Electronic Letters on Computer Vision and Image Analysis 3(1):1-12, 2004

A Novel Watermarking Technique for Tampering Detection in Digital Images

Chen-Kuei Yang* and Chang-Sheng Huang+
Department of Computer Science and Information Engineering*
Department of Computer and Information Management+
Ming Chuan University
No. 250, Chung Shan N. Rd., Shih Lin 111, Taipei, Taiwan, ROC

Received 18 January 2003; revised 10 March 2003; accepted 20 May 2003

Abstract

A novel fragile watermarking technique is proposed for hiding logo information into an image by tuning block pixels based on a bitmap parity checking approach. A secure key and a random number generator are used to hide the logo information in a secret, undetectable, and unambiguous way. The characteristics of the mean gray value and the bitmap in a block are exploited for performing the embedding work efficiently and for hiding a logo into an image imperceptibly. The logo can be extracted without referencing the original image. The proposed method is useful for authentication of original digital products. The extracted logo not only can be used to identify tampered locations in digital images but also can resist JPEG compression to a certain degree. Good experimental results have been conducted and resulting images show the feasibility and effectiveness of the proposed method.

Keyword: Fragile Watermarking, Bitmap, Tampering Detection, Digital Images, Parity Check.

1 Introduction

Tampering proof is one of the important applications of image watermarking which is a technique of embedding invisible signals or logos into images. Any tampering operations onto a watermarked image can be detected by verifying the watermarked image in a certain way. It is generally preferred to design a verification process that proceeds without referencing the original logo or the non-tampered watermarked image. Such watermarking techniques may be used for medical images, news images, commercial images, etc. Since such techniques used in these applications are not robust to any slight modification of the watermarked image, we call them *fragile watermarking*.

Correspondence to: ckyang@mcu.edu.tw

Recommended for acceptance by Francisco Perales
ELCVIA ISSN: 1577-5097
Published by Computer Vision Center / Universitat Autonoma de Barcelona, Barcelona, Spain

Digital watermarking techniques that may be employed to embed a digital logo or signature help asserting the ownership or intellectual property rights of media creators or owners. For a technique to be useful for tampering detection of digital media, some criteria must be considered which are reported as follows.

(1) Perceptual transparency

The logo should be perceptually invisible by the human visual system.

(2) Security

When a watermarking technique is applied, the logo cannot be extracted without a correct user key. And the location of the embedded logo is undetectable.

(3) Un-ambiguity

An extracted logo should be effective for identifying the original owner unambiguously. In addition, the accuracy of owner identification will degrade gracefully under tampering attacks.

(4) Tampering verification

Any tampering operations onto a watermarked image can be detected for copyright protection.

(5) Robustness

The logo must still be present if the watermarked media are attacked by common lossy signal processing operations.

In the last few years, several fragile watermarking techniques have been proposed for copyright protection by means of authentication and tampering detection. Yeung and Mintzer [1] developed a method that can be employed to detect modifications of individual pixels in images. It requires a look-up table (LUT) that maps RGB tri-stimulus values into the binary values in the table. They used a pixel color value as an index to look up the table in order to embed logo information by applying the XOR operation on the values of the LUT. In the method of Fridrich [2], a key-dependent binary value function is used for encoding a binary logo image and error diffusion is also employed to preserve the original colors. Kutter et al. [3] embedded logo information in the blue color channel because it is less sensitive to the human visual system. They randomly selected a pixel value and modified it according to the logo bit by adding or removing some constant luminance. Chang and Chen [4] improved the method of [3] by modifying the pixel value based on the mean value of six pixels at the left corner of the modified pixel. Using the mean value of the neighborhood pixels as a predict value to embed logo information by adding or removing some constant luminance, the method can be seen as a block based embedded process. Lee and Lee [5] proposed a fragile watermarking technique by using a classifier to classify pixels into higher and lower intensity areas according to whether the values of the pixels in the block are larger or smaller than the mean value of the block. Then the logo information was embedded into higher or lower intensity areas, depending on the value of logo bit. Wong [6] proposed a fragile watermarking

technique and embedded the logo information in the least significant bit plane of each block. The logo information is generated by a cryptographic hash function with the values of all pixels in the block. Wolfgang and Delp [7] proposed a technique that can give a relative measure of the tampered image block by using a spatial cross-correlation function. In [8-9], a number of fragile watermarking techniques have been proposed, which can be employed to embed logos in the discrete wavelet domain of signals by quantizing the corresponding coefficients with user-specified keys. A main advantage of these techniques is that they can be used to localize alterations both in the spatial and in the frequency domains and provide spatial and frequency domain information on how the signal is modified.

In the proposed fragile watermarking technique, block bitmaps are first obtained by comparing the pixel values with the block mean value. The pixel values may be modified to embed logos by checking the even or odd parity of the block bitmap. A secret key and a random number generator are used to enhance the logo in a secret, undetectable, and unambiguous way. By choosing a modified pixel near to the mean value of the block and diffusing the errors of modifications the neighborhood pixels, we can obtain the smallest value of the mean absolute error in the watermarked image. The main objectives of this paper are:

1) to introduction some criteria for tampering detection of digital media.
2) to present a novel tamper-proofing and authentication technique for copyright protection.
3) to demonstrate an easy way to diffuse errors after watermarked proceeded.

The remainder of this paper is organized as follows. In Section 2, the proposed logo embedding and extraction processes are presented. Several experimental results are shown in Section 3. Finally, some concluding remarks are given in Section 4.

2 Proposed Fragile Watermarking Technique

Assume that an original gray-level image F with size $W \times H$ and a binary logo image ω with size $w \times h$ are available. These two images will be combined to yield a watermarked image.

2.1 Logo Embedding Method

The proposed logo embedding process can be described by two steps: (1) bitmap calculation; and (2) pixel modification, described as follows.

(1) **Bitmap Calculation**

First, we divide the given gray-level image F into $n \times n$ non-overlapping blocks. In each block, a bit of the logo will be embedded. The logo will be embedded multiply until all the blocks are processed. For each block β of F, the mean value μ and the standard deviation σ of β are computed as two features of β. Then, a block bitmap b is generated according to Equation (1) below:

$$b_{i,j} = \begin{cases} 0 & \text{if} \quad f_{i,j} < \mu \\ 1 & \text{if} \quad f_{i,j} \geq \mu, \end{cases} \quad 1 \leq i \leq n,\ 1 \leq j \leq n, \tag{1}$$

where $f_{i,j}$ is the value of a pixel of β at coordinates (*i*, *j*), and $b_{i,j}$ is the bit value of b. Finally, a summation S is calculated by adding up the binary values in b as described by Equation (2) below:

$$S = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{i,j}. \tag{2}$$

(2) **Pixel Modification**

A. Random Selection of Candidate Pixels

We want to embed one bit of the logo in β, and the embedding process proceeds by adjusting the gray values of some pixels in β according to the parity theory to meet the following requirement:

$$(S + R) MOD\, 2 = \omega_{i,j} \tag{3}$$

where $R \in \{0, 1\}$ is produced by a random number generator with a secret key *K* and $\omega_{i,j}$ is the pixel value of the given binary logo image ω at coordinates (*i*, *j*).

If S does not satisfy the above requirement of (3), a pixel with value $f_{x,y}$ of β is randomly selected and modified to be $f'_{x,y}$ as follows:

$$f'_{x,y} = \begin{cases} \mu + \delta, & if \quad b_{x,y} = 0; \\ \mu - \delta, & if \quad b_{x,y} = 1, \end{cases} \tag{4}$$

where δ is a pre-selected random value between 0 and σ.

The pixel at coordinates (x, y) of β is randomly selected to modify its value in order to embed a logo bit in it. This modification can result in a new bit-plane b' that has the odd parity or even parity results. However, the replacement of $f_{x,y}$ with $f'_{x,y}$ may result in changes of the mean value and the bitmap. This means that the new mean value μ' of β in the watermarked image might be different from the original one μ. And the resulting block bitmap b' might not be the same as the original one b. Therefore, we need to perform an

error diffusion operation.

B. Error Diffusion

In order to preserve the original mean value and the block bitmap of each block, the difference between $f_{x,y}$ and $f'_{x,y}$ expressed in Equation (5) below must be properly *diffused*:

$$\varepsilon = | f'_{x,y} - f_{x,y} |; \tag{5}$$

The error ε can be equally diffused onto the pixels of β which have the same corresponding bit values as $b_{x,y}$. Since the bit value $b'_{x,y}$ is the complement values of $b_{x,y}$, if we plus some δ value on the pixel $f_{x,y}$ ($f_{x,y}$ is modified and it become $f'_{x,y}$ and $b_{x,y}$ will be changed to $b'_{x,y}$), then we must minus the same δ value form the pixels in β on which have the same corresponding bit value as $b_{x,y}$. Hence, we can preserve the same mean value of β in the watermarked image. The value of k is the number of pixels that need to diffuse the equal quantum errors ($\frac{\varepsilon}{k}$) on them. The switch ($2b_{i,j} - 1$) is applied to yield the positive or negative values for errors diffusion. So, after modifications, the mean value μ' of β in the watermarked image is the same as the original one μ and the resulting bit plane b' may only has one bit different with the original one *b*. However, the summation of the new bit-plane b' should be matched our parity requirement in order to embed a logo bit in a block. They can be described in Equation (6) below.

$$\begin{aligned} & f'_{i,j} = f_{i,j} + (2b_{i,j} - 1) \cdot \frac{\varepsilon}{k} \quad , \forall b_{i,j} = b_{x,y} \quad and \quad (i,j) \neq (x,y), \\ & where \quad k = \begin{cases} n \cdot n - (S+1) & if \quad b_{i,j} = 0; \\ S - 1 & if \quad b_{i,j} = 1. \end{cases} \end{aligned} \tag{6}$$

If we want to obtain a better quality of the watermarked image, the value of pixel modification should be tuned to be as small as possible. Therefore, we choose a candidate pixel $f_{x,y}$ of β for modification whose pixel value is nearest to the mean value μ of β. The smallest value of δ is tuned in β for embedding a bit of the logo and the smallest errors are need to diffuse in the watermarked block.

2.2 Logo Extraction Method

Let *G* be an image in question with size *W*×*H*. In the proposed logo extraction algorithm, *G* is first divided into *n*×*n* non-overlapping blocks. Then, each block bitmap *b* and the summation *S* of *b* can be obtained by the same procedures as mentioned in Section 2.1. Finally, the logo bit ω_{bj} can be obtained from Equation 7 described in the following:

$$\omega_{i,j} = \begin{cases} 0 & if \quad (S+R)\, MOD\ 2 = 0; \\ 1 & if \quad (S+R)\, MOD\ 2 = 1, \end{cases} \tag{7}$$

where $R \in \{0, 1\}$ is produced by a random number generator with the original secret key K.

The block size and the secret key K must be identical to those used in the logo embedding steps. If the size of the block or the secret key K is different, the extracted logo will be out of order.

3. Experimental Results

A number of gray-scale images were tested by the proposed watermarking technique. The tested image were divided into 3×3, 4×4, or 5×5 non-overlapping blocks and the sizes of the tested images and the logo are 512×512 and 100×50, respectively. Figure1 shows a tested image and the logo. Figure 2 shows the watermarked image and the extracted correct logo. We are able, as can be seen, to embed a logo into the test image without visible artifacts. The computed values of the peak-signal-to-noise ratio (PSNR) [10], bit-error rate (BER) and normalized cross correlation (NC) measures to evaluate the quality of the watermarked images are reported in Table 1 and Figure 5. Their formulas are listed below:

$$PSNR = 10 \cdot \log_{10} \frac{255^2}{MSE}(dB), \tag{8}$$

$$MSE = \frac{1}{W \cdot H} \sum_{i=1}^{W} \sum_{j=1}^{H} (F[i, j] - G[i, j])^2, \tag{9}$$

$$NC = \frac{\sum_i \sum_j w_{ij} \cdot w'_{ij}}{\sum_i \sum_j \left[w_{ij}\right]^2} \tag{10}$$

$$BER = \frac{\sum_{i=1}^{w} \sum_{j=1}^{h} \omega(i, j) \oplus \omega'(i, j)}{w \times h} \tag{11}$$

where $F[i, j]$ and $G[i, j]$ are the gray values of the original and the watermarked images at position (i, j), respectively; (W, H) and (w, h) specify the widths and the heights of the tested image and the logo image, respectively.

From Table 1, we can see that the PSNR values of the watermarked image “Mandrill” are smaller than those of the other two watermarked images. Because of the larger variance of the gray values in the image blocks of “Mandrill”, the quantity of gray value modifications of pixels for logo embedding is also bigger. Even though the PSNR values are smaller than the others, the watermarked image “Mandrill” is still almost identical to the original one perceptually and the extracted logo with a better BER value than that of Lee and Lee method [5].

Figure 3(a) shows a tampered image of the watermarked “Lena” , which results from manipulations of inserting a text, blurring the face, adding noise to the shoulder, and cutting an area on the top-right corner. The result of verifying the tampered image is shown in Figure

3(b). The noise regions indicate the altered blocks of the tampered watermarked image. The figure shows that the logo extraction steps are effective in tampered region identification.

In Figure 4, a tampered watermarked image attacked by 90% JPEG compression and the logo image extracted from the tampered image are shown. These experimental results demonstrate that the proposed watermarking approach indeed can localize the regions that have been altered on the watermarked image. In other fragile watermarking techniques, when the watermarked image is attacked by JPEG compression, the extracted logo is a scrambled image that cannot be used to verify the logo information. Therefore, our fragile watermarking technique not only can verify the ownership of the watermarked image but also can resist attacks by JPEG compression to a certain degree that illustrates in Figure 5. In Figure 6, the color image "Lena" with the RGB color model is tested. In this experiment, a logo is embedded in the blue color channel since it is less sensitive to human eyes [3]. Although the watermarked image shown in Figure 6(c) is manipulated by adding some shining stars, we still can identify the altered regions of the tampered watermarked image by the proposed method as shown in Figure 6(d).

4. Conclusions

We have proposed a novel fragile watermarking method for tampering detection by tuning block pixels based on the theory of bitmap parity checking. The method provides an easy way to embed a logo into digital images. The extraction process is conducted without referencing the original image. Any alteration of the watermarked image can be detected and localized for copyright protection. Beside, ownership authentication can also be performed by our method. It is more important to verify the modified locations than just to extract the logo from the watermarked image. And our method is also useful for this. The method is easy to implement with lower computation load and small MAE values obtained from the watermarked image. It not only can be applied to gray images but also to color ones. Although the proposed method is a fragile watermarking technique, each logo bit can be multiply embedded into the watermarked image. Hence, the method of majority voting [11] can be applied to enhance the robustness of the embedded logo. All the experimental results show the effectiveness and feasibility of the proposed method.

Acknowledgements

The research is supported by National Science Council, Republic of China under Grant NSC 89-2218-E-130-003. The authors would like to thank Prof. Wen-Hsing Tsai (Department of Computer and information Science, National Chiao Tung University, ROC) and the referees for their insightful comments and suggestions.

References

[1] M. M. Yeung and F. C. Mintzer, "Invisible Watermarking for Image Verification," *Journal of Electronic Imaging*, Vol. 7(3), pp. 578-591, July 1998.

[2] J. Fridrich, "A Hybrid Watermark for Tamper Detection in Digital Images," *Proceedings*

of the Fifth International Symposium on Signal Processing and Its Applications, Vol. 1, pp. 301-304, 1999.

[3] M. Kutter, F. Jordan, and F. Bossen, "Digital Watermarking of Color Images Using Amplitude Modulation," *Journal of Electronic Imaging*, Vol. 7(2), pp. 326-332, April 1998.

[4] H. M. Chang and L. H. Chen, "An Invisible Watermarking Method for Color Image,"*1999 13th IPPR Conference on Computer Vision Graphics and Image Processing*, pp. 22-29, 1999.

[5] C. H. Lee and Y. K. Lee, "An Adaptive Digital Image Watermarking Technique for Copyright Protection," *IEEE Transactions on Consumer Electronics*, Vol. 45, No. 4, pp. 1005-1015, November 1999.

[6] Ping Wah Wong, "A Public Key Watermark for Image Verification and Authentication," *Proceedings of IEEE International Conf. on Image Processing*, Vol. 2, pp. 455-459, 1998.

[7] R. B. Wolfgang and E. J. Delp, "A Watermark for Digital Images," *Proceedings of IEEE International Conf. On Image Processing*, Vol. 3, pp. 219-222, 1996.

[8] D. Kundur and D. Hatzinakos, "Towards a Telltale Watermarking Technique for Tamper Proofing," *Proceedings of IEEE International Conf. on Image Processing*, Vol. 2, pp. 409-413, 1998.

[9] D. Kundur and D. Hatzinakos, "Digital Watermarking for Telltale Tamper Proofing and Authentication," *Proceedings of The IEEE*, Vol. 87, No. 7, pp. 1167-1180, July 1999.

[10] William K. Pratt, *Digital Image Processing*, 2nd Ed., NY: Prentice Hall, 1991.

[11] C. K. Yang, D. C. Wu, C. S. Huang, "A study of Robust Watermarking Using Block Based and Determinant Concept," *Proceedings of 2001 IPPR Conf. On CVGIP*, Ken-Ding, Taiwan, ROC, E3-3, 2001

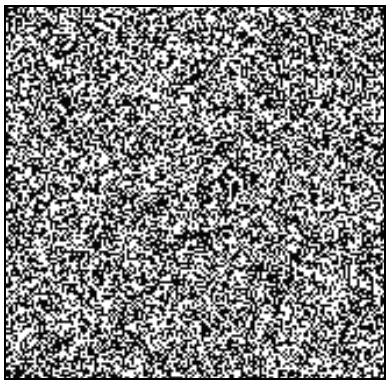
(a) Lena.

(b) Logo.

Figure 1: One of the tested images and the logo.

(a) Watermarked Lena.

(b) Extracted logo.



(c) The extracted logo with an incorrect key.

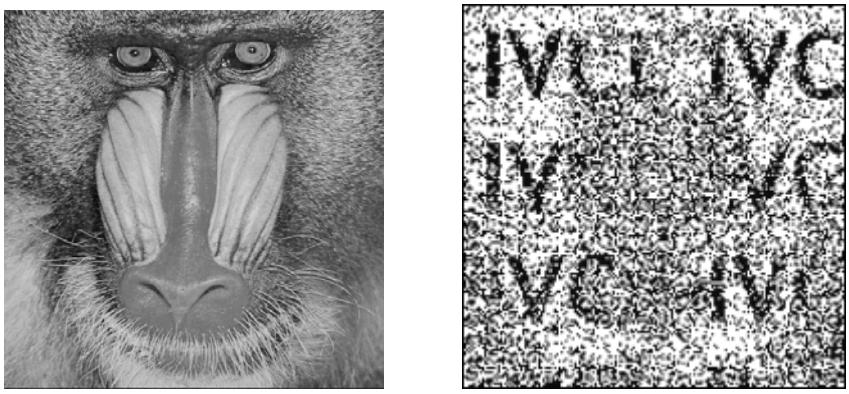
Figure 2: The watermarked image and the extracted logo of Figure 1.

Table 1: The values of PSNR and BER of three watermarked images are obtained from the proposed and Lee & Lee methods, respectively.

		Lena	Mandrill	Pepper
Randomly selecting a candidate pixel as $f_{x,y}$	PSNR	37.74	31.03	38.74
	BER	0.03%	0.07%	0.04%
Selecting a pixel $f_{x,y}$ with pixel value nearest to the mean value of β	PSNR	42.17	35.56	43.21
	BER	0.02%	0.06%	0.03%
Lee and Lee Method	PSNR	41.25	34.58	41.28
	BER	0.03%	0.08%	0.04%

(a) (b)

Figure 3: (a) The watermarked image manipulated by inserting a text, blurring the face, adding noise to the shoulder and cutting an area in the top-right corner. (b) The resulting logo obtained by the proposed method using (a) as input.



(a) (b)

Figure 4: (a) A watermarked image with 90% JPEG compression quality. (b) The logo extracted from (a) by the proposed method with BER=37.28%.

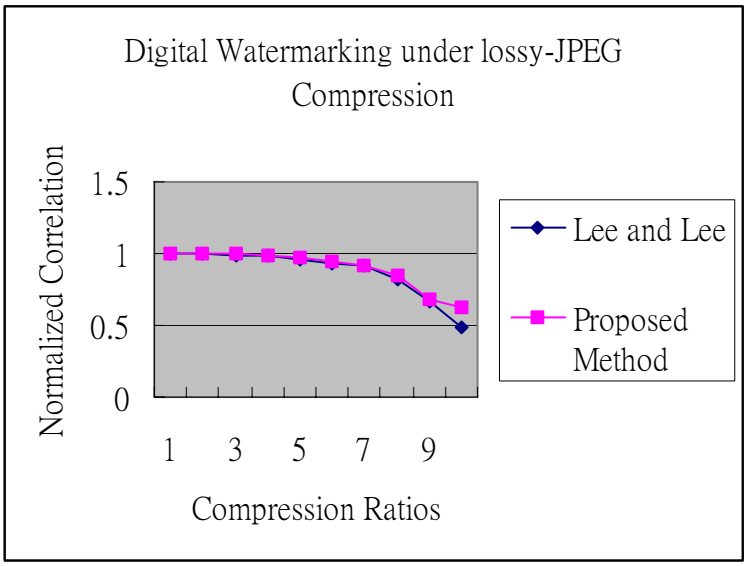


Figure 5: A comparison of normalized cross correlation values for methods of the proposed and Lee & Lee under variant lossy-JPEG compression with the tested image “Baboon”.

(a) (b)

(c) (d)

Figure 6: (a) The original color image “Lena”. (b) The watermarked color image. (c) The watermarked color image manipulated by adding some shining starts. (d) The resulting logo obtained by the proposed method using (c) as input.