1-1-2005

A new QIM-based image watermarking method and system

Wenming Lu
University of Wollongong, wl86@uow.edu.au

Wanqing Li
University of Wollongong, wanqing@uow.edu.au

Reihaneh Safavi-Naini
University of Wollongong, rei@uow.edu.au

Philip O. Ogunbona
University of Wollongong, philipo@uow.edu.au

A NEW QIM-BASED IMAGE WATERMARKING METHOD AND SYSTEM

WENMING LU, WANQING LI, REI SAFAVI-NAINI, PHILIP OGUNBONA

University of Wollongong, Australia
Email: wl86, wanqing, rei, philipo@uow.edu.au

Abstract:
Embedding capacity, distortion and resilience to attacks are three key indicators to the performance of any image watermarking systems. Study has shown that Quantization Index Modulation (QIM) can in general achieve higher embedding capacity than Spread Spectrum (SS) systems at the same level of distortion. However, QIM is more sensitive to simple attacks such as Additive White Gaussian Noise (AWGN), Uniform Distribution Noise (UDN), and JPEG compression. This paper proposed a system to decompose the host signal into the embedding signal that is designed to be resilient to attacks and the correction signal. QIM is applied on the embedding signal for improving robustness against attacks. A simple yet very effective implementation of the system, Local Average-based QIM (LAQIM) for image watermarking is developed. Theoretical analysis and experimental results have shown that LAQIM substantially reduce the error decoding rate against AWGN, UDN and JPEG compression at no extra cost of embedding distortion in comparison to conventional QIM.

1 Introduction

Image watermarking is a technique to hide data on a host image without noticeable distortion and address the security concerns on images, such as copyright protection and image authentication [1]. Many systems have been proposed in the past decade and most can be classified as Spread Spectrum (SS) based system. Recently, some Quantization Index Modulation (QIM) based systems are proposed.

A typical SS-based image watermarking system [2] is to add a secret pseudo-random sequence as the watermark to the colors/intensities or transform coefficients of a set of pixels chosen from the host image. The watermark is detected by correlating the corresponding colors/intensities or transform coefficients from the received image and the secret watermark. Since the correlation will be interfered by the image itself in the case that the original image is unknown to the decoder, SS is classified as a host interference non-rejecting system [3, 4].

On the other hand, QIM adapts watermark embedding by utilizing the information about the host image known to the encoder and therefore eliminates the interference by the host image in the process of decoding the message. This adaption makes QIM to be host interference rejecting and allows it to be able to embed more information in an image than SS. Although QIM potentially has a high embedding capacity, it suffers from a number of drawbacks. Typically, it is sensitive to simple attacks like Additive White Gaussian Noise (AWGN), Uniformly-Distributed Noise (UDN) [5], and fragile to amplitude (valumetric) scaling [7, 8]. When applied on images, QIM tends to be less robust against JPEG compression in comparison with SS [6].

Application of QIM to image watermarking requires 1) to construct from the images an appropriate host signal on which the watermark shall be embedded; and 2) to design a quantizer. Conventionally, the host signal is selected either from the spatial domain (*e.g.*, a group of pixels) or transform domains (*e.g.*, a set of DCT or wavelet coefficients) and a uniform quantizer is adopted.

Recently, Li *et al.* [7] proposed a non-uniform quantizer based on Watson't perceptual model to address the issue of amplitude scaling. Licks *et al.* [8] introduced an Angular QIM (AQIM) and presented an exact expression for the Bit Error Probability in AQIM watermarking under simultaneous amplitude scaling and AWGN attacks. However, how AQIM is going to behave under UDN and JPEG attacks was not reported. In fact, UDN tends to destroy the QIM-based watermark with much less energy than AWGN [5].

While the design of an appropriate quantizer is essential to QIM, construction of the embedding host signal is also critical to its robustness against various attacks. In the case where the target attacks are known or can be modelled, the embedding host signal should be constructed so as to be more resilient to the attacks. In this paper, we explore how such a host signal can be selected and constructed from the spatial domain of an image (We shall focus on spatial domain and the same concept can be extended to transform domain.) Specifically, we suggest to decompose the original signal selected from a group of pixels into the embedding signal and the correction signal. The decomposition is designed based on the nature of a set of target attacks. In the paper, we select AWGN, UDN with zero means and JPEG compression as our target attacks and proposed a simple yet effective Local Average based QIM (LAQIM) where local averages are used as the embedding signal. Theoretical analysis and experimental results have shown that LAQIM incurs no more embedding distortion in comparison to conventional QIM, but substantially reduce the error decoding rate against the chosen attacks.

The rest of the paper is organized as follows. Section II first present the diagram of the proposed system and then describes the general

principle on how to select and construct a host signal from image pixels for QIM such that QIM is more robust against certain attacks. Local-Average based QIM (LAQIM) method is then proposed and theoretical analysis on the properties of LAQIM with respect to its distortion and robustness against the noise attacks is given. In Section III, experimental results of LAQIM are presented and comparison between the LAQIM and the commonly used SS scheme is also made. Section IV concludes the paper with some remarks.

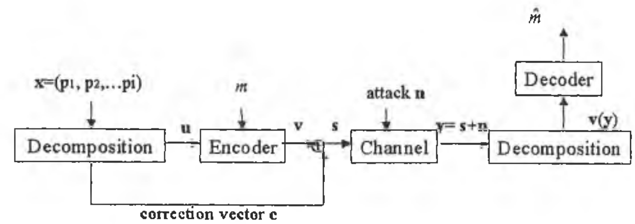


Figure 1: Proposed QIM-based image watermarking system with decomposition

2 Proposed QIM-Based Image Watermarking System

Figure 1 shows the block diagram of our proposed QIM-based image watermarking system with decomposition, where signal **x** is composed of a group of pixels and the block of "decomposition" is inserted into the conventional QIM embedding and decoding process.

2.1 Construction of the signal x

Unlike the correlation-based SS in which the embedding host signal must have certain dimension , the host signal of QIM can be as few as one dimension (*i.e.*, one pixel) to embed one bit. However, embedding a bit on a host signal representing a group of pixels strengthens the robustness, since the minimum distance, d_{min}, between different reconstruction points in the quantizer which decides the robustness, increases as the dimension of the host signal (*i.e.*, number of pixels) increases. However, there is a trade-off between the robustness and the embedding rate; the more pixels the host signal **x** contains, the stronger the robustness is; but the lower the embedding rate is.

In addition, there are many ways to select a group of pixels to form the host signal, including random selection, pixels from same row or column, or pixels from a square block. For AWGN and UDN attacks, it is expected that the robustness of QIM is independent of the selection of the pixels assuming each pixel is independently attacked by the noise. However, this is not true for JPEG attack due to the nature of block-based coding in JPEG. Matching the selection of pixels to the coding block has been proved to be effective [6]. In this paper, we consider the group of pixels from an 8x8 block as the host signal **x**. In other words, we consider a case to embed one bit into an 8×8 block.

2.2 Signal decomposition

It is expected that the formation and representation of the host signal on which the watermark is to be embedded is as essential as the design of the quantizer in QIM systems. This has been demonstrated by the work of [6], which compared the performance of QIM on the spatial and transform domains, and [8], which introduced the angular QIM. Specifically, when the potential attacks are known or can be modelled by known attacks, the host signal should be constructed in such a way that it is resilient to the attacks. Therefore, we propose to decompose the signal **x** into two components: **u**, an embedding signal that is resilient to the attacks and **c**, a correction signal. On the encoding side, messages are embedded into **u** to become **v** and the composite signal **s** is the combination of signal **v** and the correction signal **c**. On the decoding side, same decomposition is applied to extract the embedded signal possibly disrupted by the attacks, for decoding the embedded message.

The design of the decomposition system should be optimized to the potential attacks. Due to the properties of AWGN and UDN with zero means, and less loss in lower frequency components in JPEG, a local average based decomposition would be expected to be a good choice for those attacks. We denote this method as Local Average-based QIM (LAQIM).

2.3 LAQIM

The host signal **x**, which consists of a group of pixels, is partitioned into p blocks (vectors), $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p$ (the dimensions $L_{\mathbf{x}_i}$ of $\mathbf{x}_i$ need not all be identical). The average a_i of block $\mathbf{x}_i$ is calculated and all average values together construct a new p-dimensional vector $\mathbf{u}$ $(a_1, a_2, \ldots, a_p)$. For each $\mathbf{x}_i$, a correction vector $\mathbf{c}_i$ is produced as,

$$c_{ij} = x_{ij} - a_i, \quad \sum_{j}^{L_{\mathbf{x}_i}} c_{ij} = 0 \tag{1}$$

QIM is applied on the vector **u** and obtain the vector $\mathbf{v}(b_1, b_2, \ldots, b_p)$. The corresponding watermarked pixels $\mathbf{s}_i$ for $\mathbf{x}_i$ is produced as,

$$s_{ij} = c_{ij} + b_i, \; j = 1, 2, \ldots, L_{\mathbf{x}_i} \tag{2}$$

The set, $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_p$ constitute the watermarked pixels **s** for **x**. At the decoder, the received signal **y** is similarly decomposed and QIM decoding is applied on the vector of the averages of the blocks, $\mathbf{v}(\mathbf{y})$, as shown in Figure 1.

2.4 Properties of LAQIM

We present two properties of LAQIM, (i) the expected embedding distortion of LAQIM is only dependent on the quantization step size Δ, (ii) LAQIM reduces the decoding error rate against AWGN and UDN with zero mean.

Property 1. *The expected overall embedding distortion D_{emb} of LAQIM is only dependent on the quantization step size Δ.* *Proof*. For a block $\mathbf{x}_i$ and the corresponding watermarked block $\mathbf{s}_i$, the expected embedding distortion D_i is ,

$$E(D_i) = \frac{1}{L_{\mathbf{x}_i}} E(\sum (s_{ij} - x_{ij})^2,) j = 1, 2, \ldots, L_{\mathbf{x}_i} \tag{3}$$

From (1) and (2) we can write,

$$s_{ij} - x_{ij} = b_i - a_i \tag{4}$$

and upon substitution in (3) we have,

$$E(D_i) = E((a_i - b_i)^2) \tag{5}$$

Assuming that the average value a_i is uniformly distributed in the quantization interval, we have,

$$E((a_i - b_i)^2) = \frac{1}{\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} x^2 dx = \frac{\Delta^2}{12} \tag{6}$$

The expected embedding distortion D_i on $\mathbf{x}_i$ is $\frac{\Delta^2}{12}$. The expected overall distortion D_{emb} on $\mathbf{x}$ is

$$E(D_{emb}) = \frac{1}{L_{\mathbf{x}}} \sum_{i=1}^{p} (L_{\mathbf{x}_i} E(D_i)) = \frac{1}{L_{\mathbf{x}}} \sum_{i=1}^{p} (L_{\mathbf{x}_i}) \frac{\Delta^2}{12} = \frac{\Delta^2}{12} \tag{7}$$

E(D_{emb}) is only dependent on the quantization step size Δ. And, E(D_{emb}) is exactly same as that derived for QIM embedding [3, 4]. Hence, this property implies that LAQIM does not incur more embedding distortion.

Property 2. *When the watermarked image is distorted by the noise with zero mean, LAQIM can achieve more robustness than QIM.*
Proof. When QIM embedding is applied directly on $L_{\mathbf{x}}$-dimensional $\mathbf{x}$, the Signal-to-Noise Ratio SNR_{QIM}, defined by [3], is

$$SNR_{QIM} = \frac{d_{min}^2}{\sigma_n^2} = \frac{L_{\mathbf{x}}(\frac{\Delta^2}{4})}{\sigma_n^2} = \frac{L_{\mathbf{x}}\Delta^2}{4\sigma_n^2} \tag{8}$$

σ_n^2 is the energy of the noise $\mathbf{n}$ with zero mean;d_{min} measures the robustness to noise by the size of the quantization cells.

In LAQIM, embedding takes place in the reduced p-dimension space. But the change on the local average spreads over all dimensions. The property 1 also implies that the embedding strength is same in LAQIM and QIM. Therefore, the square of the minimum distance d_{min} is equal to that in QIM. However, the energy of the noise is reduced, let $\sigma_{n'}^2$ represent the energy of $\mathbf{n}$ in the reduced p-dimension space, and the SNR becomes

$$SNR_{LAQIM} = \frac{L_{\mathbf{x}}\Delta^2}{4\sigma_{n'}^2} \tag{9}$$

The improvement achieved by LAQIM over QIM can be measured by

$$\frac{SNR_{LAQIM}}{SNR_{QIM}} = \frac{\sigma_n^2}{\sigma_{n'}^2} \tag{10}$$

$\mathbf{n}$ is a zero mean noise and therefore, the energy of $\mathbf{n}$, σ_n^2, is, $\sum_{i=1}^{L_{\mathbf{x}}} n_i^2$; n_i is the i-th element of $\mathbf{n}$. In LAQIM, the reduced p-dimension noise vector $\mathbf{n}'$ is also with zero mean and its energy $\sigma_{n'}^2$ is $\sum_{i=1}^{p} n_i'^2$. Each element n_i' is an average of a group of elements in $\mathbf{n}$,

$$n'_i = \frac{1}{L_{\mathbf{x}_i}} \sum_{j=1}^{L_{\mathbf{x}_i}} n_j \tag{11}$$

Assume that each block contains same number of pixels(i.e. $L_{\mathbf{x}_i}$ is same for all blocks), we have

$$\sigma_{n'}^2 = \frac{1}{L_{\mathbf{x}_i}} \sigma_n^2 \tag{12}$$

Equation (10) becomes

$$\frac{SNR_{LAQIM}}{SNR_{QIM}} = L_{\mathbf{x}_i} \tag{13}$$

LAQIM is able to obtain improved performance over QIM by $L_{\mathbf{x}_i}$ when the noise is with zero mean. Actually, one can see LAQIM as a special case of QIM; when $L_{\mathbf{x}_i}$ is 1, the ratio becomes 1 and LAQIM degrades to QIM.

3 Experimental Results

A total of 15 gray-scale images, including the popular *lena*, *boat*, *banboo* and *pepper*, of varying sizes (from 512×512 to 1024×1024) are used in the experiments. Those images are carefully selected to cover a wide range of texture, contrast, edge strength and directions. From 15 images, there are totally 81920 8×8 blocks. Since we evaluate the robustness of the LAQIM in terms of embedding one bit into an 8×8 block (*i.e.*, fixed embedding rate for given images) against the targeted attacks and error decoding rate is calculated as average over all blocks from the 15 images. The characteristics of the blocks are more critical than that of images. We expected 81920 8×8 blocks would be large enough to evaluate the true performance of LAQIM.

Images are partitioned into 8×8 blocks and one bit is embedded into each block by respectively using

QIM No decomposition is applied, *i.e.*, $\mathbf{u} = \mathbf{x}$ and $\mathbf{c}$=0. This is equivalent to conventional QIM

LA1 The 8×8 block is further partitioned into 2×2 blocks and $\mathbf{u}$ is the averages of the 16 2×2 blocks

LA2 The 8×8 block is further partitioned into 4×4 blocks and $\mathbf{u}$ is the the averages of the 8 4×4 blocks

SS SS is applied to embed one bit into $\mathbf{x}$ using

$$\mathbf{s} = \mathbf{x} \pm \alpha \mathbf{w} \tag{14}$$

where + is for bit 1 and - is for bit 0; $\mathbf{w}$ is the watermark sequence drawn from the Normal distribution $N(0, 1)$; the strengthening parameter α is used to produce equivalent embedding distortion to QIM by letting $\alpha = \frac{\Delta}{2\sqrt{3}}$; the bit is decoded by the correlation coefficient $\mathscr{L}_{cc}$ between $\mathbf{w}$ and the received signal, bit 1 is decode if $\mathscr{L}_{cc} \geq 0$, and bit 0 is decoded otherwise.

Attacks are UDN with zero mean drawn from the interval $[-\xi, \xi]$, AWGN from $N(0, \sigma)$ and JPEG compression at different levels of quality. The error-decoding rate (Bit-Error Rate, *i.e.*, BER) is employed to measure the decoding performance; lower rate implies stronger robustness.

Images are watermarked with different embedding strengths by using different quantization step. In all cases, LAQIM consistently outperforms QIM, and matches or outperforms SS against the three attacks.

Figure 2 shows the results of four embedding experiments under the regime of (a)AWGN, (b) UDN, and (c) JPEG compression respectively; the quantization step size Δ is 10 in QIM and LAQIMs, and the strengthening parameter α is 2.88 in SS; all watermarked images have about the same PSNRs around 38.8dB. The embedded message randomly and uniformly consists of bit 0 and 1s.

In the noise-free case, QIM and LAQIMs have 0% BER while SS still has positive BER that implies lower embedding rate of SS because of the embedding failures on some blocks. As the noise level rises, SS outperforms QIM, but LAQIMs either matches or outperforms SS.

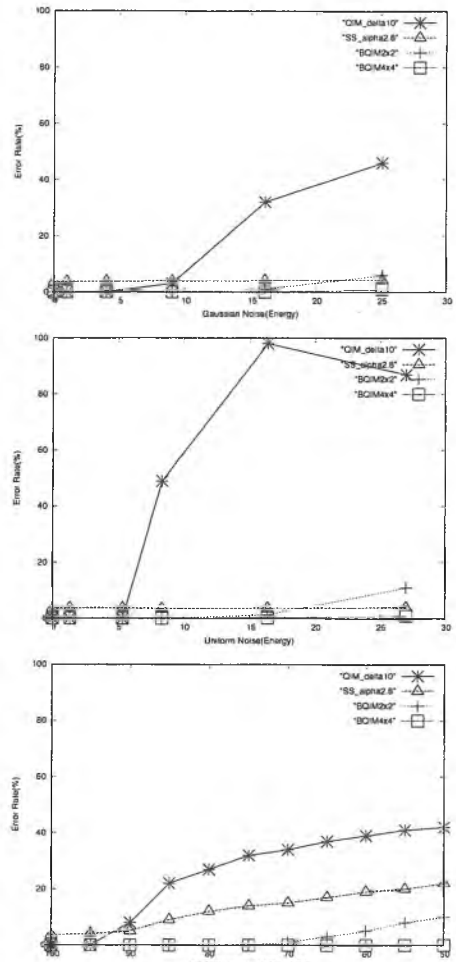
4 Discussion and Conclusion

LAQIM does not introduce more embedding distortion for the improved robustness. Both theoretical analysis and experiments have shown that the PSNR between the host and watermarked images are identical for QIM and LAQIMs.

Figure 2 also shows that QIM decoding is more sensitive to UDN than AWGN. When both noise levels (*i.e.*, variance) are around 10, error-decoding rates are 2.4% and 50% for AWGN and UDN respectively. This result is consistent with the the theoretical prediction [5] that UDN disrupts QIM more effectively than AWGN. In addition, QIM's error decoding rate reaches over 80% for UDN at levels above 15. However, LAQIM remains very low error decoding rate.

One disadvantage of LAQIM is that it may degrade the perceptual quality when the averages are taking from blocks larger than 4×4 although the PSNR remains high. Our experiment has shown that LAQIM with 2×2 does not produce any noticeable perceptual degradation. LAQIM with 4×4 still produces comparable quality images to QIM. In practice, this may be avoided by randomly averaging pixels from a 8×8 block rather than from 2×2 or 4×4 square blocks.

In all, this paper has demonstrated the importance of the construction of host signal when QIM is applied to image watermarking. The proposed LAQIM has been proved to be substantially more robust against AWGN, UDN and JPEG compression in comparison to conventional QIM. Although this paper is limited to the spatial domain, the method can be extended to transform domains as well.



(a)

(b)

(c)

Figure 2: QIM, LAQIMs (LA1-BQIM2×2 and LA2-BQIM4×4) and SS in face of Gaussian noise(a), Uniform noise(b): LA2 outperforms SS, while LA1 marginally outperforms SS except the noisiest case ; JPEG compression(c) both LAQIMs greatly outperform SS and QIM

References

[1] I.J Cox, M.L Miller, J.A Bloom, Watermarking Applications and their Properties, *Intl.Conf.on Information Technology,* pp6-10, Las Vegas, 2000

[2] I.J.Cox, J.Killian, T.Leighton, and T.Shamoon. Secure spread spectrum watermarking for multimedia . *IEEE Transactions on Image Processing*, 6(12):1673-1687, December 1997.

[3] B.Chen. Design and analysis of digital watermarking, information embedding, and data hiding systems. *Ph.D. Dissertation*, MIT, Cambridge, MA, June 2000.

[4] B.Chen and G.W.Wornell, Quantization index modulation: a class of provably good methods for digital watermarking and information embedding, *IEEE Trans. Information Theory*, vol.47,pp1423-1433, May 2001.

[5] F.Perez-Gonzalez, F.Balado, J.R. Hernandez, 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*, vol. 51, no. 4, April 2003.

[6] C.Fei, D.Kundar, and R.H.Kwong, Analysis and design of Watermarking algorithms for improved resistance to compression, *IEEE Transactions on Image Processing*, vol.,13, No.2, Febru-ary 2004.

[7] Q.Li, I.Cox, Using Perceptual Models to Improve Fidelity and Provide Invariance to Valumetric Scaling for Quantization Index Modulation Watermarking, ICASSP05, Philadelphia, PA, USA.

[8] V.Licks, R.Jordan, Prez-Gonzlez, An Exact Expression for the Bit Error Probability in Angle QIM Watermarking Under Simultaneous Amplitude Scaling and AWGN Attacks, ICASSP05, Philadelphia, PA, USA.