A LOCATION-MAP FREE REVERSIBLE DATA HIDING METHOD USING BLOCK-BASED SINGLE PARAMETER

Masaaki FUJIYOSHI, Shuji SATO†, Hong Lin JIN*, and Hitoshi KIYA**

*Dept. of Information and Communication Systems Engineering, Tokyo Metropolitan University
6–6 Asahigaoka, Hino-shi, Tokyo 191–0065, Japan
†Dept. of Computer Sciences, Tsukuba University
1–1–1 Tennodai, Tsukuba-shi, Ibaraki 305–8573, Japan

ABSTRACT

This paper proposes a reversible data hiding method that embeds an L-level data sequence to images in the spatial domain. Though reversible data hiding once distorts the image to hide data into it, the distorted image is completely separated to the original image and the hidden data. The proposed method uses only one parameter to embed and extract data, and it extracts data without any location map. In addition, it can control embedding capacity according to payload for suppression of embedding distortion. Simulation results show the effectiveness of the proposed method.

Index Terms— Image processing, Medical information systems, Military communication, Indexes

(a) A pixel block. (b) Placement of pixel blocks (filled pixels represents pixels where data element is embedded).

Fig. 1. A pixel block and its placement.

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 [7–9]. *Reversible* data hiding techniques that restore the original image have been proposed [7–12]. Reversible data hiding is classified into two groups [7]; One embeds data for restoration as well as data to be hidden [7, 10–12], the other only embeds data to be hidden [8, 9]. This paper focuses the latter.

A novel reversible data hiding method that only hides data to be embedded in the spatial domain is proposed in this paper. The proposed method embeds and extracts data by using only one threshold parameter that is derived based on simple statistics of pixels. Moreover, this method extracts hidden data without any *location map* which indicates the pixels where data are actually hidden. Furthermore, by utilizing several degrees of freedom, it controls the embedding *capacity* and the quality of a stego image.

2. CONVENTIONAL METHODS

Several conventional methods *replace* the original state, e.g., pixel values [7, 10] or transformed coefficients [11], to the new state to hide data. These methods, thus, have to embed not only data to be hidden but also data for restoration of the original image and/or extraction of the hidden data.

Other conventional methods that do not have to embed data for restoration also have disadvantages: a location map indicating pixel or coefficient positions where data are embedded is required [8, 12]. Two parameters are used and inverse transformations often move pixel values to the outside of the dynamic range of pixel values [9].

In the next section, a reversible data hiding method that embeds data in the spatial domain is proposed. It does not embed either data for restoration or location map. Though only one parameter is used in this method, hidden data and the original image are completely restored from a stego image. Moreover, it controls the embedding capacity by multi-degree-of-freedom.

3. PROPOSED METHOD

A $X \times Y$-sized gray scale image in which each pixel is represented by K bits, i.e., $\mathbf{f} = \{f(x,y) \mid 0 \le f(x,y) \le 2^K - 1, 0 \le x \le X-1, 0, \le y \le Y-1\}$, is assumed as an original image. The proposed method embeds N-length L-level sequence $\mathbf{w} = \{w_n \mid w_n \in \{0, 1, \ldots, L-1\}, n = 0, 1, \ldots, N-1\}$ to $\mathbf{f}$ in the spatial domain. Here, the *payload* is N. The pixel in which data element w_n is hidden in the central pixel of a 3×3-sized diamond-shape block, g_b, $(b = 0, 1, \ldots, B-1)$, as shown in Fig. 1 (a), and blocks are overlapped as shown in Fig. 1 (b). This method, thus, embeds elements up to

$$B = \left\lfloor \frac{X-1}{2} \right\rfloor \left\lfloor \frac{Y-1}{2} \right\rfloor + \left\lfloor \frac{|X-2|}{2} \right\rfloor \left\lfloor \frac{|Y-2|}{2} \right\rfloor, \qquad (1)$$

where $\lfloor r \rfloor$ rounds real-value r to the nearest integer towards minus infinity. Here, B is the *ideal* embedding capacity of $\mathbf{f}$. However, because some g_b cannot convey w_n, *actual* embedding capacity M that is image-dependent holds $M \le B$. Thus, payload N must be $N \le M$.

The block diagram of the proposed method is shown in Fig 2. First, original image f is divided into B of 3×3-sized diamond-shaped overlapped blocks, and statistics are derived in each block. An image dependent parameter s is obtained based on these statis-

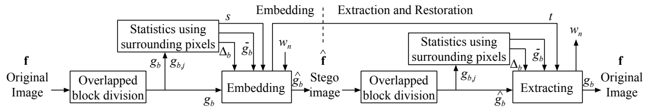


Fig. 2. Proposed method.

tics, and s is the maximum value of possible threshold t. In each block, this method decides whether g_b is capable to convey w_n or not, i.e., *embeddable* or not, based on t, and stego image $\hat{\mathbf{f}} = \left\{\hat{f}(x,y) \mid 0 \leq \hat{f}(x,y) \leq 2^K - 1\right\}$ is obtained.

3.1. Derivation of s

The algorithm to derive parameter s from original image $\mathbf{f}$ is following.

1. $b := 0$.
2. From surrounding pixels, $g_{b,p}$ ($p = 0,1,2,3$), the average is obtained by Eq. (2). Difference between g_b and $\bar{g}_b$ is derived.

$$\bar{g}_b = \left\lfloor \frac{1}{8} \sum_{p=0}^{3} g_{b,p} \right\rfloor \tag{2}$$

$$d_b = g_b - \bar{g}_b \tag{3}$$

3. Δ_b is obtained by Eq. (4).

$$\Delta_b = \begin{cases} g_{\max,b} - \bar{g}_b, & \mathrm{sgn}\,(d_b) \geq 0 \\ g_{\min,b} - \bar{g}_b, & \mathrm{sgn}\,(d_b) < 0 \end{cases}, \tag{4}$$

where

$$g_{\max,b} = \max_p g_{b,p} \quad \text{and} \quad g_{\min,b} = \min_p g_{b,p},$$

$$\mathrm{sgn}(r) = \begin{cases} 1, & r > 0 \\ 0, & r = 0 \\ -1, & r < 0 \end{cases}.$$

4. Parameter s_b that is a candidate of s is derived by Eq. (5).

$$s_b = \begin{cases} |\Delta_b|, & \bar{g}_b + Ld_b < 0 \quad \text{or} \quad 2^K - L < \bar{g}_b + Ld_b \\ \infty, & \text{others} \end{cases}. \tag{5}$$

5. $b := b + 1$. Continue to Step 2 unless $b = B$.
6. The minimum of s_b's becomes s. That is,

$$s = \min_b s_b. \tag{6}$$

3.2. Data Embedding

This algorithm decides whether g_b is embeddable based on s described above. L-level data element w_n is hidden into embeddable g_b.

1. $t := s$, $b := 0$, and $n := 0$.
2. By Eq. (7), $\hat{g}_b$, the pixel with hidden data, is derived from embeddable g_b.

$$\hat{g}_b = \begin{cases} \bar{g}_b + Ld_b + w_n, & |\Delta_b| < t \\ g_b, & \text{others} \end{cases}. \tag{7}$$

3. If $|\Delta_b| < t$, $n := n + 1$.
4. $b := b + 1$. Continue Step 2 until $b = B$.
5. Stego image $\hat{\mathbf{f}}$ is generated.

3.3. Data Extraction and Image Restoration

The following algorithm is applied to stego image $\hat{\mathbf{f}}$ to extract hidden data $\mathbf{w}$ and restore original image $\mathbf{f}$. Threshold parameter t is transmitted from the embedding side to this side in this method, but neither the original image nor any location map is required.

1. $b := 0$, $n := 0$.
2. Δ_b is obtained by Eq. (8).

$$\Delta_b = \begin{cases} g_{\max,b} - \bar{g}_b, & \mathrm{sgn}\,(\hat{g}_b - \bar{g}_b) \geq 0 \\ g_{\min,b} - \bar{g}_b, & \mathrm{sgn}\,(\hat{g}_b - \bar{g}_b) < 0 \end{cases}. \tag{8}$$

3. Data element w_n is extracted by the following equation, if $|\Delta_b| < t$.

$$w_n = (\hat{g}_b - \bar{g}_b) \bmod L \tag{9}$$

4. Pixel g_b of the original image is restored by Eq. (10).

$$g_b = \begin{cases} \frac{\hat{g}_b + (L-1)\bar{g}_b - w_n}{L}, & |\Delta_b| < t \\ \hat{g}_b, & \text{others} \end{cases}. \tag{10}$$

5. If $|\Delta_b| < t$, $n := n + 1$.
6. $b := b + 1$. Continue to Step 2 unless $b = B$.
7. N-length L-level data sequence $\mathbf{w}$ and original image $\mathbf{f}$ are obtained.

3.4. Features

3.4.1. Reversible

The proposed method restores the original image as well as extracts hidden data from a stego image without embedding any extra data for restoration. To realize these properties, the following are required:

- All stego pixels, $\hat{g}_b$'s, are in the dynamic range of pixels, i.e., must not either *overflow* or *underflow*.
- Stego pixel $\hat{g}_b$ is completely split into hidden data value w_n and original pixel g_b.

The former is represented by

$$0 \leq \hat{g}_b \leq 2^K - 1, \quad \forall b. \tag{11}$$

From Eq. (7) and $w_n \in \{0, 1, \ldots, L-1\}$, Eq. (11) is satisfied by skipping g_b's that satisfy any of

$$\bar{g}_b + Ld_b > 2^K - L \quad \text{or} \quad \bar{g}_b + Ld_b < 0 \tag{12}$$

in an embedding process. Since g_b that satisfies Eq. (12) tends to have large $|\Delta_b|$ or s_b in Eq. (5), the minimum s_b becomes s, that works as the maximum value of possible threshold t, to hold Eq. (11). Consequently, this method guarantees Eq. (11).

The latter is achieved as following. Eq. (7) is transformed into

$$\hat{g}_b = \bar{g}_b + L(g_b - \bar{g}_b) + w_n = Lg_b - (L-1)\bar{g}_b + w_n, \tag{13}$$

and it results in Eq. (10). In Eq. (10), $\hat{g}_b$ is the stego pixel itself and $\bar{g}_b$ is given by Eq. (2), and those are known in any extraction process.

Since surrounding pixels $g_{b,p}$ in a stego image are the same as that in the original image, $\bar{g}_b$ in the original and a stego images are the same. The proposed method, thus, completely separates $\hat{g}_b$ into g_b and w_n, if w_n is obtained. The mechanism to extract w_n from $\hat{g}_b$ is described in the next section.

3.4.2. *Location Map-Free Data Extraction*

A data hiding method that requires no reference image, e.g., the original image, to extract hidden data is referred to as a *oblivious* method. The proposed method is oblivious, it, moreover, extracts hidden data from stego image $\hat{\mathbf{f}}$ without any knowledge of the position of pixels that actually convey data, i.e., without any location map.

The proposed method introduces modulo arithmetic to obliviously extract w_n from $\hat{g}_b$. Eq. (7) is transformed into

$$\hat{g}_b - \bar{g}_b = Ld_b + w_n. \tag{14}$$

Since $w_n \in \{0, 1, \ldots, L-1\}$,

$$(Ld_b + w_n) \bmod L = w_n \tag{15}$$

is guaranteed. Eq. (9), thus, is guaranteed by Eqs. (14) and (15). Consequently, the proposed method is oblivious.

In Eq. (7), the proposed method leaves g_b as is, unless g_b is embeddable. That is, the proposed method has to know whether stego pixel $\hat{g}_b$ conveys w_n or not to extract $\mathbf{w}$ completely. According to Eqs. (3), (4), and (7), the positive and negative sign of d_b that is based on original pixel g_b is required to obtain Δ_b for knowing the position of embeddable pixels. The proposed method solves this problem as following. Under the conditions that $w_n \in \{0, 1, \ldots, L-1\}$ where L is a positive integer and d_b is an integer,

$$\begin{array}{lcl} \operatorname{sgn}(d_b) \geq 0 & \Rightarrow & \operatorname{sgn}(Ld_b + w_n) \geq 0 \\ \operatorname{sgn}(d_b) < 0 & \Rightarrow & \operatorname{sgn}(Ld_b + w_n) < 0 \end{array} \tag{16}$$

is guaranteed. From Eq. (7),

$$\hat{g}_b - \bar{g}_b = \begin{cases} \bar{g}_b + Ld_b + w_n - \bar{g}_b = Ld_b + w_n, & |\Delta_b| < s \\ g_b - \bar{g}_b = d_b, & \text{others} \end{cases} \tag{17}$$

is derived. Thus, from Eqs. (16) and (17),

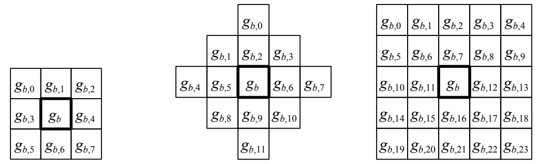
$$\begin{array}{lcl} \operatorname{sgn}(d_b) \geq 0 & \Rightarrow & \operatorname{sgn}(\hat{g}_b - \bar{g}_b) \geq 0 \\ \operatorname{sgn}(d_b) < 0 & \Rightarrow & \operatorname{sgn}(\hat{g}_b - \bar{g}_b) < 0 \end{array} \tag{18}$$

is introduced, and this property is used in Eq. 8 to obtain Δ_b without original pixel g_b. Consequently, the proposed method frees itself from location maps.

3.4.3. *Capacity Controllable*

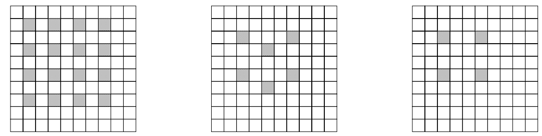
The proposed method has several degrees of freedom. It utilizes degrees to control actual embedding capacity M according to payload N, and this control suppresses distortion to stego image $\hat{\mathbf{f}}$. Two stages for controlling M are described in this section. One is the size and the shape of blocks, and the other is decreasing threshold t.

Employing other block rather than a 3×3-sized diamond-shaped can change the ideal capacity, B, and it also controls actual capacity M. Fig. 3 shows three possible blocks, and corresponding placements are shown in Fig. 4. For those blocks, ideal capacity B's are summarized in Table 1.



(a) 3×3 square. (b) 5×5 diamond. (c) 5×5 square.

Fig. 3. An example of pixel blocks.



(a) 3×3 square. (b) 5×5 diamond. (c) 5×5 square.

Fig. 4. Placement of pixel blocks.

The other stage decreases threshold t according to payload N. If actual capacity M is far larger than N under the condition s is used as t, decreasing t gets M closer to N. This strategy is reasonable from the perspective of the quality of stego images. The distortion introduced to stego image $\hat{\mathbf{f}}$ is derived from Eq. (7) as

$$\hat{g}_b - g_b = (\bar{g}_b + Ld_b + w_n) - g_b = (L-1)\,d_b + w_n. \tag{19}$$

Eq. (19) implies hiding data element into embeddable g_b whose corresponding $|d_b|$ is small is desired to suppress the embedding distortion. From Eqs. (3) and (19), g_b in a flat area, for example, is good for embedding data. In such flat area, Δ_b is also small. Consequently, it is determined that suppressing s reduces the embedding distortion rather than a lazy way: Using s as t, leaving $(M-N)$ of embeddable g_b's as is, and a location map or a further algorithm to decide the positions.

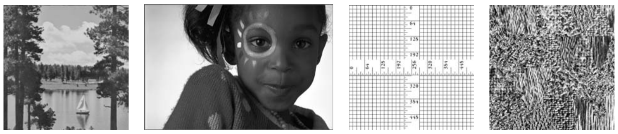
4. EXPERIMENTAL RESULTS

Actual capacity M's are investigated using 69 natural images including "sailboat," and "kodak15" from RPI-CIPR image database and nine artificial images including "ruler.512" and "texmos3b.p512" from USC-SIPI image database (Fig. 5). All images are eight bits quantized, i.e., $K = 8$. For its simplicity, L is set to two, i.e., binary data sequence is embedded to images in this paper.

Table 2 shows that actual capacity M is controlled by choosing the size and the shape of blocks. The table also shows that for images having flat area in it, e.g., "kodak15" and "ruler.512," gives quite small Δ_b, then s results in being quite small and M is remarkably decreased. In "texmos3b.p512" consisting of high frequency components, d_b tends to be large, and simultaneously surrounding pixels are similar each other, i.e., Δ_b is small. This also numerously decreases M. An improvement to increase M for such images will

Table 1. Pixel blocks and those ideal capacities B's.

Pixel block	B
3×3 square	$\left\lfloor \frac{X-1}{2} \right\rfloor \left\lfloor \frac{Y-1}{2} \right\rfloor$
5×5 diamond	$\left\lfloor \frac{X-1}{4} \right\rfloor \left\lfloor \frac{\lvert Y-2\rvert}{3} \right\rfloor + \left\lfloor \frac{\lvert X-3\rvert}{4} \right\rfloor \left\lfloor \frac{\lvert Y-3\rvert}{3} \right\rfloor$
5×5 square	$\left\lfloor \frac{\lvert X-2\rvert}{3} \right\rfloor \left\lfloor \frac{\lvert Y-2\rvert}{3} \right\rfloor$



(a) sailboat. (b) kodak15. (c) ruler.512. (d) texmos3b.p512.

Fig. 5. Images for evaluation ($X = Y = 512$ for (a), (c), and (d). $X = 768$, $Y = 512$ for (b)).

Table 2. Actual capacity M.

Image	3×3		5×5	
	diamond	square	diamond	square
sailboat	130050	64851	41837	25240
kodak15	0	0	0	0
ruler.512	0	0	0	0
texmos.3b.p512	0	3	1	0

be reported [13].

Figure 6 shows that actual capacity M is controlled by decreasing t for image "sailboat" with using 3×3-sized diamond-shaped block. Moreover, Fig. 7 shows that suppressing t reduces distortion of stego image $\hat{\mathbf{f}}$ rather than using s as t and leaving $(N - M)$ embeddable g_b's as is.

5. CONCLUSIONS

This paper has proposed a reversible data hiding method in the spatial domain. The proposed method uses only one parameter to embed and extract data, and extracts data without any location map. Moreover, it has several degrees of freedom, and this characteristic can control the embedding capacity.

Further works include information-theoretical analysis [14] of the proposed method.

ACKNOWLEDGMENT

This work has been partly supported by the Grant-in-Aid for Young Scientists (B), No.17700119, from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

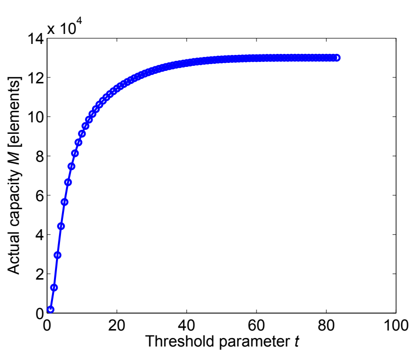


Fig. 6. Controlling actual capacity M by decreasing s (lena-y. 3×3-sized diamond-shaped blocks. Initially $s = 19$ and $M = 120748$).

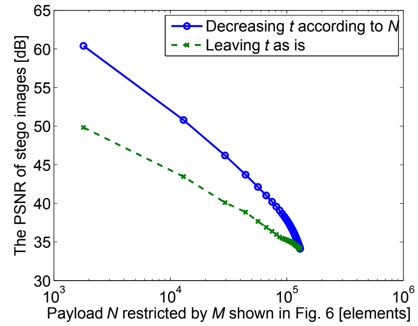


Fig. 7. Decreasing s according to N improves the PSNR of stego images (lena-y. 3×3-sized diamond-shaped blocks. N is limited to M shown in Fig. 6).

REFERENCES

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

[2] M. Barni, F. Pérez-González, M.L. Miller, F. Bartolini, J.J. Eggers, I.J. Cox, P. Moulin, N. Memon, and T. Kalker, "What is the future for watermarking? (Part I)," *IEEE Signal Processing Mag.*, vol.20, no.5, pp.55–59, 2003.

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

[4] M. Barni, F. Pérez-González, M.L. Miller, F. Bartolini, J.J. Eggers, I.J. Cox, P. Moulin, N. Memon, and T. Kalker, "What is the future for watermarking? (Part II)," *IEEE Signal Processing Mag.*, vol.20, no.6, pp.53–59, 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] M.U. Celik, G. Sharma, A.M. Tekalp, and E. Saber, "Lossless generalized-LSB data embedding," *IEEE Trans. Image Processing*, vol.14, pp.253–266, Feb. 2005.

[8] C. De Vleeschouwer, J.-F. Delaigle, and B. Macq, "Circular interpretation of bijective transformations in lossless watermarking for media asset management," *IEEE Trans. Multimedia*, vol.5, pp.97–105, Mar. 2003.

[9] F. Uchiyama and A. Miyazaki, "An image watermarking method using the lifting wavelet transform," *IEICE Technical Report*, vol.105, no.149, pp.79–84, 2005.

[10] J. Fridrich, M. Goljan, and R. Du, "Lossless data embedding — New paradigm in digital watermarking," *EURASIP J. Applied Signal Processing*, vol.2002, pp.185–196, Feb. 2002.

[11] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Trans. Circuits Syst. Video Technol.*, vol.13, pp.890–896, Aug. 2003.

[12] D. Coltuc and J.-M. Chassery, "High capacity reversible watermarking," in *Proc. IEEE ICIP*, 2006, pp.2656–2568.

[13] H.L. Jin, M. Fujiyoshi, and H. Kiya, "Improvement in the image quality and the capacity for the lossless data hiding," in *Proc. ITC-CSCC*, 2007, accepted.

[14] T. Kalker and F.M. Willems, "Capacity bounds and constructions for reversible data-hiding," in *Proc. IEEE DSP*, 2002, pp.71–76.