Improving robustness of visible image watermarks

S.-C. Shie*[a] and **S. D. Lin[b]**

[a]Department of Computer Science and Information Engineering, National Formosa University, Yunlin, Taiwan

[b]Department of Computer Science and Information Engineering, National Dong Hwa University, Hualien, Taiwan

Abstract: The visible watermarking is an important intellectual property rights (IPR) protection technique for digital images. In some circumstances such as multimedia contents used in learning websites or digital libraries, digital images have to be released but illegal reproductions of them are prohibited. Digital images embedded with visible watermarks contain perceptible but unobtrusive patterns. The embedded patterns should be difficult to remove unless expensive human labours are involved. Recently, Huang and Wu have proposed an efficient attacking scheme against visible watermarks. Based on the attacking scheme, the watermark patterns can be almost completely eliminated and a perceptually satisfying recovered image can be obtained by means of limited human interventions. In the present article, a novel scheme is proposed to improve the robustness of visible watermarks. For counterattacking the removal of the visible watermark pattern, the significant information of watermark patterns is extracted and embedded in the protected image imperceptibly by the information hiding techniques. In addition, to specify the positions of destroyed areas of watermarks, a fragile watermarking technique is designed and applied in the proposed scheme. Simulation results demonstrate that the proposed scheme is robust to the up to date attacks.

Keywords: fragile watermarking, information hiding, visible watermarking, watermark attacks

INTRODUCTION

Image watermarking is the process of embedding data called watermark into an image such that the embedded watermark can be detected or extracted later to verify the image. Generally, a watermarking scheme consists of three parts, the watermark, the encoder and the decoder. The watermark embedding algorithm incorporates the watermark into the host image, whereas the verification algorithm extracts and authenticates the watermark determining the ownership and integrity of the image. The watermarking techniques can be applied either in spatial domain or in frequency domain. According to human perception, digital watermarking can be divided into three categories: visible watermarking, invisible watermarking and dual watermarking.[1–6]

Visible watermarking schemes are used to protect digital images that have to be released for some purposes, such as the contents used in learning websites or digital libraries, while illegal copying or reproduction is prohibited. A visible watermark is a secondary translucent overlay on the primary image and appears visible to a viewer on careful inspection. Visibly watermarked images often contain recognisable but unobtrusive copyright patterns indicating the identity of intellectual property rights (IPR) owners. The main advantage of using visible watermarks is that it conveys an immediate claim of ownership. It also prevents or at least discourages unauthorised use of copyrighted high quality

The MS was accepted for publication on 9 October 2006.

* *Corresponding author: S. C. Shie, Department of Computer Science and Information Engineering, National Formosa University, Yunlin, Taiwan; scshie@nfu.edu.tw*

images.[7,8] A functional visible watermarking scheme should meet the following requirements:

(i) perceptibility of host image details: all the details in the original host images should remain perceptible after watermark embedding
(ii) visibility of watermark patterns: the embedded watermark should be easily recognised from the watermarked images by the naked eye
(iii) adaptive spreading over host image: the visible watermark should be adaptively spread over a large or important area of host image to prevent its deletion by clipping
(iv) robustness: embedded watermarks should be difficult or impossible to remove unless exhaustive and costly human interventions are involved
(v) efficiency: the watermark patterns should be applied automatically with little human labour.

Recently, Huang and Wu[9] proposed an attacking scheme against visible image watermarks. For watermarks purely composed of thin patterns, basic image recovery techniques such as averaging and inpainting can completely remove the embedded patterns. For more general watermarks consisting of thick patterns, the information in the surrounding unmarked areas and within the watermarked areas is utilised to recover the host image. By few human interventions to select the watermarked areas, the structure of the embedded visible watermark can be seriously destroyed and a perceptually satisfying recovered image can be obtained by this attacking scheme. Therefore, they concluded that the robustness of the current visible image watermarking schemes is doubtful and needs to be improved.

To improve the robustness of the current visible image watermarking techniques, a novel scheme, which incorporates information hiding and fragile watermarking techniques with visible watermarking techniques, is studied and proposed in the present article. The information of visible watermark patterns is embedded in the protected image to strengthen the robustness of the watermark by the information hiding techniques. Besides, an efficient fragile watermarking scheme is designed and applied to detect and specify the positions of the attacked areas of watermarks.

PROPOSED SCHEME

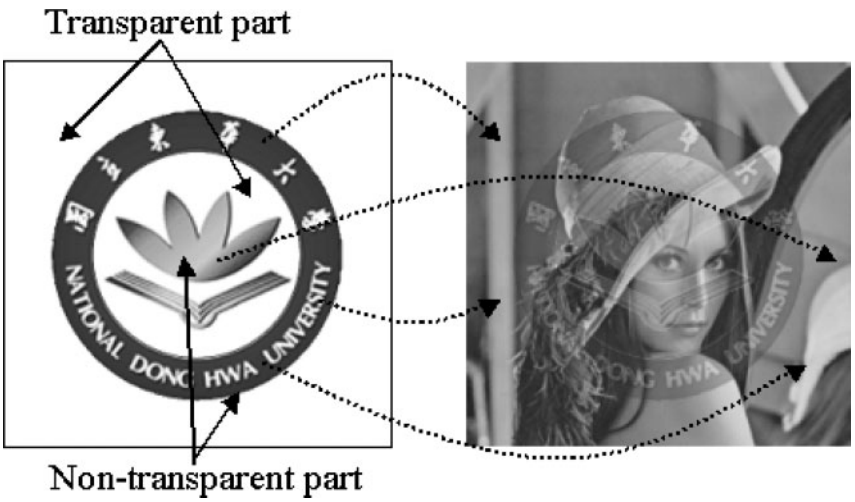
The proposed visible image watermarking scheme consists of the following stages: the information hiding stage, the attack detecting stage and, if necessary, the watermark reconstructing stage. In addition, the information hiding stage can be further divided into two processes: the watermark information embedding process and the detection code embedding process. Assume that the input image is X with $v \times h$ pixels, where v and h are multiples of w, and the original watermarked image is X′. Generally, the visible watermark image W embedded in X′ has the same size as X. Moreover, this visible watermark can be divided into two parts: the non-transparent part W_{NT} (the visible information of W) and the transparent part W_T (the transparent background of W). The non-transparent part of W is the watermark pattern itself. It can be easily recognised on X′. On the other hand, the transparent part of W appears transparent on X′. In this scheme, a new idea is proposed to further embed the information of W_{NT}, together with detection codes, into X′ to improve the robustness of the visible watermark. The information of W_{NT} is embedded into the region of X′ corresponding to the transparent region of W. Figure 1 illustrates the transparent and non-transparent parts of the watermark and the region of the original watermarked image X′ where the extra watermark information is embedded. In addition, the notations used in the present article are listed in Table 1. After the information hiding stage is accomplished, the original watermarked image X′ has been slightly modified and another marked image X″ obtained. In case that the visible watermark pattern on X″ was destroyed or even removed by attackers, the attacked region can be detected and the destroyed watermark pattern can be repaired by the proposed scheme.

Embedding visible watermark information

Before describing the proposed scheme, the set of blocks on W_{NT} (S_{WNT}) and the set of blocks on X′

Table 1 Notations used in this article

X	Input image
X′	Original watermarked image
X″	Final modified watermarked image
W	Watermark image
W_T	The transparent part of watermark image
W_{NT}	The non-transparent part of watermark image
S_{WNT}	The set of blocks of W_{NT}
S_{EMB}	The set of blocks on X′ for embedding W_{NT}



1 Idea of proposed scheme

for embedding W_{NT} (S_{EMB}) have to be defined. That is, S_{WNT} will be embedded into S_{EMB}. Let the size of all blocks be $w \times w$ pixels. In the watermark information embedding process, all blocks in S_{WNT} are transformed into frequency domain by the discrete cosine transformation (DCT). After DCT transformation, each coefficient in the DCTed block is quantised based on the luminance normalisation matrix defined in the JPEG sequential baseline system. To obtain the most important information of a W_{NT} block, some important coefficients are selected and various numbers of bits allocated to encode them based on the zonal mask and the zonal bit allocation map[10] respectively. In fact, the first k coefficients of one W_{NT} block are scanned and collected in zigzag scan order and these coefficients are transformed into one bit string. Each bit string stands for the major information of the corresponding W_{NT} block and is totally embedded into one block, selected by a predefined mapping function f, of S_{EMB}. The bits in the bit string are sequentially embedded into the pixels of the selected block by modifying the second least significant bits (LSB) in zigzag scan order. In addition, the last w pixels of each S_{EMB} block are reserved for embedding detection codes.

Based on the properties of W_{NT}, such as the location and the covering region of W_{NT} in W, the mapping function f can be designed and defined by applying hash functions or permutation techniques. For example, the visible watermark pattern used in the present article is located in the centre of the watermark image W. In addition, the covering region of this pattern on the original watermarked image X′ is large. This is one of the most complex cases in defining S_{WNT} and S_{EMB}. Here, a simple scheme is provided to collect S_{WNT} blocks and S_{EMB} blocks respectively.

The blocks of S_{WNT} are selected based on the following steps:

(i) let c be the coordinate of the centre of the visible watermark pattern W_{NT} and let r be the minimum radius such that the circle O formed by c and r covers W_{NT}
(ii) collect the blocks that locate on or within the circle O from left to right and top to bottom
(iii) let the number of the collected blocks be n and let the set of these blocks be S_{WNT}.

The blocks of S_{EMB} are selected based on the following steps:

(i) select n blocks that are not located on or within the circle O from the original watermarked image X′. These blocks are sequentially selected from the outer region to the inner region of X′ and from left to right and top to bottom
(ii) let the set of these blocks be S_{EMB}.

The main reason for selecting S_{EMB} blocks from the outer region to the inner region of X′ is that the blocks in the neighbouring regions of watermark pattern are more likely to be modified by the existing attacking schemes. After defining S_{WNT} and S_{EMB}, the numbers of blocks in S_{WNT} and S_{EMB} are the same and these blocks are ordered respectively. To map S_{WNT} blocks to S_{EMB} blocks equally and randomly, the two-dimensional torus automorphism is applied to permute the blocks in S_{WNT}[11] and then

the permuted S_{WNT} blocks are mapped to the blocks in S_{EMB}. Note that the numbers of blocks in S_{WNT} and S_{EMB} have to be the power of an integer m for performing the torus automorphism. In the proposed scheme, the insufficient m^2-n blocks are filled by duplicating the first m^2-n blocks in S_{WNT} and S_{EMB} respectively. Finally, the S_{WNT} blocks and S_{EMB} blocks are respectively organised as a matrix of $m \times m$ blocks. The matrix of S_{WNT} blocks is permuted by the torus automorphism defined as equation (1)

$$\begin{pmatrix} x^* \\ y^* \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ z & z+1 \end{pmatrix} \times \begin{pmatrix} x \\ y \end{pmatrix} \bmod m \tag{1}$$

where (x, y) and $(x^*, y^*) \in [0, m-1] \times [0, m-1]$, and $z \in [0, m-1]$ is a security parameter. Based on equation (1), the S_{WNT} block with location index (x, y) is mapped to the S_{EMB} block with location (x^*, y^*). The relative parameters, w, k, c, r, n and z, have to be preserved for future use in the watermark reconstructing stage.

Embedding detection codes

The second process in the information hiding stage is the process of detection codes embedding. For specifying the positions of destroyed areas of watermarks, an efficient fragile watermarking technique is designed and applied in the proposed scheme. After all the information of W_{NT} is embedded into the blocks of S_{EMB}, a series of detection codes are first generated and then embedded into the blocks of the W_{NT} embedded watermarked image. These codes can be used to find out and locate the positions of the destroyed blocks in the attack detecting stage. In this process, one w bits detection code $D\,(=d_1, d_2, \ldots, d_w)$ is embedded into each block of the W_{NT} embedded watermarked image. In addition, the detection codes are different for different blocks. For each block b in the W_{NT} embedded X′ (b may contain the information of W_{NT} block or not), a pseudorandom w bits string $P\,(=p_1, p_2, \ldots, p_w)$ is first generated based on a predefined parameter s and then, the most significant bits (MSBs) of b are extracted in zigzag scan order. Let the extracted MSBs be $L\,(=l_1, l_2, \ldots, l_{w*w})$. The detection code D for block b is obtained based on equation (2) and is embedded into the second LSBs of the last w pixels of b (the lower right corner of b, selected based on the zigzag scan order). After all the blocks of the W_{NT} embedded X′ are embedded with detection codes, the final watermarked image X″ is obtained

$$\begin{aligned} d_1 &= p_1 \oplus (l_1, l_2, l_3, \cdots, l_w) \\ d_2 &= p_2 \oplus (l_{w+1}, l_{w+2}, l_{w+3}, \cdots, l_{2w}) \\ d_3 &= p_3 \oplus (l_{2w+1}, l_{2w+2}, l_{2w+3}, \cdots, l_{3w}) \\ d_w &= p_w \oplus (l_{(w-1)*w+1}, l_{(w-1)*w+2}, l_{(w-1)*w+3}, \cdots, l_{w*w}) \end{aligned} \tag{2}$$

where $\oplus$ is a bit operation such that if the number of bits '1' in p_i and $[l_{(i-1)*w+1}, l_{(i-1)*w+2}, l_{(i-1)*w+3}, \ldots, l_{i*w}]$ is even, $d_i=1$; otherwise, $d_i=0$.

Detecting attacked regions of watermark pattern

The stage for detecting the attacked or removal regions of watermark pattern is quite simple. If the visible watermark pattern has been destroyed or even thoroughly removed from the protected image, the destroyed part of the watermark pattern can be easily specified and located by examining the embedded detection codes. Based on the predefined parameter s, the detection code for each block of X″ can be directly computed for checking. The process of detection code computation is fully the same as that in the information hiding stage. Each attacked block can be found by comparing the computed code with the detection code directly obtained from the pixels of the block itself. For each block, if the computed detection code is different from the code directly extracted from the block, it means this block has been attacked. Meanwhile, the position of this attacked block is specified. Consequently, the information of the watermark pattern on this block has been destroyed or removed.

Reconstructing destroyed watermark

After all the attacked blocks are found and located, the proposed scheme repairs the visible watermark pattern by first extracting the watermark information from the other undamaged blocks of X″ and then restoring the destroyed watermark using the extracted watermark information. In the watermark reconstructing stage, the S_{WNT} blocks and S_{EMB} blocks are collected using the parameters c, r and n. The value of m can be derived from the value of parameter n. Therefore, the two $m \times m$ matrices of blocks for S_{WNT} and S_{EMB} can be organised respectively. Finally, the mapping between S_{WNT} blocks and S_{EMB} blocks can be found by applying the torus automorphism using the security parameter z. The destroyed watermark pattern is repaired and therefore, the image protected by the visible watermark is reconstructed.

2 *a Lena*, *b* watermark pattern, *c* watermarked *Lena* (before information hiding) and *d* final watermarked *Lena* (after information hiding)

3 *a* resulting *Lena* after being attacked by attacking scheme[9] and *b* repaired watermarked *Lena*

RESULTS

The proposed scheme has been applied on several standard images to test its feasibility. The original test image *Lena* and the visible watermark pattern used in the experiments are shown in Fig. 2*a* and *b* respectively. Note that both the test image and the watermark pattern have a resolution of 512×512 pixels and 256 levels per pixel, and the background of the watermark image is transparent. In this experiment, the block size is 8×8 ($w=8$). The first nine coefficients ($k=9$) of each W_{NT} block are selected and the corresponding bit string embedded into the blocks of S_{EMB}. Figure 2*c* illustrates the watermarked *Lena* before information hiding and Fig. 2*d* shows the final watermarked *Lena* where a fragile watermark (the detection codes) and the compact information of watermark pattern have been embedded.

To verify the robustness of the present scheme, the recently proposed attacking algorithm against visible watermarks has been performed on the final watermarked images produced by the proposed scheme. Figure 3*a* illustrates the resulting image *Lena*, which has been attacked by the attacking scheme. The watermark pattern on the test image has been almost removed. After detecting the attacked areas and reconstructing the destroyed areas of the visible watermark, the repaired watermarked *Lena* is obtained. As shown in Fig. 3*b*, it demonstrates that the proposed scheme improves the robustness of visible watermarking against the newest attacks proposed by Huang and Wu.[9]

CONCLUSIONS

An improved visible watermarking scheme for digital images has been introduced in the present article. The proposed scheme takes advantage of information hiding and fragile watermarking techniques to improve the robustness of current visible watermarking schemes. The information of visible watermark is embedded in the image by the information technique. In addition, the destroyed areas of watermark are detected and located by the fragile watermarking technique. Simulation results demonstrate that the proposed scheme is robust to the most recent attacking scheme against visible watermarks. Therefore, it can be concluded that the robustness of visible watermarking has been improved by the proposed scheme.

ACKNOWLEDGEMENTS

The authors are supported by the National Science Council (Taipei, Taiwan) and National Dong Hwa University (Hualien, Taiwan) under project no. NSC 93-2213-E-259-005.

REFERENCES

1 G. W. Braudaway, K. A. Magerlein and F. C. Mintzer: Proc. Int. Conf. on 'Electronic imaging', San Jose, CA, USA, February 1996, SPIE, Vol. 2659, 126–133.
2 S. P. Mohanty, K. R. Ramakrishnan and M. S. Kankanhalli: Proc. 7th Int Conf. on 'Multimedia', Orlando, FL, USA, November 1999, ACM, Vol. 2, 49–51.
3 M. S. Kankanhalli, Rajmohan and J. R. Ramakrishnan: Proc. Int. Conf. on 'Multimedia computing and systems', Florence, Italy, June 1999, IEEE, 568–573.
4 S. P. Mohanty, J. R. Ramakrishnan and M. S. Kankanhalli: Proc. Int. Conf. on 'Multimedia and expo', New York, USA, August 2000, IEEE, 1029–1032.
5 S. P. Mohanty, N. Ranganathan and R. K. Namballa: Proc. 17th Int. Conf. on 'VLSI design', Mumbai, India, January 2004, IEEE, 1063–1068.
6 G. R. Feng, L. G. Jiang and H. Chen: *IEICE Trans. Fundam. Electron. Commun. Comput. Sci.*, 2005, **E88-A**, (1), 374–377.
7 A. R. Rao, F. C. Mintzer, G. W. Braudaway and M. Yeung: Proc. 1st Workshop on 'Multimedia signal processing', Princeton, NJ, USA, June 1997, IEEE, 357–362.
8 Z. M. Lu, H. T. Wu, D. G. Xu and S. H. Sun: *IEICE Trans. Inform. Syst.*, 2003, **E86-D**, (9), 1931–1933.
9 C. H. Huang and J. L. Wu: *IEEE Trans. Multimedia*, 2004, **6**, (1), 16–30.
10 R. C. Gonzalez and R. E. Woods: 'Digital image processing', 2nd edn; 2002, USR, NJ, Prentice Hall.
11 G. Voyatzis and I. Pitas: Proc. Int. Conf. on 'Image processing', Lausanne, Switzerland, September 1996, IEEE, Vol. 1, 237–240.