Megha Kansal, Sukhjeet k. Ranade, Amandeep Kaur / International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622 www.ijera.com
Vol. 2, Issue4, July-August 2012, pp.1759-1763

Fragile Watermarking For Image Authentication Using A Hierarchical Mechanism

Megha Kansal*, Sukhjeet K. Ranade, Amandeep Kaur*****
*(Department of Computer Science, Punjabi University, Patiala)
** (Department of Computer Science, Punjabi University, Patiala)
*** (Department of Computer Science, Punjabi University, Patiala)

ABSTRACT

With ubiquitous computing, safeguarding the creative content and intellectual property in a digital form has become essential necessity. Watermarking is one such technique for protecting the digital data like image, video and audio signal. In this paper, we process a novel fragile watermarking using the Hierarchical Mechanism which is a combination of Block-wise and Pixel-wise Approach. The embedded watermark data are derived both from pixels and blocks. On the receiver side, one can first identify the blocks containing the tampered content, and then use the watermark hidden in the rest blocks to exactly locate the tampered pixels. By combining the advantages of both block-wise and pixel-wise techniques, this scheme is capable of finding the detailed tampered positions even if the modified area is more extensive. Moreover, after localizing the tampered-pixel, the original watermarked version can be perfectly restored using exhaustive attempts.

Keywords - Multimedia Security, Content Authentication, Fragile Watermarking, Tampered-pixel localization, Image restoration

I. INTRODUCTION

The world has been witnessing a rapid growth of internet technologies, multimedia distribution and e-commerce for better quality, use and services. Because of these advancements, a large amount of digital data is easily accessible to some individual or group without the permission of the owner. This digital data can be easily manipulated, tampered and distributed with the help of powerful image processing tools. The digital data image, audio, video canbe altered, has needed a requirement for techniques that decide integrity of the information. Cryptography was suggested as an effective tool to prevent illegal distribution. But in order to avoid the digital access, a special hardware should be merged with cryptography tool, which make it costly. Authentication and recovery of tampered localization of a digital data are two critical requirements. So, watermarking technique is widely used to protect the digital information.

II. BASICS ON DIGITAL WATERMARKING

A digital image watermark is a distinguishing piece of information that may be visible or invisible, is stuck to the data intended to be protected. Digital watermarking is capable in copyright protection, data authentication, integrity checks in multimedia contents [1]. A watermarking system is usually divided into three distinct steps, embedding, attack and detection. In embedding, an algorithm accepts the host and the data to be embedded and produces a watermarked signal. The watermarked signal is then transmitted or stored, usually transmitted to another person. If this person makes a modification, this is called an attack. There are many possible attacks. Detection is an algorithm which is applied to the attacked signal to attempt to extract the watermark from it. If the signal was not modified during transmission, then the watermark is still present and it can be extracted. If the signal is copied, then the information is also carried in the copy. The embedding takes place by manipulating the content of the digital data, which means the information is not embedded in the frame around the data, it is carried with the signal itself. A secret key is used during the embedding and the extraction process in order to prevent illegal access to the watermark as shown in Fig.1.

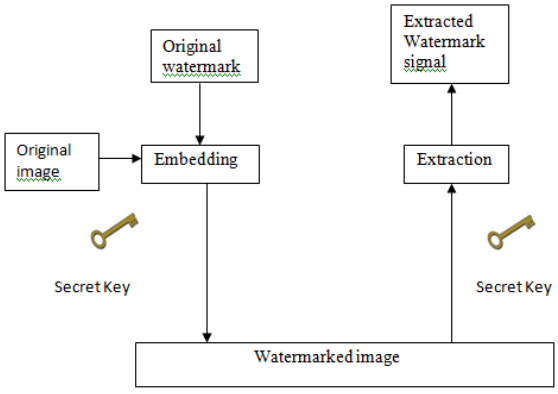

Fig. 1 Block Diagram of Watermarking System

III. CLASSIFICATION OF WATERMARKING TECHNIQUES

Watermarking algorithms can be classified on several criteria as given below:

- According to Visibility

i) Visible Watermark
ii) No Visible Watermark

- According to ability of Watermark to resist attack

i) Fragile Watermarking
ii) Semi-fragile Watermarking
iii) Robust Watermarking

- According to domain of Watermark insertion and extraction

i) Spatial domain Watermarking Technique
ii) Transform domain Watermarking Technique

- According to watermark detection and extraction

i) Blind Watermarking
ii) Non-blind Watermarking

The non-blind watermarking requires that original image to exist for detection and extraction whereas blind techniques do not require original image.

A visible watermark is easily detected by the observation. A invisible watermark is not seen to the observer and detected by the algorithms. İnvisible watermark technique is broadly categorized into three classes: Fragile, Semi-fragile and Robust watermarking.

Robust watermarking is used for copyright protection. The invisible robust watermark is embedded in such a way that alterations made to the pixel value are perceptually not noticed and it can be recovered only with appropriate decoding mechanism. Semi fragile watermarking has features of both robust and fragile watermarking. It is used for data authentication. It is sensitive to signal modification.

Fragile watermarking is used for tamper detection. The fragile watermark is a mark that is readily altered or destroyed when a host image is modified through a linear or non linear transformation. The digital watermark is fragile to any kind of distortion; the watermark image may go through. E.g.: the image after lossy compression processing could be found to be authentic by “robust” image authentication but it would fail “fragile” image authentication. For “fragile” image authentication, one bit error in a message leads to a totally different authenticator.

Fragile image authentication is highly sensitive and dependent on the exact value of the image pixels.

IV. BACKGROUND OF FRAGILE WATERMARKING TECHNIQUES

A fragile marking system detects any tampering in the marked image. Fragile Watermark is easily distorted when host image is modified by small transformations. Block Diagram of Fragile Watermarking is shown in Fig.2.

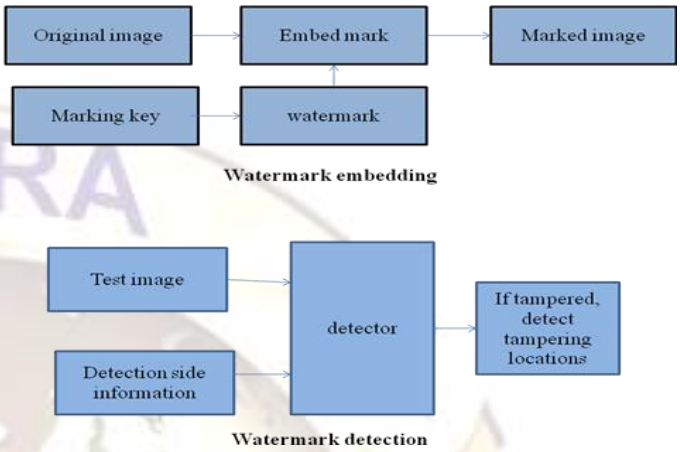

Fig. 2 Watermark embedding and Watermark detection process of Fragile Watermarking System

Fragile watermarking systems are categorized into two categories according to the working domain.

- Spatial Domain: The technique that works directly on the pixel values of the host image.
- Transform Domain: In this technique, watermark is inserted into transformed coefficients of image.

First, fragile watermarking that works directly in the spatial domain and second, that works in a transform domain. Spatial domain techniques embed the watermark in the least significant bit plane for perceptual transparency. Transform domain alters the frequency coefficients of the data elements to hide the watermark data. Transform domain has proved to be more robust than spatial domain technique [2]. Possible image transformations include the Discrete Fourier transformation DFT, Discrete cosine transformation DCT, Discrete wavelet transformation DWT by modifying the coefficients of global or block transform. Most fragile watermarking systems embed the mark in spatial domain described in Lin and Delp [3].

One of the first fragile watermarks for authentication was proposed by Walton [4]. In his scheme, key-dependent checksums of the seven most significant bits of grayscales along pseudo-random walks hides in the least significant bits of pixels forming the walk. First, in this scheme, attacker can modify image content while keep their LSB unchanged. Second, this cannot determine the exact regions of modification in verification process. Yeung

and Mintzer [5] proposed a scheme that authenticates individual pixels. This uses the secret key to generate a unique mapping that randomly assigns a binary value to grey levels of the image. This scheme has very good localization property but it is not easy to develop a feasible key management infrastructure without introducing security gaps.

Wong [6] described a scheme in which image is divided into blocks and each block contains the hash calculated from the MSBs in the LSBs of the pixels forming that block. The localization properties of this scheme are limited. Different attacks have been proposed to either break or increase the security of Wong's algorithm. A counterfeiting attack was proposed by Holliman and Memon [7]. This attack belongs to the class of vector quantization counterfeiting and defeats any fragile technique targeting localization accuracy by watermarking small image blocks independently. By making each watermark block dependent upon other blocks in watermarked image, the problem of watermark counterfeiting becomes infeasible. Another attack derived by Fridrich [8] that can be mounted against the Yeung-Mintzer scheme is that the lookup table and the binary logo can be inferred when the same lookup table and logo are reused for multiple images.

Chang [9] proposed a block-based watermarking scheme which divides the protected image into 3*3 image blocks and makes use of the cryptographic hash function for feature extraction of the blocks. The feature of image block is extracted from the eight neighboring pixels of the central pixel of the image block. The extracted feature is hidden into the central pixel of that image block. This scheme still allows pixels to be tampered because the feature of each image block does not depend on the central pixel of the image block. The block features computed for the original central pixel and the tampered central pixel are the same.

Zhang [10] used a scheme in which both pixel-derived and block-derived watermark data are carried by LSBs of all bits. During the extraction of watermark data, first blocks containing the tampered data is identified then the watermark data hidden in the rest blocks are checked to identify the tampered pixel. Disadvantage of this scheme is that if the percentage of ratio between the numbers of tampered blocks and that of all blocks is more than 5% then this will show that one half of the image is tampered but cannot locate the tampered pixels.

V. PROPOSED METHOD

This paper proposes the scheme in which the embedded watermark data are derived from pixels and blocks. On the receiver side, one can first identify the blocks containing the tampered content, and then use the watermark hidden in the rest blocks to exactly locate the tampered pixels.

The implementation process of fragile watermarking can be divided into following stages as

1. Watermark embedding in the image
2. Tampered- pixel localization
3. Restoration of original image

1. WATERMARK EMBED PROCEDURE

In the watermark embedding procedure, the 5 most significant-bit (MSB) planes in the host image are kept unchanged, and the 3 least-significant-bit (LSB) planes are replaced with watermark data. Here, the watermark data are determined by the MSBs and made up of two parts, which are respectively used to identify tampered blocks and to locate tampered pixels.

The detailed steps are as follows:

Denote the numbers of rows and columns in an original image as N1 and N2, the total number of pixels as N (N = N1 X N2), and the gray pixel-values $P_n \in [0, 255]$, n = 1, 2… N. Each P_n can be represented with 8 bits, $B_{n,7}, B_{n,6}, \ldots, B_{n,0}$, where

$$B_{n,u} = \lfloor P_n / 2^u \rfloor \bmod 2, \quad u = 0, 1, \ldots, 7 \qquad (1.1)$$

For each pixel, generate M authentication bits according to its 5most significant bits.

$$\begin{bmatrix} a_{n,1} \\ a_{n,2} \\ \vdots \\ a_{n,M} \end{bmatrix} = A_n \begin{bmatrix} B_{n,7} \\ B_{n,6} \\ \vdots \\ B_{n,3} \end{bmatrix}, \quad n = 1,2, \ldots \ldots, N \qquad (1.2)$$

Where $\mathbf{A_n}$ are pseudo-random binary matrices derived from a secret key, and their size is M X 5. To ensure security, the matrices $\mathbf{A_n}$ should be mutually different. The arithmetic in (2) is modulo-2, meaning that, if there is any change in the 5 MSBs of a pixel, the authentication bits will be flipped with a probability1/2.

1. According to a secret key, pseudo-randomly divide the M X N authentication bits into a series of subsets, each of which contains K bits. Then, calculate modulus-2 sums of the K authentication bits in each subset, and call the (M x N/K) results the sum-bits. Here, we let M be a multiple of 5and K = 2M/5 so that the number of sum-bits is 5N/2.

2. Assuming that both N1 and N2 are multiples of 8, we divide the original image into N/64 non-overlapped blocks sized 8x8. In each block, we pseudo-randomly select 160 positions from the 3 LSB-layers according to the secret key. Also, the LSB-selection in different blocks should be mutually different. Then, a total number of selected LSB is

5N/2, and replace the original bits at the selected positions with the sum-bits.

3. For each block, we collect the 320 original bits in the 5 MSB-layers and the 160 sum-bits used to replace the selected LSBs. Then, feed the 480 bits into a hash function to compute 32 hash-bits. Here, the hash function must have the property that any change on an input would result in a completely different output. Put the hash-bits into the 32 remaining positions in the 3 LSB-layers, and combine the original MSBs and the substituted LSBs to produce a watermarked image.

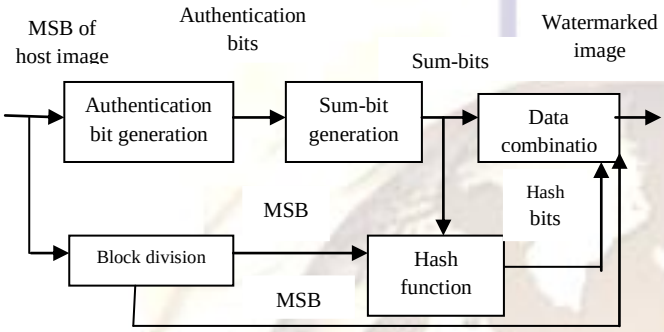

Fig. 3 Watermark Embedding Procedure

2. TAMPERED-PIXEL LOCALIZATION

Assume that an attacker may alter the gray values of some pixels without changing the image size. After receiving the image, we want to locate the tampered pixels and restore the original content. Here, ‘‘tampered pixels’’ are those with changes in their 5 MSBs. The tampered-pixel localization procedure is made up of two stages.

1. The first is to identify the tampered blocks. After dividing the received image into non-overlapped 8x8 blocks, we select 160 positions from the 3 LSB-layers in each block according to the same secret key. For each block, if the hash result of the 320 bits in the 5 MSB-layers and the 160 bits at the selected positions in 3 LSB-layers is identical to the 32 bits at the other LSB positions, the block is judged as “not tampered’’.

2. In the second stage, we locate the tampered pixels in the tampered blocks. Considering a pixel in tampered blocks, its M authentication bits are distributed into M subsets, each of which contains K elements. For each subset, if all the other (K-1) authentication bits in it are derived from pixels in ‘‘not tampered’’ blocks and its sum-bit is also hidden in a ‘‘not tampered’’ block, we say the subset is usable for the pixel. So, a receiver can derive the (K-1) authentication bits in usable subset from their corresponding pixels, and extract the sum-bit from its embedding position. Denote a ratio between the numbers of tampered blocks and that of all blocks as **α**.

The probability of a subset being usable for a certain pixel is

$$p_u = (1-\alpha)^K \quad (1.3)$$

Then, for a given pixel in tampered blocks, the number of the usable subsets, $\mathbf{n_u}$, obeys the following distribution:

$$P(n_u = t) = \binom{M}{t}.(1-p_u)^{M-t}.p_u^t \quad (1.4)$$

For each usable subset, we check whether or not the extracted sum-bit is consistent with the modulus-2 sum of the pixel’s authentication bit and other (K-1) authentication bits. If, and only if, the consistency is satisfied in all usable subsets, the pixel is judged as ‘‘not tampered’’, indicating that there is no alteration in its 5 MSBs. Otherwise, it Is judged as a ‘‘tampered’’ pixel. This way, a pixel without any alteration in its 5 MSBs must be judged as ‘‘not tampered’’, and probability with which a pixel containing modified MSBs is falsely judged as ‘‘not tampered’’ is

$$p_E = \sum_{t=0}^{M}[P(n_u = t).2^{-t}] \quad (1.5)$$

3. RESTORATION OF ORIGINAL IMAGE

After finding a ‘‘tampered’’ pixel, we can further recover its original MSBs. The number of possible patterns of 5 MSBs is 32. We attempt to use 31 other patterns different from the received pattern of the pixel to check consistency between the extracted sum-bit and the modulus-2 sum of the pattern’s authentication bit and other (K-1) authentication bits. When consistency is arrived in all usable subsets, the attempted pattern is regarded as the original MSBs. If more than one pattern satisfies the consistency condition in all usable subsets, restoration of original pattern will be failed since we do not know which one is the true original pattern. The true original pattern must satisfy the consistency condition in all usable subsets, and the probability of the other pattern satisfies the consistency condition in all usable subsets is 2^{-nU}. So, probability of failure is $1-(1-2^{-nU})^{30}$. Considering the distribution of n_u, probability for the original MSBs of a pixel with ‘‘tampered’’ judgment being unable to find is

$$p_c = \sum_{t=0}^{M}\{P(n_u = t).[1-(1-2^{-t})^{30}]\} \quad (1.6)$$

Denoting the number of tampered pixels labeled ‘‘tampered’’ but being unable to be restored as N_C, its average is

$$E(N_C) = p_C.(1-p_E).N_T \quad (1.7)$$

If $N_C = 0$, the receiver can obtain the original MSBs of all pixels, leading to restoration of the original watermarked image without any error.

VI. PERFORMANCE EVALUATION OF DIGITAL IMAGE WATERMARKING ALGORITHM

Performance evaluation of the algorithm is done with the perceptual transparency. Perceptual transparency means the quality of the image should not be destroyed by the presence of watermark. The quality of the watermark is measured by PSNR (Peak Signal to Noise Ratio). Bigger the PSNR better the quality of watermarked image.
PSNR is measued with the formula:

$$\text{PSNR} = 10 \log_{10} \frac{255^2}{\text{EQM}}$$

(1.8)

Where, EQM is the quadratic mean error between two images, I be an input image and I_w be an approximation that result from digital image watermarking. İmages are of size N*M.

$$\text{EQM} = \frac{1}{N*M} \sum_{i=1}^{N} \sum_{j=1}^{M} (I(i,j) - I_w(i,j))^2$$

(1.9)

where, $I(i,j)$ & $I_w(i,j)$ are the intensities of gray level images I and I_w at position (i,j).

İn this paper to improve the picture quality of restored image we use Gaussian filter. This filter does sharpening and smoothing of image gray value. By using Gaussian filter we get better image with fragile watermark. As we know bigger the PSNR, better the quality watermarked image.

Where, M is the authentication bits for each pixel according to its 5most significant bits.
The results are as follows:

Value of M	PSNR of watermarked image(dB)	PSNR of Gaussian filtered image
40	37.87	39.89
50	38.03	39.91
60	38.35	39.96

TABLE showing the comparsions of PSNR values of Watermarked image with Gaussian filtered image.

VII. CONCLUSION AND FUTURE WORK

The proposed fragile watermarking is based on hierarchical mechanism in which watermark data derived from the MSBs and directly replace all the LSBs of a host image. Watermarking embedding procedure is based on spatial domain. So, it's easy to implement. On the receiver side, first the tampered block is identified then the watermarks hidden in the rest blocks are used to exactly locate the tampered pixel and then restore the original watermarked version. By using Gaussian filter we get the better image with fragile watermark. So, improves the PSNR value. This scheme is also capable of recovering the original content and regenerating the watermarked version on the receiver side. This scheme have the property of both the methods block-wise and pixel-wise fragile watermarking.

In future work, we eliminate the disadvantage of [10]. So, in comparison with scheme described in [10], the proposed scheme will exactly locate the tampered pixel if the percentage of ratio between the numbers of tampered blocks and that of all blocks is more than 5%.

REFERENCES

[1] N. Memon and P. W. Wong, "Protecting digital media content: Watermarks for copyrighting and authentication," *Communications of ACM*, July 1998.
[2] I. Cox, J. Kilian and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia", *IEEE Transaction on Image Processing, vol. 6*, Dec. 1997, 1673-1687.
[3] Lin and E. Delp, "A Review of Fragile Image Watermarks," *in Proc. of the Multimedia and Security Workshop*, 1999, 25-29.
[4] S. Walton, "Information authentication for a slippery new age", *Dr. Dobbs J., vol. 20*, 1995, 18–26.
[5] M. M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification", *in Proc. ICIP, Santa Barbara, CA*, 1997.
[6] P. W. Wong, "A public key watermark for image verification and authentication", *in Proc. ICIP, Chicago, IL*, Oct. 1998.
[7] M. Holliman and N. Memon, "Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes", *IEEE Trans. Image Processing, vol. 9*, 2000, 432–441.
[8] M. Fridrich and A. Baldoza, "New Fragile Authentication Watermark For Images," *in Proc. of the IEEE International Conference on Image Processing, Vancouver, Canada*, 2000, 446–449.
[9] C. C. Chang, Y. S. Hu, and T. C. Lu, "A watermarking-based image ownership and tampering authentication scheme", *Pattern Recognition Letters, vol. 27*, April 2006, 439-446.
[10] X. Zhang and S. Wang, "Fragile Watermarking scheme using a hierarchical mechanism", *Signal Processing, vol. 89*, 2009, 675-679.