IMPROVE SPREAD TRANSFORM DITHER MODULATION BY USING A PERCEPTUAL MODEL TO PROVIDE RESISTANCE TO AMPLITUDE SCALING AND JPEG COMPRESSION

Qiao Li and Ingemar J. Cox

University College London Adastral Park Postgraduate Campus
Email: q.li@adastral.ucl.ac.uk , ingemar@ieee.org

ABSTRACT

It is well known that Spread Transform Dither Modulation (STDM) is more robust to re-quantization such as JPEG compression than regular Quantization Index (QIM) does. However, STDM introduces relatively higher perceptual distortion and is quite sensitive to amplitude scaling, which is actually a quite common processing. In this paper, we propose algorithms to reduce perceptual distortion and to provide robustness against amplitude scaling. In doing that, perceptual fidelity is improved by introducing Watson's model into the framework of STDM. Watson's model is then modified to gain resistance to amplitude scaling.

***Index Terms*—** Digital watermarking, perceptual model, quantization index modulation

1. 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 (STDM).

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, 4, 5, 6, 7, 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 STDM.

Pérez-Gonzàlez *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 STDM and it remains unclear whether this method is superior to STDM.

In this paper, we describe preliminary work to introduce a perceptual model within the STDM framework. Section 2 provides a brief introduction to quantization index modulation and particularly STDM and provide experimental results demonstrating the sensitivity to re-quantization and the relative robustness of STDM. Section 3 then describes how the projection vector used in STDM can be determined so as to minimize the perceptual distortion. The experimental results of Section 5 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. Moreover, if the perceptual distance rather than DWR is held fixed, then the new algorithm demonstrates a very significant improvement in BER. The proposed Spread Transform algorithm is very vulnerable to valumetric scaling attack. we then propose adaptive STDM schemes to overcome this problem in section 4. Section 6 summarizes our results and describes directions for future work.

2. SPREAD TRANSFORM DITHER MODULATION

The basic quantization index modulation (QIM) algorithm quantizes each signal sample, x, using a quantizer, $\mathrm{Q}(.)$, 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:

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

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

$$\mathrm{Q}(x, \Delta, m, \delta) = \Delta.\mathrm{Round}\left(\frac{x - \delta - m\frac{\Delta}{2}}{\Delta}\right) + \delta + m\frac{\Delta}{2} \tag{2}$$

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.

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

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 $\{x_1, \ldots, x_N\}$. The detection equation then becomes:

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

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. Our prior paper [12] shows experimental results to illustrate this point of DM. There, 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]

2.1. Adaptive QIM

QIM and DM are also very sensitive to valumetric distortion. A number of algorithms have recently been proposed to counter this [3, 4, 5, 7, 6, 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 requantization. However, [12], which looks at the robustness of RDM-MW to JPEG compression reveals that, perhaps surprisingly, it is actually slightly less robust.

2.2. Spread transform Dither Modulation

Figure 1 illustrates the basic framework for spread transform QIM.

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

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

and the corresponding detection is given by:

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

[1] Similar performance was observed in the pixel domain.

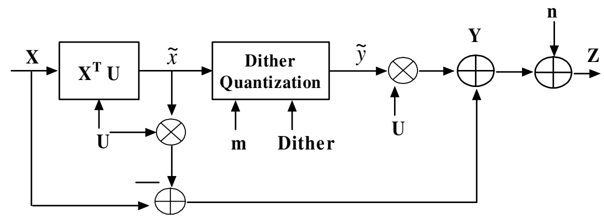

Fig. 1. Block diagram of spread transform Dither Modulation

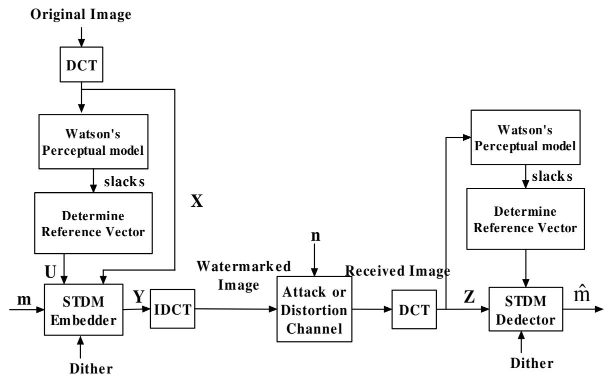

Fig. 2. Block diagram of STDM watermark embedder and detector with a perceptual model.

3. SPREAD TRANSFORM DITHER MODULATION 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 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 2.

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 2. 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 considered an alternative algorithm termed as STDM-RW in [12] which does not require knowledge of Watson's perceptual slacks at the decoder.

4. STDM BASED ON MODIFIED WATSON MODEL TO PROVIDE RESISTANCE TO VALUMETRIC SCALING

The proposed Spread Transform algorithms, while exhibiting improved performance to re-quantization, are not invariant to valumetric scaling. This is because that in traditional STDM scheme, the step size used in detector for quantization is not scaled linearly with β, which is the scaling factor to received signal. To make Spread Transform algorithm to be robust to valumetric scaling, we need to ensure two things: (i) make the step sizes to scale linearly with valumetric scaling, i.e. we want the estimated $\hat{\Delta}$ to be multiplied by β when the amplitude of the signal is scaled by β. (ii) the reference vector $\mathbf{u}$ used in embedder as Equation (5), onto which host signal is projected should be in the same direction with $\hat{\mathbf{u}}$ used in detector as Equation (6), onto which received signal is projected. please note that they don't have to be the same vector.

In our previous work [5], we proposed a modified Watson's model. Modified "slack",s^M, scales linearly with β. Based on that, we design two schemes as following to provide invariance to valumetric scaling. Motivated by all above in this section, we examine performances of two new methods.

STDM-MW

This scheme is an extension with regular STDM. Instead of using randomly generated vector, we use modified "slack" itself as reference vector to do projection onto. In this case, all elements of reference vector $s^{\hat{M}}$ in detector scale linearly when received signal is scaled by a factor of β. Thus the reference vector keeps same direction no matter received signal is scaled or not. However, this scheme still use fixed step size to do quantization.

STDM-MW-SS

Also, this scheme use "slack" from Modified Watson model to determine reference vector to do projection onto. Moreover, step sized here also depends on that. Given a $length - L$ vector of DCT coefficients $\{x_i; i = 1, 2, \ldots, L\}$ and its corresponding vector of Modified "slack" $\{S_i^M; i = 1, 2, \ldots, L\}$, we calculate step size as following:

$$\Delta = G_fac \times \sum_{i=1}^{L} S_i^M, i = 1, 2, \ldots, L. \qquad (7)$$

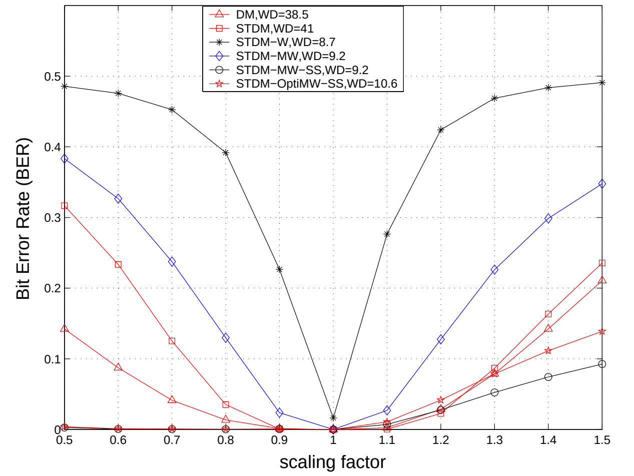

Fig. 3. Bit error rate (BER) as a function of valumetric scaling using an embedding rate of 1/320 and at a fixed DWR of 35 dB

Then we use this step size as Δ in Equation (5) to do STDM embedding. On the other hand in the detector, we firstly calculate modified slack according to received signal and then get $\hat{\Delta}$ in the same way as Equation (7). Detected bit is finally determined by Equation (6).

STDM-OptiMW-SS

In order to be as robust as STDM-W against JPEG compression, we try to optimize modified watson slack to make its value closer to regular watson slack. The way we optimize slacks is to divide the 43 higher frequency ones by 4, and keep the 21 lower frequency ones the same with modified watson.

5. EXPERIMENTAL RESULTS

Experiments were performed on 1000 images from the Corel image database. Each image has dimensions 768×512. 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.

6. CONCLUSIONS

7. 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-Gonzàlez, 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.

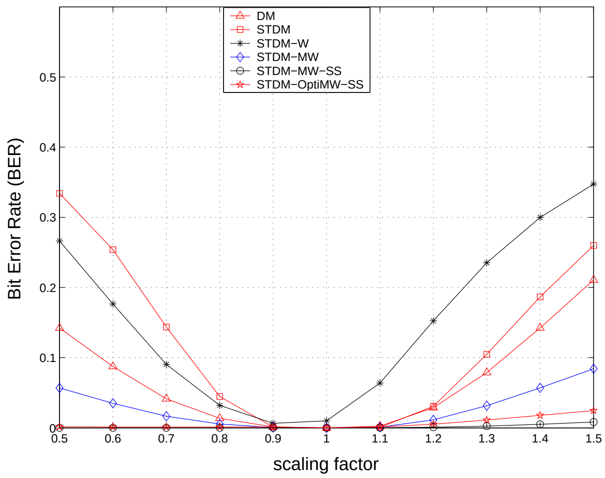

Fig. 4. Bit error rate (BER) as a function of valumetric scaling using an embedding rate of 1/320 and at a fixed Watson distance of 39

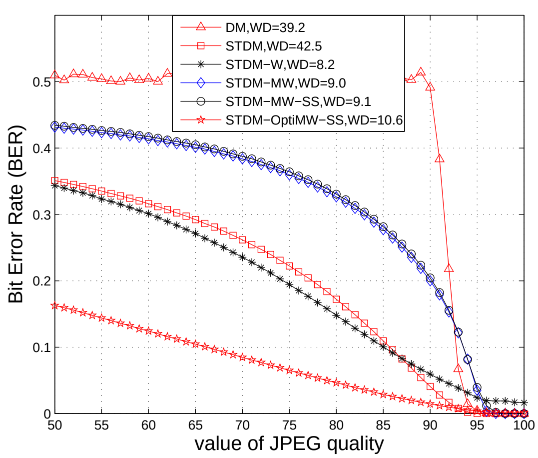

Fig. 5. Bit error rate (BER) as a function of JPEG Compression STDM-based schemes using an embedding rate of 1/320 and a DWR of 35 dB

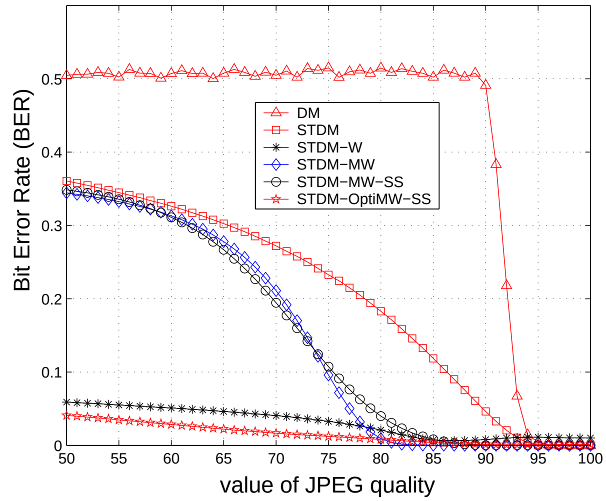

Fig. 6. Bit error rate (BER) as a function of JPEG Compression using an embedding rate of 1/320 and at a fixed Watson distance of 39

[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. Lagendijk, "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-Gonzàlez, 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] Q. Li, G. Doerr, and I. J. Cox, "Spread transform dither modulation using a perceptual model," in *IEEE International Workshop on Multimedia Signal Processing(MMSP)*, Victoria, Canada, Oct. 2006.