the watermarked signal for the (a) SCS, (b) TCQ and (c) STTCQ.

5. ACKNOWLEDGMENT

The authors would like to thank Professor Pierre Duhamel for his help and collaboration to this paper.

6. REFERENCES

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

[2] G. D. Forney, Jr., *The Viterbi algorithm*, Proc. IEEE, vol. 61, pp.268-278, Mar. 1973.

[3] G. Ungerboeck. *Channel Coding with Multilevel/Phase Signals*, IEEE Transaction on Information Theory, vol. IT-28 no. 1 pp. 55-67, janvier 1982.

[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. J. Eggers, R. Baüml, R. Tzchoppe and B. Girod. *Scalar Costa scheme for information embedding* , IEEE Trans. on Signal Processing, Apr. 2003.

[6] G. Le Guelvouit, A. Ould Bouya, J. Bourgeois, C. Delpha, and R. Boyer, *Analyse stéganographique du schéma scalaire de Costa*, Gresti 2007, 2007

[7] Max H. M. Costa. *Writing on dirty paper*, IEEE Transactions on Information Theory, vol. 29, no. 3, pp. 439-441, May 1983.

[8] M.W. Marcellin, T.R. Fischer. *Trellis coded quantization of memoryless and Gauss-Markov sources*. In IEEE Trans. on Com., Vol. 38, pp. 83-93 Jan 1990.

[9] P. Guillon, T. Furon, P. Duhamel, *Applied public-key steganography* , Proc. SPIE, San Jose, CA, 2002.

[10] S. Katzenbeisser, F.A.P Petitcolas. *Information Hiding Techniques for Steganography and Digital Watermarking* .