MVA2011 IAPR Conference on Machine Vision Applications, June 13-15, 2011, Nara, JAPAN

A Grayscale Image Authentication Method with a Pixel-level Self-Recovering Capability against Image Tampering

Che Wei Lee
Department of Computer Science
National Chiao Tung University, Hsinchu, Taiwan
paradiserlee@gmail.com

Wen Hsiang Tsai
Department of Computer Science
National Chiao Tung University, Hsinchu, Taiwan
whtsai@cis.nctu.edu.tw

Abstract

A new grayscale image authentication method with a pixel-level self-recovering capability for tampered region repairing is proposed. By dividing the grayscale range into bins, a 3-bit bin code is generated as the authentication signal for each pixel in the input cover image. The authentication signals then are embedded randomly into the image pixels for the double purposes of tampering localization and data repairing in the image authentication process. This leads to great saving of storage space for embedding the signals and recovery data, and so results in the possibility of pixel-level authentication. Tampered pixel repairing is conducted by retrieving the embedded bin code to obtain a corresponding representative value for use as the new gray value of the tempered pixel. Good experimental results show the effectiveness of the proposed method.

1. Introduction

With the advance of digital technologies, it is easy to make imperceptible modifications to digital image contents for illegal purposes. One way to solve this problem is to use image authentication techniques based on the data hiding approach [1]. Such techniques determine whether an input image is tampered with or not by comparing the authentication signals embedded in the original image with those computed from the current image content. There are two approaches to image authentication: *fragile* and *semi-fragile* [1]. Methods of the former approach [1-3] are characterized by the use of fragile authentication signals which are sensitive to image content modifications, and so are suitable for protection of images requiring high precision, such as those used in medical or military applications.

On the contrary, methods of the second approach [4-6] are characterized by their resistance to common image modification operations without destroying the embedded *semi-fragile* authentication signals. An advantage of such methods is the allowance for a legitimate user to carry out desired image manipulations in applications without causing failures in later authentication of the resulting image.

In this study, a method for grayscale image authentication using fragile signals with an additional self-recovering capability for repairing attacked image parts is proposed. The method is based on the concept of *compressing* a number of most significant bits (MSBs) of a pixel's gray value into a shorter "bin code" for use both as an authentication signal for the pixel and as an index for retrieving the data for repairing the pixel when it is authenticated as being tampered. The bin code is generated from a bin-mapping scheme which transforms each pixel's gray value into one of eight "bins," coded by three bits.

The proposed method has at least three merits. (1) Different from other methods [7-8] which usually generate authentication signals and recovery data as two separate items, the proposed method uses the above-mentioned bin code for the two purposes *simultaneously*, leading to the use of less storage to embed the two types of data in the image. (2) And the use of less storage leads further to the possibility of conducting more precise *pixel-level* authentication because allowance of every pixel to have its own data both for authentication and content recovery has now become possible. Note that most related methods with data recovery capabilities authenticate images at the *block level* [9-10], yielding *coarser* tampering localization and data recovery results. (3) The proposed method is characterized by the *blind* process of image data recovery, that is, no additional information is needed for conducting the data recovery process. Note that extra information like codebooks is required in some existing methods for the data recovery purpose [11].

The remainder of this paper is organized as follows. In Section 2, the details of the proposed method are described. In Section 3, some measures for enhancing the proposed method are proposed. Experimental results are shown in Section 4. Finally, conclusions are made in Section 5.

2. Idea of proposed method

In the proposed method, the 8-bit gray value g of each pixel in the input image is divided into two parts. The first is the five MSBs of g and the second the remaining three least significant bits (LSBs). The former is used to generate an authentication signal for the pixel itself, with the signal also working as an index for generating the data for repairing the pixel's gray value when the pixel is authenticated to have been tampered with. The five MSBs *ideally* are expected to be embedded

directly in a randomly-selected pixel elsewhere and can be retrieved later for use in the two previously-mentioned purposes of authentication signal and repairing data generations. However, due to the limited data hiding capacity in the image, it is difficult to embed such MSBs of all the pixels in the input image without creating noticeable distortions. Consequently, we propose to use a bin-mapping scheme for the purpose of compressing these MSB data before embedding them. Specifically, we map the five MSBs into a 3-bit bin code, which is then embedded into the second part of the gray value of a pixel randomly chosen by a secret key.

During the authentication process, an authentication signal, i.e., a bin code, is computed from the first part of the gray value of every pixel p. Also, another authentication signal embedded in the second part of the gray value of a corresponding previously-randomly-chosen pixel p' is retrieved. The two authentication signals then are compared with each other. If mismatching occurs, pixel p is regarded as having been tampered with. In this case, we use further the second part of the gray value of pixel p' as an index to generate the data for repairing the altered gray values of p. The data item corresponding to each bin for repairing a tampered pixel' value is the middle value of the grayscale interval of the bin. Specifically, if the bin interval is $[a, b]$, then the value used in the repairing is just $\left\lceil \frac{a+b}{2} \right\rceil$ where $\lceil \cdot \rceil$ is the integer ceiling function. We call this value the *representative value of the bin*. Detailed algorithms implementing these ideas are described subsequently.

Algorithm 1: authentication signal generation and embedding.

Input: a grayscale image I, a random number generator f, and a secret key K.

Output: a stego-image I_s with authentication signals embedded.

1 *(Beginning of looping)* Divide the range of the 5-bit grayscale 0 through 31 into eight equal-length intervals, each called a bin.

2 *(Authentication generation)* For each pixel p of I selected in a raster-scan order, perform the following steps.

- 2.1 Transform the gray value of p into an 8-bit string $b_7b_6\ldots b_0$, denoted as S.
- 2.2 Transform the five MSBs $b_7b_6b_5b_4b_3$ of S into a decimal number d.
- 2.3 Determine the bin into which d falls with the bin number B computed by the mapping function $B = \left\lfloor \frac{d}{4} \right\rfloor$ where $\lfloor \cdot \rfloor$ is the integer floor function.
- 2.4 Transform B into a 3-bit string $c_2c_1c_0$, called the bin code, as the authentication signal for p.
- 2.5 Select randomly a pixel p' and take it as corresponding to p, using the random number generator f and the input key K as the seed.
- 2.6 Embed the bin code $c_2c_1c_0$ into p' by replacing the three LSBs of the gray value of p' with the bin code.

3 *(End of looping)* Take the final I as the desired stego-image I_s.

In the next algorithm below for image authentication, to show the pixel-level authentication results, we create additionally a binary image initially with all white pixels, and called it an *authentication image*. And if a pixel in a stego-image is authenticated as having been tampered with, we mark the corresponding pixel on this authentication image as a black point.

Algorithm 2: Image authentication, tampering detection, and data recovery.

Input: a stego-image I_s generated by algorithm 1, an originally-white authentication image I_a, and the random number generator f and the secret key K used in Algorithm 1.

Output: an image I with tampered pixels, if any, repaired.

Steps*:*

1. *(Beginning of looping for authentication)* For each pixel p of I_s selected in a raster-scan order, perform the following steps.

Stage 1 --- computation of authentication signals

- 1.1 Transform the gray value of p into an 8-bit string $b_7b_6\ldots b_0$, denoted as S.
- 1.2 Transform the five MSBs $b_7b_6b_5b_4b_3$ of S into a decimal number d.
- 1.3 Determine the bin into which d falls with the bin number B computed by the mapping function $B = \left\lfloor \frac{d}{4} \right\rfloor$.
- 1.4 Transform B into a 3-bit bin code $c_2c_1c_0$, called the *computed authentication signal*.

Stage 2 --- extraction of the hidden authentication signal.

- 1.5 Use the random number generator f and the input key K to select from I_s randomly a pixel p' corresponding to p, where presumably a previously-embedded authentication signal for p is located.
- 1.6 Transform the gray value of p' into an 8-bit string, denoted as S'.
- 1.7 Extract the three LSBs $c_2'c_1'c_0'$ of S', called the *extracted authentication signal*.

Stage 3 --- authentication signal matching and tampered pixel marking.

- 1.8 Match the computed authentication signal $c_2c_1c_0$ and the extracted one $c_2'c_1'c_0'$ bit by bit; and if mismatching occurs, regard p as having been tampered with and mark its corresponding pixel on the authentication image I_a as a black point.
- 1.9 *(End of looping)* Take the final I_a as a new authentication image I_a' for use in the next stage of the algorithm for image repairing.

Stage 4 --- tampered image part repairing.

2. *(Beginning of looping for pixel repairing)* For each black point p_a in I_a' selected in the raster-scan order, perform the following steps.

- 2.1 For the pixel p' in I_s corresponding to p_a, use the random number generator f and the input key K to select randomly a corresponding pixel p'', where presumably a previously-embedded authentication signal for p' is located.

2.2 Transform the gray value of p'' into an 8-bit string, denoted as S''.
2.3 Extract the three LSBs $c_2''c_1''c_0''$ of S''.
2.4 Transform $c_2''c_1''c_0''$ into a decimal number B'', which specifies the bin into which p' falls.
2.5 Repair the tampered pixel p' by performing the following steps.
2.5.1 Derive the representative value R of bin B''.
2.5.2 Transform R into a 5-bit binary string $r_7r_6r_5r_4r_3$.
2.5.3 Pad three 0's to the end of $r_7r_6r_5r_4r_3$ as LSBs to get an 8-bit string $T = r_7r_6r_5r_4r_3000$.
2.5.4 Transform T into a decimal number d' and replace the gray value of p' with d'.

3. (*End of looping*) Take the final I_s as the desired output image I.

3. Security Enhancement Measures

As seen previously, a secret key is used in Algorithm 1 to select randomly pixels for embedding the authentication signals. Without the key, these signals cannot be retrieved correctly for image authentication in Algorithm 2. In order to enhance further security protection of the proposed method against authentication signal counterfeiting, two additional measures are proposed in the following.

(1) *Randomizing the bin number mapping* --- We adjust the mapping function $f(b) = \left\lfloor \frac{b}{4} \right\rfloor$ for generating the bin number to be $f'(b) = (\left\lfloor \frac{b}{4} \right\rfloor + k)_{mod\,8}$ with the value k controlled by a second key.

(2) *Randomizing the bin code bit order* --- Before embedding the binary bin code, we may rotate the order of the code bits by random shifting operations controlled by a third secret key. For example, we may shift the binary bin code 110 randomly to be 011.

4. Experimental results

Many experiments have been conducted to test the proposed method and two results are shown in Figs. 1 and 2. Fig. 1(a) is the input image, Lena, of size 512×512. The resulting watermarked image of applying Algorithm 1 to generate and embed authentication signals into Fig. 1(a) is shown in Fig. 1(b), in which every pixel has its own authentication signal, i.e., the bin code, embedded in a randomly-chosen pixel in the image. No perceptible distortion can be seen in this stego-image. In order to show the effectiveness of the proposed method, we selected an area of Lena's hair and modify it by superimposing an object, a rose flower, on it. The modification result is shown in Fig. 1(c).

After the authentication process described by Algorithm 2 was applied to the tampered image Fig. 1(c), we obtained the resulting authentication image as shown in Fig. 1(d). In this authentication image, we can see that the modified area covered by the rose flower has been detected correctly. However, there also exist a lot of noise points which indicate that the pixels in the original image corresponding to these noise points are also erroneously authenticated to have been tampered with. The reason for the undesirable noise phenomenon is explained in the following.

If a pixel A is authenticated as having been tampered with, the authentication signal of a pixel B, which is embedded in pixel A, is also damaged. That means that B will also be authenticated as having been tampered with even when B was in fact not so. This effect of *mutual affection* leads to many erroneous single points being marked in the authentication image, creating a pepper-and-salt noise phenomenon as seen in Fig. 1(d).

To remove this effect, we applied median filtering to eliminate the isolated noise points before performing the pixel repairing operations described in Stage 4 of Algorithm 2. Accordingly, the new authentication image resulting from Fig. 1(d) is shown in Fig. 1(e), from which we can see that the tampered area is now presented more clearly. The repairing result in this case is shown in Fig. 1(f), and it is noted that the PSNR value compared with the watermarked image is 52.16.

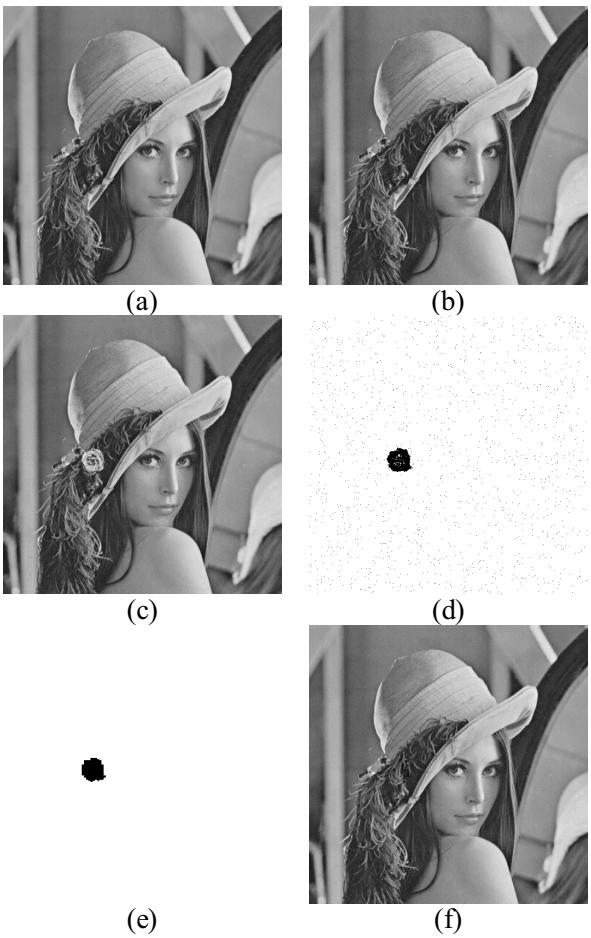

Fig. 1. Authentication result of a grayscale image. (a) Input image Lena. (b) Stego-image with PSNR 39.34. (c) Image with modification. (d) Authentication image with noise effect. (e) Modified authentication image after applying median filtering.(f) Final repairing result with PSNR 52.16 relative to watermarked image.

The proposed method is able to be used in many applications such as surveillance images shown in Fig. 2(a). Assume that two numbers printed on the license plate in Fig. 2(b), the watermarked version, were tampered with other numbers as shown in Fig.2(c). Fig. 2(e) shows the authentication result obtained after applying the median filtering to Fig. 2(d). As can be seen in Fig. 2(e), the *final* authentication result adequately reveals the tampered area in Fig. 2(c). Finally, a satisfied repairing result with a PSNR value of 45.49 is shown in Fig. 2(f) in which two original numbers printed on the license plate were successfully reappeared. These experimental results have proven the capabilities of image authentication and self-recovering of the method is effective.

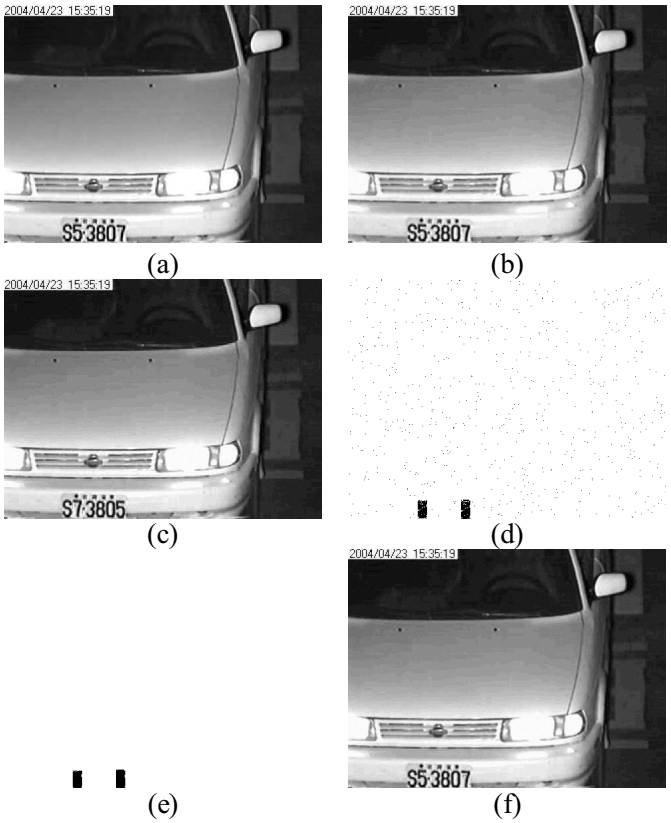

Fig. 2. Authentication result of a surveillance image taken by a monitor with tampered area. (a) Input image. (b) Stego-image with PSNR 37.53. (c) Image with modification. (d) Authentication image with noise effect. (e) Modified authentication image after applying median filtering. (f) Final repairing result with PSNR 45.49 relative to watermarked image.

5. Conclusions

A grayscale image authentication method with a capability of localizing tampered regions and repairing them has been proposed. Based on dividing the 5-bit grayscale into eight bins, an authentication signal is generated for each input image pixel according to a bin number computed by a mapping function. The authentication signals are embedded into the pixels randomly. The signals are used not only for detecting and localizing tampered pixels but also as indices for generating the representative values for repairing the tampered pixels. This merit of the proposed method leads to the possibility of pixel-level tampering detection. Experimental results show the effectiveness of the proposed method.

References

[1] M. U. Celik, G. Sharma, E. Saber and A. M. Tekalp, "Hierarchical watermarking for secure image authentication with localization," *IEEE Trans. Image Process*., vol. 11, no. 6, pp. 585-595, 2002.

[2] P. W. Wong and N. Memon, "Secret and public key image watermarking schemes for image authentication and ownership verification," *IEEE Trans. Image Process*., vol. 10, no. 10, pp. 1593-1601, 2001.

[3] H. Yang and A. C. Kot, "Binary Image Authentication with tampering localization by embedding cryptographic signature and block identifier," *IEEE Signal Process. Lett.*, vol. 13, no. 12, pp. 741-744, 2006.

[4] C. S. Lu and H. Y. M. Liao, "Structural digital signature for image authentication," *IEEE Trans. Multimedia*, vol. 5, no. 2, pp. 161–173, 2003.

[5] C. H. Tzeng and W. H. Tsai, "A new technique for authentication of image/video for multimedia applications," *Proc. ACM Multimedia Workshops --- Multimedia & Security: New Challenges*, Ottawa, Ontario, Canada, pp. 23-26, 2001.

[6] M. P. Queluz, "Content-based integrity protection of digital image," *Proc. of SPIE, Security & Watermarking of Multimedia Contents II*, vol. 3971, Bellingham, WA, pp. 85–93, 2000.

[7] S. S. Wang and S. L. Tsai, "Automatic image authentication and recovery using fractal code embedding and image inpainting," *Patt. Recog.*, no. 41, pp. 701-712, Feb., 2008.

[8] P. L. Lin, P. W. Huang and A. W. Peng, "A fragile watermarking scheme for image authentication with localization and recovery," *Proc. IEEE 6th Int'l Symp. on Multimedia Software Eng.,* pp. 146-153, Dec., 2004.

[9] P. L. Lin, C. Hsieh and P. Huang, "A hierarchical digital watermarking method for image tamper detection and recovery," *Patt. Recog.,* no. 38, pp. 2519-2529, 2005.

[10] Y. Park, H. Kang, K, Yamaguchi and K. Kobayashi, "Watermarking for tamper detection and recovery," *IEICE Electronic Express*, vol. 5, no. 17, pp. 689-696, Sept. 2008.

[11] C. W. Yang and J. J. Shen, "Recover the tampered image based on VQ indexing," *Signal Processing*, vol. 90, pp. 331-343, July, 2009.