Reversible Authentication Watermark for Image

Xiaoping Liang, *Student Member*, *IEEE*

***Abstract*—Reversible (invertible, lossless) authentication draws much attention recently for its ability to restore the original image from marked image without any distortions upon verification. In this paper, we propose a reversible water-marking scheme for image authentication based on histogram modification in integer wavelet transform (IWT) domain. The proposed scheme has features: i) Reversibility, high embedding capacity, and good perceptual invisibility; ii) Tamper localization and discerning; iii) Detection without requiring explicit knowledge of the original image; iv) Verification before reconstruction of the original image, therefore computational require- ments are reduced when reconstruction isn't needed. To our best knowledge, none of the reversible authentication schemes reported in the literature has all of the features. Those features make the proposed scheme practical, effective and appealing for strict content integrity authentication system. Experimental results demonstrate the feasibility and validity of the proposed scheme.**

***Index Terms*—Digital Watermarking, Fragile Authentication, Image Integrity, Reversible Data Hiding, Tamper Localization.**

I. Introduction

Since digital images are widely used nowadays and even an amateur can easily modify an image and create "perfect" forgeries with powerful image processing software, the need of originality and integrity or credibility check for images is raised in law, commence, defense, and journalism desirably. Fortunately, image authentication technique achieves such a goal.

Digital signature and digital watermarking are the two approaches for image authentication reported in the literature. The former is the traditional authentication methods and has some drawbacks. As the signature is appended to a digital image, it not only increases the file size but also can be removed easily. When facing format conversion, the authentication code will be lost. Moreover, it can not locate the tampered area of an image with high accuracy. The latter overcomes the above drawbacks and provides additional functionality. By embedding a watermark into digital image, the file size keeps unchanged. This watermark is very sensitive to any modifications imposed upon an image and can be used for tamper localization with high accuracy. However, conventional watermarking techniques distort the original image permanently. That is, the original image can not be recovered from the marked image when the marked image is deemed authentic. These distortions are not allowed in some sensitive applications, such as law enforcement, medical and military image systems. Reversible water- marking is a solution to those cases. As long as the marked image is authentic, the original image can be reconstructed without any distortion. Reversible watermarking techniques can be classified into three categories [1]: 1st, those for fragile authentication, 2nd, those for high embedding capacity, and 3rd, those for semi-fragile authentication. The second category may be applied to secret communication, and the third category to media such as for entertainment, which allow the marked media be stored in lossy-compression format, and be deemed authentic though it underwent common signal processing which keeps the content of the disturbed media. The first category is suitable for content that every bit is too important to neglect, i.e. require bit-by-bit exactness in special scenarios. We call the first category reversible authentication watermarking (RAW), where reversibility of the original media bit-by-bit makes sense.

An effective authentication scheme basically should have the desirable features [2]–[4]: tamper detection and localization, good perceptual invisibility, and detection without requiring explicit knowledge of the original image. Some RAW schemes for image have been reported in the literature [5]–[11], whereas rare of the schemes satisfy all the basic desirable features. Except the scheme proposed by Celik et al [10]–[11], none of the schemes in the literature provides tamper localization capability, which plays very important role in image authentication. Celik et al. utilized the hierarchical authentication watermark [12] in conjunction with the lossless generalized-LSB data embedding algorithm [13] to offer localized lossless authentication watermark. Beside tamper localization, another distinct advantage of the scheme proposed by Celik et al. over all the other existing RAW schemes is that it allows validation of the marked image before recovery of the original image, which is a new framework, hence reduce computational requirements in situations when either the verification step fails or the lossless restoration is not needed. However, the scheme relies on the context-based, adaptive, lossless image codec (CALIC) lossless image compression algorithm [14]–[15] to make space for embedding payload to achieve reversibility, thus it adds complexity to the implementation of authentication system, and may be inapplicable in case the original image has complex texture. Moreover, the quantized values of the original image are required as side information to reconstruct the original image. Tamper localization of the scheme is provided by the block-based nature of the algorithm, and the block size couldn't be too small in order to perform the CALIC lossless image compression successfully. Hence, the accuracy of tamper localization is restricted. In

Manuscript received July 21, 2008.

Xiaoping Liang was with the Guangdong Province Key Laboratory of Information Security, Sun Yat-Sen University, Guangzhou, China, now is with the Information Security Laboratory, Department of Information Engineering, The Chinese University of Hong Kong, Hong Kong (phone: 852-2609-8482; fax: 852-2609-5282; e-mail: lxp007@ie.cuhk.edu.hk).

addition, the scheme couldn't discern tamper originally done in frequency domain from that originally done in spatial domain, there- fore, any tamper originally done in frequency regions results in tamper localization in spatial regions of the image, i.e. it is deemed that the tamper is originally done in spatial domain. An example for illustration will be given in Section III.

Based on our prior work in the 2nd and the 3rd category of reversible watermarking [16]-[18], we present a RAW scheme for image to cater for the need of strict content integrity authentication in this paper. The proposed scheme satisfies all the basic desirable features, and allows validation before reconstruction of the original image. The proposed scheme has three advantages over the scheme proposed by Celik et al. Firstly, reversibility and high embedding capacity are guaranteed by the statistical property of coefficients in high frequency sub-bands and histogram modification in IWT domain, and don't rely on any lossless compression technique, hence the proposed scheme is easy to implement and has low computational complexity. Secondly, by exploiting the space-frequency localization property of IWT and combining a bi-level image and hashes of IWT coefficients as watermark, the scheme can discern tamper originally done in frequency domain from that originally done in spatial domain. Thirdly, the detector doesn't require explicit knowledge of the original image, and therefore, verification and reconstruction of the original image can be performed only with secret key.

The rest of this paper is organized as follows. The proposed RAW scheme for image is described in Section II. Tamper discerning and security consideration are presented in Section III. Some experimental results are presented in Section IV. The conclusion is drawn in Section V.

II. The Proposed RAW Scheme for Image

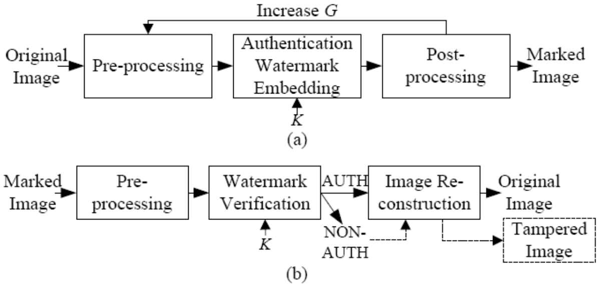


Fig. 1 Reversible authentication watermarking (RAW) framework. (a) Reversible authentication watermark embedding; (b) Watermark verification and image reconstruction.

The whole framework of the proposed RAW scheme is shown in Fig. 1. The reversible authentication watermark embedding phase at sender end is shown in Fig. 1(a), which is composed of three parts: pre-processing, authentication watermark embedding, and post-processing. The watermark verification and image reconstruction phase at receiver end is shown in Fig. 1(b), which is also composed of three parts: pre-processing, watermark verification, and image reconstruction, where the final part is optional. In the former phase, authentication watermark should be formed and embedded reversibly into subbands of IWT. In the latter phase, authentication watermark should be extracted, and the original image may be reconstruction. Reversible watermarking method adopted in the proposed scheme to embed and extract authentication watermark is based on our prior work [16], which belongs to the 2nd category of the reversible watermarking technique.

A. Reversible Authentication Watermark Embedding

Pre-processing. In order to avoid overflow/underflow, first we narrow the range of pixel value of the original image before IWT decomposition is done to it. Let G be narrowing value we set, x and x' be pixel of 8-bits depth before and after the following modification

$$\begin{cases} x' = x + G, \text{if } x \in [0, G] \\ x' = x - G, \text{if } x \in [255 - G, 255]' \end{cases} \tag{1}$$

Hence the range of pixel value is changed from [0, 255] to [G, 255-G]. We save the modified pixels as S_M.

In the light of [20]–[21], we construct IWT of CDF 9/7 biorthogonal wavelet based on lifting scheme as shown in table I, which includes decomposition and reconstruction.

Table I IWT of CDF 9/7 biorthogonal wavelet based on lifting scheme

Decomposition	Reconstruction
$\begin{cases} s_l^{(0)} = x_{2l} \\ d_l^{(0)} = x_{2l+1} \end{cases}$	$\begin{cases} s_l^{(4)} = s_l \\ d_l^{(4)} = d_l \end{cases}$
$\begin{cases} d_l^{(1)} = d_l^{(0)} + Int(\alpha(s_l^{(0)} + s_{l+1}^{(0)})) \\ s_l^{(1)} = s_l^{(0)} + Int(\beta(d_l^{(1)} + d_{l-1}^{(0)})) \end{cases}$	$\begin{cases} s_l^{(3)} = s_l^{(4)} - d_l^{(4)} \\ d_l^{(3)} = d_l^{(4)} - Int((\zeta - 1)s_l^{(3)}) \end{cases}$
$\begin{cases} d_l^{(2)} = d_l^{(1)} + Int(\gamma(s_l^{(1)} + s_{l+1}^{(1)})) \\ s_l^{(2)} = s_l^{(1)} + Int(\delta(d_l^{(2)} + d_{l-1}^{(2)})) \end{cases}$	$\begin{cases} s_l^{(2)} = s_l^{(3)} - Int((-1/\zeta)d_l^{(3)}) \\ d_l^{(2)} = d_l^{(3)} - Int((\zeta - \zeta^2)s_l^{(2)}) \end{cases}$
$\begin{cases} d_l^{(3)} = d_l^{(2)} + Int((\zeta - \zeta^2)s_l^{(2)}) \\ s_l^{(3)} = s_l^{(2)} + Int((-1/\zeta)d_l^{(3)}) \end{cases}$	$\begin{cases} s_l^{(1)} = s_l^{(2)} - Int(\delta(d_l^{(2)} + d_{l-1}^{(2)})) \\ d_l^{(1)} = d_l^{(2)} - Int(\gamma(s_l^{(1)} + s_{l+1}^{(1)})) \end{cases}$
$\begin{cases} d_l^{(4)} = d_l^{(3)} + Int((\zeta - 1)s_l^{(3)}) \\ s_l^{(4)} = s_l^{(3)} + d_l^{(4)} \end{cases}$	$\begin{cases} s_l^{(0)} = s_l^{(1)} - Int(\beta(d_l^{(1)} + d_{l-1}^{(0)})) \\ d_l^{(0)} = d_l^{(1)} - Int(\alpha(s_l^{(0)} + s_{l+1}^{(0)})) \end{cases}$
$\begin{cases} s_l = s_l^{(4)} \\ d_l = d_l^{(4)} \end{cases}$	$\begin{cases} x_{2l} = s_l^{(0)} \\ x_{2l+1} = d_l^{(0)} \end{cases}$
α= -1.586134342; β=-0.05298011854; γ=0.8829110762; δ=0.4435068522; ζ=1.149604398	

In table I, $\{x_l\}_{l \in Z}$ is pixel sequence of the image. s_l and d_l are generally referred to as lower frequency and detail coefficients (i.e. high frequency coefficients), respectively. $s_l^{(i)}$, $d_l^{(i)}$ (i=0, 1, 2) are mid-output. $Int(x)$ means integer part of x. α, β, γ, δ and ζ are parameters. After the 3rd-level IWT decomposition is done on the pre-processed image, we get ten subbands that are labeled as HH$_1$, HL$_1$, …, LH$_3$ (HH, HL and LH stand for the horizontal, vertical, and diagonal detail subband respectively, and subscript number stands for resolution level), and LL$_3$ which denotes approximation coefficients. Coefficient value is denoted as c.

Authentication Watermark Embedding. The authentication watermark is composed of bi-level image I that has visual meaning, four hashes of IWT coefficients of different subbands, overhead information O_{inf} that indicates information relating to reversible data hiding in the 1st-level detail subbands, and other data for lossless image restoration successfully at receiver end. Different parts of the authentication watermark are embedded into different subbands.

Firstly, we calculate hashes on four sets {HH$_1$, HL$_1$, LH$_1$}, {HH$_2$, HL$_2$, LH$_2$}, {HH$_3$, HL$_3$, LH$_3$}, and {LL$_3$} using MD5, and get h_1, h_2, h_3, and h_4 respectively. We concatenate the four hashes to form Hs.

Secondly, we process the bi-level image I with size adjustment and encryption, and then embed it into one of the 2nd-level subbands. We adjust I to the same size with the 2nd-level subband by image scaling in order to maintain high accuracy of tamper localization, and then we encrypt I using stream cipher with a secret key K as *seed* and get I*

$$\begin{cases} K \to seed \\ I^* = I \oplus PRNG(seed) \end{cases}, \tag{2}$$

where *PRNG* is *pseudo-random number generator* [21]. Then we embed I* into one 2nd-level subband by replacing LSBs. The 2nd-level subband is selected by the same key K for simplicity. Note that the same secret key K will be used in this authentication watermark embedding part for simplicity. The original LSBs replaced by I are recorded as O_2. Note that if the high frequency coefficient value c is -1, and the to-be-embedded bit is '0', c changes to 0 after bit replacement, and the restored c will be 1 according to O_2. We create a one-bit bitmap BM_2 as the location map of c to solve the problem. When c has value -1 and is replaced by '0', value 1 is assigned to the corresponding bit location of BM_2, otherwise value 0 assigned. Because BM_2 has many sequential 0 values, it can be losslessly compressed by run-length encoding (RLE) to spare embedding capacity, and denoted as BM_{2C} after compression.

Thirdly, because all data except I is embedded into the 1st-level subbands, we evaluate the embedding capacity Cap by summing up all the peak amplitude P and/or sub-peak amplitude P' of the three 1st-level subbands

$$\begin{cases} Cap_{\text{mode}-1} = P_{HH_1} + P_{HL_1} + P_{LH_1} \\ Cap_{\text{mode}-2} = Cap_{\text{mode}-1} + P'_{HH_1} + P'_{HL_1} + P'_{LH_1} \end{cases}, \tag{3}$$

where the subscript stands for embedding mode or the label of the 1st-level subband. According to bit-length *len* of all the data to be embedded, e.g. G, S_M, Hs, O_2, BM_{2C}, and *header* that will appear in the next step, we choose embedding mode-1 if $len \le Cap_{\text{mode-1}}$, otherwise we choose embedding mode-2. As a rule of thumb, embedding mode-1 or mode-2 is enough for the most cases. We record the chosen embedding mode that is denoted by several bits, and the coefficient values corresponding to P and P' of the 1st-level subbands as overhead information O_{inf}.

Fourthly, we embed O_{inf} into selected coefficients of selected sub-band from {HH$_3$, HL$_3$, LH$_3$, LL$_3$} by replacing LSBs. The sub-band and the coefficients in the sub-band are selected according to the secret key K. The original selected LSBs replaced by O_{inf} are recorded as O_3, and the location bitmap as BM_3. Because the length of BM_3 is short, lossless compression isn't needed. If LL$_3$ is selected, bitmap only has value 0 because all low frequency coefficient values are larger than 0.

Fifthly, we form bit-stream Bs as

$$Bs = G \cup S_M \cup Hs \cup O_2 \cup BM_{2C} \cup O_3 \cup BM_3, \tag{4}$$

and encrypt it using the secret key K as follows

$$\begin{cases} K \to seed \\ Bs^* = Bs \oplus PRNG(seed) \end{cases}. \tag{5}$$

We form B as

$$B = header \cup Bs^*, \tag{6}$$

where *header* indicates bit-length of every part of Bs*. In our experiment, the bit-length of *header* is 80 bits.

Finally, we embed B into the 1st-level subbands using reversible data hiding algorithm based on histogram modification of wavelet coefficients.

Reader may refer to [16] for detail on the above embedding capacity evaluation and reversible data hiding, which includes both reversible data embedding that has been exploited above, and data extraction and restoration that will be exploited in the following watermark verification and image reconstruction phase.

Post-processing. We perform the 3rd-level IWT reconstruction to form the authenticable marked image, and check whether any over/under-flow happens. If it does, we increase the value of G, and go back to pre-processing part, and go through the phase again; if it doesn't, we get the authenticable marked image.

B. Watermark Verification and Image Reconstruction

As opposed to previous reversible authentication schemes that required reconstruction of the original image prior to validation, the proposed RAW scheme allows validation of the marked image before reconstruction of the original image (see Fig. 1(b)), hence reduces computational requirements in situations when either the verification step fails or the lossless restoration is not needed.

Pre-processing. The 3rd-level IWT decomposition is done on the to-be-authenticated image, and we get ten subbands that are labeled as HH'$_1$, HL'$_1$, …, LH'$_3$.

Watermark Verification. This part includes two stages, i.e. bi-level image extraction and verification, shown in Fig. 2, and further discerning/tamper localization.

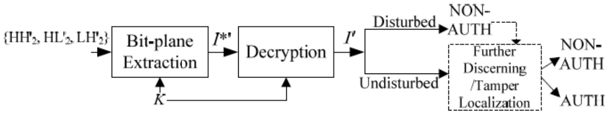


Fig. 2 Bi-level image extraction and verification.

In the first stage, the bi-level image I*' is extracted from LSBs of one 2nd-level subband according to the secret key K, i.e. bit-plane extraction, and decrypted as follows

$$\begin{cases} K \to seed \\ I' = I^{*\prime} \oplus PRNG(seed) \end{cases}. \tag{7}$$

We check I' by observation or by auto-detector. The latter method is out of range of this paper, and we use the former method for simplicity. If I' is disturbed, it means that the image under check has been disturbed, and the image is non-authentic. In order to find whether the tamper is originally done in frequency region or in spatial region, we go to the second stage, i.e. further discerning/tamper localization. If I' is not disturbed, it means that there are no tamper in spatial domain according to the space-frequency localization of IWT. In order to find whether there is any tamper in frequency region, we go to the second stage. If there are no tamper found in frequency domain in the second stage, the image is authentic, otherwise non-authentic. If the image is deemed authentic, we go to the final part, i.e. image reconstruction, and the original image is reconstructed without any distortion. If the image is deemed non-authentic, the procedure is ended, or the tampered image is restored with distortion according to user's choice. Hence, computational requirements may be reduced.

In the second stage, firstly, overhead information O'_{inf} is extracted from LSBs of one 3rd-level subband according to the secret key K, and we obtain the embedding mode and the coefficient values corresponding to P and/or P' of the 1st-level subbands. Secondly, by the overhead information we extract bitstream B' from the 1st-level subbands, and restore the 1st-level subbands simultaneously. Thirdly, we decrypt $Bs*'$ and get Bs' as follows

$$\begin{cases} K \rightarrow seed \\ Bs' = Bs*' \oplus PRNG(seed) \end{cases}, \tag{8}$$

and parse Bs' with *header* information, and obtain G', S'_M, Hs', O'_2, BM'_{2C}, O'_3, and BM'_3. Fourthly, we restore the 3rd-level subband which O'_{inf} is extracted from by replacing LSBs with O'_3 by the aid of BM'_3, and restore the 2nd-level subband which $I*'$ is extracted from by replacing LSBs with O'_2 by the aid of BM'_2, which is decompressed from BM'_{2C} according to RLE. Coefficients with value 1 are restored to -1 when their position marked by the location bitmap is '1'. Fifthly, we recalculate hashes on the four sets {HH'$_1$, HL'$_1$, LH'$_1$}, {HH'$_2$, HL'$_2$, LH'$_2$}, {HH'$_3$, HL'$_3$, LH'$_3$}, and {LL'$_3$} using MD5, and get h_{r1}, h_{r2}, h_{r3}, and h_{r4} respectively, then compare them with the extracted h'_1, h'_2, h'_3, and h'_4. If

$$h_{ri} = h'_i, \text{ for all } i \in [1,4], \tag{9}$$

the image is deemed authentic without any distortion, and we go to the final part, i.e. image reconstruction, else if

$$h_{ri} \neq h'_i, \text{ for at least one } i \in [1,4], \tag{10}$$

the image is deemed a forgery and non-authentic.

Image Reconstruction. If the image has passed the above authentication, we perform the 3rd-level IWT reconstruction on the ten subbands, and restore the pixels with the extracted G' and S'_M, and therefore get the original image without distortion.

III. Tamper Discerning and Security Consideration

One advantage of the proposed RAW scheme is that it can discern tamper originally done in frequency domain from that originally done in spatial domain. Here we give an example to show the discerning capability of the proposed RAW scheme in comparison with that of the scheme proposed by Celik et al [10]-[11]. We set the LSBs of a rectangle region with size 50x50 coefficients to zero in the center of subband {HH$_1$} of Lena, see Fig. 3(a), i.e. tamper in one high frequency subband artificially. Tamper done in {HH$_1$} is discerned by hash comparison in the proposed RAW scheme. In addition, discerning precision can be improved by calculating hash of every subband instead of subbands with the same resolution level. However, tamper originally done in spatial region instead of frequency region is the detection result of the scheme proposed by Celik et al [10]-[11], see Fig. 3(b), though the tampered image has no artifacts. In Fig. 3(b), the white pixels denote the pampered region in spatial domain of the image under authentication, and the black pixels denote the regions without any tamper.

As to security, we choose stream cipher instead of block cipher, because stream cipher is more suitable than block cipher in this application, i.e. bit-stream encryption. In the proposed RAW scheme, the secret key is used to encrypt the to-be-embedded bi-level image and bit-stream, and to select sub-band of IWT to embed data. Using the same secret key is for convenience of the user, simplicity and practicability of the authentication system. To enhance security, one secrete key is used for one image or a group of images in the proposed RAW scheme. This is a bit like *One-time Pad*, hence it is very difficult for a would-be attacker to get the right key. Moreover, more than one secret key may be used for one image in order to enhance system security. In addition, hashes are used to discern tamper. Any change made to the to-be-authenticated image will change the recalculated hash and/or the extracted hash. The probability of obtaining a match of the recalculated hash and the extracted hash is comparable to finding a collision for the hash. Hence it is difficult for attacker to get the right secret key and a match simultaneously, i.e. it is difficult to forge an image passed the authentication system.

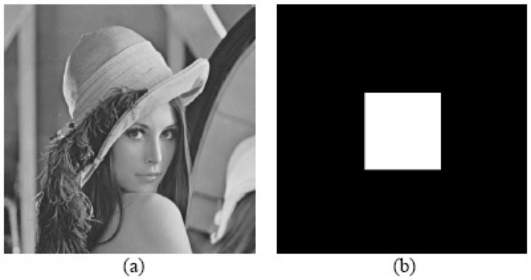


Fig. 3 (a) The tampered image; (b) Detection result where white pixels denote the tampered regions.

IV. Experiment Results

We applied the proposed RAW scheme to lots of grayscale images, and some examples that are frequently used are listed in Fig. 4. Table II contains experimental results of the listed images. In this table, embedding capacity means the total length of bit-stream including *header*, and PSNR is defined as

$$PSNR = 10\log\left(255^2 \Big/ \tfrac{1}{N\times M}\sum_{n=1}^{N}\sum_{m=1}^{M}(x'(n,m)-x(n,m))^2\right)(dB)$$

where $x(n,m)$ is pixel of original image with size of $N\times M$, and $x'(n,m)$ the pixel of the marked image. Embedding capacities of every image in table II are different, because overhead information is different for every test image generally. Table II indicates that the proposed RAW scheme keeps low distortion between marked image and the original one measured by PSNR.

Fig. 5 shows the visual impact of the proposed RAW scheme on grayscale images Pentagon and Airplane. It indicates that the marked image of the proposed RAW scheme has good visual quality.

Fig. 6 shows the tamper localization capability of the proposed RAW scheme. The disturbed regions of the extracted bi-level image denote the tamper regions of the marked images. Fig. 6(a) shows the cut and replacement operation using commercial image editing software in the center section of the marked image. Fig. 6(b) shows the modification on text printed on airplane body. Fig. 6(d) and (e) show the detection results. The figure indicates that the proposed RAW scheme can detect malicious modifications and locate the tampered regions with high probability.

V. Conclusion

A new reversible authentication watermark scheme is presented in this paper. The proposed RAW scheme exploits the statistical property of coefficients in high frequency sub-bands of IWT to ensure enough reversible embedding capacity. The proposed RAW scheme utilizes the space-frequency localization of IWT, and combining a bi-level image and hashes on IWT coefficients to discern and localize tamper originally done in spatial or frequency regions. In addition, the detector doesn't require explicit knowledge of the original image to verify the marked image and reconstruct the original one. Moreover, the proposed RAW scheme performs verification before reconstruction of the original image, in that way computational requirements can be reduced. To improve the security of the proposed scheme, stream cipher with secret key and hashes are used simultaneously. With those advantages, the proposed RAW scheme is practical and effective, and can be applied to strict content integrity authentication system in law, commence, defense, and journalism.

Acknowledgment

X. Liang thanks Prof. Jiwu Huang for his valuable comments, and thanks student Zhenhua Qu for carrying out part of the experiment. Both Jiwu Huang and Zhenhua Qu are with Guangdong Province Key Laboratory of Information Security, Sun Yat-Sen University, Guangzhou, China.

References

[1] Y. Q. Shi, "Reversible data hiding," *Third International Workshop on Digital Watermarking*, *Lecture Notes in Computer Science*, vol. 3304, 2005, pp. 1–12.

[2] E. T. Lin and E. J. Delp, "A Review of Fragile Image Watermarks," *Proceedings of the Multimedia and Security Workshop (ACM Multimedia'99) Multimedia Contents*, Oct. 1999, pp. 25–29.

[3] M. Wu and B. Liu, "Watermarking for image authentication," *Proceedings of the IEEE International Conference on Image Processing*, vol. 2, Oct. 1998, pp. 437–441.

[4] D. Kundur and D. Hatzinakos, "Digital watermarking for telltale tamper proofing and authentication," *Proceedings of the IEEE*, vol. 87, no. 7, July 1999, pp. 1167–1180.

[5] J. M. Barton, "Method and apparatus for embedding authentication information within digital data," US Patent 5,646,997, 1997.

[6] C. W. Honsinger, P. Jones, M. Rabbani, and J. C. Stoffel, "Lossless recovery of an original image containing embedded data," US Patent 6,278,791, 2001.

[7] J. Fridrich, M. Goljan, and R. Du, "Invertible Authentication," *Proceedings of SPIE*, vol. 3971, Jan. 2001, pp. 197–208.

[8] J. Fridrich, M. Goljan, and R. Du, "Invertible Authentication Watermark for JPEG Images," *Proceedings of IEEE ITCC*, April 2001, pp. 223–227.

[9] J. Tian, "Wavelet-based Reversible Watermarking for Authentication," *Proceedings of SPIE*, vol. 4675, 2002, pp. 679-690.

[10] M. U. Celik, G. Sharma, E. Saber, and T. A. Murat, "Localized Lossless Authentication Watermark," *Proceedings of SPIE*, vol. 5002, 2003, pp. 689-698.

[11] M. U. Celik, G. Sharma, and A. M. Tekalp, "Lossless watermarking for image authentication: A new framework and an implementation," *IEEE Trans. Image Processing*, vol. 15, no. 4, April 2006, pp. 1042 -1049.

[12] M. U. Celik, G. Sharma, and E. Saber, A. M. Tekalp, "Hierarchical Watermarking for Secure Image Authentication with Localization," *IEEE Trans. Image Processing*, vol. 11, no. 6, June 2002, pp. 585-595.

[13] M. U. Celik, G. Sharma, A. M. Tekalp, and E. Saber, "Lossless generalized-LSB data embedding," *IEEE Trans. Image Processing*, vol.14, no.2, Feb. 2005, pp. 253-266.

[14] X. Wu and N. Memon, "Context-based, adaptive, lossless image codec," *IEEE Trans. Commun.*, vol. 45, no. 4, April 1997, pp. 437-444.

[15] X. Wu, "Lossless compression of continuous-tone images via context selection, quantization, and modeling," *IEEE Trans. Image Processing*, vol. 6, no. 5, May 1997, pp. 656–664.

[16] X. Liang, X. Wu, and J. Huang, "Reversible Data Hiding for Image Based on Histogram Modification of Wavelet Coefficients", *International Conference on Computational Intelligence and Security (CIS)*, *Lecture Notes on Artificial Intelligence (LNAI)*, vol. 3802, Dec. 2005, pp. II 573-580.

[17] X. Wu, X. Liang, H. Liu, J. Huang and G. Qiu, "Reversible Semi-fragile Image Authentication using Zernike Moments and Integer Wavelet Transform,", *Int. Conf. on Digital Rights Management: Technologies, Issues, Challenges and Systems*, Lecture Notes on Computer Science (LNCS), vol. 3919, July 2006, pp. 135-145.

[18] X. Liang, W. Liang and W. Zhang, "Reversible Semi-fragile Authentication Watermark," Proceedings of the IEEE International Conference on Multimedia and Expo (ICME'07), July 2007, pp. 2122-2125.

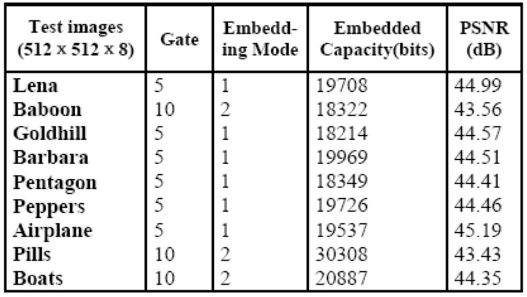
[19] I. Daubechies and W. Sweldens, "Factoring wavelet transform into lifting step," *Journal of Fourier Analysis*, vol. 4, 1998, pp. 245-267.

[20] R. Calderbank, I. Daubechies, W. Sweldens, and B. L. Yeo, "Wavelet transforms that map integers to integers," *Journal of Applied and Computational Harmonic Analysis*, vol. 5, 1998, pp. 332-369.

[21] D. R. Stinson, *Cryptography Theory and Practice* (3rd ed.). Boca Raton: Chapman & Hall/CRC, 2006, pp. 323-349.

Fig. 4 Some of test images with size of 512×512.

Table II Embedding capacity and distortion (in PSNR) of some test image



Test images (512 x 512 x 8)	Gate	Embedding Mode	Embedded Capacity(bits)	PSNR (dB)
Lena	5	1	19708	44.99
Baboon	10	2	18322	43.56
Goldhill	5	1	18214	44.57
Barbara	5	1	19969	44.51
Pentagon	5	1	18349	44.41
Peppers	5	1	19726	44.46
Airplane	5	1	19537	45.19
Pills	10	2	30308	43.43
Boats	10	2	20887	44.35

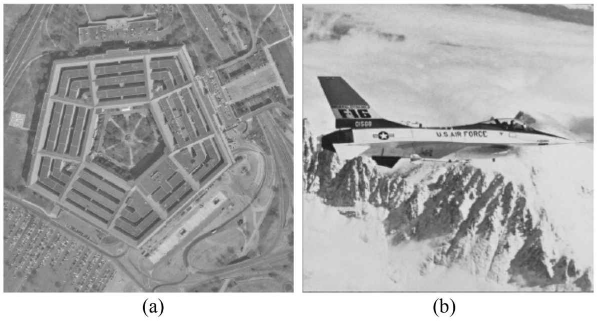


Fig. 5 Marked grayscale images (a) Pentagon, PSNR=44.41dB, 18349bits embedded; (b) Airplane, PSNR=45.19dB, 19537bits embedded.

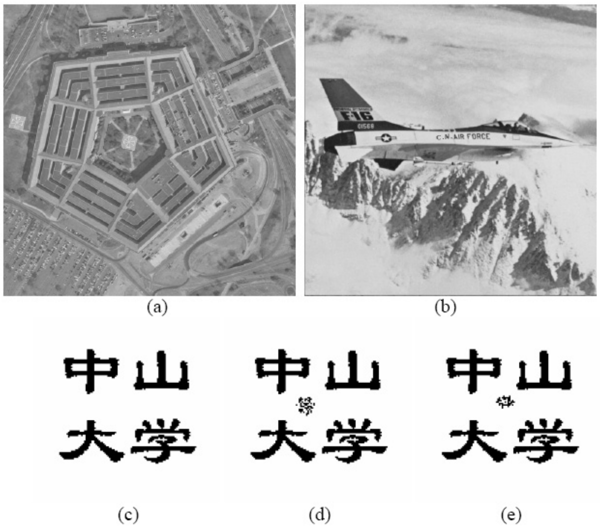


Fig. 6 Tamper Detaction. (a)(b) Tampered images; (c) The original bi-level image; (d)(e) Extracted bi-level images.