IMPROVED SPREAD TRANSFORM DITHER MODULATION USING A PERCEPTUAL MODEL: ROBUSTNESS TO AMPLITUDE SCALING AND JPEG COMPRESSION

Qiao Li and Ingemar J. Cox

ABSTRACT

Spread transform dither modulation (STDM) is a form of quantization index modulation (QIM) that is more robust to re-quantization. However, the robustness of STDM to JPEG compression is still very poor and it remains very sensitive to amplitude scaling. Here, we show how a perceptual model that scales linearly with amplitude scaling can be used to (i) provide robustness to amplitude scaling, (ii) reduce the perceptual distortion at the embedder and (iii) significantly improve the robustness to re-quantization.

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 common and occurs when any multimedia digital signal undergoes digital-to-analog conversion and subsequent analog-to-digital conversion. A major application of watermarking is provide protection from this "analog hole". Thus, if STDM is to be used for this application it is imperative that it be robust to re-quantization. Even in the absence of D-to-A and A-to-D conversion, requantization will occur whenever a multimedia 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]. In contrast, there has been surprisingly little research focused on the issue of re-quantization.

Fei *et al.* [7] 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 to both re-quantization and valumetric scaling.

Pérez-Gonzàlez *et al.* [8] 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, there is no comparison with STDM and it remains unclear whether this method is superior to STDM.

In this paper, we describe how a perceptual model that scales linearly with amplitude (valumetric) scaling can be incorporated within the STDM framework. This improves both the fidelity of STDM and provides robustness to valumetric scaling. We then demonstrate that such an approach can be extended to provide significant improvements in robustness to re-quantization. This is accomplished by adaptively adjusting the quantization step size based on the perceptual model.

Section 2 provides a brief introduction to STDM. Section 3 then describes how the projection vector used in STDM can be chosen 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 [9] is reduced from 42.5 to as little as 8.2, 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 STDM algorithm (STDM-MW) remains sensitive to valumetric scaling attack. In Section 4, we then propose an adaptive STDM methods (STDM-MW-SS and STDM-OptiMW-SS) to overcome this problem. Section 6 summarizes our results and describes directions for future work.

2. SPREAD TRANSFORM DITHER MODULATION

Both QIM and DM are very sensitive to re-quantization. Our prior paper [10] shows experimental results to illustrate this point. 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, 6], specifically rational dither modulation (RDM) [3], adaptive QIM using a modified Watson distance

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

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

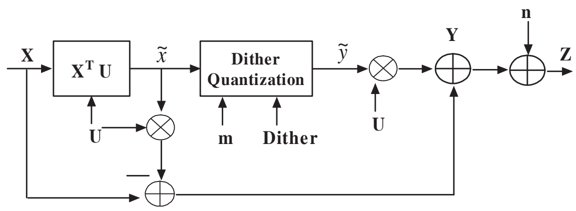

Fig. 1. Block diagram of spread transform dither modulation

(QIM-MW) [4] and adaptive RDM using a modified Watson distance (RDM-MW) [5]. 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, [10], which looks at the robustness of RDM-MW to JPEG compression reveals that, perhaps surprisingly, it is actually slightly less robust if comparison is made based on a constant document-to-watermark ratio (DWR).

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:

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

and the corresponding detection is given by:

$$\hat{m} = \arg\min_{b\in\{0,1\}} |\mathbf{z}^{\mathrm{T}}\mathbf{u} - \mathrm{Q}(\mathbf{z}^{\mathrm{T}}\mathbf{u}, \Delta, b, \delta)| \qquad (2)$$

3. SPREAD TRANSFORM DITHER MODULATION WITH PERCEPTUAL MODELING

From Equation (1) 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 [9]. 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 an element of this vector, the more we may change the corresponding DCT coefficient before the change becomes noticeable.

To incorporate a perceptual model within the STDM framework, the projection vector, $\mathbf{u}$, is assigned the slack values corresponding to each DCT coefficient, rather than pseudo-random values. Note that the vector magnitude is normalized

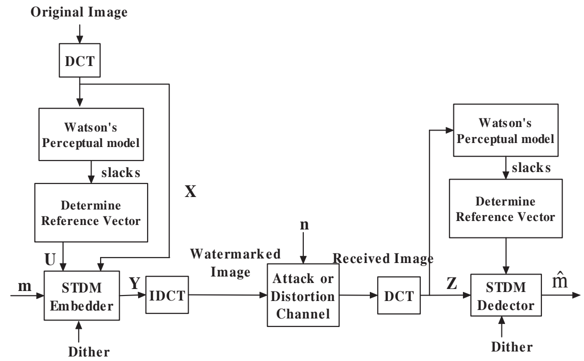

Fig. 2. Block diagram of STDM watermark embedder and detector with a perceptual model.

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 slacks. As a result, the perceptual distortion introduced by STDM will be substantially reduced, as is confirmed by experiments described shortly.

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 [10], which does not require knowledge of Watson's perceptual slacks at the decoder.

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

The proposed STDM algorithm (STDM-W), while exhibiting improved fidelity, is not invariant to valumetric scaling. This is because the quantization step size for both STDM and STDM-W is fixed, i.e. it does not scale linearly with the valumetric scaling factor, β.

To provide robustness to valumetric scaling, we need to ensure: (i) that the step size scales 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, and (ii) the reference vector $\mathbf{u}$ used in embedder (Equation (1)), is (approximately) the same as $\hat{\mathbf{u}}$ used in detector (Equation (2)). Note, however, that $\mathbf{u}$ and $\hat{\mathbf{u}}$ do not have to be identical, though some small degradation in performance may occur.

In our previous work [4], we proposed a modified Watson's model such that the modified slack, S^M, scales linearly with β. Based on this perceptual model, we designed two STDM algorithms. The first, STDM-MW, simply replaces the Watson model of STDM-W with the modified perceptual model and is provided for evaluation purposes. The second, STDM-MW-SS, also adaptively modifies the quantization step size thereby providing invariance to valumetric scaling. A further modification to Watson's model, provides us with a third STDM method, STDM-OptiMW-SS. All these new methods are described in this section.

STDM-MW

This method, STDM-MW, is the similar to STDM-W, except that the projection vector is now determined by the modified perceptual model. We provide this in order to examine the perceptual impact of the modification.

STDM-MW-SS

The STDM-W and STDM-MW methods do not provide invariance to valumetric scaling. The STDM-MW-SS not only uses the perceptual model to determine the projection vector, but also uses the same model to select the quantization step size.

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:

$$\Delta = G_{fac} \times \sum_{i=1}^{L} S_i^M, i = 1, 2, \ldots, L. \qquad (3)$$

Where G_{fac} is a global factor to adjust watermarking strength. Then we use this step size as Δ in Equation (1) 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 (3). Finally, the detected bit is determined by Equation (2).

STDM-OptiMW-SS

Our experimental results, described next, revealed that the performance of STDM-MW-SS did not perform as well as expected with respect to JPEG compression. Further investigation revealed that this was due to the fact that our modified perceptual model was (i) generally producing large slack estimates than Watson's model and (ii) that this error was larger for high frequency DCT coefficients. Thus, much more of the watermark signal was being placed in the very high frequency DCT coefficients which are the first to be eliminated by JPEG compression.

To remediate this, we altered our perceptual model to more closely follow the original Watson model yet retain the necessary linear scaling characteristic. This is accomplished by creating a piecewise linear model in which the modified slacks calculated for the 43 highest frequency coefficients are divided by 4. The modified slacks for the 21 low frequency coefficients were not modified.

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. Our embedding rate is 1/320, i.e. one bit in 5 8×8 blocks. Strictly, the embedding rate is 1/310 since the highest and lowest DCT coefficients if each block were unmodified, i.e. the number of modified coefficients is 62 rather than 64 and thus the total number of quantized DCT coefficients per bit is $5 \times 62 = 310$.

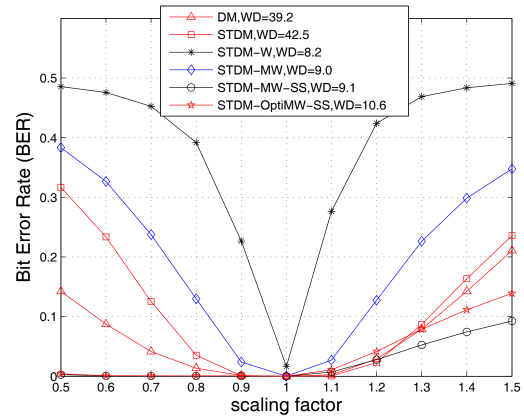

Fig. 3. Bit error rate(BER) vs. valumetric scaling using an embedding rate of 1/320 and at a fixed DWR of 35 dB

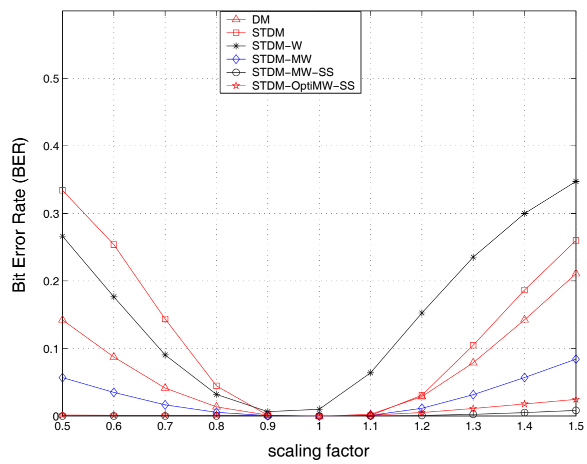

Fig. 4. Bit error rate(BER) vs. valumetric scaling using an embedding rate of 1/320 and at a fixed Watson distance of 39

Figure 3 shows the bit error rate (BER) as a function of valumetric scaling for DM, STDM and the set of modified STDM algorithms. As expected, DM, STDM, STDM-W and STDM-MW are sensitive to valumetric distortion. However, notice that the perceptual distortions, as measured by Watson distance, are 38.5, 41, 8.7 and 9.2 respectively. That is, both STDM-W and STDM-MW provide a significant improvement in fidelity. Further, the perceptual degradation due to our modified perceptual model is slight. Both STDM-MW-SS and STDM-OptiMW-SS both demonstrate very good robustness to valumetric scaling together with improved fidelity.

In Figure 3 all watermarks were embedded at a fixed DWR of 35dB. If, instead, the watermarks are embedded at a fixed fidelity, e.g. a Watson distance of 39, then the robustness to valumetric scaling is even better, as shown in Figure 4.

Figure 5 illustrates the robustness to JPEG compression. As expected, DM is the most sensitive, exceeded a 20% BER for a quality factor of about 92%. Standard STDM is considerably better, not exceeding a 20% BER until QF=77%. Both

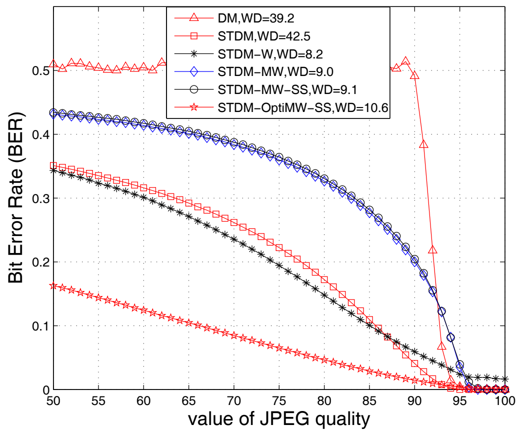

Fig. 5. Bit error rate(BER) vs. JPEG Compression using an embedding rate of 1/320 and a DWR of 35 dB

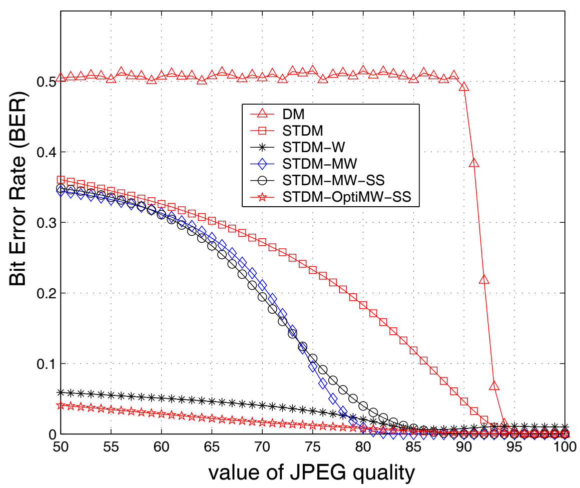

Fig. 6. Bit error rate(BER) vs. JPEG Compression using an embedding rate of 1/320 and at a fixed Watson distance of 39

STDM-MW and STDM-MW-SS have almost identical performance that is better than DM but worse than STDM. As discussed earlier, this is due to that fact that the modified perceptual model over estimates the slacks for high frequency coefficients which are most sensitive to JPEG compression. By using the piecewise linear model, which reduces the slack values for high frequencies, we see a very significant improvement in performance for STDM-OptiMW-SS. And this is achieved with a lower perceptual distortion of 10.6 as compared to 42.5 for STDM.

In Figure 5, all watermarks were embedded at a fixed DWR of 35dB. If, instead, we fix fidelity, i.e. a Watson distance of 39, then it is observed from Figure 6 that DM and STDM perform worse and that STDM-W and STDM-OptiMW-SS are the best performing methods with bit error rates that never exceed 5% even for JPEG Quality Factors of 50%. Most importantly, STDM-OptiMW-SS achieves this whilst also providing robustness to valumetric scaling.

6. CONCLUSIONS

We described how a perceptual model can be introduced into the STDM framework. This is accomplished by choosing the projection vector such that the watermark changes are directed to regions that have large perceptual slack. By using a modified perceptual model based on Watson's model and using this not only to select the projection vector but also to determine the quantization step size, we are able to significantly improve both the fidelity and robustness to valumetric scaling. Experimental results on 1000 images confirmed this.

Experiments revealed an unexpected sensitivity to JPEG compression. This was due to the modified perceptual model over estimating the slack values for the high frequency DCT coefficients. As a results, more of the watermark energy was placed in these regions, which are very sensitive to JEPG compression. This problem was resolved by introducing a piecewise linear model that attenuated the slack estimates in the higher frequencies. The resulting algorithm, STDM-OptiMW-SS exhibits very good fidelity, and is very robust to both JPEG compression and valumetric scaling.

We believe these results are important as it is imperative that watermarks be robust (preferably invariant) to valumetric scaling and re-quantization if they are to be applied in copy protection applications where D-to-A and A-to-D conversion and changes in brightness/volume are common.

Further improvements may be possible by considering a version of STDM based on rational dither modulation and a perceptual model, similar to that described in [5].

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] 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.

[4] Q. Li and I. J. Cox, "Using perceptual models to improve fidelity and provide invariance to valumetric scaling for quantization index modulation watermarking," in *IEEE ICASSP*, Philadelphia, USA, March 2005.

[5] ——, "Rational dither modulation watermarking using a perceptual model," in *IEEE MMSP*, Shanghai, China, Oct. 2005.

[6] I. Shterev and R. Lagendijk, "Amplitude scale estimation for quantization-based watermarking," *IEEE Transactions on Signal Processing*, to appear.

[7] 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.

[8] 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.

[9] 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.

[10] Q. Li, G. Doerr, and I. J. Cox, "Spread transform dither modulation using a perceptual model," in *IEEE MMSP*, Victoria, Canada, Oct. 2006.