Reversible Data Hiding for Sparse Histogram Images

Masaaki FUJIYOSHI and Hitoshi KIYA
Dept. Inform. & Commun. Syst., Tokyo Metropolitan Univ., Hino, Tokyo 191–0065, Japan
mfujiyoshi@m.ieice.org, kiya@sd.tmu.ac.jp

Abstract

This paper reports a primary study on reversible data hiding (RDH) for sparse histogram (SH) images. This paper studies a histogram modification-based RDH (HM-RDH) method in which pixel values between the maximum frequency and the zero frequency in the histogram of an original image are increased (or decreased) to hide data to the image. Whereas most of the HM-RDH methods project hidden information to pixel values based on binary pulse modulation (PPM) manner due to dense histogram property, the method studied in this paper uses M-ary PPM to increase the hidden data capacity, thanks to the SH property of images. As an example of SH images, high bit depth images are used to analyze the sparseness of the histogram as well as to evaluate the capacity and watermarked image quality.

keywords: Image processing, Medical information systems, Military communication, Indexes

1. Introduction

Data hiding technology has been diligently studied, for not only security-related problems [1, 2], in particular, intellectual property rights protection of digital contents [3], but also non security-oriented [1,4] such as broadcast monitoring [5]. A data hiding technique embeds data into a target signal referred to as the *original* signal. It, then, generates a slightly distorted signal that is referred to as a *stego* signal. Many of data hiding techniques extract hidden data but leave a stego signal as it is [6].

In military and medical applications, restoration of the original signal as well as extraction hidden data are desired [7]. *Reversible* data hiding (RDH) techniques that restore the original image have been proposed [7–12]. Among many RDH methods, this paper focuses *histogram modification*-based RDH (HM-RDH) methods [9–12]. HM-RDH methods modify the histogram of an original image [9] or a processed image [10–12] to hide data into the image.

On the other hand, high bit depth (HBD) images such as high dynamic range images and its image processing are in progress [13–15]. Though HBD images have different properties from ordinary images, RDH methods for HBD images have not been studied well yet.

This paper investigates the performance of a simple HM-RDH method for HBD images. First, the histogram of HBD images are reviewed to clarify the property of the images. Then, an effective way to apply a simple HM-RDH method to HBD images is considered, and the hidden data capacity and the image quality of stego images are evaluated.

(a) Bride (12 bits/pixel).

(b) Cars (16 bits/pixel).

Figure 1: Examples of high bit depth images. Both images have 512×512 pixels.

2. Preliminaries

This section reviews the histogram of the HBD images to clarify the property of the HBD images and mentions the original HM-RDH method.

2.1. Histogram of HBD Images

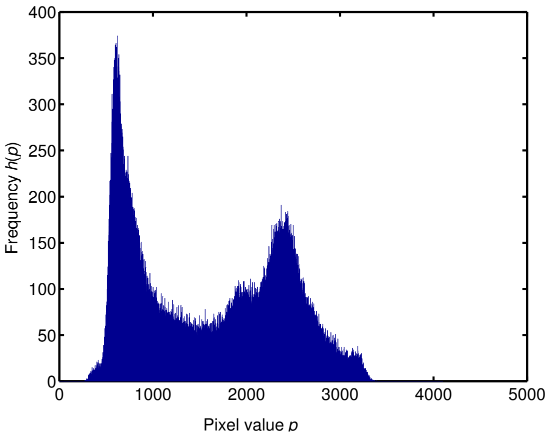
Figure 1 shows two of 512×512-sized natural HBD images which the images have 12-bits quantized pixels from zero to 4095 and 16-bits quantized pixels from zero to 65535, respectively. In addition, the histogram of images shown in Fig. 1 are available in Fig. 2.

It is confirmed that these HBD images have the sparse histogram (SH) in which the number of the nonzero frequency bins is small and the number of the zero frequency bins is quite large. Moreover, it is found that the smallest luminance from zero to a certain level are not used in these images. These imbalanced histogram properties hold in many natural HBD images. This paper focuses the latter property which the lowest luminance levels are not used to imperceptibly hide data to HBD images.

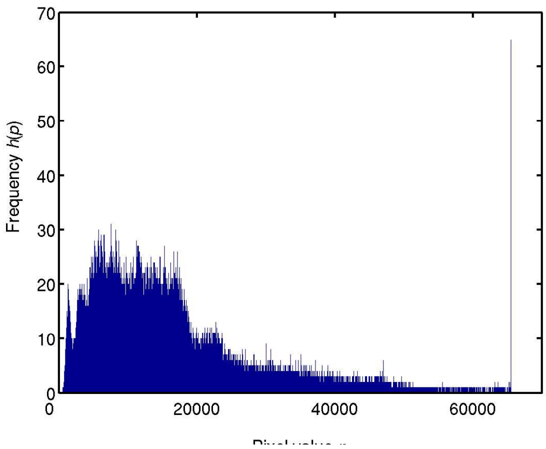
2.2. HM-RDH

Though several improved HM-RDH methods have been proposed to increase the hidden data capacity in which the methods shift the histogram of processed image, e.g., divided image tiles [10] or differences between pixels [11] or between a pixel and a prediction error [12], and so on, this paper focuses the simple original HM-RDH [9].

The method derives the histogram of an original image



(a) Bride (3053 nonzero bins over 4096 levels).



(b) Cars (38629 nonzero bins over 65536 levels).

Figure 2: Histogram of the images shown in Fig. 1.

in which the frequency of pixel value $p = 0, 1, \ldots, 2^Q - 1$ is represented by $h(p)$, where Q is the quantization bits of pixels. The method, then, finds pixel value $p_{\max}$ corresponding to maximum frequency $h_{\max} = h(p_{\max})$ and pixel value p_0 corresponding to zero frequency and satisfying $p_0 < p_{\max}$, i.e.,

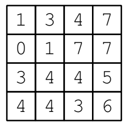
$$h_{\max} = \max_p h(p), \quad p = 0, 1, \ldots, 2^Q - 1, \tag{1}$$

$$p_{\max} = \arg\max_p h(p), \tag{2}$$

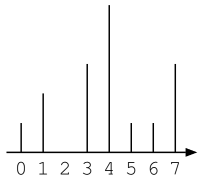
$$h(p_0) = 0, \quad p_0 < p_{\max}. \tag{3}$$

The method subtracts one from all pixels having pixel values from p_0 to $(p_{\max} - 1)$ to prepare for data hiding. According to data bit to be hidden, the pixel having $p_{\max}$ is decreased by one to distinguish a hidden data bit. So, the hidden data capacity is $h_{\max}$. Figure 3 shows an example of data hiding by the HM-RDH in which $p_{\max} = 4$, $h(p_{\max}) = 5$, and $p_0 = 2$. In this example, 5-bits binary data '10110' is hidden to a 3-bits quantized original image.

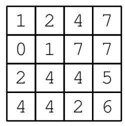
In the next section, a modified HM-RDH method is proposed for SH images. The proposed method utilizes the sparseness of the lower bins of histogram to increase the hidden data capacity and to simultaneously reduce the distortion introduced by hiding data. Moreover, to utilize such successive space, M-ary pulse position modulation (PPM) is introduced to further increase the hidden data capacity.



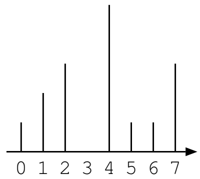
(a) Original image.



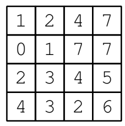
(b) Original histogram.



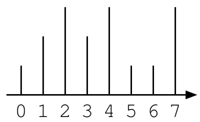
(c) Shifted image.



(d) Shifted histogram.



(e) Stego image.



(f) Stego histogram.

Figure 3: An example of data hiding by the HM-RDH [9]. Five bits data '10110' is hidden to a 3-bits quantized image.

3. Proposed Method

This section proposes a HM-RDH method for SH images. The proposed method hides data to successive zero frequencies of the histogram for smaller pixel values, based on the properties described in Sect. 2.1. This strategy increases the hidden data capacity as well as reduces the distortion in stego images.

3.1. Algorithms

3.1.1. Data Hiding Algorithm

As the conventional HM-RDH method described in Sect. 2.2, the proposed method firstly derives the histogram of an original SH image. Then, the method finds $p_{\max}$ by Eqs. (1) and (2). This method also finds the longest successive zero frequencies of the histogram from $p_{0_{\min}}$ to $p_{0_{\max}}$ which are smaller than $p_{\max}$, i.e.,

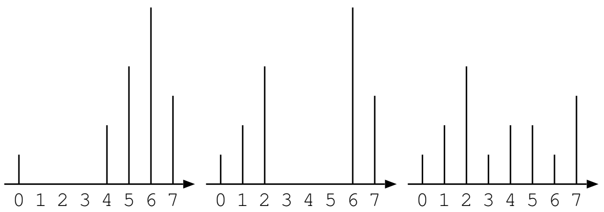
$$h(\pi) = 0, \qquad \forall \pi : p_{0_{\min}} \leq \pi \leq p_{0_{\max}} < p_{\max}. \tag{4}$$

It is noted that this method assumes

$$p_{0_{\min}} \leq p_{0_{\max}} < p_{\max}, \tag{5}$$

and this assumption is reasonable from the description given in Sect. 2.1.

The proposed method, then, subtracts r in pixel values from pixels having value from $(p_{0_{\max}} + 1)$ to $(p_{\max} - 1)$,



(a) Original. (b) Modified. (c) Stego.

Figure 4: An example of the histogram modification in the proposed method ($p0_{\min} = 1$, $p0_{\max} = 4$, $p_{\max} = 6$, $h(p_{\max}) = 6$, $b = 2$, and $r = 3$). The hidden data capacity is $bh(p_{\max}) = 36$.

where

$$r = 2^b - 1, \qquad b = \left\lfloor \log_2 \left(p0_{\max} - p0_{\min} + 2\right) \right\rfloor . \qquad (6)$$

The histogram of the image now has $(b-1)$ successive zero frequencies followed by the maximum frequency. According to a b-bits data substring to be hidden, the pixel value of a pixel having $p_{\max}$ is change to the value between $(p_{\max} - r)$ to $p_{\max}$. Through this process, the proposed method embeds $bh_{\max}$-bits data to the original SH image. Figure 4 show an example of the data hiding of the proposed method.

3.1.2. Hidden Data Extraction and Image Recovery

The proposed method requires $p0_{\min}$, $p0_{\max}$, and $p_{\max}$ to extract hidden data and to recovery the original image from the stego image. These three parameters are stored and/transmitted. Once these parameters are given with the corresponding stego image, this method easily extracts hidden data and recovers the original image.

By Eq. (6), b and r are derived. With $p_{\max}$ and r, a b-bits data portion is extracted from one pixel having the value between $(p_{\max} - r)$ and $p_{\max}$. All pixels with the pixel value between $(p_{\max} - r)$ and $p_{\max}$ are then set to those pixel values to $p_{\max}$, and this process recovers the histogram shown in Fig. 4 (b). The method, then, easily recovers the original pixel values by using $p0_{\min}$ and r, and its corresponding histogram is the identical to which shown in Fig. 4 (a).

3.2. Features

This section summarizes the features of the proposed method; An utilization of the properties of SH images and data hiding utilizing Mary PPM.

The proposed method modified pixels with lower pixel values which have zero frequency bins in the histogram of the image. As mentioned in Sect. 2.1, HBD images have a large number of zero frequencies in lower pixel values in the histogram. So, this strategy of the proposed method helps the proposed method to increase the hidden data capacity. Simultaneously, modified pixel values are quite small, and this achieves the high imperceptively of the proposed method.

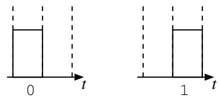


Figure 5: Binary pulse position modulation. The pulse position differentiated by time shifting carries a bit.

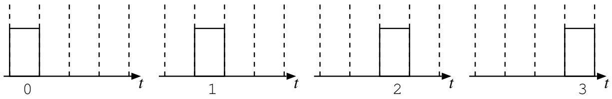


Figure 6: An example of M-ary pulse position modulation in which $M = 2$. Data with $\log_2 M$-bits are carried with a single pulse over M of possible positions.

The proposed method modulates hidden data bits by M-ary PPM to further increase the hidden data capacity. The conventional HM-RDH method described in Sect. 2.2 decides the pixel value according to a hidden data bit That is, it can recognize that hidden data are treated as a signal modulated by binary PPM which is shown in Fig. 5. So, the conventional method hides a bit to a pixel to be watermarked. On the other hand, the data hiding based on the M-ary PPM (c.f., Fig. 6) which the hiding approach is introduced for some RDH for palette images [16, 17], $\log_2 M$-bits data is able to hidden to a pixel to be watermarked.

4. Experimental Results

The proposed method hides data to 512×512-sized grayscale images shown in Fig. 1; (a): 12-bits quantized pixel values from zero to 4095, $p_{\max} = 619$, $h_{\max} = 374$, $p0_{\min} = 0$, and $p0_{\max} = 272$, and (b): 16-bits quantized pixel values from zero to 65535, $p_{\max} = 65535$, $h_{\max} = 65$, $p0_{\min} = 0$, and $p0_{\max} = 538$.

Table 1 shows the hidden data capacity and the stego image quality in terms of the peak signal-to-noise ratio (PSNR). The proposed method improves the hidden data capacity b times, where $b = 8$ for the image shown as Fig. 1 (a) and $b = 9$ for the image shown as Fig. 1 (b).

Figure 7 shows examples of stego images for images shown in Fig. 1. It is confirmed that it is difficult to perceive the difference between the original and stego images. This thanks to that the modification for data hiding is applied to pixels having the lowest value, i.e., pixels in quite dark regions in which the distortion is imperceptible.

Table 1: The hidden data capacity and the stego image quality.

Image	Capacity [bits]	PSNR [dB]
Bride	2992	33.34
Cars	585	42.16

(a) Bride (12 bits/pixel) with 2992 hidden data bits. PSNR: 33.34 dB

(b) Cars (16 bits/pixel) with 585 hidden data bits. PSNR: 42.16 dB.

Figure 7: Examples of stego images.

5. Conclusions

This paper proposes a histogram modification-based reversible data hiding (HM-RDH) method for sparse histogram (SH) images. The proposed method hides data by utilizing the sparseness of the histogram of SH images. This strategy increases the hidden data capacity and keeps the imperceptively of stego images. To further increase the hidden data capacity, the M-ary pulse position modulation (PPM) is introduced to the proposed method, whereas the conventional HM-RDH method uses the binary PPM.

Further works include the performance investigation in using larger zero frequencies of higher pixel values in the histogram to hide data and the development of parameter memorization-free HM-RDH.

Acknowledgment

This work has been partly supported by the Grant-in-Aid for Young Scientists (B), No.20336522, from the Ministry of Education, Culture, Sports, Science and Technology of Japan and from the Japan Society for the Promotion of Science.

References

[1] G.C. Langelaar, I. Setyawan, and R.L. Lagendijk, "Watermarking digital image and video data," *IEEE Signal Process. Mag.*, vol.17, no.5, pp.20–46, Sep. 2000.

[2] M. Barni, "What is the future for watermarking? (Part I)," *IEEE Signal Process. Mag.*, vol.20, no.5, pp.55–59, Sep. 2003.

[3] C.-C.J. Kuo, T. Kalker, and W. Zhou, eds., "Digital rights management," *IEEE Signal Process. Mag.*, vol.21, no.2, pp.11–117, Mar. 2004.

[4] M. Barni, "What is the future for watermarking? (Part II)," *IEEE Signal Process. Mag.*, vol.20, no.6, pp.53–59, Nov. 2003.

[5] T. Tachibana, M. Fujiyoshi, and H. Kiya, "A removable watermarking scheme retaining the desired image quality," in *Proc. IEEE ISPACS*, 2003, pp.538–542.

[6] M. Fujiyoshi, Y. Seki, H. Kobayashi, and H. Kiya, "Modulo arithmetic-based image watermarking and its theoretical analysis of image-quality," in *Proc. IEEE ICIP*, 2005, pp.969–972.

[7] R. Caldelli, F. Filippini, and R. Becarelli, "Reversible watermarking techniques: an overview and a classification," *EURASIP J. Inf. Security*, vol.2010, 2010.

[8] M. Fujiyoshi, S. Sato, H.L. Jin, and H. Kiya, "A location-map free reversible data hiding method using block-based single parameter," in *Proc. IEEE ICIP*, 2007, pp.257–260.

[9] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Trans. Circuits Syst. Video Technol.*, vol.16, pp.354–362, Mar. 2006.

[10] M. Fallahpour, D. Megias, and M. Ghanbari, "High capacity, reversible data hiding in medical images," in *Proc. IEEE ICIP*, 2009, pp.4241–4244.

[11] C.-C. Lin, W.-L. Tai, and C.-C. Chang, "Multilevel reversible data hiding based on histogram modification of difference images," *Pattern Recognition*, vol.41, pp.3582–3591, Dec. 2008.

[12] P. Tsai, Y.-C. Hu, and H.-L. Yeh, "Reversible image hiding scheme using predictive coding and histogram shifting," *Signal Process.*, vol.89, pp.1129–1143, Jun. 2009.

[13] E. Reinhard, W. Heidrich, P. Debevec, S. Pattanaik, G. Ward, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, 2nd ed. Morgan Kaufmann, Jun. 2010.

[14] B. Hoefflinger, ed., *High-Dynamic-Range (HDR) Vision*, SSAM 26. Springer, Feb. 2007.

[15] Y. Bandoh, G. Qiu, M. Okuda, S. Daly, T. Aach, and O. Au, "Recent advances in high dynamic range imaging technology," in *Proc. IEEE ICIP*, 2010, pp.3125–3128.

[16] L. Hongmei, Z. Zhefeng, H. Jiwu, H. Xialing, and Y.Q. Shi, "A high capacity distortion-free data hiding algorithm for palette image," in *Proc. IEEE ISCAS*, 2003, pp.II–916–II–919.

[17] N.A. Saleha, H.N. Boghdadya, S.I. Shaheenb, and A.M. Darwish, "High capacity lossless data embedding technique for palette images based on histogram analysis," *Digital Signal Process.*, vol.20, pp.1629–1636, Dec. 2010.