

# SET THEORETIC QUANTIZATION INDEX MODULATION WATERMARKING

Okta Altun, Gaurav Sharma, Mark Bocko

Department of Electrical and Computer Engineering  
 University of Rochester, Rochester, NY, 14627-0126  
 {altun,gsharma,bocko}@ece.rochester.edu

## ABSTRACT

We introduce a set theoretic framework for quantization index modulation (QIM) watermarking and illustrate its potency by designing a semi-fragile watermark that is both visually adaptive and tolerant to compression. We determine the watermarked image to satisfy the multiple constraints of watermark detectability, imperceptibility and robustness to compression using the method of projections onto convex sets (POCS). Mark embedding is performed through *implicit* quantization of statistical features, specifically the mean, of randomly selected pixel locations from the image. This is accomplished by defining a *detectability* constraint set that imposes the quantization constraint. We present experimental results demonstrating the efficacy of the technique in the presence of JPEG compression.

## 1. INTRODUCTION

Quantization index modulation (QIM) watermarking methods have attracted considerable attention since their introduction by Chen and Wornell [4] because of the superior capacity-distortion performance of these methods. Incorporation of additional constraints such as visual adaptability and robustness against signal processing operations, however, poses special challenges for these techniques in comparison to the spread spectrum watermarking methods introduced earlier. In this paper, we address this issue by extending the recently introduced set-theoretic watermarking framework [9, 10, 11] to QIM embedding. Set-theoretic watermarking allows the flexible incorporation of multiple requirements inherent in watermarking systems as feasibility constraints on the watermarked image. The embedding methods proposed within this framework thus far have addressed spread spectrum watermarks and have demonstrated the utility of the framework for multiple watermark insertion, linear transform domain watermark insertion, host-interference cancellation and visual adaptation. Here we introduce additional constraint sets and demonstrate that the same benefits can be extended to QIM watermarks.

Our watermark insertion randomly selects pixel locations within the image and quantizes the average of these pixel values to embed information. We deliberately choose spatial averages to exploit the mean preserving property of compression operations. The quantization is implicitly performed by *projecting* onto the set of images for which the mean matches the desired embedding value. The embedded watermark “noise” is similarly shaped implicitly by projecting on to suitable sets that define imperceptibility and robustness to compression. The method of successive projections

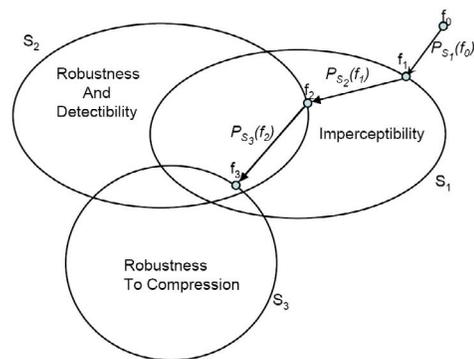
is utilized to obtain a solution that meets all constraints simultaneously. In prior work, information embedding in statistical features has been proposed for robust watermark insertion [13]. Compared to the optimization algorithm employed there for the quantization of image regions, our method exploits the power of the simple set-theoretic framework to employ more of the constraints and more detailed constraints that encapsulate the imperceptibility and robustness criteria.

## 2. OVERVIEW OF SET THEORETIC WATERMARKING AND POCS

Set theoretic watermarking [9, 10, 11] represents each property desired of the watermarked image as a *constraint set*. A watermarked image is then determined using an iterative algorithm that determines a point in the intersection of all the sets. [14] When the sets are convex the method of POCS provides a robust algorithm for this purpose. Given  $n$  convex sets  $\{S_i\}_{i=1}^n$  the POCS method determines a point in their intersection in the limit of projection sequence. If the intersection set is non-empty, the sequence  $\{f_k\}_{k=0}^{\infty}$  generated by successive (relaxed) projections onto the sets converges to a point in the intersection, where

$$f_{k+1} = (T_{S_n}(T_{S_{n-1}} \dots T_{S_1}(f_k) \dots)), k = 0, 1, \dots \quad (1)$$

$T_{S_i} = (1 - \lambda_{S_i})I + \lambda_{S_i}P_{S_i}$ ,  $0 < \lambda_i < 2$  is the relaxed projection operator onto set  $S_i$ . For unity relaxation  $T_{S_i}(f_k)$  will be equal to  $P_{S_i}(f_k)$ , which is set to be in the rest of this paper.



**Fig. 1.** A generic illustration of semi-fragile watermarking by POCS. [10]

This work was supported by the Air Force Research Laboratory/IFEC under grant number F30602-02-1-0129.

Figure 1 illustrates the general notion of watermark insertion by POCS. The domain is a Euclidean space with dimensions of the cover file where watermark is inserted. Successive projections onto detectability, imperceptibility and robustness to compression provides image adaptive semi-fragile watermark embedding.

### 3. CONSTRAINTS FOR QIM WATERMARK EMBEDDING

QIM is an efficient high capacity watermarking method [4]. In its general form, QIM embeds a  $M$ -ary message by quantizing the signal (or a suitable statistics of the signal) using one of  $M$  quantizers. Dithering by a key-dependent pseudo-random sequence chosen from a uniform distribution improves the security and efficiency of the system further.

Various modifications have been proposed for QIM based watermarking algorithms. Most recently an adaptive QIM scheme is proposed that adapts the quantization bins to the visual complexity of the image [15]. The adaptation of the bins, however, trades-off robustness against noise particularly in smooth image regions.

Here we propose another method for image adaptive QIM watermark embedding using the set theoretic watermarking framework [9, 10, 11]. The watermarking is implicitly performed by determining the watermarked image as a feasible point meeting detectability constraints and fidelity constraints, the latter being formulated as both pixel-wise and overall constraints. Successive projections on to convex sets (POCS) is utilized for this purpose. The individual components of the quantized metric adapt according to the constraints imposed by the fidelity criteria.

A general framework for set theoretic watermarking has already been proposed [9, 10, 11], where the framework was demonstrated with spread-spectrum (SS) embedding. In this paper, we extend the method to QIM. The current work differs from set-theoretic spread-spectrum watermark embedding in the watermark detectability set and robustness to compression set. We define these sets for QIM below and provide the associated projection operation operators. The remaining constraints are same as spread-spectrum scenarios explained in [11]. Readers are referred to these references for details on the projections for these sets.

#### 3.1. Detectability and robustness

If  $X^* \in R^{M \times N}$  denotes the image with dimensions  $M \times N$ , we denote  $X = \text{vec}(X^*) \in R^{MN}$  as the vectors obtained by stacking together the columns of each. We will adopt the notation in terms of 1-dimensional vectors throughout and assume that any image operators are also represented as matrices/functions conformal with the vector representation.

We randomly choose  $L$  locations within the image to form a sample vector  $Y$  for each of  $N$  bits. For every bit embedding, the mean of the selected locations are quantized. Let  $Y_0$  denote the value of the vector  $Y$  in the original image. We use subtractive dithered scalar quantizer to quantize the mean values  $\mu_i$  to obtain  $\mu_i^q$ :

$$\mu_i = \frac{1}{L} \mathbf{1}^T Y_0 \quad (2)$$

where  $\mathbf{1}$  is an  $L \times 1$  vector of 1's.

$$\mu_i^q = Q(\mu_i + d_i + b_i \Delta, \Delta) - d_i \quad i = 1, \dots, N \quad (3)$$

Here  $b_i$  represents the message bit to be embedded,  $Q$  is the integer scalar quantizer with  $\Delta$  as scaling parameter and  $d_i$  is pseudo-random scalar dither which is assumed iid uniformly-distributed between  $-\Delta/2$  and  $\Delta/2$ .

The embedding process outlined above is readily incorporated into a set-theoretic watermarking method using a *Quantization-based detectability* set defined as:

$$S_1^i \equiv \{Y : \frac{1}{L} \mathbf{1}^T Y = \mu_i^q\} \quad i = 1, \dots, N \quad (4)$$

The definition of (4) assumes watermark embedding in the spatial domain. However, method is generic in nature and can be applied to most transform domains for possible attractive properties of them. For any linear transform domain, the convexity will be preserved and a similar set can be defined in that context.

Projection of  $Y$  onto  $S_1$  is given by

$$P_1(Y) = \arg \min_{Z \in S_1^i} \|Z - Y\|^2$$

The Lagrangian [16] for this constrained optimization problem can be written as:

$$L = \|Z - Y\|^2 + \lambda \left( \frac{1}{L} \mathbf{1}^T Y - \mu_i^q \right) \quad (5)$$

The Lagrange parameter is readily shown to be

$$\lambda = \begin{cases} 2Y^T Y - 2\mu_i^q L & \frac{1}{L} \mathbf{1}^T Y \neq \mu_i^q \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

The projection can then be expressed in terms of the Lagrange parameter as

$$P_1(Y) = Y - \frac{\lambda}{2L} \quad (7)$$

Thus the projection operation modifies each of the chosen pixel locations by  $\lambda/2L$ . However, note that in the overall algorithm this represents a single step and the imperceptibility set (to be defined) will shape the watermark according to the other constraint sets (e.g. visual and compression). Since compression is a mean preserving operation, some robustness to compression is assured as a by-product of the embedding in the mean of a selected spatial region. Additionally, we incorporate robustness to compression as a set in Section 3.3.

#### 3.2. Watermark imperceptibility

An important characteristic of watermarks is their imperceptibility within the cover. The watermarked image should look identical (or similar) to the original image. Set theoretic framework allows us to employ multiple visual models at the same instance. We incorporate the visual fidelity requirement of image through two constraints: a frequency domain based overall visual fidelity constraint captured by a linear visual system model and a pixel-wise image fidelity constraint determined in terms of a noise visibility function.

*Overall Image Fidelity Constraint:*

Human observers are more sensitive to changes at low spatial frequencies than high frequencies. We can approximate this behavior by assuming that the perceived image by assuming that the *perceived image* in response to an input image  $X$  is given by

$HX$  where  $H$  represents a 2-D spatial low-pass filter<sup>1</sup>. Euclidean distance between the perceived watermarked image  $HX$  and the perception of the original image  $HX_0$  should be small (where  $X_0$  denotes the original image). Using the Euclidean norm for quantifying the difference, the constraint therefore becomes,

$$S_2 \equiv \{X : \|HX - HX_0\| \leq \theta\} \quad (8)$$

where  $\|v\|$  represents the Euclidean norm of  $v$  and  $\theta$  is a suitably chosen threshold. The spatial filter  $H$  is determined by the specific visual system model employed. In particular, we employ the model proposed by Mannos et al [19], which has been extensively utilized in image processing research [11].

*Pixel-wise Image Fidelity Constraint:*

Since the overall fidelity constraint (8) is based primarily on psychophysical data for individual sinusoidal stimuli, it does not adequately handle localized perturbations of the image in a small area. Therefore, we use an additional model to limit local perturbations to ensure imperceptibility. Our model exploits the perceptual phenomenon of spatial masking [20] in which perturbations introduced in an image region at a frequency are masked by stronger image content at similar frequencies. In particular, we use the spatial domain texture masking model proposed by Pereira et al [17]. Given an original image, the model predicts the allowable distortion at each pixel level that is visually tolerable, leading in turn to pixel-wise upper and lower bounds for the difference from the original image. The resulting constraint can be expressed as:

$$S_3 \equiv \{X : D_L \leq (X - X_0) \leq D_U\} \quad (9)$$

where  $D_U$  and  $D_L$  are the same size as  $X$  and represent the pixel-wise upper and lower bounds on the distortion,  $X_0$  is the original image, and the inequalities in (9) apply term-wise. The constraint can alternately be expressed as  $S_2 = \{X : U \leq X \leq L\}$  where  $U = X_0 + D_U$  and  $L = X_0 + D_L$  form for pixel wise upper and lower bounds [11].

### 3.3. Robustness to compression constraint

Motivated by the observation that typical transform coding schemes provide coding gain through the compaction of signal energy into a few coefficients, we approximate the robustness to compression set by the following :

$$\widehat{S}_4 \equiv \{X : \overline{S(\mathcal{T}_I(Q_0[\mathcal{T}_F(X)]))} = \mu_i^q\} \quad (10)$$

where  $\mathcal{T}_F$  denotes the operator that transforms from the spatial domain to the transform domain (e.g. DCT for JPEG),  $\mathcal{T}_I$  denotes the corresponding inverse transform (e.g. IDCT for JPEG),  $S$  refers to the random selection matrix such that  $Y = SX$ , overline refers to mean of the variable and  $Q_0[\ ]$  refers to the ‘‘quantizer’’ determined from the original image  $X_0$  by defining its constituent scalar quantizers as

$$Q_0^k[t] = \begin{cases} 0 & \text{if } Q^k[(\mathcal{T}_F(X_0))_k] = 0 \\ t & \text{otherwise} \end{cases} ; \quad k = 0, \dots, MN - 1. \quad (11)$$

where  $(\mathcal{T}_F(X_0))_k$  denotes the  $k^{th}$  transform coefficient of the original image  $X_0$ . Thus the quantizer  $Q_0[\ ]$  sets the transform

<sup>1</sup>The visual system also includes a significant point-wise non-linearity. Common digital image representations, however, already include a compensation for this nonlinearity [18] and its effect can therefore be ignored with minimal error.

coefficients that are zero in  $Q[\mathcal{T}_F(X_0)]$  to zero and leaves other coefficients unchanged. This approximation has the underlying assumption that the transform coefficients that is quantized to zero cause the major loss of watermark information. We can readily see that for our definition of  $Q_0[\ ]$ , we have  $Q_0(Z + Y) = Q_0(Z) + Q_0(Y)$ . Hence from the linearity of the transformation  $\mathcal{T}_F$  the convexity of the set follows immediately.

## 4. EXPERIMENTAL RESULTS

We use the USC image set to illustrate the performance of the set-theoretic QIM watermarking scheme. The set consists of eight 8-bit gray-scale image of size  $512 \times 512$  pixels. For our set theoretic watermarking method, the parameters for the constraint sets are set as follows. The quantization parameter  $\Delta$  is chosen to be 4 and sample size  $L$  for mean computation is chosen as 100. The bound for the overall image fidelity threshold [19] is set at  $\theta = 10$  and values of  $P_0 = 30$  and  $P_1 = 3$  are used for the pixel-wise image fidelity parameters [17]. For our examples, we embedded 1000 bits, each utilizing QIM of the mean of  $L = 100$  independently and randomly selected locations.

The watermarked images are obtained by using the iterative POCS method to satisfy all the imposed constraints. The visual quality of the watermarked images obtained is illustrated using the *Peppers* image in figures 2(a) and 2(b), where the former represents the original image and the latter the image with the 100 bits embedded using the proposed technique. It can be seen that the watermarked image maintains a high degree of visual fidelity to the original as required by the imperceptibility constraints.



(a) Original peppers image



(b) Watermarked peppers image

**Table 1.** Summary of watermark detection results across 8 images in the USC Image database at various JPEG compression levels. In each image, 1000 bits are embedded and total number of bit errors (out of 8000) are reported.

Samp. $L$	$Q = 90$	$Q = 80$	$Q = 70$	$Q = 60$	$Q = 50$	$Q = 40$	$Q = 30$	$Q = 20$	$Q = 10$
10	232	842	1182	1431	1648	1914	2347	2840	3750
50	0	26	66	133	218	313	519	984	2345
100	0	2	8	24	49	88	166	434	1541

For the parameters and the image from the USC database, the POCS iterations for the watermark embedding converge to a feasible solution in typically 200 iterations. This takes approximately 40 minutes on a Pentium M 1.70 GHz machine with a Matlab implementation.

Table 1 illustrates the performance of the watermark in the presence of compression. Note that the method gains robustness against compression both from the embedding within the mean and from the robustness to compression constraint set. In the absence of the constraint set, significantly poorer performance is obtained.

## 5. CONCLUSION

In this paper we extended the set theoretic watermarking framework [9, 10, 11] to semi-fragile QIM based watermarking. We introduced new constraint sets for detectability of QIM and robustness to compression. Combining these with constraint sets for imperceptibility defined in earlier work, we formulated QIM watermark embedding as a feasibility problem with multiple constraints and developed an algorithm for watermark embedding using the method of projections onto convex sets (POCS). Specifically, the watermarking scheme automatically and implicitly handles quantization of statistics of the image, ensures visual fidelity, and provides robustness against mild JPEG compression. The technique further illustrates the power and flexibility of the set-theoretic watermarking framework.

## 6. REFERENCES

- [1] A. Menezes, P. van Oorschot, and S. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Florida, USA, 1997.
- [2] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data-embedding and watermarking technologies," *Proceedings of the IEEE*, vol. 86, no. 6, pp. 1064–1087, June 1998.
- [3] I. J. Cox, J. Killian, F. T. Leighton, T. Shamoan, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol.6, pp. 1673–1687, Dec. 1997.
- [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, no. 4, pp. 1423–1443, May. 2001.
- [5] M. Yeung and F. Mintzer, "Invisible watermarking for image verification," in *Journal of Electronic imaging*, Jul 1998, vol. 7, no. 6, pp. 578–591.
- [6] P. W. Wong, "A public key watermark for image verification and authentication," in *Proceedings of IEEE International Conference on Image Processing*, Chicago, USA, October 4–7, 1998, pp. 425–429.
- [7] D. Kundur and D. Hatzinakos, "Digital watermarking for telltale tamper proofing and authentication," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1167–1180, July 1999.
- [8] J. Eggers and B. Girod, "Blind watermarking applied to image authentication," in *Proceedings of IEEE ICASSP*, Salt Lake City, Utah, USA, May 2001.
- [9] O. Altun, G. Sharma and M. Bocko, "Informed Watermark Embedding In Fractional Fourier Domains," *EUSIPCO*, Antalya, Turkey, Sep.2005.
- [10] O. Altun, G. Sharma, M. Celik and M. Bocko, "Semifragile Hierarchical Watermarking In A Set Theoretic Framework," *ICIP*, Genoa, Italy, Sep.2005.
- [11] O. Altun, G. Sharma, M. Celik and M. Bocko, "Set Theoretical Watermarking And Its Application To Semi-fragile Tamper Detection" *Submitted to IEEE Trans. Information Forensics And Security*, August, 2005.
- [12] T. Liu and P. Moulin, "Error exponents for one-bit watermarking," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Hong Kong, April, 2003.
- [13] M. K. Mihcak, R. Venkatesan and M. Kesal, "Optimization Algorithms for Quantizing Randomized Statistics of Image Regions," in *Proceedings of the Fortieth Annual Allerton Conference on Communication, Control and Computing*, Monticello, IL, Oct. 2002.
- [14] P. L. Combettes, "The foundations of set theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 182–208, Feb. 1993.
- [15] Q. Li and I. Cox, "Using Perceptual Models To Improve Fidelity And Provide Invariance To Valumetric Scaling For Quantization Index Modulation Watermarking" *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing*, Philadelphia, Apr. 2005.
- [16] D. G. Luenberger *Linear and Nonlinear Programming*, 2nd. Ed., Addison Wesley, Reading, MA, 1989.
- [17] S. Pereira, S. Voloshynskiy, and T. Pun "Optimal transform domain watermark embedding via linear programming," *Signal Processing*, vol. 81, no. 6, pp. 1251–1260, Jun. 2001.
- [18] G. Sharma, "Color Fundamentals for Digital Imaging," Chapter 1 in *Digital Color Imaging Handbook*, CRC Press, Boca Raton, FL, 2003.
- [19] J. L. Mannos and D. L. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," *IEEE Transactions on Information Theory*, vol. 20, no. 4, pp. 525–536, Jul. 1974.
- [20] A. B. Watson *Digital Images and Human Vision*, MIT Press, 1993.