September 2003

Fragile watermark based on polarity of pixel points

C. Kailasanathan
University of Wollongong

Fragile watermark based on polarity of pixel points

Abstract

A fragile watermarking scheme for authenticating images based on the Yeung-Mintzer scheme is proposed in this paper. This scheme does provide a better protection against all the attacks proposed for Yeung-Mintzer scheme. A polar set derived from the image blocks is used in the embedding process. The center pixel values of image blocks are perturbed by small quantities in such a way that the perceptual quality of the image is not modified. This paper also analyse the security level with respect to other attacks.

Disciplines

Physical Sciences and Mathematics

Publication Details

This article was originally published as: Kailasanathan, C, Fragile watermark based on polarity of pixel points, ISPA 2003. Proceedings of the 3rd International Symposium on Image and Signal Processing and Analysis, 18-20 September 2003, vol 2, 860-865. Copyright IEEE 2003.

Fragile Watermark Based on Polarity of Pixel Points

C.Kailasanathan
Centre for Computer Security Research,
Department of Computer Science, University of Wollongong,
Northfield Avenue,NSW 2522, Australia.
E-mail ck12@uow.edu.au

Abstract

We propose a fragile watermarking scheme for authenticating images based on the Yeung-Mintzer [1] scheme. This scheme does provide a better protection against all the attacks proposed for Yeung-Mintzer scheme. A polar set derived from the image blocks is used in the embedding process. The center pixel values of image blocks are perturbed by small quantities in such a way that the perceptual quality of the image is not modified. We also analyse the security level with respect to other attacks.

1 Introduction

The aim of authentication is to protect the information being transmitted. Cryptography provides the means for accomplishing the authenticity of the information being received/sent and guarantees that the information has not been tampered with. Hash function protects the integrity of the information by impacting much on the hash digest for even a single bit change. For images this level of protection is not so important because a change in a single pixel value will virtually have no impact on the impression of the image. Watermarking as opposed to cryptographic tool provides a sufficient amount of protection while localising the tampered areas without having to store additional information. This is accomplished at the expense of slightly modifying the pixel values.

2 Previous Schemes

An early proposal for a fragile watermark is found in [3]. In this, Walton proposes to hide key dependent checksums of the seven most significant bits (MSB) of gray scales along pseudo random walks in the least significant bits (LSB) of pixels forming the walk. The main drawback of this scheme is that if the random walk is guessed it is easy to attack the image.

Wong's method of fragile watermarking scheme based on cryptographic hash function is found in [2]. In this he proposes a fragile water marking scheme which divides an image into blocks, calculates the hash value from MSB of the pixels and embeds the hash sequence in the LSB of the pixels. The main drawback of this scheme is that it is subject to collage attack if the block sizes are known.

In Honsinger's scheme [6] the original image can be recovered if the watermark image has been found to be authentic.

3 Our Proposal

We propose a new fragile watermarking scheme which tries to avoid the attacks proposed for Yeung-Mintzer scheme. Their scheme is subject to two types of attacks. The first one assumes that two or more images have been embedded with a watermark using the same binary logo and look up table [4] [5]. The second one which is also known as the collage attack assumes that portions of two or more images that have been embedded watermarks using the same binary logo and look up table can be put together to construct a new watermark image. To eliminate the weak points we embed the watermark on the center pixel values of image blocks in such a way that the polarity of those pixel points are not disturbed during the verification process unless the image has been tampered with. Even if the same logo and table look up are used on different images, choosing the polarity unique to each image will avoid the first attack to greater extend. This will happen because the polarity of the center pixel of the image blocks which differ from image to image are taken into consideration during embedding. This method of embedding marks unique to each image will also avoid the combination of marked pixels in determining or guessing the logo and look up table, and also the collage attack. The collage attack is somewhat protected because combining portions of different images wouldn't give the same sequence of polarity as the original one.

The idea of this fragile watermarking scheme come from Yeung-Mintzer's scheme and our critical set based fragile watermarking scheme. In our proposed watermarking scheme [8] critical points are derived from taking the difference between the original image and the low pass image and keeping the co-ordinates which have the differences above a certain threshold as critical points.

In this scheme, we determine the polarity of the center pixel of an image block by determining if the difference between the center pixel of the image block and the mean of the image block pixels is greater than zero or not. If the difference is greater than zero, we say it is a peak point, having positive polarity, else we say it is a valley point, having negative polarity. If the difference is a positive one increasing the center pixel value of the image block wouldn't change the polarity of the center pixel. If the difference is a negative one decreasing the center pixel value of the image block wouldn't change the polarity of the center pixel. The idea of polarity of image pixel come from [9], where the polarity of VQ indices are XORed with watermarked image to obtain the key. The properties mentioned above will allow the polarity of center pixels of image blocks to be in tact during watermark recovery. We construct an algorithm preserving the above properties.

3.1 Watermark Embedding

1. Given an image I, divide it into blocks of size $a \times b$.
2. Let $B_1, B_2, ..., B_n$ be the blocks of size $a \times b$ ordered according to raster scan of the image I.
3. Let $C_1, C_2, ..., C_n$ be the center pixels of the image blocks $B_1, B_2, ..., B_n$ and let $\overline{x}_1, \overline{x}_2, ..., \overline{x}_n$ be the mean of image blocks $B_1, B_2, ..., B_n$.
4. The polarity of the center pixels $C_1, C_2, ..., C_n$ are determined by finding the difference between C_i and $\overline{x}_i$ for each block B_i. If $C_i - \overline{x}_i > 0$, then C_i is said to have positive polarity ('+'), else C_i is said to have negative polarity ('-').
5. Let $P_1, P_2, ..., P_n$ be the polarity of the center pixels $C_1, C_2, ..., C_n$ of the blocks $B_1, B_2, ..., B_n$.
6. Let $L = \{l_1, l_2, ..., l_n\}$ be the raster scanned binary map of the logo.
7. Let $f_{pos} : \{0, 1, .., 255\} \rightarrow \{0, 1\}$ defines the look up table of positive polar center pixels for a particular key.
8. Let $f_{neg} : \{0, 1, .., 255\} \rightarrow \{0, 1\}$ defines the look up table of negative polar center pixels for a particular key.
9. For each center pixel from C_1 to C_n
   if $f_{pos}(C_i) = l_i$ and $P_i = " + "$ then do not modify C_i
   else if $f_{neg}(C_i) = l_i$ and $P_i = " - "$ then do not modify C_i
   else if $P_i = " + "$ then increase C_i by a value of δ until $f_{pos}(C_i + \delta) = l_i$
   else if $P_i = " - "$ then decrease C_i by a value of δ until $f_{neg}(C_i - \delta) = l_i$

3.2 Verification Algorithm

1. Upon receiving an image I^r, divide it into blocks of size $a \times b$.
2. Let $B_1^r, B_2^r, ..., B_n^r$ be the blocks of size $a \times b$ ordered according to raster scan of the image I^r.
3. Let $C_1^r, C_2^r, ..., C_n^r$ be the center pixels of the image blocks $B_1^r, B_2^r, ..., B_n^r$ and let $\overline{x}_1^r, \overline{x}_2^r, ..., \overline{x}_n^r$ be the mean of image blocks $B_1^r, B_2^r, ..., B_n^r$.
4. Let $P_1^r, P_2^r, ..., P_n^r$ be the polarity of the center pixels $C_1^r, C_2^r, ..., C_n^r$ of the blocks $B_1^r, B_2^r, ..., B_n$ as defined above.
5. Let $l_1^r, l_2^r, ..., l_n^r$ be the logo sequence extracted using the procedure defined in the next step.
6. For each center pixel from C_1 to C_n
   if $P_i = " + "$ then extract l_i^r as $f_{pos}(C_i^r)$ and
   if $P_i = " - "$ then extract l_i^r as $f_{neg}(C_i^r)$
   For each center pixel from C_1 to C_n
   if $l_i^r = l_i$ then the received image I^r has not been tampered with. if for some blocks $l_i^r = l_i$ do not hold, that portion of the image has been tampered with.

4 Experimental Results

We performed experiments on Lena, Peppers, Pills, Paper, and Corrosion images to verify if the polarities of the center pixels of the original image blocks remain the same as the polarity of the center pixels of the watermarked image blocks. As expected polarity of the image blocks remained the same. We also looked at the quality of the image after embedding by computing PSNR values. Our experiments revealed that exact recovery of the logo was possible on all five images. PSNR values never became less than 32 on all images. Table 1 shows the block sizes chosen for embedding and the PSNR values that measure the image fidelity. For each block size, appropriate size logos were chosen to incorporate entire blocks of the image.

We also embedded the same logo (Figure 1(a)) on all five images and extracted them to see the visual quality degradation caused to the logo due to embedding. Figures 1(b)-(f) show the extracted logo from Lena, Peppers, Pills, Paper, and Corrosion images.

5 Drawbacks and Possible Remedies

This method is subject to similar attacks proposed for Yeung-Mintzer scheme. Some possible protection methods against this attacks are as follows.

1. *Randomize pixel selection within each block*:
   This could be done by choosing a random pixel in each block rather than choosing a center pixel.

Block size	Lena (200x200)	Peppers (200x200)	Pills (130x200)	Paper (132x200)	Corrosion (137x200)
5x5	65.002	65.090	34.468	41.877	32.514
10x10	71.318	71.229	34.471	41.893	32.516
15x15	75.231	74.875	34.472	41.896	32.516
20x20	77.339	76.908	34.472	41.896	32.517
25x25	78.13	79.837	34.472	41.897	32.517

Table 1. PSNR computed on images choosing five different block sizes

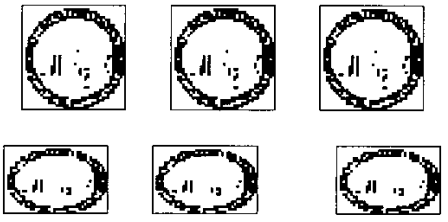

Figure 1. (a) Original Logo (b) Logo on Lena (c) Logo on Peppers (d) Logo on Pills (e) Logo on Paper (f) Logo on Corrosion

2. *Randomize block selection in embedding*:
   In embedding, instead of choosing the blocks in a raster scan of an image, choose the blocks in a random order.

3. *More than one Look Up Tables(LUTs)*:
   To enhance protection, choose more than one LUT. For example, LUT one for center pixels with positive polarity and LUT two for center pixels with negative polarity.

6 Advantages of the Scheme

Some advantages of the scheme are as follows:

1. Embedding information in this way on the center pixels of image blocks would not disturb or influence the neighbouring blocks at all. This is because non-overlapping blocks are chosen in the embedding/extraction process.

2. Localization of the manipulations is possible because the blocks used in the embedding/extraction process are uniformly spread over the image.

3. There is no syncronization loss during logo recovery due to embedding information or manipulation of the watermarked image. The size of the embedded logo sequence will always be the same as the size of the recovered logo sequence.

7 Possible Improvements: Multiple Watermarks

7.1 Variance Based

Since the variance of image blocks can be at various ranges, it should be possible to divide the image into various regions based on block variances. For example, if a set of blocks with variances ranging from 0 to l_1 forms a region R_1 and a set of blocks with variances ranging from $l_1 + \delta$ to l_2 forms a region R_2, it should be possible to embed two logos by choosing l_1, l_2 and δ appropriately. Note that a gap or a buffer of δ must be kept between adjacent ranges of variances inorder to avoid bock variances moving from one range to another due to change in variance caused by embedding. Ranges for variances must be chosen in such a way that the number of blocks falling in those ranges is equal to the logo sizes.

7.2 K-mean Based

Since the center pixels of image blocks can be obtained at various regions, it should be possible to embed multiple watermarks. For example, if a set of center pixels of image blocks are captured in a region R_1 forming a set one, and another set of center pixels of image blocks are captured in a region R_2 forming a set two, it should be possible to embed two logos. Using k-means on feature vectors of pixel points for segmenting watermarked images might end up in wrong region recovery, and hence lead to syncronization loss. Voronoi diagram as used in [7] is a favorable one for segmenting watermarked images, because the regions would not change due to watermark embedding. This is because the centroids on pixel positions, rather than the pixels values, are used as keys.

7.3 A Comparison between Variance and K-means Based Methods

Our experiments show both the methods perform equally well. The first method involves computation of variances of each block, finding the ranges for block grouping, and finally embedding a logo in each group of blocks. The second one involves segmenting the image into various regions using k-means and choosing each region blocks for embedding a logo. To decide on performance level of the above

two methods, one has to take the following into consideration.

1. Amount of computation involved
2. Security level

Figure 2 shows the two logos that are embedded in Peppers image, choosing variances of image blocks at two different ranges and the same logos extracted from the watermarked image.

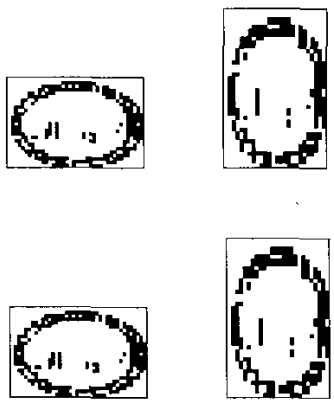

Figure 2. (a) and (b) are embedded logos and (c) and (d) are extracted logos

Figure 3 shows the three logos that are embedded in Peppers image after the image has been segmented into three regions by kmeans algorithm using the centroids (23,88), (115,39) and (93,162) trained on image pixel locations. The segmented regions and the three logos that have been recovered are also shown.

8 Security Analysis

8.1 Block Size as a Key

Since center pixel of adjacent blocks are chosen for embedding, block sizes could be chosen as keys. By choosing varying size blocks for embedding, the security of the scheme could be enhanced,

8.2 Seed of PRNG as a key

To accomplish randomization of pixel selection, a pseudo random number generator (PRNG) could be used with a distinct seed for each image.

8.2.1 Randomize Pixel Selection within a Block

Instead of choosing the center pixel of image blocks for embedding a binary bit of a logo, an arbitrary pixel in a block could be chosen. This randomization of pixel selection will avoid all the attacks proposed for Yeung-Mintzer scheme.

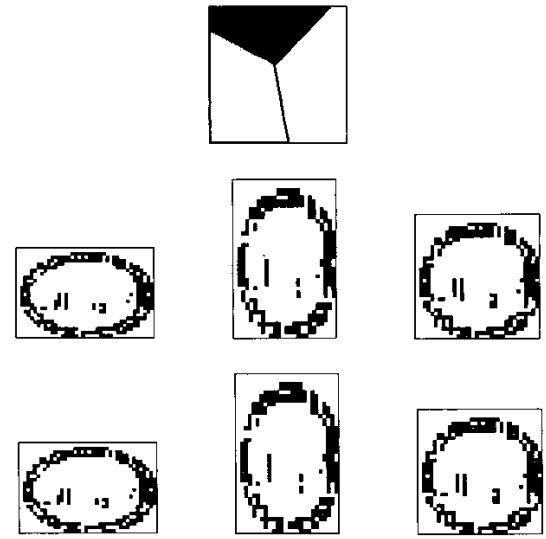

Figure 3. (a) Three Voronoi regions chosen (b) embedded logos and (c) extracted logos

8.2.2 Randomize Block Selection within an Image

Instead of choosing the blocks in a raster scan of an image for embedding binary bits of a logo, an arbitrary block within the image could be chosen. This randomization of block selection will avoid all the attacks proposed for Yeung-Mintzer scheme.

8.3 Binary Look Up Table as a Key

It has already been proposed that if the embedded pixel points of two or more images with a same logo are known, the binary look up table can easily be guessed. Therefore, major security of our scheme must come from hiding the embedded pixel points rather than hiding the binary lookup table. Unlike our scheme in [8], we do not need alternating zeros and ones in the binary look up table to reduce the syncronization loss during recovery. That is, there is no major restriction on the type of LUTs. As in Yeung-Mintzer scheme, as long as the LUT has randomized zeros and ones, this scheme should perform well.

We did experiments to verify how much watermarked images and extracted logos would change from the original ones for different sequences of 0's and 1's in the look up table. As expected highest alternating look up tables (ie. 1010... or 0101...) produced the best PSNR values for watermarked images. This was because the highest alternating look up tables produced the least changes in pixel values during embedding. (Table 2 and Figure 4)

8.4 Centroids as keys

In our region based method, there can be several regions chosen depending on the number of centroids used. The

LUT	PSNR (Image)	PSNR (Logo)
Random	57.995	Infinity
010101...	65.035	Infinity
101010...	65.208	Infinity
001100...	61.271	Infinity
110011...	60.902	Infinity
000111...	57.043	Infinity
111000...	53.175	Infinity

Table 2. Look up tables, PSNR values of watermarked image and extracted logo

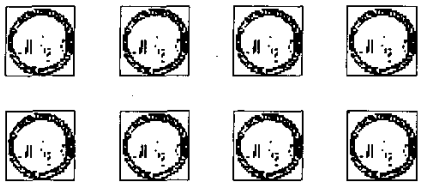

Figure 4. (a) Original logo and the logo extracted from watermarked image when (b) random LUT (c) LUT with 0101.. (d) LUT with 1010.. (e) LUT with 0011.. (f) LUT with 1100.. (g) LUT with 000111 and (h) LUT with 111000 are used

locations of those centroids can be numerous. The number of centroids determines the number of logos that could be embedded. The locations of those centroids determine the regions in which the logos are embedded. Since the regions chosen are arbitrary, the centroids of those regions can be used as keys.

9 Size of Key Space

9.1 Block Sizes as Keys

If the image size is $m \times n$, the number of blocks that could be chosen can range from 1 to $m \times n$, assuming the size of blocks range from 1×1 to $m \times n$. If we consider only square blocks of size $p \times p$, the maximum number of square blocks within an image of size $m \times n$ will be equal to $m/p \times n/p$. In this case, the size of the key space will be equal to n (if $n < m$) or m (if $m < n$).

9.2 Seed of a PRNG as a Key

9.2.1 Arbitrary Pixel within a Block

Suppose equal size blocks are chosen for embedding, the number of possible choices for choosing a pixel within a block of size $p \times p$ will be equal to $p \times p$. Therefore, the size of the key space will be equal to $(p \times p)^{(m/p \times n/p)}$. To miximize the size of the key space for an image of size $m \times n$, an appropriate value for p must be chosen.

9.2.2 Arbitrary Block within an Image

Assuming the block sizes are $p \times p$, there can be maximum of $m/p \times n/p$ blocks in an image of size $m \times n$. Suppose a logo of size $m/p \times n/p$ is to be embedded, choosing the blocks in a random order, the size of the key space will be equal to $(m/p \times n/p)!$.

9.3 Binary Look Up Table as a Key

There are altogether 2^{256} possible look up tables. Out of these tables, The best once for the least modification of the images are 0101.... and 1010.... If the maximum number of consecutive 0's and 1's are kept within 3, there will be several possible tables to choose from.

10 A Comparison Between Bhattacharjee's Scheme and Ours

Bhattacharjee's scheme some what resembles our multiple watermaring scheme and hence we give a comparison between our scheme and theirs.

1. Their scheme uses Maxican-hat wavelet for feature point detection, whereas ours uses the center pixels of image blocks for feature extraction.

2. In their scheme, Voronoi regions are formed choosing the feature points determined as centroids. In our scheme, Voronoi regions are determined based on centroids trained on arbitrary initial points.

3. Spread spectrum watermarking has been recommended for marking each region in their scheme. We use Yeung-Mintzer scheme for embedding marks on center pixels of image blocks. We use polarity of center pixels as additional information in embedding.

4. They do not have sufficient explanation for choosing the feature points as centroids of Voronoi regions. We choose arbitrary points as centroids of Voronoi regions.

5. Their main reason for watermarking is copyright protection, thus robustness plays a vital role than fragility. Our main reason for watermarking is authentication, thus fragility plays a vital role than robustness.

6. Since their scheme is designed for copyright protection, it has not been made key dependent. Since ours is for authentication, it has been made key dependent by hiding the pixel points chosen within each block as well as the blocks chosen for embedding.

7. In their case, multiple spread spectrum can be embedded by choosing each one for each region, whereas in ours, multiple logos can be embedded by choosing each logo for each region.

11 Performance evaluation

To evaluate the performance of our scheme, after embedding the logo, we performed the following attacks on the watermarked image: JPEG compression at quality level of 90, blurring, embossing, horizontal flipping, oil-painting, rotation to one degree, vertical flipping, sharpening, cutting-1, cutting-2, copy and paste-1, copy and paste-2. Unlike our scheme in [8], we were able to recover the embedded logo in a distorted form from the watermarked images for all the attacks mentioned above. This was because the size of the image blocks and the size of the image remained the same during the process of embedding and extraction. The recovered logo for all attacks mention above are shown in Figure 6. Some geometric attacks on Peppers image and the recovered logos for those attacks have also been shown in Figures 5 and 6 to illustrate the localization of manipulation. The recovery of logos from all attacked images and the localization of manipulations show the superior performance of our scheme.

Figure 5. (a) cutting-1 (b) cutting-2 (c) copy and paste-1 and (d) copy and paste-2 attacks on Peppers image

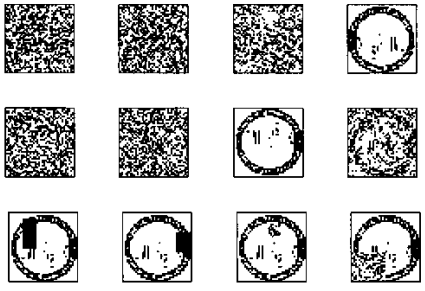

Figure 6. Logo recovered from (a) JPEG compression at quality level 90 (b) blurring (c) embossing (d) horizontal flipping (e) oilpainting (f) rotation (g) vertical flipping (h) sharpening (i) cutting-1 (j) cutting-2 (k) copy and paste-1 and (l) copy and paste-2

12 Conclusion

In this paper, we have proposed a new fragile watermarking scheme which embeds the marks on the center pixels of image blocks based on the polarity of the pixels. The polarity of these points are in tact even after embedding the marks unless the image has been tampered with. This scheme prevents the two main attacks proposed for Yeung-Mintzer's scheme. The security level of the scheme and the possible extension to multiple watermarking scheme are also investigated.

References

[1] M.Yeung and F.Mintzer, "An Invisible Watermarking Technique for Image Verification", Proc. ICIP'97, Santa Barbara, California, 1997.

[2] P.Wong, "A Watermark for Image Integrity and Ownership Verification", Proc. IS and T PIC Conference, Portland, Oregon 1998

[3] S.Walton, "Information Authentication for a Slippery New Age", Dr. Dobbs Journal, vol. 20, no. 4, pp. 18-26, April 1995.

[4] N.Memon, S.Shende and P.Wong, "On the Security of the Yeung-Mintzer Authentication Watermark", Proc. of the IS and T PICS Symposium, Savannah, Georgoa, March 2000

[5] J.Fridrich, M.Goljan, N.Memon, "Further Attacks on Yeung-Mintzer Watermarking Scheme", Proc. SPIE Electronic Imaging 2000, San Jose, Jan 24-26, 2000

[6] C.WHonsinger, P.Jones, M.Rabbani, J.C.Styyoffel, "Lossless Recovery of an Original Image Containing Embedded Data", US Patent application, Docket No: 77102/E-D, 1999.

[7] M.Kutter, S.K.Bhattacharjee, T.Ebrahimi, "Towards Second Generation Watermarking Schemes", Image Processing, 1999, ICIP99, Proceedings, 1999 International Conference on, Volume 1, 1999 pp 320-323

[8] C.Kailasanathan, R.Safavi Naini, and P.Ogunbona, "Fragile Watermark on Critical Points", International Workshop on Digital Watermarking Seoul, Korea, (Springer-Verlag), November 21-23, 2002, pp 196-210.

[9] Pan Jeng-Shyang, Wang Feng-Hsing, Jain Lakhmi, and Ichalkaranje Nikhil, " A Multistage VQ Based Watermarking Technique with Fake Watermarks", International Workshop on Digital Watermarking Seoul, Korea, (Springer-Verlag), November 21-23, 2002, pp 104-114.