ANGLE QIM: A NOVEL WATERMARK EMBEDDING SCHEME ROBUST AGAINST AMPLITUDE SCALING DISTORTIONS

Fabrício Ourique, Vinicius Licks, Ramiro Jordan

The University of New Mexico
Electrical & Computer Engineering Dept
Albuquerque, USA, 87106
fourique,vlicks,rjordan@ece.unm.edu

Fernando Pérez-González

University of Vigo
Dept. Teoría de la Señal y
Comunicaciones, Vigo, Spain, 36200
fperez@tsc.uvigo.es

ABSTRACT

Quantization index modulation (QIM) watermarking has received a great deal of attention ever since the rediscovery of Costa's result on codes with host-interference rejecting properties. While such embedding scheme exhibit considerable improvement in watermark capacity over its earlier predecessors, (e.g. spread-spectrum), their fragility to even the simplest attacks soon became apparent. Among such attacks, *amplitude scaling* has received special attention. In this paper, we introduce a quantization scheme that is provably insensitive to amplitude scaling attacks, named Angle QIM (AQIM). Instead of embedding information by quantizing the amplitude of pixel values, AQIM works by quantizing the *angle* formed by the host-signal vector with the origin of a hyperspherical coordinate system. Hence, AQIM's invariance to amplitude scaling can be shown by construction. Experimental results are presented for the bit error rate performance of AQIM under additive white Gaussian noise attacks.

1. INTRODUCTION

The rediscovery of Costa's original results on dirty paper codes by Chen and Wornell, in 1999, marked the beginning of a new stage in watermarking research [1, 2]. The idea of using host-signal state information at the encoder side in order to guarantee host-interference rejection influenced the creation of embedding schemes based on the quantization of the original image, namely quantization index modulation (QIM) methods. In these schemes, the amplitudes of one single pixel or of a vector of pixels are quantized using one of a series of quantization lattices, chosen accordingly to the symbol to be embedded. While such methods exhibit a significant gain in terms of watermark capacity over known-host statistics schemes such as the spread-spectrum (SS), they were shown in turn to be easily defeated by even the simplest attacks. This limitation of pure quantization based embedding motivated the creation of hybrid schemes (e.g., quantized projection, QP [3]) that merged concepts from both SS and QIM to simultaneously increase robustness and capacity. This accounts for quantizing a diversity projection of the host signal, in a much similar way to what is done for spread transform dither modulation (STDM), proposed earlier by Chen and Wornell. While such different amends to Costa's original idea helped to mitigate the effects of attacks, at the same time they turned out to be suboptimal in terms of capacity, lying far away from the originally targeted achievable rate for the watermark channel modeled after the AWGN channel.

In this paper, we present a novel technique that is shown by construction to be insensitive to amplitude scaling, named Angle QIM (AQIM). Instead of embedding information by quantizing the amplitude of pixel values, AQIM works by quantizing the *angle* formed by the host-signal vector with respect to the origin of a hyperspherical coordinate system. We present a detailed description of this method by building upon a simple example in two dimensions in order to construct angle quantizers in arbitrarily higher dimensions. Finally, we present experimental results that evidence AQIM's bit error performance under additive white Gaussian noise attacks.

2. PRELIMINARIES

In this paper, we follow the usual watermarking notation, where: k is a secret key used during the embedding/decoding process; x are samples taken from the original image, which can be pixels, DCT coefficients, DWT coefficients, or any other transformed domain coefficients used for embedding; m is a message that needs to be transmitted to the receiver end; w is the watermark to be added to the original image samples; y is the watermarked image; n represents an additive noise source contaminating y and z is the possibly

This work has been partially funded by the Brazilian Ministry of Education under CAPES grant 1423-00/2, and by the Ibero American Science and Technology Education Consortium (ISTEC), under the Los Libertadores initiative.

attacked watermarked image received at the decoder. Based on z the decoder attempts to obtain an estimate of the embedded message, $\hat{m}$.

Additionally, the following definitions are in place: $x \in \mathbb{R}^L$ is a L-dimensional vector in hyperspace; m is a vector with dimension $R_m L$ chosen from the set of possible messages, with cardinality 2^{LR_m}, where R_m is the code rate. $w \in \mathbb{R}^L$ is a vector, and it is added element-by-element to w to generate y. In order to account for perceptual quality, i.e. to guarantee that the distortion level lays bellow visual perception, the norm of w needs to be bounded. For now, the watermark distortion is defined as

$$D_w = \frac{1}{L}\|w\|^2 = \frac{1}{L}\|y - x\|^2. \tag{1}$$

By means of this distortion measure, the Document-to-Watermark Ratio (DWR) is defined as

$$DWR = 10\log_{10}\frac{\sigma_x^2}{D_w}. \tag{2}$$

where σ_x^2 is the variance of the original signal.

Additionally, the distortion introduced by the channel needs to be quantified as well. Let the channel distortion be defined as

$$D_c = \frac{1}{L}\|n\|^2. \tag{3}$$

In the same fashion, the Watermark-to-Noise Ratio (WNR) is defined as

$$WNR = 10\log_{10}\frac{D_w}{D_c}. \tag{4}$$

Known-host state schemes such as QIM use prior knowledge about the host signal state in order to generate a watermark vector "in the direction" of x. One simple way to accomplish this without resorting to large randomly generated codebooks is to quantize the host signal with one among several distinct lattices in the set $\{\Lambda_0, \Lambda_1, \cdots, \Lambda_M\}$. The cardinality of this set is equal to the dimension of the M-ary symbol alphabet. In addition, there is a direct dependence between the choice of the lattice Λ_i and the i-th symbol belonging to the message vector $m = (m_1, m_2, \ldots, m_{LR_m})$. Having said that, we will restrict our attention to binary alphabets, such that the corresponding lattices are given by

$$\begin{aligned} \Lambda_0 &= 2\Delta\mathbb{Z} \\ \Lambda_1 &= 2\Delta\mathbb{Z} + \Delta \end{aligned} \tag{5}$$

where Δ is the quantization step chosen to be small enough in order to ensure perceptual quality. The watermarked signal is then given by

$$y = \mathcal{Q}_{m[i]}(x, \Delta) = x + w(x, m[i]). \tag{6}$$

where $m[i]$ is the i-th symbol belonging to the message vector $m = (m_1, m_2, \ldots, m_{LR_m})$. The resulting watermark signal w is then the quantization error $w(x, m[i]) = y - x = \mathcal{Q}_{m[i]}(x, \Delta) - x$.

At the decoder end, the estimated message symbol $\hat{m}[i]$ is chosen from the set $\{0, 1\}$ in a way that minimizes the distance between the extracted vector z and its respective centroid, i.e.

$$\hat{m}[i] = \arg\min \|z - \mathcal{Q}_{m[i]}(z, \Delta)\|^2,\ m[i] \in \{0, 1\} \tag{7}$$

2.1. QIM under amplitude scaling attack

Under an amplitude scaling attack the received signal z becomes

$$z = \beta y \tag{8}$$

where β is a scalar. As one may notice, this has the effect of scaling the watermark vector by β in such a way that it can be eventually moved away from its original quantization cell. This leads to a potentially devastating effect on the watermark decoder, since it might incur in systematic decoding failure. Obviously, the ideal condition for perfect reception is that $\beta = 1$, in which case there is no attack at all. Therefore, for $\beta \neq 1$ the encoder may lose its ability to correctly estimate the embedded message $\hat{m}$, depending on the existing relation among Δ (the quantization step) and β (the scaling factor). In fact, the probability Pe of committing a bit error upon decoding is given by

$$Pe = 1 - Pr\left\{y - \frac{\Delta}{2} < \beta y \leq y + \frac{\Delta}{2}\right\} \tag{9}$$

where y represents here one single sample of the watermarked image, which is scaled by β by assumption. As long as βy is between $y - \frac{\Delta}{2}$ and $y + \frac{\Delta}{2}$ the probability of error is zero. It is not difficult to see that the expression for the probability of error is

$$Pe = Pr\left\{|y| > \frac{\Delta}{2(\beta - 1)}\right\} \tag{10}$$

$$= 2Q\left\{\frac{\Delta}{2(\beta - 1)\sigma_y}\right\} \tag{11}$$

where σ_y^2 is the variance of the watermarked signal and $Q(\cdot)$ is the Q-function. From the equation above, the dependence between P_e and β for QIM becomes evident. Therefore, a small scale attack can lead to a large probability of error. We shall see in the next section that this is not the case for AQIM.

3. ANGLE QUANTIZATION INDEX MODULATION

In communications theory, there are some situations in which substantial performance improvement can be obtained by modulating the carrier's phase instead of amplitude. This is evidenced by the superior noise performance of FM over

AM techniques in analog communications and by the performance of PSK methods over multipath fading environments, for example. The relevance of phase modulation schemes to the data hiding problem under amplitude scaling attacks was first identified by Chen in [4], although this idea has not been pursued any further in connection with quantization to date.

Consider a point in the two dimensional Euclidean space. In QIM, this point would be quantized to the closest centroid of the lattice defined in 5. Instead of using Cartesian coordinates, let this point be represented by the tuple (r, θ) in polar coordinates. Then, let the *angle* θ be quantized to the nearest centroid associated with message symbol $m[i]$. These centroids, for a binary alphabet, are defined by the following lattices

$$\Lambda_{0\mathcal{A}} = 2\Delta_\theta \mathbb{Z} \qquad mod(2\pi) \tag{12}$$

$$\Lambda_{1\mathcal{A}} = 2\Delta_\theta \mathbb{Z} + \Delta_\theta \qquad mod(2\pi). \tag{13}$$

This modulation scheme, which goes by the name of Angle QIM (AQIM), can be easily extended to L-dimensions, where a point in the hyperplane is represented by its hyperspherical coordinates, as we shall see in a moment. In the following, we exemplify this idea in a simple two-dimensional case and in section 3.2 we generalize its construction to the general L-dimensional case.

3.1. 2-Dimensional case

Let $x_i \in \mathbb{R}$ for $i = 1, 2$ be two samples taken from an arbitrary domain of the original image. These samples belong to the set $\mathcal{S}_i$, where $\mathcal{S}_i \cap \mathcal{S}_j = \emptyset,\ \forall i \neq j,\ i, j \in \{1, 2, \cdots, LR_m\}$. The assignment of a pixel to the set $\mathcal{S}_i$ is made key dependent and resemble the interleaving process used in spread-spectrum watermarking. The two samples x_1, x_2 may be viewed as a point in a two dimensional plane. This point can be described by its polar coordinates representation (r, θ). For that end, the angle r and radius θ are given by (14) and (15) respectively, as indicated bellow

$$\theta = \arctan\left(\frac{x_2}{x_1}\right) \tag{14}$$

$$r = \sqrt{x_1^2 + x_2^2} \tag{15}$$

Then, the angle θ is quantized as follows

$$\theta^{\mathcal{Q}} = \mathcal{Q}_{m[i]}(\theta, \Delta_\theta) = \left\lfloor \frac{\theta + m[i]\Delta_\theta}{2\Delta_\theta} \right\rfloor 2\Delta_\theta + m[i]\Delta_\theta \tag{16}$$

where $m[i] \in \{0, 1\}$ is one of the LR_m bits necessary to represent message m, and Δ_θ is the size of the quantization step. Note that while the angle is quantized, the radius coordinate remains unchanged.

Now, converting $(r, \theta^{\mathcal{Q}})$ back to its Cartesian coordinate representation yields the new amplitude values for the pixels in the set $\mathcal{S}_i$, i.e. $y_1 = r\cos\theta$ and $y_2 = r\sin\theta$. This process is illustrated in Figure 1.

3.2. L-Dimension case

Let x be a vector in the L-dimensional hyperplane with Cartesian coordinates given by $(x_1, \cdots, x_L)$. Then, let x be represented in hyperspherical coordinates by its radius r and angle vector $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_{L-1})$. These quantities can be obtained from $(x_1, \cdots, x_L)$ as follows:

$$\theta_1 = \arctan\frac{x_2}{x_1} \tag{17}$$

$$\theta_i = \arctan\frac{x_{i+1}}{\left(\sum_{k=1}^{i} x_k^2\right)^{1/2}},\ \forall\ i = 2, 3, \cdots, L-1. \tag{18}$$

$$r = \left(\sum_{k=1}^{L} x_k^2\right)^{1/2} \tag{19}$$

Then, the quantization in (16) is applied to the components of $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_{L-1})$, where the appropriate lattice is chosen according to the value of $m[i]$.

Mapping the radius and the quantized angle vector back to its representation in Cartesian coordinates yields the watermarked pixels, as in the following

$$y_1 = r\prod_{k=1}^{L-1}\cos\theta_k \tag{20}$$

$$y_i = r\sin\theta_{i-1}\prod_{k=i}^{L-1}\cos\theta_k,\ \forall\ i = 2, 3, \cdots, L \tag{21}$$

3.3. Robustness of AQIM under amplitude scaling attacks

Intuitively speaking, when all the (Cartesian) coordinates of a vector y are scaled by the same arbitrary factor β, its angular coordinates $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_{L-1})$ in the equivalent hyperspherical representation do not change at all. The invariance of AQIM to amplitude scaling attacks follows directly from this observation.

In a more formal way, let y be scaled by an arbitrary constant β and let the receiver use $z = \beta y$ in order to obtain an estimate of the embedded message symbol $\hat{m}[i]$. Then, to compute the probability of incurring in a decoding error, we will assume without loss of generality that $m[i] = 0$. This assumption is valid for two reasons: (i) the symbols $m[i]$ are equally likely; (ii) the quantization lattice is uniform. Since the angle vector $\hat{\boldsymbol{\theta}}$ is estimated based on z, it follows that

$$\hat{\theta}_1 = \arctan\frac{z_2}{z_1} \tag{22}$$

$$\hat{\theta}_i = \arctan\frac{z_{i+1}}{\left(\sum_{k=1}^{i} z_k^2\right)^{1/2}},\ \forall\ i = 2, 3, \cdots, L-1. \tag{23}$$

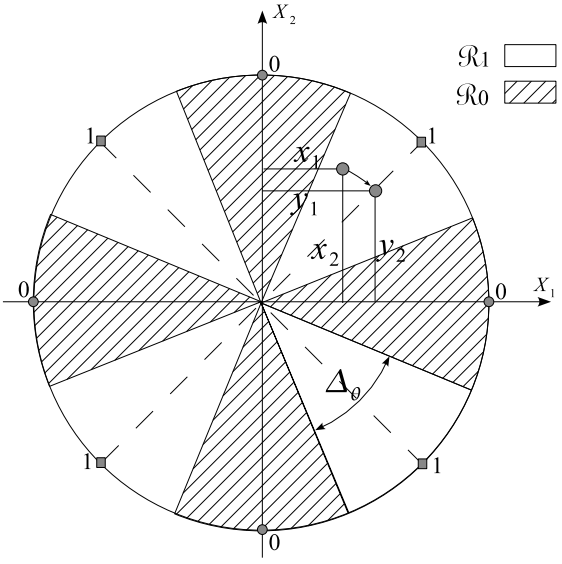

Fig. 1. Angle quantization index modulation for $L = 2$.

Noticing that $z = \beta y$, (i.e., each element $z_i = \beta y_i,\ \forall i = 1, \cdots, L$) and replacing it into (22), after some algebraic manipulation we have that

$$\hat{\theta}_1 = \arctan \frac{y_2}{y_1} \tag{24}$$

$$\hat{\theta}_i = \arctan \frac{y_{i+1}}{\left(\sum_{k=1}^{i} y_k^2\right)^{1/2}},\ \forall\, i = 2, 3, \cdots, L-1. \tag{25}$$

Therefore, for a pure amplitude scaling attack, $\hat{\theta} = \theta^{\mathcal{Q}}$ and

$$\begin{aligned} Pe &= Pr\left\{\hat{\theta} \in \mathcal{R}_1 / m = 0\right\} \\ &= Pr\left\{\theta^{\mathcal{Q}} - \frac{\Delta_\theta}{2} < \hat{\theta} \leq \theta^{\mathcal{Q}} + \frac{\Delta_\theta}{2}\right\} \\ &= 1 - Pr\left\{\left|\frac{\Delta_\theta}{2}\right| \leq 0\right\} \\ &= 0 \end{aligned}$$

In this way, we have shown that the AQIM method is insensitive to amplitude scaling attacks for any β, except of course for $\beta = 0$ when the watermarked image is completely erased.

3.4. Bit Error Rate for Additive White Gaussian Noise Channel

In addition to amplitude scaling robustness, AQIM's performance to AWGN attacks was assessed experimentally by means of Monte Carlo simulations. The family of curves shown in Figure 3.4 was obtained by arbitrarily varying the quantization step $\Delta_\theta = \pi/M$ for $M = 16, 32, 64, 128$, the dimensionality of the quantization lattices for $L = 2, 4, 8, 16, 32, 64$, and $WNR = 19, 21, 22dB$.

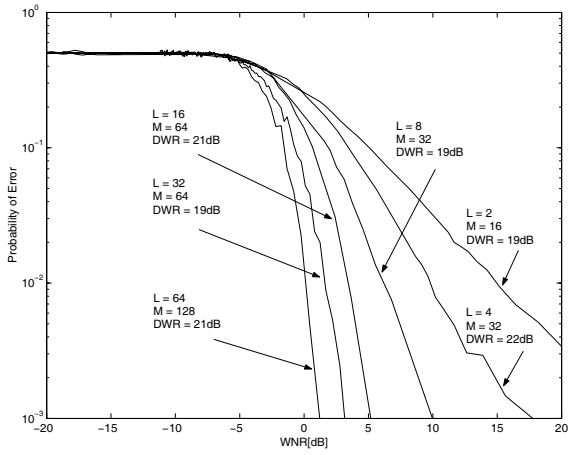

Fig. 2. Bit error rate performance of AQIM for an additive white Gaussian noise attack.

4. CONCLUSIONS

In this paper a new QIM scheme is presented, which is provably insensitive to amplitude scaling attacks. It is shown that AQIM is robust against any amplitude scaling parameter, except possibly for $\beta \neq 0$. This robustness problem was the main drawback of classical QIM based modulation, where even small scaling parameters could severely compromise correct message decoding. In addition to amplitude scaling robustness, the AQIM presented a performance under AWGN attack comparable to classical QIM. Work is underway to provide a rigorous analytical characterization of AQIM, including expressions and/or bounds for symbol error probability in L-dimensions.

5. REFERENCES

[1] Brian Chen, Gregory W. Wornell, "Provably Robust Digital Watermarking," *Proceeding of SPIE: Multimedia Systems and Applications II*, vol. 3845, no. 2, pp. 43–54, 1999.

[2] Max H. M. Costa, "Writing on Dirty Paper," *IEEE Transactions on Information Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.

[3] Fernando Prez-Gonzlez, Flix Balado, "Quantized Projection Data Hiding," *Proceeding of the IEEE International Conference on Image Processing, ICIP*, vol. 2, pp. 889–892, September 2002.

[4] Brian Chen, *Design and Analysis of Digital Watermarking, Information Embedding, and Data Hiding Systems*, PhD Thesis, Massachusetts Institute of Technology, Massachusetts, USA, June 2000.