

Spread Transform Dither Modulation using a Perceptual Model

Qiao Li, Gwenaël Doërr and Ingemar J. Cox

University College London Adastral Park Postgraduate Campus

Email: {q.li,g.doerr}@adastral.ucl.ac.uk , ingemar@ieee.org

Abstract—In previous work, we demonstrated how perceptual modeling can be applied to dither modulated quantization index modulation and rational dither modulation, to improve both robustness and fidelity. These algorithms were shown to be significantly more robust to valumetric scaling. However, they, and their predecessors, remain extremely sensitive to re-quantization which commonly occurs due to JPEG compression, numerical rounding and analog-to-digital conversion. It is well known that spread transform dither modulation (STDM) is more robust to re-quantization. In this paper we describe how to incorporate a perceptual model into this framework and present two algorithms based on Watson’s perceptual model. Experimental results of robustness to JPEG compression are reported for 1000 images at embedding rates of 1/32 and 1/320. At the high embedding rate, the robustness of the two algorithms is the same as STDM but the perceptual distortion is reduced from 23 to about 4, based on Watson’s perceptual distance. At the lower embedding rate, we simultaneously observed superior robustness to STDM as well as improved fidelity. If the perceptual distance rather than the document-to-watermark ratio (DWR) is held fixed, then the two adaptive spread transform methods exhibit significant improvements in robustness to JPEG compression.

I. INTRODUCTION

Quantization index modulation (QIM) is a popular form of digital watermarking based on the framework of communications with side information [1]. In their original paper, Chen and Wornell [2] described a number of variants of the basic QIM algorithm, namely dither modulation QIM (DM), distortion compensated dither modulation (DC-DM) and spread transform dither modulation (ST-DM).

The popularity of QIM is, in part, due to its ease of implementation, computational flexibility and amenability to theoretical analysis. Nevertheless, there are practical limitations of the approach due to its extreme sensitivity to valumetric scaling and re-quantization. Valumetric scaling is a very common signal processing operation and occurs whenever the volume of an audio signal or the brightness of an image is changed. Re-quantization is also commonly occurring, for example, when a signal undergoes lossy compression or numerical rounding.

The problem of valumetric scaling has received widespread attention and a number of solutions have been proposed [3]–[8]. In contrast, there has been only limited research focused on the issue of re-quantization, among which JPEG compression is a typical one.

Fei *et al.* [9] analyzed the performance of two popular classes of watermark embedding techniques, spread spectrum

watermarking and quantization-based embedding, in the presence of JPEG compression. They also proposed a hybrid watermarking scheme to exploit the theoretically predicted advantages of spread spectrum and quantization-based watermarking to achieve superior performance. In contrast, this paper is focused on improving the fidelity and/or robustness of ST-DM.

Pérez-González *et al.* [10] examined the performance of Distortion Compensated Dither Modulation (DC-DM) against JPEG compression and proposed a new method for detection based on a weighted Euclidean distance. Experimental results demonstrated improved performance over traditional DC-DM. However, here is no comparison with ST-DM and it remains unclear whether this method is superior to ST-DM.

In this paper, we describe preliminary work to introduce a perceptual model within the ST-DM framework. Section II provides a brief introduction to quantization index modulation and particularly ST-DM and provide experimental results demonstrating the sensitivity to re-quantization and the relative robustness of ST-DM. Section III then describes how the projection vector used in ST-DM can be determined so as to minimize the perceptual distortion. The experimental results of Section IV show that for a document-to-watermark ratio (DWR) of 35 dB, the perceptual distortion as measured by Watson’s distance [11] is reduced from 23 to as little as 4, while the bit error rate (BER) is the same or better. However, if the perceptual distance rather than DWR is held fixed, then the new algorithm demonstrates a very significant improvement in BER. Section V summarizes our results and describes directions for future work.

II. SPREAD TRANSFORM DITHER MODULATION

The basic quantization index modulation (QIM) algorithm quantizes each signal sample, x , using a quantizer, $Q(\cdot)$, that is chosen from a family of quantizers based on the message bit, m , that is to be embedded. The watermarked signal sample, y is given by:

$$y = Q(x, \Delta, m, \delta), \quad m \in \{0, 1\} \quad (1)$$

where Δ is a fixed quantization step size and δ a random dither. The quantizer $Q(\cdot)$ is defined as follows:

$$Q(x, \Delta, m, \delta) = \Delta \cdot \text{Round} \left(\frac{x - \delta - m \frac{\Delta}{2}}{\Delta} \right) + \delta + m \frac{\Delta}{2} \quad (2)$$

At the detector, the received signal sample, z , a corrupted version of y , is re-quantized using the family of quantizers to determine the embedded message bit, i.e.

$$\hat{m} = \arg \min_{b \in \{0,1\}} |z - Q(z, \Delta, b, \delta)| \quad (3)$$

Note that the re-quantization at the detector is *not* a source of noise and does *not* refer to the re-quantization due to say JPEG compression, that we will discuss shortly.

Equations (1) and (3) assumed that one message bit is embedded in one sample. To improve robustness, it is common to embed the same message bit across several input signal samples $\{x_1, \dots, x_N\}$. The detection equation then becomes:

$$\hat{m} = \arg \min_{b \in \{0,1\}} \sum_{i=1}^N |z_i - Q(z_i, \Delta, b, \delta)| \quad (4)$$

where we have assumed the use of a soft decoder.

Note that a dither signal is used because (i) it improves the perceptual fidelity of the watermarked signal, (ii) the quantization noise can then be modeled as independent from the cover signal and (iii) the pseudo-random noise can be considered a key, without which, detection is not possible.

Both QIM and DM are very sensitive to re-quantization. Figure 1 illustrates this point for DM. Here, robustness to JPEG compression is examined for DM in the discrete cosine transform (DCT) domain, i.e. we quantize the DCT coefficients rather than the pixel value.¹ Two embedding rates are depicted of 1/32 and 1/320, i.e. embedding 2 bits per 8×8 block and 1 bit per 5 blocks. Testing was performed on 1000 images. If we arbitrarily consider when the BER exceeds 20%, this occurs at a JPEG quality factor of about 96%. Even when the embedding rate is reduced by a factor of 10, the BER exceeds 20% for a quality factor of about 92%. These quality factors are extremely high and in many scenarios much better robustness may be needed.

A. Adaptive QIM

QIM and DM are also very sensitive to volumetric distortion. A number of algorithms have recently been proposed to counter this [3]–[8], specifically rational dither modulation (RDM) [4], adaptive QIM using a modified Watson distance (QIM-MW) [5] and adaptive RDM using a modified Watson distance (RDM-MW) [7]. These latter methods are based on adaptively changing the quantization step size.

Since the step size varies in these systems, we had hoped that they would exhibit some improved robustness to re-quantization. However, Figure 2, which looks at the robustness of RDM-MW to JPEG compression reveals that, perhaps surprisingly, it is actually slightly less robust.

B. Spread transform Dither Modulation

Figure 3 illustrates the basic framework for spread transform QIM.

ST-DM differs from regular QIM in that the signal, \mathbf{x} is first projected onto a randomly generated vector, \mathbf{u} , and the

¹Similar performance was observed in the pixel domain.

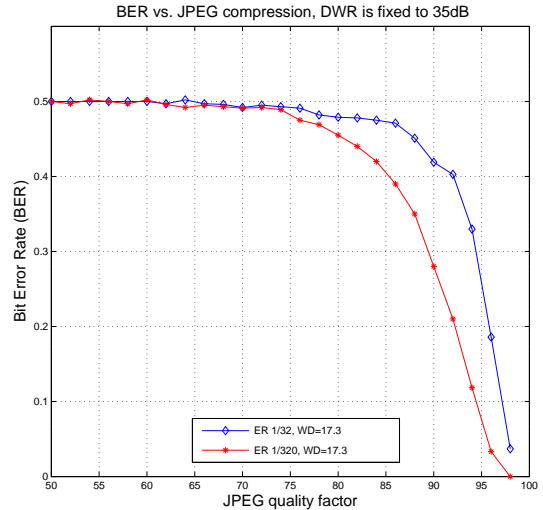


Fig. 1. Bit error rate (BER) as a function of JPEG compression for DM with embedding rates of 1/32 and 1/320 and fixed DWR of 35 dB

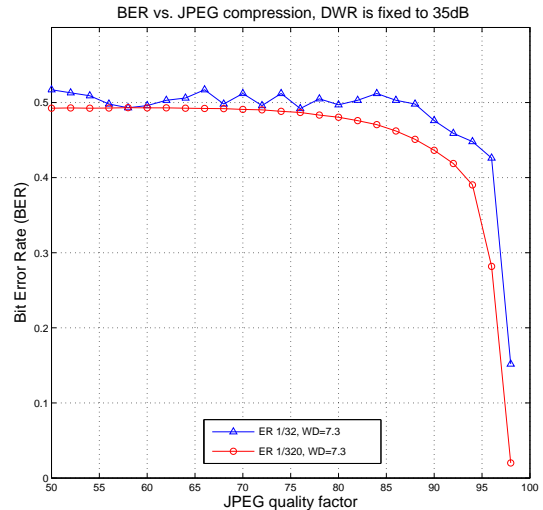


Fig. 2. Bit error rate (BER) as a function of JPEG compression for RDM-MW with embedding rates of 1/32 and 1/320 and a fixed DWR of 35 dB

resulting scalar value is then quantized before being added to the components of the signal that are orthogonal to \mathbf{u} . The equation for embedding is thus:

$$\mathbf{y} = \mathbf{x} + (Q(\mathbf{x}^T \mathbf{u}, \Delta, m, \delta) - \mathbf{x}^T \mathbf{u}) \mathbf{u}, \quad m \in \{0, 1\} \quad (5)$$

and the corresponding detection is given by:

$$\hat{m} = \arg \min_{b \in \{0,1\}} |z^T \mathbf{u} - Q(z^T \mathbf{u}, \Delta, b, \delta)| \quad (6)$$

III. SPREAD TRANSFORM QIM WITH PERCEPTUAL MODELING

From Equation (5) we see that the change to the signal \mathbf{x} is in the direction of the random vector \mathbf{u} , and the magnitude

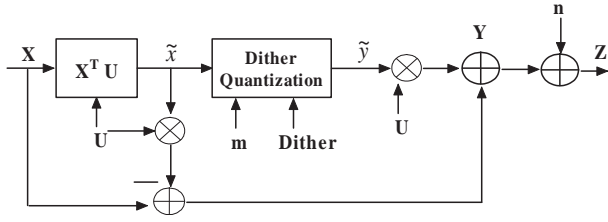


Fig. 3. Block diagram of spread transform Dither Modulation

of the change is controlled by the quantization error. Since \mathbf{u} is random, no consideration is given to the perceptual qualities of the signal \mathbf{x} .

In principle, a perceptual model can provide an estimate of the smallest change that each component of the signal \mathbf{x} accepts before becoming just noticeable. In prior work, we have referred to the change needed to introduce a just noticeable distortion (JND) as the “slack”.

In practice, for image signals, Watson provides a perceptual model for calculating the slack associated with each discrete cosine transform (DCT) coefficient within an 8×8 block [11]. Thus, given an image, \mathbf{x} , and its block-based DCT coefficients, \mathbf{X} , we can apply Watson’s model to compute the corresponding slack vector associated with each DCT coefficient. The larger a element of this vector, the more we may change the corresponding DCT coefficient before the change becomes noticeable.

For DM, we propose assigning to the projection vector, \mathbf{u} , the slacks computed by Watson’s perceptual model, \mathbf{s} , rather than pseudo-randomly generating the vector. Note that the vector magnitude is normalized to unity and quantization is performed in the DCT domain, as illustrated in Figure 4.

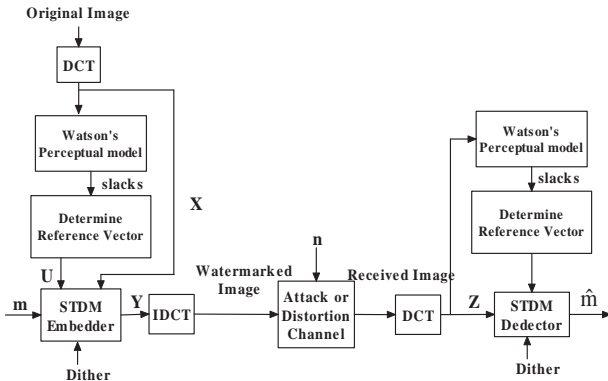


Fig. 4. Block diagram of STDM watermark embedder and detector with a perceptual model.

In this arrangement, which we refer to as STDM-W (STDM Watson), the change in \mathbf{x} is no longer randomly distributed, but is arranged based on the perceptual properties of the

signal - more change is directed to coefficients with larger slack. As a result, the perceptual distortion introduced by DM should be substantially reduced, as is confirmed by subsequent experiments.

Since the projection vector is now a function of the signal (image), it is unique for each image. Consequently, a blind watermark detector must be able to estimate the projection vector from the received, watermarked signal, as illustrated in Figure 4. However, since watermark embedding alters the signal, the detector’s estimate of the projection vector may not be exact.

In order to overcome this potential weakness, we also consider an alternative algorithm which does not require knowledge of Watson’s perceptual slacks at the detector. When we look at the original STDM algorithm depicted in Figure 3, the embedding algorithm simply reduces to “modify the cover Work \mathbf{x} so that its projection onto some random vector \mathbf{u} is equal to a given value”. An obvious and straightforward solution is to move the cover Work along the projection direction \mathbf{u} . However this is not mandatory! One may prefer to move the cover Work along a direction, close to \mathbf{u} , but which introduces significantly less perceptual distortion. In this perspective, previous work (E_PERC.OPT in [12]) suggests to use a perceptually shaped version, \mathbf{u}_s , defined for each sample as follows:

$$\mathbf{u}_s(i) = (\mathbf{u}(i)s(i)^4)^{\frac{1}{3}} \quad (7)$$

where s are Watson’s perceptual slacks. The embedding process can then be written:

$$\mathbf{y} = \mathbf{x} + \frac{(Q(\mathbf{x}^T \mathbf{u}, \Delta, m, \delta) - \mathbf{x}^T \mathbf{u})}{\mathbf{u}_s^T \mathbf{u}} \mathbf{u}_s, \quad m \in \{0, 1\} \quad (8)$$

Simple manipulation of Equation (8) reveals that $\mathbf{y}^T \mathbf{u} = Q(\mathbf{x}^T \mathbf{u}, \Delta, m, \delta)$. This implies that the detector given in Equation (6) can be reused. The key point is that this detector does not require the knowledge of Watson’s perceptual slacks. We refer to this algorithm as STDM-RW (STDM random Watson).

IV. EXPERIMENTAL RESULTS

Experiments were performed on 1000 images from the Corel image database. Each image has dimensions 512×512 . Quantization was performed on the DCT coefficients. We considered two embedding rates of 1/32 and 1/320. Thus, the 1/32 rate code embeds two bits in each 8×8 block. However, the number of modified coefficients is 62 rather than 64 since we ignore both the lowest and highest frequency coefficients of each block. Similarly for the rate 1/320 code.

Figures (5) and (6) illustrate the robustness of the original STDM together with the two perceptually adaptive methods, STDM-W and STDM-RW for the two embedding rates. At a rate of 1/32 we observe a small improvement in robustness compared with Figures 1 and 2. The point at which the BER exceeds 20% is now at about a JPEG quality factor of 92% compared with about 96% for regular DM and 97% for RDM-MW. It should also be noted that the BER does not increase

as rapidly for the spread transform methods. A BER of 50% is not reached until the JPEG quality factor is almost 50%, while for DM and RDM-MW, this occurs for JPEG quality factors of 75% and 85% respectively.

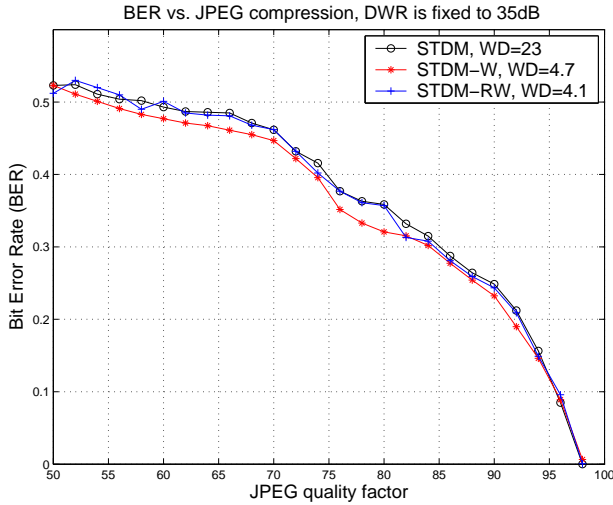


Fig. 5. Bit error rate (BER) as a function of JPEG compression for STDM-based schemes using an embedding rate of 1/32 and a DWR of 35 dB

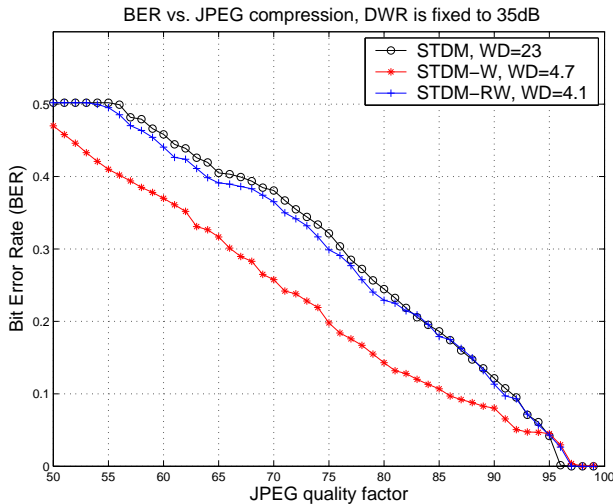


Fig. 6. Bit error rate (BER) as a function of JPEG compression for STDM-based schemes using an embedding rate of 1/320 and a DWR of 35 dB

At the rate of 1/320, we observe significant improvements, as shown in Figure 6. First, the point at which the BER exceeds 20% is now at about a JPEG quality factor of 84% for STDM and STDM-RW and about 75% for STDM-W. In comparison, DM and RDM-MW exceed the 20% BER for JPEG quality factors of 92% and 97% respectively. Once again, the BER does not increase as rapidly. A BER of 50% occurs at a JPEG quality factor of 55% for STDM and STDM-RW and at about 50% for STDM-W.

The performance of the STDM and STDM-RW methods is

approximately the same. This is to be expected. The perceptual modeling was not intended to improve robustness, but rather improve fidelity. And indeed, the perceptual distortion is considerably reduced. For a constant DWR of 35 dB, the average Watson distance between the original and watermarked images was 23 for STDM but only 4.7 for STDM-W and 4.1 for STDM-RW. However, surprisingly, for STDM-W, we not only observe an improvement in fidelity but also an improvement in robustness at low embedding rates.

Finally, Figure 7 shows the robustness to JPEG compression for the three ST methods when the Watson distance is fixed at 23 (rather than fix the DWR to 35 dB). In this case, we observe significant improvements in robustness for a fixed fidelity.

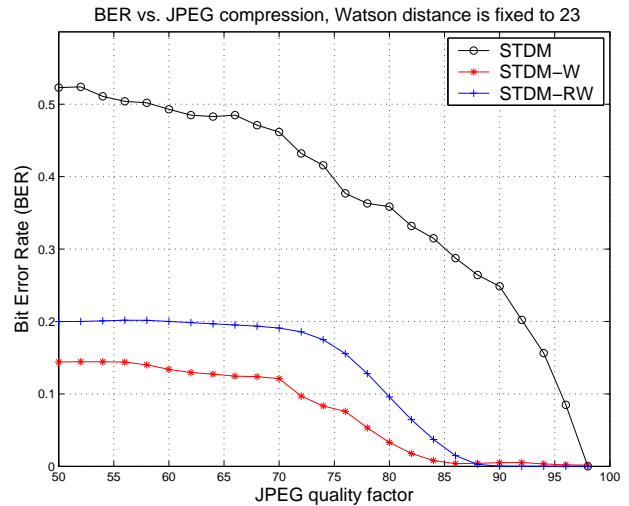


Fig. 7. Bit error rate (BER) as a function of JPEG compression for STDM-based schemes using an embedding rate of 1/32 and a Watson distance of 23

V. CONCLUSIONS

Quantization index modulation methods are fragile to val-umetric scaling and re-quantization. While considerable work has been directed to improving the robustness to scaling, the issue of re-quantization has received much less attention. This is surprising since re-quantization commonly occurs due to lossy compression, numerical rounding and analog-to-digital conversion.

Of the number of QIM variants, spread transform dither modulation exhibits most robustness to re-quantization. We proposed two modifications to STDM to incorporate perceptual modeling. In the first method, referred to as STDM-W, the random projection vector is replaced with the vector representing the magnitudes of the slacks associated with each DCT coefficient. This has the effect of directing the change to the cover signal to those areas of the signal where the changes will be least noticeable. Since the projection vector is now image dependent, the detector must also be able to estimate the slack for each image. Although the very act of watermark embedding may slightly change these slack values,

experimental results indicate that these changes are sufficiently small to permit very good detector performance.

Nevertheless, a second method, called STDM-RW, was also proposed, which avoided the need for the detector to determine the perceptual slacks for each image.

Experimental results were performed on 1000 images from the Corel database. We examined the robustness of the spread transform algorithms for embedding rates of 1/32 and 1/320. At high embedding rates, all three ST algorithms performed similarly. However, the perceptual distortion, as measured by the Watson distance, was 23 for the original ST algorithm, but only 4.7 for STDM-W and 4.1 for STDM-RW. For a JPEG quality factor of 92% or less, all three STDM algorithms exhibited a BER or over 20%. In comparison, regular DM and RDM-MW exceeded the 20% BER for JPEG quality factors of 96 and 97% respectively.

At the lower rate of 1/320 the improvement over DM and RDM-MW is much greater. Here a BER of 20% is exceeded for a JPEG quality factor of approximately 84% for STDM and STDM-RW and at 75% for STDM-W. This compares with quality factors of 92% and 96% for DM and RDM-MW.

Surprisingly, at the lower code rate, we also observe significant improvements in robustness as well as fidelity for STDM-W. It is unclear why this is so.

If the Watson distance rather than the DWR is held fixed, then the two adaptive methods show much more robust to re-quantization. Experimental results showed that for a fixed Watson distance of 23, and an embedding rate of 1/32 (i.e. 2 bits per block), the STDM-RW is robust to JPEG compression with a quality factor of about 65 before the BER exceeds 20%. For STDM-W, even at a JPEG quality factor of 50%, the BER did not exceed 15%.

The proposed ST algorithms, while exhibiting improved performance to re-quantization, are not invariant to valumetric scaling. This remains a direction of future work.

REFERENCES

- [1] M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [2] B. Chen and G. Wornell, "Quantization index modulation: A class of provably good methods for digital watermarking and information embedding," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423–1443, May 2001.
- [3] K. Lee, D. Kim, T. Kim, and K. Moon, "EM estimation of scale factor for quantization-based audio watermarking," in *Proceedings of the 2nd International Workshop on Digital Watermarking*, ser. Lecture Notes in Computer Science, vol. 2939, October 2003, pp. 316–327.
- [4] F. Pérez-González, F. Mosquera, M. Barni, and A. Abrardo, "Rational dither modulation: A high-rate data-hiding method invariant to gain attacks," *IEEE Transactions on Signal Processing, Supplement on Secure Media*, vol. 53, no. 10, pp. 3960–3975, October 2005.
- [5] Q. Li and I. Cox, "Using perceptual models to improve fidelity and provide invariance to valumetric scaling for quantization index modulation watermarking," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. II, March 2005, pp. 1–4.
- [6] P. Bas, "A quantization watermarking technique robust to linear and non-linear valumetric distortions using a fractal set of floating quantizers," in *Proceedings of the 7th International Workshop on Information Hiding*, ser. Lecture Notes in Computer Science, vol. 3727, June 2005, pp. 106–117.
- [7] Q. Li and I. J. Cox, "Rational dither modulation watermarking using a perceptual model," in *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, Shanghai, China, Oct. 2005.
- [8] I. Shterev and R. Legendijk, "Amplitude scale estimation for quantization-based watermarking," *IEEE Transactions on Signal Processing*, to appear.
- [9] C. Fei, D. Kundur, and R. Kwong, "Analysis and design of watermarking algorithms for improved resistance to compression," *IEEE Transactions on Image Processing*, vol. 13, no. 2, pp. 126–144, February 2004.
- [10] F. Pérez-González, P. Comesaña, and F. Balado, "Dither-modulation data hiding with distortion-compensation: Exact performance analysis and an improved detector for JPEG attacks," in *Proceedings of the IEEE International Conference on Image Processing*, vol. II, September 2003, pp. 503–506.
- [11] A. Watson, "DCT quantization matrices optimized for individual images," in *Human Vision, Visual Processing and Digital Display IV*, ser. Proceedings of SPIE, vol. 1913, September 1993, pp. 202–216.
- [12] I. Cox, M. Miller, and J. Bloom, *Digital Watermarking*. Morgan Kaufmann Publishers, 2001.