SSIM-based Rate Distortion Optimal Differential Energy Watermarking

Guixing Wu and Yudong Xiang
School of Software Engineering, Suzhou Institute for Advanced Study, USTC Suzhou, Jiangsu, P. R. China
Email: gxwu@ustc.edu.cn, sa514005@mail.ustc.edu.cn

Abstract. In this paper, we propose a rate distortion optimal method to maximize the compression performance of differential energy watermarking. First, the compression performance of differential energy watermarking is analyzed, which shows the deterioration of compression performance due to digital marking. Second, we developed an SSIM-based rate distortion optimal method to optimize the compression performance. Experimental results show that the proposed method can achieve better compression performance or big rate reduction while visual quality is not affected.

1. Introduction

Digital watermarking, as a proposed solution to digital rights management, has been studied extensively during the past decades [1], [2], [3], [4]. Digital watermarks are usually embedded into multimedia contents such that they are imperceptible. The embedded watermarks should be robust against various attacks. In addition, the watermark embedding capacity and security are desired to be high.

Since multimedia signals are usually transmitted in compressed formats, watermarks are usually embedded in the compressed domain[5]. In [6], [7], the authors investigated watermark embedding within the compressed domain from an information-theoretic perspective. In [8], [9], [10], [11], [12], the authors embedded information into the compressed bitstreams, the tradeoff between the robustness and encoding distortion was investigated. Nevertheless, the resulting distortion or compression rate may increase after information embedding. Hence the embedding of digital information may result in a degradation of compression performance.

In some applications it is desirable to maintain compression performance after digital information is embedded into the compressed bit streams. Consider, for example, the application of digital camera. In a digital camera image, date information is often visibly inserted, which degrades the quality of the image. One way to address this problem is to hide the date or similar type of information invisibly into the image. The embedded information is tied to the image, and remains intact during mild distortion and image format conversion. Later, the embedded information can be extracted using a watermark decoder whenever necessary.

Due to the limited storage space of digital camera, the resulting image file size is desired to be maintained or slightly increased after information embedding. However, it is usually difficult to achieve this goal without degrading the quality of the image. In [13], [14], the authors embedded watermarks into the H.264 compressed bit streams while preserving the compression bit rate, but the video quality optimization was not taken into consideration.

In this paper, we look into the optimization of compression performance while embedding watermarks into the compressed bit streams. As JPEG is the most widely used image compression
format and differential energy watermarking is a well known watermarking method [10], [15], we shall use a JPEG compressed and differentially energy watermarked system as an example to demonstrate how to maximize the compression performance.

The rest of the paper is organized as follows. Section II analyzes the compression performance of differential energy watermarking. In Section III we present an SSIM-based rate distortion optimal algorithm to maximize the compression performance of a JPEG compressed and differentially energy watermarked system. Simulation results are given in Section V.

2. Analysis Of Differential Energy Watermarking
In differential energy watermarking, watermarks are em-bedded in DCT energy differences between DCT blocks of an image. That is, a watermark bit is embedded into $2^n$ DCT blocks which are divided into two regions: region A and region B. Each of these two regions contains $n$ DCT blocks. Then the high frequency DCT coefficients in one chosen region are discarded such that the DCT energy difference between these two regions are greater than a desired threshold. The chosen region could be region A or region B depending on whether a bit “1” or a bit “0” is embedded.

Figure 1 compares the compression performance of differential energy watermarking (DEW) with that of standard JPEG compression for the $512 \times 512$ image Lena. In Fig. 1 STD-JPG stands for standard JPEG compression. In DEW, 64 bits are embedded, the quality factor of JPEG recompression is set to 25. PSNR means peak signal to noise ratio, it is an image quality metric. From the figure, we can see, the watermark embedding degrades the compression performance. For instance, at PSNR=38dB, the compression bit rate after watermark embedding increases nearly 70%. Similar results were observed for other test images. Therefore, it is necessary to optimize the compression performance to compensate for the rate degradation due to watermark embedding. In this paper, we develop a rate distortion optimal method to maximize the compression performance of JPEG compressed images while DEW is applied to the compression process.

![Figure 1 Comparison of performance between the standard JPEG compression and DEW](image)

3. Ssim-Based Rate Distortion Optimization

3.1. Problem Formulation
To maximize the compression performance, we want to solve the following constrained optimization problem

$$\min_{r} \quad \text{subject to } d \leq d_{budget} \text{ and } (1-2m)E \geq D$$

(3.1)
where \( r \) means the compression bit rate, \( d \) means the encoded distortion, \( d_{budget} \) is the distortion constraint, \( m \) is watermark bit 1 or 0, \( E \) is the energy difference, \( D \) is the energy threshold, and

\[
(1 - 2m)E \geq D
\]

is the constraint due to watermark embedding.

The above constrained optimization problem can be solved by minimizing

\[
m \in J(\lambda) = d + \lambda r
\]

where \( \lambda \) is a fixed parameter that represents the tradeoff of rate for distortion, and \( J(\lambda) \) is the Lagrangian cost.

In this paper, we use a widely used quality metric—structural similarity index (SSIM) as the visual quality metric [16]. In the case of high rate quantization, Yeo et al approximate SSIM as [17]

\[
SSIM = \frac{2\sigma_x^2 + \epsilon_2}{2\sigma_x^2 + \text{MSE} + \epsilon_2}
\]

where \( \sigma_x^2 \) is the variance of the original image signal \( x \), \( \epsilon_2 \) is a constant for numerical stability, MSE is the mean square error between an original image and the degraded image.

The distortion can then be approximated as

\[
d \approx 1 + \frac{\text{MSE}}{2\sigma_x^2 + \epsilon_2}
\]

The Lagrangian cost can be defined as

\[
J(\lambda) = \text{SSE} + 2(\sigma_x^2 + \epsilon_2)\lambda r
\]

where SSE means the sum of square error between an original image and the degraded image.

### 3.2. Algorithm Design

Before presenting the developed algorithm, let’s first have a brief review of JPEG compression [15]. A JPEG encoder first partitions an input image into 8x8 blocks. Each of these 8x8 blocks is DCT transformed, and the resulting DCT coefficients are then uniformly quantized using an 8x8 quantization table. The coefficient indices from the quantization are finally entropy coded using zero run-length coding and Huffman coding.

The JPEG syntax leaves the selection of the quantization table, the Huffman codewords, and DCT indices to the encoder provided that the encoding process is compatible with the JPEG standard[18]. Since the coefficient indices can be equivalently represented as run-size pairs followed by in-category indices through run-length coding, in [18] Yang and Wang proposed a JPEG-compatible joint optimization algorithm to maximize the compression performance over all possible sequences of run-size pairs \((R, S)\) followed by in category indices \(ID\), all possible Huffman tables \(H\), and all possible quantization tables \(Q\).

In this paper, we investigate how to maximize the JPEG compression performance while subjecting to the water-marking constraint (3.2). Therefore, the minimization problem can be converted to

\[
\min_{(R, S, ID), H, Q} \{J(\lambda) = \text{SSE} + 2(\sigma_x^2 + \epsilon_2)\lambda r\}
\]
where \((R, S, I, D)\) and \(Q\) are chosen in the set where the watermarking constraint (3.2) is satisfied. Since the Huffman table \(H\) is completely determined by the run-size probability distribution \(P\), we use \(P\) in the following to replace the Huffman table \(H\) in (3.7).

The search for the minimum of \(J(\lambda)\) through optimizing \((R, S, I, D)\), \(P\) and \(Q\) simultaneously entails too much computational complexity. To reduce the computational complexity, the following alternating minimization procedure is used to solve the problem (3.7):

1) Fix \(P\) and \(Q\). Find an optimal sequence \((R, S, I, D)\) that achieves the following minimum

\[
\min_{(R, S, I, D)} \{ J(\lambda) = SSE + 2(\sigma_1^2 + c_2)\lambda r \} \tag{3.8}
\]

while satisfying the watermarking constraint (3.2);

2) Fix \((R, S, I, D)\). Update \(P\) and \(Q\) respectively to achieve the following minimum

\[
\min \{ J(\lambda) = SSE + 2(\sigma_1^2 + c_2)\lambda r \} \tag{3.9}
\]

while satisfying the watermarking constraint (3.2).

Note that in (3.9) \(P\) can be selected as the empirical run-size distribution of \((R, S)\). The two steps are repeated until the decrease of \(J(\lambda)\) is below a prescribed threshold \(\epsilon\).

### 3.3. Optimal DCT Indices Updating

Fix \(P\) and \(Q\). The optimal DCT indices (or \((R, S, I, D)\) sequence) can be determined independently for each \(8 \times 8\) image block as \(J(\lambda)\) is block-wise additive. Following a similar approach as in [18], we solve the problem using a graph based optimization algorithm while satisfying the watermarking constraint.

Let the cut off index in differential energy watermarking be equal to \(C\). That is, the DCT coefficients after this position \(C\) within a DCT block are either discarded or included in the computation of DCT energy. When a DCT block belongs to the region where high frequency DCT coefficients are discarded, the graph utilized in the searching of the optimal result has \(C + 1\) states as shown in Fig. 2. The first \(C\) states correspond to the first \(C\) DCT coefficient indices of a DCT block in zigzag order. Each state may have incoming connections from its previous 16 states, which correspond to the run \(R\), in an \((R, S)\) pair. The last state is called end state. The end state may have incoming connections from all the other states where the indices are not equal to zeros.

![Figure 2 Optimization graph 1](image1)

![Figure 3 Optimization graph 2](image2)
Associated with each transition \((r, s)\) in Fig. 2 is a cost defined as the incremental Lagrangian cost of going from state \(i - r - 1\) to state \(i\) when the \(i\)th quantized DCT coefficient index \(e\)eds \(s\) bits to represent its amplitude and all the \(r\) DCT coefficients \(\theta_j\) before the \(i\)th DCT coefficient are quantized to zeros. Specifically, this incremental cost is equal to

\[
\sum_{j=r+1}^{i} \theta_j^2 + |\theta_i - q_i| + \mu D_s^2 + 2(q_j^2 + c_j)\lambda \cdot (-\log P(r,s) + s)
\]

where \(I D_i\) is an index in the size group \(s\) that gives rise to the minimum square error to the DCT coefficient \(\theta_i\), \(q_i\) is the \(i\)th quantization step size in the quantization table, and \(P(r,s)\) is the probability of the pair \((r,s)\).

Let \(L_i\) be the minimum incremental Lagrangian costs from state 0 to state \(i\). Let \(SSE_i\) be the cost of dropping all DCT coefficients from \(i + 1\) to 63. Find a certain value \(k\) which minimizes the sum of \(L_i\) and \(SSE_i\)

\[
\min_{0 \leq i \leq C} \{ J(i) = L_i + SSE_i \}
\]

Then backtrack from state \(k\) to find the optimal sequence \((R, S, I D)\) with the minimum Lagrangian cost.

When the DCT block belongs to the region where high frequency DCT coefficients are not discarded, the optimization graph has 65 states as shown in Fig. 3. The first 64 states correspond to the DCT coefficient indices of an image block. The calculation of the minimum incremental Lagrangian cost is similar to that in Fig. 2, but with the following exceptions. To preserve the DCT energy difference, the DCT indices after the cut off index should satisfy the constraint

\[
ID_i \geq \frac{[\theta_i]_Q \cdot r_{pre}}{q_i}
\]

where \([\theta_i]_Q \cdot r_{pre}\) is the pre-quantized DCT coefficient using standard JPEG recompression. After the cutoff index no transition is allowed across state \(i\) where \([\theta_i]_Q \cdot r_{pre} \neq 0\). Assume \(t\) is the position of last nonzero pre-quantized DCT coefficient within a DCT block, only states from \(t\) to 63 are allowed to go to state \(end\).

After the calculation of the incremental Lagrangian cost, find a certain value \(k\) which minimizes the sum of \(L_i\) and \(SSE_i\) sequence \((R, S, I D)\) with the minimum Lagrangian cost.

\[
\min_{0 \leq i \leq 63} \{ J(i) = L_i + SSE_i \}
\]

Then backtrack from state \(k\) to state 0 find the optimal sequence \((R, S, I D)\) with the minimum Lagrangian cost.

3.4. Optimal Huffman and Quantization Table Updating

Fix \((R, S, I D)\). The Huffman table \(H\) is determined by the empirical run-size distribution \(P\). As the compression rate does not depend on the quantization table \(Q\) under this condition, we only need to minimize the sum of square error in the Lagrangian cost \(J(\lambda)\). That is, we need to solve the following minimization problem

\[
\min_Q \{ SSE \}
\]
Among all possible $Q$, let $Q = (q_{ij}, q_{ij}, \ldots, q_{ij})$. Let $\theta_{ij}$ and $ID_{ij}$ denote the $i$th DCT coefficient and the $i$th DCT index of the $j$th DCT block in zigzag order respectively.

Then SSE can be expressed as

$$SSE = \sum_{i=1}^{N_{num\_blk}} \sum_{j=1}^{8N_{num\_blk}} (\theta_{ij} - ID_{ij} \cdot q_{ij})$$

(3.15)

Where $N_{num\_blk}$ is the number of 8x8 blocks in an image.

The minimization of $SSE$ can be achieved by minimizing independently the inner summation of (3.15) for each $i = 1, 2 \ldots 63$. Taking the derivative with respect to $q_{ij}$, we get the minimum of (3.15) with a set of quantization step sizes

4. Experimental Results

Having described the proposed optimization algorithm, in this section, we evaluate the compression performance of the proposed algorithm. In the experiments, the constant $c_2 = (0.03 \times L)^2$, where $L$ is the peak value of an image. We use test images from [17] to make the experiments.

![Figure 4 SSIM comparison between the proposed algorithm and DEW](image1)

![Figure 5 PSNR comparison between the proposed algorithm and DEW](image2)

Figure 4 compares the SSIM values of the proposed algorithm with the original DEW algorithm. The test image is 512x512 Lena, The quality factor of JPEG recompression is set to 25. 64 bits of watermark information are embedded. It can be seen from the figure that the proposed algorithm could achieve better compression performance. For example, at $SSIM = 0.94$, the proposed algorithm can reduce nearly 50% of the bit rate of a DEW encoder.

Figure 5 compares the PSNR values of the proposed algorithm with the original DEW algorithm. The test image is 512x512 Lena, The quality factor of JPEG recompression is set to 25. We can see that the proposed algorithm could achieve better compression performance. For example, at $PSNR = 32.5dB$, the proposed algorithm can reduce nearly 50% of the bit rate of a DEW encoder.

Similar results can be observed for other test images and/or other quality factors of JPEG recompression. As the DCT energy difference is preserved in the optimization, the watermark robustness was not affected.

5. Conclusion

In this paper, we developed an SSIM-based rate distortion optimal method to maximize the compression performance of a differential energy watermarking system. We first analyzed the
differential energy watermarking, which showed the necessity for the optimization of compression performance. Next we proposed a rate distortion optimal method to optimize the compression performance. From the experimental results we can see that the proposed method can boost the compression performance of a differential energy watermarking system.

6. Acknowledgment
The authors would like to thank General Program of Natural Science Foundation of Jiangsu Province (BK20141209) and Suzhou Fundamental Research Project (SYG201543) for supporting this work.

7. References
[1] A. K. Khanna, N. R. Roy and B. Verma, “ Digital image watermarking and its optimization using genetic algorithm,” 2016 International Conference on Computing, Communication and Automation, pp. 1140 - 1144, 2016.
[2] Y. Erfani, R. Pichevar and J. Rouat, “ Audio Watermarking Using Spikegram and a Two-Dictionary Approach,” IEEE Transactions on Information Forensics and Security, Vol.12, No. 4, pp.840 - 852, 2017.
[3] I.J. Cox, M. Miller and J. Bloom, Digital Watermarking, October 2001, 576 pages Cloth.
[4] B. Chen and G. W. Wornell, “Quantization index modulation: A class of provably good methods for digital watermarking and information embedding,” IEEE Trans. Inform. Theory, Vol. 47, No. 4, pp.1423–1443, May 2001.
[5] G. Wu and E.-H. Yang, “Joint watermarking and compression using scalar quantization for maximizing robustness in the presence of additive gaussian attacks,” IEEE Trans. Signal Processing, Vol. 53, No. 2, pp. 834-844, Feb. 2005.
[6] D. Karakos and A. Papamarcou, “A relationship between quantization and watermarking rates in the presence of additive Gaussian attacks,” IEEE Trans. Inform. Theory, Vol. 49, No. 8, pp. 1970–1982, Aug. 2003.
[7] A. Maor and N. Merhav, “On joint information embedding and lossy compression in the presence of a stationary memoryless attack channel,” IEEE Trans. Inform. Theory, Vol. 51, No. 9, pp. 3166–3175, Sept. 2005.
[8] C. -S. Lu, J. -R. Chen, M. Liao, K. -C. Fan, “Real-time MPEG2 video watermarking in the VLC domain,” Proc.16th IAPR Int. Conf. on Pattern Recognition, Vol. 2, pp.552-555, Aug. 2002.
[9] A. M. Alattar, E. T. Lin and M. U. Celik “Watermarking low bit-rate advanced simple profile MPEG-4 bitstreams,” in Proc. Int. Conf. on Acoustics, Speech and Signal Processing, pp. 513-516, 2003.
[10]G.-C. Langelaar and R. L. Lagedijk, “Optimal differential energy watermarking of DCT encoded images and video,” IEEE Trans. Image Processing, Vol. 10, No. 1, pp. 148-158, Jan. 2001.
[11] S. A. PARAH,J. A. SHEIKH , N. A. LOAN, “Robust and blind watermarking technique in DCT domain using inter- block coefficient differencing,” in Digital Signal Process- ing, pp.11-24, Vol. 53, 2016.
[12]L. Wang, H. Ling, F. zou, Z. Lu, “Real-Time Compressed- Domain Video Watermarking Resistance to Geometric Dis- tortions,”IEEE MultiMedia, Vol 19, No. 1. pp. 70 - 79,2012.
[13]D. Xu, R. Wang, and Y. Q. Shi, “Data hiding in Encrypted H.264/AVC Video Streams by Codeword Substitution ,” IEEE Trans. on Information Forensics and Security, Vol.9, No. 4, pp. 596–606, April 2014.
[14]Z. Ma, J. Huang, M. Jiang, X. Niu, “A Video Watermarking DRM Method Based on H.264 Compressed Domain with Low Bit-Rate Increaseasement,”Chinese Journal of Electr- onics, Vol. 25, No. 4, pp. 641-647, April 2016.
[15]W. Pennebaker and J. Mitchell, JPEG Still Image Data Compression Standrad, Kluwer Academic Publishers,1993.
[16]Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, “ Image quality assessment: from error visibility to structural similarity ,” IEEE Transactions on Image Processing, Vol.13, No. 4, pp. 600–612, 2004.
Circuits and Systems for Video Technology, Vol. 23, No. 7, pp. 1170–1181, July 2013.

[18] E.-H. Yang and L. Wang, “Joint optimization of run-length coding, Huffman coding and quantization table with complete baseline JPEG decoder compatibility,” IEEE Trans. Image Processing, Vol. 18, No.1, pp. 63-74, Jan. 2009.

[19] http://sipi.usc.edu/database/.