ON THE DETERMINATION OF CAPACITY PARAMETERS IN PEE-BASED REVERSIBLE IMAGE WATERMARKING

Jiantao Zhou and Oscar C. Au

Department of Electrical and Computer Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong SAR
Email: {eejtzhou@gmail.com and eeau@ust.hk}

ABSTRACT

In the existing prediction-error expansion (PEE)-based reversible image watermarking schemes, the capacity parameters are determined in a recursive manner by gradually turning these parameters to fit the payload. This class of method needs multiple rounds of embedding iterations, and hence, it is computationally inefficient. In addition, when multiple capacity parameters need to be handled, the previous methods are generally not capacity-distortion optimized. In this work, we formulate the task of determining the capacity parameters as a capacity-distortion optimization problem, which can be shown to be convex. We also prove that under some conditions, even simple analytical solutions exist.

Index Terms— Reversible Watermarking, prediction-error expansion, difference expansion, capacity-distortion optimization

1. INTRODUCTION

Reversible watermarking is a special class of data hiding technique, which ensures perfect reconstruction of the original cover signals after watermark extraction. Reversible image watermarking is an ideal solution for many critical applications, e.g., protecting the image integrity in medical imaging and remote sensing systems.

To achieve reversible image watermarking, early methods were based on lossless compression of certain image features, creating vacancies for embedding the payload [1, 2]. However, these methods can only provide rather low capacity, and usually lead to severe degradation of the image quality. Among the existing reversible image watermarking schemes, the prediction-error expansion (PEE)-based approaches [3–5] have received growing attention due to its excellent embedding capacity and good controllability of the distortions caused by watermarking embedding. PEE is actually a generalization of the difference-expansion (DE) technique initially developed by Tian [6]. Compared with DE-based methods, PEE exploits the prediction error, instead of the difference of two adjacent pixels, for expansion embedding. The superior de-correlating ability of image predictors such as GAP [7] and MED [8] enables efficient exploitation of spatial redundancy of images. To reduce the overhead caused by location map, Thodi and Rodriguez also suggested to incorporate histogram shifting with expansion embedding [3]. In [4], Hu *et al.* proposed to use payload dependent location map to achieve further compression of the location map, and hence, achieved better capacity-distortion performance.

In PEE-based schemes, upon getting the prediction errors, we can obtain the prediction-error histogram, which is typically modeled as a zero-mean Laplacian distribution. Then the so-called capacity parameters are determined to divide the histogram into inner region and outer region. For the pixels in the inner region, expansion operations are conducted to embedded 1 bit for each pixel. For the pixels in the outer region, shifting operations are applied to avoid ambiguity, while no message bits are embedded. Within this framework, a key issue is to determine the capacity parameters which control the embedding capacity and influence the distortions caused by watermark embedding. In the existing methods, the capacity parameters are determined in a recursive manner by graduating turning these parameters to fit the payload. This class of method needs multiple rounds of embedding iterations, and hence, it is computationally inefficient. This is especially the case if we need to operate with several histograms simultaneously, as that done in [5], where more capacity parameters have to be determined.

In this work, we address the problem of optimally determining the capacity parameters in an efficient manner. We formulate the task of finding the capacity parameters as a capacity-distortion problem, which can be shown to be convex. We also prove that under some conditions, even simple analytical solutions exist. This analytical framework can also be readily extended to handle multiple histograms where more capacity parameters need to be determined. Preliminary experimental results are provided to validate our findings.

This work has been supported in part by the Innovation and Technology Commission of the Hong Kong Special Administrative Region, China (project no GHP/048/08).

The rest of the paper is organized as follows. Section 2 briefly introduces the PEE-based reversible image watermarking. In Section 3, we present the formulation of the capacity-distortion optimization problem and prove its convexity. Section 4 gives the experimental results that verify our theoretical findings. We conclude in Section 5.

2. PEE-BASED REVERSIBLE IMAGE WATERMARKING

A common feature of the reversible image watermarking is the use of some de-correlating operator, e.g., image predictors GAP [7] and MED [8], that creates a feature set with majority of its elements having small magnitudes. The data embedding is accomplished by expanding these features to create vacancies into which the message bits are embedded. The distortion resulting from expansion embedding mainly depends on the magnitudes of those feature elements that are expanded. Hence, it is desirable to have as many as possible feature elements with small magnitudes. In the reversible image watermarking schemes via prediction error expansion, the prediction errors are the feature elements on which expansion embedding is conducted.

More specifically, each pixel of the host image I is predicted by using the predictor, and the prediction error can be computed as

$$e = I - \hat{I} \tag{1}$$

where $\hat{I}$ is the prediction. It should be noted that the prediction $\hat{I}$ is based on the casual past pixels, and hence, the decoder can get the same prediction without any ambiguity.

Upon getting the prediction errors, the watermark embedding can be performed using the following three steps.

Step 1:) Expansion embedding

If the prediction error e belongs to the inner region $\mathcal{I} = \{e|e \in [T_l, T_r)\}$, then expand it to

$$e^w = 2e + w \tag{2}$$

where T_l and T_r are the capacity parameters controlling the embedding capacity, and $w \in \{0, 1\}$ denotes the message bit embedded. The watermarked pixel value then becomes

$$I^w = \hat{I} + e^w = I + e + w \tag{3}$$

Step 2:) Histogram shifting

If the prediction error e belongs to the outer region $\mathcal{O} = \{e|e \in (-\infty, T_l) \cup [T_r, +\infty)\}$, then the corresponding pixel will not carry any message bit, and the prediction error is simply shifted to avoid overlapping with the expanded inner region. In this case, the watermarked pixel value becomes

$$I^w = \begin{cases} I + T_r & \text{for} \quad e \geq T_r \\ I + T_l & \text{for} \quad e < T_l \end{cases} \tag{4}$$

In many existing schemes, it was assumed that $T_r = -T_l = T > 0$. It can be easily seen that larger T leads to larger embedding capacity, but at the same time incurs larger distortion.

Step 3:) Location map construction

The above expansion and shifting operations would result in some watermarked pixel values that are outside the range $[0, 255]$, causing the overflow/underflow (overflow in short) problem. A simple way of solving this problem is to record the locations of these problematic pixels, and avoid performing expansion/shifting operations on the associated prediction errors. The location map needs to be losslessly compressed and embedded into the host as auxiliary information to ensure reversibility. The employment of the location map decreases the effective embedding capacity. Fortunately, the overhead induced by the location map is negligible, especially at the low embedding rate region.

At the decoder side, upon getting the watermarked image I^w, the task is to retrieve the embedded message bits and restore the original image in a lossless fashion. Since the procedure is somewhat similar to that conducted at the encoder, we omit the details here.

3. CAPACITY-DISTORTION OPTIMIZATION FOR FINDING THE CAPACITY PARAMETERS

In the existing schemes, the capacity parameters T_l and T_r are obtained recursively by gradually decreasing T_l and increasing T_r to fit the payload. Since this class of method needs multiple rounds of embedding iterations, it is computationally inefficient. In this work, we demonstrate that the optimal capacity parameters can be well estimated by solving a capacity-distortion optimization problem. In the case that $T_l = -T_r$, even a simple close-form analytical solution exists.

As a well-known fact, the prediction error e can be satisfactorily modeled as a zero-mean Laplacian distribution

$$p(x) = \frac{1}{2b} \exp\Big(- \frac{|x|}{b} \Big) \tag{5}$$

where $b > 0$ is a scale parameter.

Since the prediction is based on the original, unwatermarked pixels, all the prediction errors of the host image are available prior to the watermark embedding. Given the prediction error samples, the scale parameter b can be optimally estimated in maximum likelihood (ML) sense as

$$b = \frac{1}{N} \sum_{i=1}^{N} |e_i| \tag{6}$$

where e_i and N denote the ith sample and the total number of samples of the prediction error. It is assumed that $b > 0.5$, which has been verified experimentally to be true for most natural images. As mentioned previously, the number of pixels with overflow problem is generally negligible compared with the payload, especially in the low embedding rate region, and hence, the watermarking embedding capacity C can be estimated as the number of pixels in the inner region $\mathcal{I}$, which in mathematical form reads

$$\begin{aligned} C &= N \cdot \int_{T_l}^{T_r} p(x)dx \\ &= N - \frac{N}{2}(L+R) \end{aligned} \tag{7}$$

where $p(x)$ is the p.d.f of the Laplacian distribution given in (5),

$$L = \exp\left\{\frac{T_l}{b}\right\} \tag{8}$$

and

$$R = \exp\left\{-\frac{T_r}{b}\right\} \tag{9}$$

As $T_l \leq 0$ and $T_r \geq 0$, we have $0 < L, R \leq 1$.

Let us then calculate the distortion caused by the watermarking embedding and histogram shifting. From (4), we compute the distortion incurred by histogram shifting as

$$D_s = (I^w - I)^2 = \begin{cases} T_r^2 & \text{for} \quad e^w \geq T_r \\ T_l^2 & \text{for} \quad e^w < T_l \end{cases} \tag{10}$$

On the other hand, it follows from (3) that the distortion due to the expansion embedding can be expressed as

$$D_e = (e+w)^2 = e^2 + 2ew + w^2 \tag{11}$$

Combining (10) and (11) gives the expected distortion

$$\begin{aligned} &\mathbb{E}_{e,w}(D) \\ = \; & N \cdot T_r^2 \cdot \int_{T_r}^{+\infty} p(x)dx + N \cdot T_l^2 \cdot \int_{-\infty}^{T_l} p(x)dx \\ & + N \cdot \int_{T_l}^{T_r} p(x)\Big[e^2 + 2e\mathbb{E}_w(w) + \mathbb{E}_w(w^2)\Big]dx \end{aligned} \tag{12}$$

where the expectation is taken over e and w, and $\mathbb{E}_w(w)$ and $\mathbb{E}_w(w^2)$ denote the first and the second moments of the random variable w. Noticing (5) and the fact that w takes 0 and 1 with equal probabilities, (12) can be further simplified as

$$\begin{aligned} &\mathbb{E}_{e,w}(D) \\ = \; & N \cdot \Big[(b^2 - 0.5b)L\ln L + (b^2 + 0.5b)R\ln R \\ & -(b^2 - 0.5b + 0.25)L - (b^2 + 0.5b + 0.25)R \\ & +2b^2 + 0.5\Big] \end{aligned} \tag{13}$$

Upon getting the embedding capacity and the expected distortion, we can formulate the problem of finding the optimal T_l and T_r as

$$\begin{aligned} & \min \ \mathbb{E}_{e,w}(D) \\ & \text{subject to: } C \geq \tau \end{aligned} \tag{14}$$

where τ is the payload. Plugging in (7) and (13) into (14) gives

$$\begin{aligned} & \min_{L,R} \ \Big\{(b^2 - 0.5b)L\ln L + (b^2 + 0.5b)R\ln R \\ & -(b^2 - 0.5b + 0.25)L - (b^2 + 0.5b + 0.25)R \\ & +2b^2 + 0.5\Big\} \\ & \text{subject to: } L + R \leq 2(N-\tau)/N \end{aligned} \tag{15}$$

Since $x \ln x$ is convex because it is the 1-D entropy function [9], and $b > 0.5$ is a constant, the objective function of (15) is convex. In addition, the constraint of (15) is linear with respect to L and R. Therefore, the capacity-distortion optimization problem in (15) is convex, which permits efficient numerical solution.

A special case of the above capacity-distortion optimization is when $T_l = -T_r$, which leads to $L = R$. In this case, (15) becomes

$$\begin{aligned} & \min_{L} \ 2b^2 L\ln L - (2b^2 + 0.5)L + 2b^2 + 0.5 \\ & \text{subject to: } L \leq (N-\tau)/N \end{aligned} \tag{16}$$

It can be easily proved that the objective function of (16) is a decreasing function with respect to L, and hence, the optimal solution is given by

$$L^\star = (N-\tau)/N \tag{17}$$

which follows that

$$T_r^\star = -T_l^\star = \Big[(b \cdot \ln \frac{N}{N-\tau})\Big] \tag{18}$$

where $[x]$ returns the nearest integer of x.

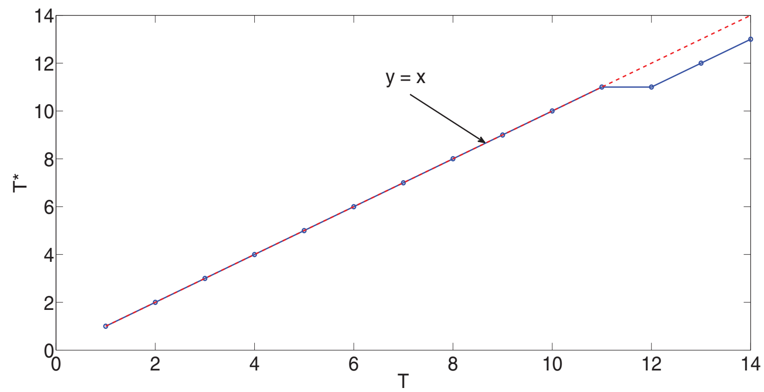

Fig. 1. Comparison between T and $T^{\star}$ for Lena image.

4. EXPERIMENTAL RESULTS

To verify the above analytical results, we compare the estimated capacity parameters with the actual optimal ones. For simplicity, we only consider the case that $T_r = -T_l = T$, where $T \geq 1$. Fig. 1 shows the comparison between the estimated parameter $T^{\star}$ obtained using (18) and the truly optimal T, for the test image Lena with different embedding payload. It can be seen that the estimation of the capacity parameter is very accurate when $T \leq 11$. When T becomes larger, the estimated $T^{\star}$ differs T by one. This is because when T is very large, the overhead of the location map becomes non-negligible, and hurts the accuracy of the capacity estimation in (7). We also have tried some other test images, and similar observations can be obtained.

5. CONCLUSIONS

The determination of the capacity parameters is an important issue in the PEE-based reversible image watermarking schemes. In this paper, we have shown that these capacity parameters can be optimally determined under a capacity-distortion optimization framework. Due to the convexity of the proposed capacity-distortion optimization problem, the capacity parameters can be obtained efficiently. Under some conditions that are commonly satisfied, even simple analytical solutions exist.

In fact, our proposed capacity-distortion optimization framework can be used to determine the capacity parameters when multiple histograms need to be handled. In such cases, the objective function and the constraint of (15) would become vector form; but the problem can be shown to be still convex. Therefore, the capacity parameters for multiple histograms can be efficiently found as well.

6. REFERENCES

[1] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber,“Lossless generalized-LSB data embedding,” *IEEE Trans. on Imag. Process.*, vol. 14, no. 2, pp. 253–266, 2005.

[2] M. U. Celik, G. Sharma, and A. M. Tekalp,“Lossless watermarking for image authentication: A new framework and an implementation,” *IEEE Trans. on Imag. Process.*, vol. 15, no. 4, pp. 1042–1049, 2006.

[3] D. M. Thodi and J. J. Rodriguez,“Expansion embedding techniques for reversible watermarking,” *IEEE Trans. on Imag. Process.*, vol. 16, no. 3, pp. 721–730, 2007.

[4] Y. Hu, H. K. Lee, and J. Li, “DE-based reversible data hiding with improved overflow location map,” *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 250–260, 2009.

[5] X. L. Li, B. Yang, and T. Y. Zeng, “Efficient reversible watermarking based on adaptive prediction-error expansion and pixel selection,” to appear in *IEEE Trans. on Imag. Process.*, 2011.

[6] J. Tian, “Reversible data embedding using a difference expansion,” *IEEE Trans. on Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, 2003.

[7] X. Wu and N. Memon, “Context-based, adaptive, lossless image codec,” *IEEE Trans. on Commun.*, vol. 45, pp. 437–444, 1997.

[8] M. J. Weinberger, G. Seroussi, and G. Sapiro, “The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS,” *IEEE Trans. on Imag. Process.*, vol. 9, no. 8, pp. 1309–1324, 2000.

[9] S. Boyd and L. Vandenberghe, “Convex Optimization,” Cambridge University Press, 2003.