

Reversible Watermarking Based on Invariability and Adjustment on Pixel Pairs

Shaowei Weng, Yao Zhao, *Member, IEEE*, Jeng-Shyang Pan, *Member, IEEE*, and Rongrong Ni

Abstract—A novel reversible data hiding scheme based on invariability of the sum of pixel pairs and pairwise difference adjustment (PDA) is presented in this letter. For each pixel pair, if a certain value is added to one pixel while the same value is subtracted from the other, then the sum of these two pixels will remain unchanged. How to properly select this value is the key issue for the balance between reversibility and distortion. In this letter, half the difference of a pixel pair plus 1-bit watermark has been elaborately selected to satisfy this purpose. In addition, PDA is proposed to significantly reduce the capacity consumed by overhead information. A series of experiments is conducted to verify the effectiveness and advantages of the proposed approach.

Index Terms—Invariability of the sum of pixel pairs, pairwise difference adjustment (PDA), reversible watermarking.

I. INTRODUCTION

A CONSIDERABLE amount of research on reversible watermarking has been done over the past few years. Some reversible watermarking algorithms [1]–[5] have been proposed in the literature since its first appearance in the patent by Hon-singer *et al.* [6]. Lee *et al.* propose a block-based reversible watermarking scheme, which can determine which blocks contain pixels that will overflow or underflow in the embedding process, and then the watermark is adaptively embedded into those blocks without overflowing or underflowing pixels [1]. Reversible watermarking based on difference expansion is proposed by Tian [2]. For the purpose of controlling embedding capacity and distortion, only those pixel pairs with absolute differences less than or equal to a predefined threshold are expanded to carry watermark information without causing overflow and underflow. However, when the threshold is small, a large part of the available capacity is consumed to store the compressed location map due to low compressibility of the location map. As a result, the embedding capacity is usually low in such cases. In Thodi's paper [3], histogram shifting is incorporated into Tian's difference expansion technique to avoid sending the location map.

In the method proposed by Coltuc *et al.*, an integer transform defined on pixel pairs is proposed to add the difference value of

pixel pair to one pixel and subtract the same value from the other [4]. If both LSBs of a pixel pair are not one in the same time, then two watermark bits are embedded into the LSBs of its transformed one by simple LSB replacement without causing overflow or underflow. A threshold-controlled embedding scheme is introduced into Coltuc's 2007 paper [5]. However, in both methods, for images with large intensity differences near edges or textures, such as "Barbara" and "Baboon," the embedding distortion is very high.

To solve the above problems, a novel reversible embedding scheme based on invariability of the sum of pairs and pairwise difference adjustment (PDA) is proposed in this letter. Considering high distortion caused by pixel modification in Coltuc's method, a smaller value than the difference value, i.e., half the difference plus 1-bit watermark, is added to one pixel and subtracted from the other. Since the modification to each pixel becomes smaller, higher PSNR value can be obtained compared with Coltuc's method. To avoid the problems in Tian's method, PDA is proposed to largely increase the number of the pairs available for embedding. Accordingly, a considerable bias between the numbers of 1s and 0s in the location map is formed. As a result, higher lossless compression ratio can be achieved. The compressibility is largely increased, especially when the threshold is small. As revealed in the experimental results, our performance exceeds that of the expansion difference in [2] and is comparable to that of the scheme in [3].

The rest of this letter is organized as follows. In Section II, the proposed scheme is described. The performance is evaluated in Section III. The experimental results are shown in Section IV, and finally, we conclude the paper in Section V.

II. PROPOSED SCHEME

In the proposed scheme, an integer transform is designed to embed 1-bit watermark into one pixel pair in a way that the sum of the pixel pair remains unchanged. Based on the invariability of sum values and the equality between the parities of sum values and difference values, the extraction of watermarks and the recovery of pixel pairs can be easily achieved.

The requirement on embedding small amount of payload at low embedding distortions can be satisfied if only those pixel pairs whose intensity differences are below a threshold are transformed. However, when the threshold is small, it is difficult to efficiently compress the location map, which is used to identify the transformed locations. To solve this problem, PDA is proposed to achieve higher compressibility by reshaping the distribution of the location map.

A. Integer Transform

The forward integer transform, $T(\cdot)$, is defined as

$$x' = x + \left\lfloor \frac{d}{2} \right\rfloor + b \quad y' = y - \left\lfloor \frac{d}{2} \right\rfloor - b \quad (1)$$

Manuscript received December 20, 2007; revised April 04, 2008. This work was supported in part by the National Natural Science Foundation of China (No. 60702013, No. 60776794, and No. 90604032), in part by the 973 Program (No. 2006CB303104), in part by the 863 program (No. 2007AA01Z175), in part by PCSIRT (No. IRT0707), in part by the Beijing NSF (No. 4073038), and in part by the Specialized Research Fund of BJTU. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Min Wu.

S. Weng, Y. Zhao, and R. Ni are with the Institute of Information Science, Beijing Jiaotong University, Beijing, China.

J.-S. Pan is with the Department of Electronic Engineering, Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan, R.O.C.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2008.2001984

where b is used to denote one bit watermark, and d is the difference between x and y , i.e., $d = x - y$. Obviously, $x + y$ equals $x' + y'$. $x + y$ and d have the same parity since $x + y = d + 2y$ and $2y$ is an even number. That is, $LSB(x + y)$ equals $LSB(d)$, where the operator $LSB(\cdot)$ denotes the least significant bit. It is similar for $x' + y'$ and d' , where $d' = x' - y'$. Hence, we have $LSB(x + y) = LSB(d) = LSB(x' + y') = LSB(d')$.

On the decoding side, the sum of x' and y' is first calculated, so $x + y$ is determined.

From (1), the difference value of x' and y' can be calculated as $d' = d + 2 \times \lfloor d/2 \rfloor + 2b$. Since d equals $2 \times \lfloor d/2 \rfloor + LSB(d)$ and $LSB(d) = LSB(d')$, d' is simplified to $2d + 2b - LSB(d')$. Accordingly, d is calculated as

$$d = \frac{d' + LSB(d')}{2} - b. \quad (2)$$

The value of d and the watermark bit can be uniquely deduced because the parity of d is known and b is a binary number. For example, if $x = 4$, $y = 3$, and $b = 1$, then $x' = 5$, $y' = 2$ after embedding. On the receiver side, $(d' + LSB(d'))/2$ is calculated as 2. The parity of d can be guaranteed to be the same as that of d' if and only if $b = 1$. As a result, watermark bit b is correctly extracted and the value of d is obtained.

Once d and $x + y$ are obtained, x is calculated as $(x + y + d)/2$ while y is $(x + y - d)/2$.

B. PDA

According to (1), if all the differences without causing overflow and underflow are used for embedding, the distortion will be very high. To solve this problem, only those pairs with intensity differences below a threshold are selected for embedding. Therefore, the embedding distortion can be constrained to a low level. However, when the threshold is small, it is difficult to efficiently compress the location map due to the nearly equal probabilities of “1” and “0,” where bit “1” indicates that a pixel pair has been selected and “0” for not. In order to solve the above problems, an adjustment to the difference named PDA is proposed to modify (1) as follows:

$$\begin{aligned} x' &= x + M(d) \\ y' &= y - M(d) \end{aligned} \quad (3)$$

where $M(\cdot)$ is the PDA, which is defined as

$$M(d) = \begin{cases} \lfloor \frac{d}{2} \rfloor + b, & |d| < T_h \\ \text{sign}(d) \times C_o, & |d| \geq T_h \end{cases} \quad (4)$$

where $b \in \{0, 1\}$, $\text{sign}(d)$ is 1 if $d > 0$, 0 if $d = 0$, and -1 if $d < 0$, T_h is a predefined threshold. The constant C_o is equal to $\lfloor (T_h + 1)/2 \rfloor$. A clear explanation will be given below on why the embedding space can be increased by using PDA.

For an 8-bit grayscale image I with size $r \times c$, every two neighboring pixels are grouped into a pair $p = (x, y)$ in raster scan order.

Each p is classified into one of two sets S_1 and S_2 .

S_1 : For all values of $b \in \{0, 1\}$, if the transformed pixel pair $p' = (x', y')$ based on (3) falls into the range of $[0, 255]$, then the pixel pair p is classified into S_1 .

S_2 : The rest of pixel pairs belongs to S_2 .

S_1 is further classified into two subsets: $S_{1 < T_h}$ including all pairs with $|d| < T_h$ and $S_{1 \geq T_h}$ including the others.

For each p in $S_{1 < T_h}$, the difference value d' after embedding will fall within a range denoted by R . If T_h is even, R is $[-2T_h + 1, 2T_h - 1]$, while R is $[-2T_h + 2, 2T_h]$ if T_h is odd. For each p in $S_{1 \geq T_h}$, since the magnitude of difference is increased by $2C_o$, the changed difference value d' will fall in another range \tilde{R} , where $\tilde{R} \cap R = \phi$, and $\tilde{R} = (-\infty, -2T_h] \cup [2T_h, \infty)$ if T_h is even, $(-\infty, -2T_h - 1] \cup [2T_h + 1, \infty)$ if T_h is odd. Hence, the marked pairs can be differentiated from those changed ones without any additional information in the extraction process.

For the pixel pairs in $S_{1 \geq T_h}$, they cannot carry watermark bits based on (4), but their application is beneficial to largely increase the number of pixel pairs in S_1 . The number of the pixel pair in S_2 is correspondingly decreased since the total number of the pixel pairs is fixed. Therefore, most locations are marked by “1” while few are marked by “0” in the location map. As a result, the location map can be efficiently compressed at a high compression ratio. The embedding capacity is considerably increased since the capacity consumed by the compressed location map is largely decreased.

C. Data Embedding

For pairs belonging to $S_{1 < T_h}$, each can carry one watermark bit, so the maximum hiding capacity is $C_{ap} = \|S_{1 < T_h}\| - L_S$ bits, where $\|\cdot\|$ represents the cardinality of a set, and L_S is the bit length of the compressed location map.

Assume the length of the desired payload \mathcal{P} is \mathcal{P}_L , and $\mathcal{P}_L \leq C_{ap}$. If $\mathcal{P}_L = C_{ap}$, then all $r \times c/2$ pairs are processed for carrying \mathcal{P} . Otherwise, only a part is needed. Assume the minimal number of pairs is ξ , which are capable of carrying $\mathcal{P}_L + L_S$ bits. ξ pairs are pseudo-randomly selected from all $r \times c/2$ pairs controlled by a key. ξ and the key are sent to the receiver so as to reconstruct the original image. Denote the set of all ξ pairs as p_1, p_2, \dots, p_ξ .

A unique EOS symbol is added at its end of the compressed map, denoted as \mathcal{L} with length L_S . The bitstream \mathcal{L} is embedded into the LSBs of the pixels after embedding the watermark by simple LSB replacement. The L_S LSBs to be replaced need to be embedded into the original image so as to recover the original image. The detailed embedding process is described as the following steps.

Step 1) Embedding the payload \mathcal{P}

For each of ξ pixel pairs, if it is in S_2 , then it is kept unaltered, i.e., $x' = x$, $y' = y$. Otherwise, for each p with $|d| < T_h$, one watermark bit is embedded into it, while for p with $|d| \geq T_h$, its x and y are incremented or decremented by C_o based on (3).

After the first $L_S/2$ pixel pairs have been processed, the LSBs of x' and y' are appended to \mathcal{P} . Hence, the watermark bits are composed of two parts: the payload \mathcal{P} and the bit sequence \mathcal{C} consisting L_S appended bits. After ξ pairs are processed, the marked image I' is obtained.

Step 2) Embedding the compressed map \mathcal{L}

The compressed location map \mathcal{L} is embedded into the LSBs of the first L_S processed pixels in I' by simple LSB replacement. Thus, a new marked image I_w is obtained.

TABLE I
COMPARISONS ON THE COMPRESSIBILITY OF THE LOCATION MAP WITH AND WITHOUT PDA

T_h	4	6	8	10	12	14	16	18	20	25	30	35	40
Size of the compressed location map with PDA (bits)	40	40	40	40	40	40	40	40	40	72	72	72	72
Size of the compressed location map without PDA (bits)	123696	103544	83072	69144	59680	52144	46216	41248	36832	27936	21472	16736	13320

D. Data Extraction and Image Restoration

An one-dimensional pair list $\{p''_1, p''_2, \dots, p''_{L_S/2}, p''_{L_S/2+1}, \dots, p''_{r \times c/2}\}$ is formed by coupling all the pixels in I_w according to the same order as in embedding. The extraction process is just applied to the pseudo-randomly selected ξ marked pairs. The detailed process consists of the following three steps.

Step 1) Extraction of the location map

The LSBs of ξ marked pairs are collected into a bitstream \mathcal{B} . By identifying the EOS symbol in \mathcal{B} , the bits from the start to the EOS, which comprise the compressed map \mathcal{L} , are decompressed by an arithmetic decoder to retrieve the location map.

Step 2) Restoration of the first L_S marked pixels

Data extraction is carried out in inverse order of the list so as to obtain the original LSBs of the first $L_S/2$ marked pairs. For each $p' = (x', y')$, if p' 's location is associated with "0" in the location map, then it is ignored. Otherwise, d is retrieved as follows:

$$d = \begin{cases} \frac{d' + LSB(x' + y')}{2} - b, & d' \in R \\ d' - 2 \times \text{sign}(d') \times C_o, & d' \in \tilde{R} \end{cases} \quad (5)$$

where $d' = x' - y'$. For $d' \in R$, the detailed description for the restoration of d and the extraction of b is referred to in Section II-A. Since $x + y = x' + y'$, x and y can be calculated by $x = (x + y + d)/2$ and $y = (x + y - d)/2$.

The extraction procedure is terminated when the number of the extracted bits reaches L_S . These bits are reversely ordered to form a bitstream $\mathcal{O} = b_1 b_2 \dots b_{L_S}$. For each pair in $p''_1, \dots, p''_{L_S/2}$, both LSBs of x'' and y'' are replaced by every two neighboring bits from \mathcal{O} to obtain the integral marked pairs $p' = (x', y')$.

Step 3) Extraction of the payload \mathcal{P}

Except for the processed pairs in Step 2, the residual ones are processed based on (5) to retrieve the corresponding original ones and extract \mathcal{P} . Finally, the original image I is obtained.

III. PERFORMANCE ANALYSIS

In this section, a complete analysis on the embedding distortion is conducted for the proposed scheme, the approaches by Tian [2], Thodi *et al.* [3], and Coltuc *et al.* [4], [5].

A. Tian's and Thodi's Methods

In Tian's method, Haar integer wavelet transform is used to calculate the average values and difference values of pairs, and the watermark is embedded in the LSBs of the expanded difference values. In Tian's method, the embedding distortion is mainly derived from the distortion caused by expanding.

After embedding, $p' = (x', y')$ is $x' = \lfloor (x + y)/2 \rfloor + \lfloor (2 \times d + b + 1)/2 \rfloor$, $y' = \lfloor (x + y)/2 \rfloor - \lfloor (2 \times d + b)/2 \rfloor$.

Let x_d be the difference between x' and x , then x_d is calculated as $x_d = \lfloor (x + y)/2 \rfloor + \lfloor (2 \times d + b + 1)/2 \rfloor - x$. If $x + y$ is even, then $x_d = \lfloor d/2 \rfloor + b$. This is the same if $x + y$ is odd. Let y_d be the difference between y' and y . If $x + y$ is even, then $y_d = -\lfloor d/2 \rfloor$. If $x + y$ is odd, then $y_d = -(d + 1)/2 = -(d - 1)/2 - 1 = -\lfloor d/2 \rfloor - 1$.

In [3], for $|d| \leq \Delta_s$, x_d and y_d are the same as those described in Tian's method, where Δ_s is a predefined threshold, whereas, for $|d| > \Delta_s$, $x_d = \lfloor (\Delta_s + 2)/2 \rfloor$, $y_d = -\lfloor (\Delta_s + 1)/2 \rfloor$ if d is even, $x_d = \lfloor (\Delta_s + 1)/2 \rfloor$, $y_d = -\lfloor (\Delta_s + 2)/2 \rfloor$ if d is odd.

B. Coltuc's Method

For each $p = (x, y)$ in Coltuc's method, after transformation, $x' = x + d$, $y' = y - d$. Therefore, $x_d = d$, $y_d = -d$.

C. Proposed Method

In this subsection, we evaluate the distortions, respectively, with $|d| < T_h$ and $|d| \geq T_h$.

1) *Distortions With $|d| < T_h$* : For each $p = (x, y)$ in S_1 , x' is equal to $x + \lfloor d/2 \rfloor + b$, and y' equals $y - \lfloor d/2 \rfloor - b$. x_d and y_d are denoted as

$$x_d = \left\lfloor \frac{d}{2} \right\rfloor + b \quad y_d = -\left\lfloor \frac{d}{2} \right\rfloor - b. \quad (6)$$

The embedding distortion caused to x in the proposed method is the same as that in Tian's. For y , the differences between the distortions of the proposed approach and Tian's are related to the parity of $x + y$. If $x + y$ is even, then the difference value is $-b$, whereas if $x + y$ is odd, the difference value is $-b + 1$. The results reveal that the proposed approach and Tian's have comparable performance in terms of embedding distortion.

Compared with the results of Coltuc's approach, the embedding distortion on x or y in the proposed approach has little difference, when $|d| \leq 2$. With $|d|$ increased, the distortion of x or y in the proposed method is further less than that of Coltuc's.

2) *Distortions With $|d| \geq T_h$* : For $|d| \geq T_h$, x_d and y_d are denoted as

$$x_d = \text{sign}(d) \times \left\lfloor \frac{T_h + 1}{2} \right\rfloor \quad y_d = -\text{sign}(d) \times \left\lfloor \frac{T_h + 1}{2} \right\rfloor. \quad (7)$$

For $|d| > T_h$, the distortions to x and y are similar to those in Thodi's method. From (7), the distortion is linearly related to T_h . Hence, the applying of PDA will also lead to the decrease in PSNR, but the degradation in the perceptual quality of images is justified considering the high embedding capacity can be achieved.

IV. EXPERIMENTAL RESULTS

The proposed scheme is tested on various 512×512 natural images ranging from smooth images (e.g., "Lena") to rough images (e.g., "Baboon"). The first experiment is done to verify the

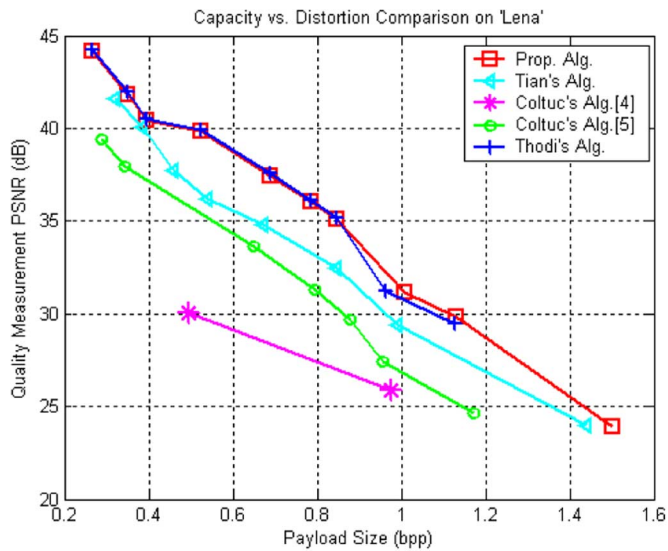


Fig. 1. Capacity versus distortion comparison on "Lena."

performance of PDA. Take "Lena" with size 512×512 for example, for once embedding, Table I shows that much higher compressibility of the location map is achieved when PDA is used. When the threshold is smaller than 25, since neither overflow nor underflow occurs with PDA, all pixel pairs are classified into S_1 . As a result, the location map can be efficiently compressed into a 40-bit sequence. If PDA is not applied, the compressibility of the location map is very low when the threshold is small. With the increase of the threshold, the compressibility of the location map gets increased.

The capacity versus distortion comparisons among the proposed method, Coltuc's, Tian's, and D2 from Thodi's are shown in Figs. 1–3. Multiple embedding has been done in order to achieve rates above 0.5 bpp (bit per pixel). The first embedding process is applied to the original image horizontally, and then the second embedding process is vertically carried out in the first-embedded image. As shown in Fig. 1, the top curve is the proposed method. When embedding the same payload, the PSNR value of the proposed method is about 0–3 dB greater than those of Tian's and Coltuc's. For example, when the embedding capacity is nearly 0.3 bpp, the PSNR achieved can be over 43 dB. However, at the same embedding rate, the PSNR value for Tian's is less than 42 dB. "Barbara" and "Baboon" are typical images with large areas of complex texture, so the obtained bit-rate is slightly lower at the same PSNR. Figs. 2 and 3 show that the proposed method achieves higher embedding capacity with lower embedding distortion than Tian's and Coltuc's. The performance of the proposed method is comparable to that of Thodi's.

V. CONCLUSIONS

A novel reversible data hiding scheme based on invariability of the sum of pixel pairs and PDA is presented in this letter. Higher PSNR can be achieved by making smaller modifications to pairs compared with Coltuc's method. PDA is proposed to largely increase the number of the pairs in S_1 , so that high compressibility of the location map is achieved. The experimental results reveal that the proposed method outperforms Tian's and Coltuc's both in hiding capacity and PSNR value.

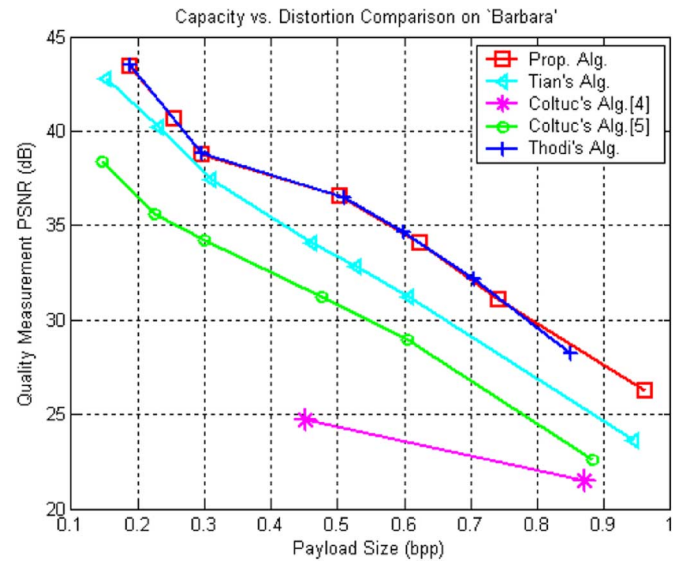


Fig. 2. Capacity versus distortion comparison on "Barbara."

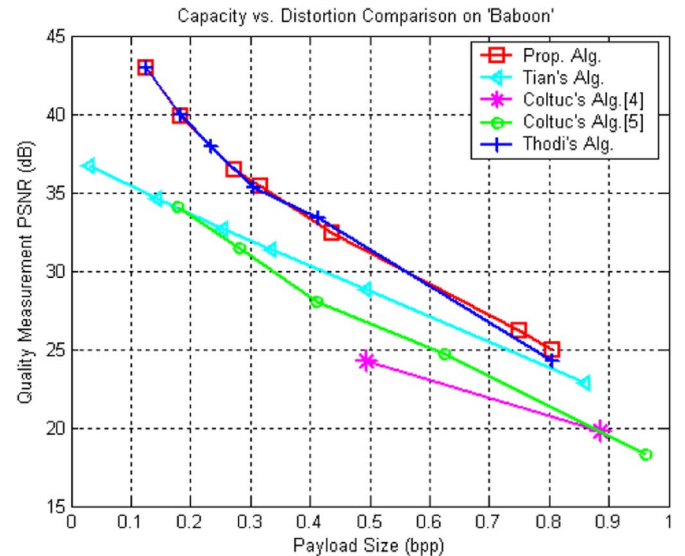


Fig. 3. Capacity versus distortion comparison on "Baboon."

REFERENCES

- [1] S. Lee, C. D. Yoo, and T. Kalker, "Reversible image watermarking based on integer-to-integer wavelet transform," *IEEE Trans. Inf. Forensics Security*, vol. 2, no. 3, pp. 321–330, Sep. 2007.
- [2] J. Tian, "Reversible data embedding using a difference expansion," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 8, pp. 890–896, Aug. 2003.
- [3] D. M. Thodi and J. J. Rodriguez, "Expansion embedding techniques for reversible watermarking," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 721–730, Mar. 2007.
- [4] D. Coltuc and A. Tremeau, "Simple reversible watermarking schemes," in *Proc. SPIE, Security, Steganography, Watermarking of Multimedia Contents*, 2005, vol. 5681, pp. 561–568.
- [5] D. Coltuc and J. M. Chassery, "Very fast watermarking by reversible contrast mapping," *IEEE Signal Process. Lett.*, vol. 14, no. 4, pp. 255–258, Apr. 2007.
- [6] C. W. Honsinger, P. Jones, M. Rabbani, and J. C. Stoffel, "Lossless Recovery of an Original Image Containing Embedded Data," US patent: 6278791, 2001.