creative commons
COMMONS DEED

저작자표시-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.
- 이 저작물을 영리 목적으로 이용할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

[UCI]I804:24011-200000239140

2010 년 2 월

석사학위 논문

A Study on the Color Image Tamper Detection and Recovery Based on BWI Watermarking Scheme

조선대학교 대학원

전자공학과

왕 나

A Study on the Color Image Tamper Detection and Recovery Based on BWI Watermarking Scheme

2010 년 2 월 25 일

조선대학교 대학원

전자공학과

왕 나

A Study on the Color Image Tamper Detection and Recovery Based on BWI Watermarking Scheme

지도교수 김 정 화

이 논문을 공학석사학위신청 논문으로 제출함.

2010 년 2 월

조선대학교 대학원

전자공학과

왕 나

왕나의 석사학위논문을 인준함

위원장 조선대학교 교수 김종빈 (인)

위 원 조선대학교 교수 이강현 (인)

위 원 조선대학교 교수 김정화 (인)

2009 년 11 월

조선대학교 대학원

ABSTRACT

A Study on the Color Image Tamper Detection and Recovery Based on BWI Watermarking Scheme

Wang Na
Advisor: Prof. Kim Chung-Hwa Ph.D.
Department of Electronics Engineering,
Graduate School of Chosun University

For the information era, there exists a kind of technology that is information hiding. However, in that era, information hiding techniques received much less attention from the research community and from industry than cryptography. Until since the growth of computer power and digital era, people start to be concerned about protecting copyright because malicious attackers usually try to modify meaningful information of an image to change its meaning. Therefore, in order to solve this problem, steganography and watermarking has been presented. Steganography is a method that establishes a covered information channel in point-to-point connections, whereas watermarking does not necessarily hide the fact of secret transmission of information from third persons. For the watermarking, it requires a certain perceptual threshold allowing the insertion of additional

information to adjust between perceptible degradation and imperceptible degradation to fulfill visible or invisible watermarking. So, in this sense, watermarking represents a specific application of steganographic techniques.

In this thesis, we propose a digital watermarking scheme about color image of tamper detection and recovery using block-based watermarking. Firstly, in our watermark embedding, each color channel (R, G, and B) is divided into 2×2 non-overlapping blocks. The embedding space is created by setting the three LSBs (least significant bits) of selected each block to zero, which will hold the authentication and recovery codes. The authentication codes are obtained through doing exclusive-or (XOR) both the extracted local features of image blocks and the generated global features of the original color image. The recovery codes are generated using block-mapping original color image and using color model conversion. Secondly, in the tamper detection, instead of examining each embedded authentication data, we get the global features of the tampered watermarked image through doing XOR both all the extracted authentication codes and the generated local features of tampered watermarked image blocks. And then we utilize a majority-selecting technique to detect the legitimacy of image blocks. Lastly, in the tamper recovery, we propose 3 stages to accomplish recovery ability. Experiment demonstrates that our scheme is oblivious, correctly localizes the tampering wherever the tampered positions are and can be successfully accomplish to recover with acceptable visual quality.

Table of Contents

List of Tables

List of Figures

I. Introduction

A. Research Overview

1. Definition of Digital Watermarking

Digital watermarking is a process of embedding unobtrusive marks or labels (which we call the watermark) into digital content. These embedded marks are typically imperceptible (invisible) that can later be detected or extracted. The concept of digital watermarking is associated with steganography. Steganography is defined as covered writing. It has a long history of being associated with methods of secret communication. Steganography does not immediately evoke in the suspicion of something secret or valuable. Instead, it hides an important message in an unimportant one. Therefore, digital watermarking is a way to hide a secret or personal message to protect a product's copyright or to demonstrate data integrity

2. Purpose of Digital Watermarking

Watermarks added to digital content serve a variety of purposes. The following list details six purposes of digital watermarking:

a. Ownership Assertion

It used to establish ownership of the content (i.e. image)

Watermarks can be used for ownership assertion. To assert ownership of an image, Alice can generate a watermarking signal using a secret private key, and then embed it into the original image. She can then make the watermarked image publicly available. Later, when Bob contends the ownership of an image derived from this public image, Alice can produce the unmarked original image and also demonstrate the presence of her watermark in Bob's image. Since Alice's original image is unavailable to Bob, he cannot do the same. For such a scheme to work, the watermark has to survive image processing operations aimed at malicious removal. In addition, the watermark should be inserted in such a manner that it cannot be forged as Alice would not want to be held accountable for an image that she does not own.

b. Fingerprinting

It used to avoid unauthorized duplication and distribution of publicly available multimedia content

In applications where multimedia content is electronically distributed over a network, the content owner would like to discourage unauthorized duplication and distribution by embedding a distinct watermark (or a

fingerprint) in each copy of the data. If, at a later point in time, unauthorized copies of the data are found, then the origin of the copy can be determined by retrieving the fingerprint. In this application the watermark needs to be invisible and must also be invulnerable to deliberate attempts to forge, remove or invalidate. Furthermore, and unlike the ownership assertion application, the watermark should be resistant to collusion. That is, a group of k users with the same image but containing different fingerprints should not be able to collude and invalidate any fingerprint or create a copy without any fingerprint.

c. Authentication and integrity verification

The authenticator is inseparably bound to the content whereby the author has a unique key associated with the content and can verify integrity of that content by extracting the watermark

d. Content labeling

Bits embedded into the data that gives further information about the content such as a graphic image with time and place information

Information in a passport or ID (e.g., passport number, person's name, etc.) can also be included in the person's photo that appears on the ID. By extracting the embedded information and comparing it to the written text, the ID card can be verified. The inclusion of the watermark provides an

additional level of security in this application. For example, if the ID card is stolen and the picture is replaced by a forged copy, the failure in extracting the watermark will invalidate the ID card.

e. Usage control

It used to add to limit the number of copies created whereas the watermarks are modified by the hardware and at some point would not create any more copies (i.e. DVD)

Watermarks can also be used for copy prevention and control. For example, in a closed system where the multimedia content needs special hardware for copying and/or viewing, a digital watermark can be inserted indicating the number of copies that are permitted. Every time a copy is made the watermark can be modified by the hardware and after a point the hardware would not create further copies of the data. An example of such a system is the Digital Versatile Disc (DVD). In fact, a copy protection mechanism that includes digital watermarking at its core is currently being considered for standardization and second generation DVD players may well include the ability to read watermarks and act based on their presence or absence.

Another example is in digital cinema, where information can be embedded as a watermark in every frame or a sequence of frames to help investigators locate the scene of the piracy more quickly and point out weaknesses in

security in the movie's distribution. The information could include data such as the name of the theater and the date and time of the screening. The technology would be most useful in fighting a form of piracy that's surprisingly common, i.e., when someone uses a camcorder to record the movie as it's shown in a theater, and then duplicates it onto optical disks or VHS tapes for distribution.

f. Content protection

Content stamped with a visible watermark that is very difficult to remove so that it can be publicly and freely distributed

When multimedia content is used for legal purposes, medical applications, news reporting, and commercial transactions, it is important to ensure that the content was originated from a specific source and that it had not been changed, manipulated or falsified. This can be achieved by embedding a watermark in the data. Subsequently, when the photo is checked, the watermark is extracted using a unique key associated with the source. The integrity of the data is verified through the integrity of the extracted watermark. The watermark can also include information from the original image that can aid in undoing any modification and recovering the original. Clearly a watermark used for authentication purposes should not affect the quality of an image and should be resistant to forgeries. Robustness is not critical as removal of the watermark renders the content inauthentic and

hence of no value.

Unfortunately, there is not a universal watermarking technique to satisfy all of these purposes. The content in the environment that it will be used determines the digital watermarking technique. The following section describes some digital watermarking techniques.

3. Digital Watermarking Techniques

The most important properties of any digital watermarking techniques are robustness, security, imperceptibility, complexity, and verification. Robustness is defined as if the watermark can be detected after media (normal) operations such as filtering, lossy compression, color correction, or geometric modifications. To improve the robustness and protection, the concept of dual watermarking [1,2,3] is introduced by Swanson et al. [1]. In these entire schemes, authors embed one visible and one invisible watermark. When the visible watermarked image is in question, the invisibility watermark can provide rightful ownership. Security implies that the embedded watermark cannot be removed beyond reliable detection by targeted attacks. Similarly, imperceptibility implies that the watermark is not seen by the human visual system. Complexity is described as the effort and time required for watermark embedding and retrieval. Lastly, verification is a procedure whereby there is a private key or public key function.

Each of these properties must be taken into consideration when applying a

certain digital watermarking techniques. The following sections describe a few of the most common digital watermarking techniques.

Spatial and frequency domain watermarking are applied to graphic images and text. Spatial domain watermarking slightly modifies the pixels of one or two randomly selected subsets of an image. Modifications might include flipping the low-order bit of each pixel. However, this technique is not reliable when subjected to normal media operations such as filtering or lossy compression. Frequency domain watermarking technique is also called transform domain. Values of certain frequencies are altered from their original. Typically, these frequency alterations are done in the lower frequency levels, since alternations at the higher frequencies are lost during compression. The watermark is applied to the whole image so as not to be removed during a cropping operation. However, there is a tradeoff with the frequency domain technique. Verification can be difficult since this watermark is applied indiscriminately across the whole image.

4. Types of Attacks

Digital watermarking does not have the same capability or level of security as data encryption. It does not prevent the viewing or listening of content, nor does it prevent accessing that content. Therefore, digital watermarking is not immune to hacker attacks. The following are some intentional attacks on watermarks.

a. Active Attacks

Hacker tries to remove the watermark or make it undetectable. An example is to crop it out.

b. Passive Attacks

Hacker tries to determine whether there is a watermark and identify it. However, no damage or removal is done.

c. Collusion Attacks

Hacker uses several copies of one piece of media, each with a different watermark to construct a copy with no watermark.

d. Forgery Attacks

Hacker tries to embed a valid watermark of their own rather than remove one.

e. Distortive Attacks

Hacker applies some distortive transformation uniformly over the object in order to degrade the watermark so that it becomes undetectable.

These intentional attacks are just one of the barriers of why authors do not put their works into digital format. However, the government has stepped in to help these authors by establishing new laws.

5. Types of Digital Watermarks

a. *Robust watermarks*

They are designed to resist against heterogeneous manipulations; all applications presupposing security of the watermarking systems require this type of watermark.

b. *Fragile watermarks*

They are embedded with very low robustness. Therefore, this type of watermark can be destroyed even by the slightest manipulations. In this sense, they are comparable to the hidden messages in steganographic methods. This can be used to check the integrity of objects.

c. *Public and Private watermarks*

They are differentiated in accordance with the secrecy requirements for the key used to embed and retrieve markings. According to the basic principle of watermarking, the same key is used in the encoding and decoding

process. If the key is known, this type of watermark is referred to as public, and if the key is hidden, it is referred as private watermarks. Public watermarks can be used in applications that do not have security-relevant requirements (e.g., for the embedding of Meta information).

d. *Visible or localized watermarks*

They can be logos or overlay images in the fields of image or video watermarking. Due to the implicit localization of the information, these watermarks are not robust.

B. Research Objective

In this thesis, we adopt the spatial domain watermarking to embed watermark, which consists of authentication data and recovery data, fulfill the color image recovery. Although most of watermarking technologies are only limited for robustness and verification before, watermark for tamper detection and recovery [4-9] has received much attention recently. They are still in developing phase and have some shortcomings, so we need to keep on improving the recovery abilities and explore new better recovery method. However, the watermark is no longer a logo or a piece of message; it is the image itself. In this paper, the embedding space for watermark, which are authentication and recovery codes, is created by setting the three LSBs of selected each block to zero and is required to hold in the LSBs plane. The

authentication codes embedded in image blocks is generated by exploiting both the local features of image blocks and the global features of the images which are from hashing function to verify the integrity of an image. At the same time, the recovery codes are generated from block-mapping original image and they are also embedded original image blocks for accomplishing the recovery abilities. In the tamper detection process, the proposed scheme use generated global features of the tampered watermarked image and utilizes a majority-selecting technique [9] to determine the legitimacy of image blocks. In the tamper recovery process, we propose 3 stages to accomplish and improve recovery abilities using the recovery data generation and average intensity of its 3×3 neighboring block values. Experimental results show that the precision of tamper detection and localization is close to 100% under different positions attacks. In addition, the results reveal that the tampered regions can be approximately recovered with satisfactory image quality.

C. Research Layout

Formally, a watermarking system can be described by a tuple (C_O, W, K), where C_O is the set of all original data, W the set of all watermarks, and K the set of all keys. The generic watermark encoder and decoder are shown in the following Figure 1-1 and Figure 1-2.

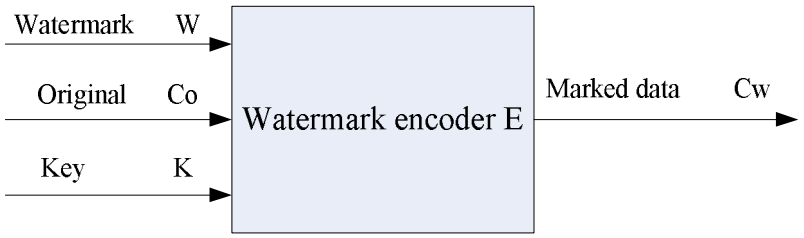


Figure 1- 1: Generic watermark encoder.

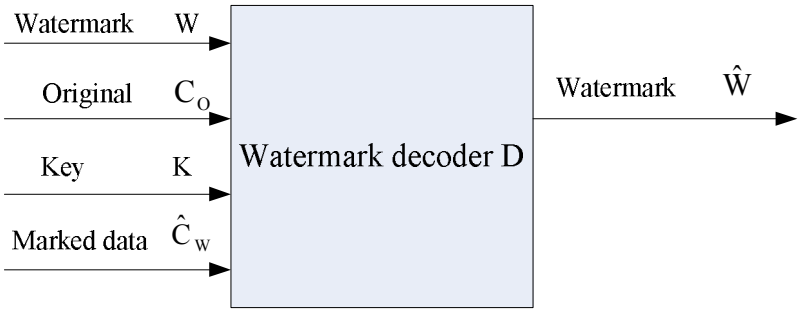


Figure 1- 2: Generic watermark decoder.

D. Research embedding algorithm

In the past several years, some watermark embedding algorithms, such as LSB(least significant bit) method for spatial domain and DWT(discrete wavelet transform) method for frequency domain, have been developed [10-13]. Recently, another embedding algorithm such as SVD (singular value decomposition) [14-15] can also accomplish the embedding.

- LSB method: LSB is also called the least significant bit. Grayscale images can be transformed into a sequence of binary images by breaking them up into their bit-planes. If we consider the gray value of each pixel of an 8-bit image as an 8-bit binary word, then the zeroth bit plane consists of the last bit of each gray value. Since this

bit has the least effect in terms of the magnitude of the value, so we can embed watermark signal into LSB to fulfill one properties of watermarking: imperceptibility. In fact, in the 8-bit binary bits, if watermark signal is too long to put in the zeroth bit plane, we can also choose the first and second bits to be embedded. In other words, watermarks embed into 3 LSBs of image can also accomplish imperceptibility of watermark. In this thesis we adopt such method to embed watermarks. Most of this watermarking embedding technology is just for grayscale image, but our objective image in thesis is 24-bit color scale image. So, we need to transform the 2 4-bit color image to 8-bit grayscale image. We choose one of the color transformation models that are (RGB->HSV) to transform to 8-bit grayscale images of 3 image matrix and it will be proposed the following part of thesis.

- DWT method: DWT is also called discrete wavelet transform. It is more popular and common technology than LSB method nowadays. The discrete wavelet transform can indeed be approached in terms of filtering; the two filters above are then known as the low-pass and high-pass filters, respectively. After discrete wavelet transform to image, we can get four images: the low-pass/low-pass image (LL), the low-pass/high-pass image (LH), the high-pass/ low-pass image (HL) and the high-pass/high-pass image (HH). Each one is half the size of original. It is applied to the original image. The watermarks embedded in low valued coefficients increases the robustness with

respect to attacks such as low pass filtering, lossy compression, and geometric distortions. The watermark inserted in middle and high valued coefficients are typically less robust to above-mentioned attacks but are highly resilient with respect to noise addition and nonlinear deformations of the gray scale. Three different formulas are suggested by the authors for embedding whose difference lies in their embedding characteristic and in their invertibility, they are shown in following part:

$$C_W(i,j) = C(i,j) + \alpha W(i,j) \quad (1\text{-}1)$$

$$C_W(i,j) = C(i,j)(1 + \alpha W(i,j)) \quad (1\text{-}2)$$

$$C_W(i,j) = C(i,j)\exp(\alpha W(i,j)) \quad (1\text{-}3)$$

- SVD method: The full name of SVD is singular value decomposition. This technique was explored for watermarking a few years ago. SVD is an important factorization of a rectangular real or complex matrix, with many applications in signal processing and statistics. The algorithm of SVD is shown in the following:

$$A = U \sum V^T \quad (1\text{-}4)$$

where, $A \in R^{M\times N}, U \in R^{M\times N}, V \in R^{M\times N}, \sum$: diagonal matrix.

Suppose A is an M-by-N matrix whose entries come from the field R, which is either the field of real numbers or the field of complex numbers. Using the singular value decomposition (SVD) algorithm, a real matrix A can be decomposed into a product of 3 matrices: they are left eigenvectors of A, singular values of A and right eigenvectors

of A, respectively. U and V are orthogonal matrices. U is an M-by-M unitary matrix over R, the matrix $\sum$ is M-by-N diagonal matrix with non-negative real numbers on the diagonal and V^T denotes the conjugate transpose of V, is an N-by-N unitary matrix over R.

Generated the singular values of matrix have three properties for watermarking:

1. Singular values can show the intrinsic characteristics of the image. Take advantage of this feature makes watermark embedding algorithm has good invisibility.
2. For singular values sequence, a common convention is to order the diagonal entries $\sum_{i,j}$ in decreasing fashion. In this case, the diagonal matrix $\sum$ is uniquely determined by image matrix, so we have chosen to embed a watermark into the singular values.
3. The most important property of singular values is that the largest of the modified singular values change very little for most types of attacks. So we use this property to fulfill the robustness.

E. Thesis Organization

The remainder of the paper is organized as follows. Section II describes the proposed block-based watermarking scheme including watermark generation and embedding procedure, tamper detection procedure and

tamper recovery procedure. In Section III, the experimental results are presented to demonstrate the effectiveness of the proposed scheme. Section IV derives the concluding remarks.

II. The Proposed Scheme

In this part, the following 3 procedures will be described in detail: (A) watermark embedding, which consists of block mapping, watermark generation and watermark embedding part. (B) tampered block detection, which can be accomplished through a majority-selecting strategy to identify the tampered blocks. (C) tampered block recovery, which can be achieved through a 3-stage error block recovery scheme.

A. Block-based Watermark Embedding

In this paper, we suppose that H is a 24-bit RGB color original image and the original image used in this paper is assumed to be of size M×M pixel, where M is assumed to be an even number. These color images are divided into non-overlapping 2×2 blocks $H_{i,j}(1 \le i, j \le \frac{M}{2})$. The flowcharts of watermark generation and watermark embedding are shown in Fig. 2-1, and the detail will be described as follows:

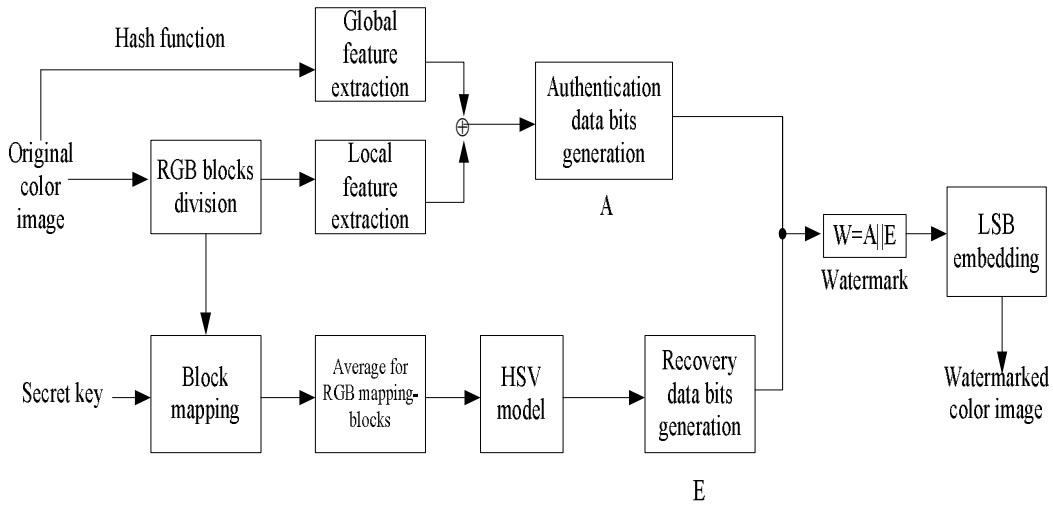


Figure 2- 1: Flowchart of watermark generation and embedding.

1. Block Mapping Sequence generation

A block-mapping sequence A→B→C→…→A is used for the recovery data embedding, where each symbol represents one block. The data for recovering block A will be embedded into block B, the data for recovering block B will be embedded into block C, and so on. Therefore, using a block mapping sequence can accomplish to encrypt watermark information.

A 1-D transformation algorithm described in Ref. [6], which is shown in Eq. (2-1), is used to obtain a one-to-one mapping sequence.

$$X' = [f(X) = (k \times X) \bmod N] + 1 . \qquad (2\text{-}1)$$

Where X, X' ($\in$ [0, N-1]) are the block number, k (a prime and $\in$ Z-{N's factors}) is a secret key and N is the total number of blocks in the image. In this paper k=37,N=128×128. Note that the secret key k must be a prime in

order to obtain a one-to-one mapping; otherwise, the period is less than N and a many-to-one mapping may occur.

Therefore, in this thesis, using a block mapping sequence can accomplish to encrypt watermark information into exchange pixel positions of block-mapping. We transfer the pixel positions of original image and do some operations for them, then generating result data bits will be recovery data bits. When attackers do some attacks for watermarked color image, some pixels of original image positions will be attacked, and then we need to extract the recovery data bits which are in mapping positions corresponding to original image positions and replace the tampered regions of original image positions. From above-mentioned content, we can know a block mapping sequence generation method is very important step for tampered image recovery.

The generation algorithm of the block-mapping sequence is as follows:

- Divide the image into non-overlapping blocks of 2×2 pixels.
- Assign a unique and consecutive integer $X \in \{1,2,\ldots,N\}$ to each block from left to right and top to bottom, where $N=(M/2)\times(M/2)$.
- Randomly pick a prime number $k \in [1, N-1]$.
- For each block number X, apply Eq.(2-1) to obtain X', the number of its mapping block.
- Record all pairs of X and X' to form the block mapping sequence

2. Block Watermark Generation

In the proposed paper, the authentication codes and the recovery codes are generated through being converted to binary from decimal data, as a 36-bit watermark is embedded into the three LSBs of all the color channels image blocks, where the authentication codes are used to identify any modification made to the watermarked image, while the recovery codes are utilized to restore those blocks which have been tampered with. For each color channels block of 2×2 pixels in R,G,B, the watermark in each color block is a 2-tuple (a , e), where a is 16-bit authentication watermark, and e is a 20-bit recovery watermark for the corresponding mapping block. The following algorithm describes how the 2-tuple watermark of length 36 bits of each color block is generated and embedded.

a. The Authentication Codes Generation

The authentication codes of length 16 bits is obtained through doing exclusive-or (XOR) operation both the 16-bit local features of each RGB channels pixel within the corresponding original image block and the 16-bit global features of the original image.

The local features used for generating the authentication codes are the color values and block indices of image blocks. For each pixel l ($1 \leq l \leq 4$) with each RGB channels block, a 16-bit binary sequence can be obtained by concatenating the five MSBs (Most Significant Bits) of all the color components of pixel l with its corresponding parity-check bit.

$$c_{i,j}^{l} = R_{i,j}^{l1} \| \cdots \| R_{i,j}^{l5} \| G_{i,j}^{l1} \| \cdots \| G_{i,j}^{l5} \| B_{i,j}^{l1} \| \cdots \| B_{i,j}^{l5} \| P_{i,j}^{l} \qquad (2\text{-}2)$$

where || denotes the concatenation operator. In addition, i and j are the indices of image blocks. The parity-check bit for pixel l is generated by

$$P_{i,j}^{l} = \begin{cases} 1, & \text{If } N_{i,j}^{l} \text{ is odd,} \\ 0, & \text{Otherwise.} \end{cases} \qquad (2\text{-}3)$$

Where $N_{i,j}^{l}$ denote the total number of 1 in the five MSBs of all the RGB components of pixel *l*.

The sequence $C_{i,j}$ that represents the local features of each block is then constructed by

$$C_{i,j} = c_{i,j}^{1} \oplus c_{i,j}^{2} \oplus c_{i,j}^{3} \oplus c_{i,j}^{4} \oplus I \oplus J, \qquad (2\text{-}4)$$

where $\oplus$ denotes the exclusive-or (XOR) operator. The indices of each block, i and j, are converted to the 16-bit binary representation sequence, I and J , respectively.

The global features applied to generate the authentication codes are the unique characteristics of an image using a cryptographic hash function to hash these unique characteristics to an image digest. In the proposed scheme, we adopt a well-known one-way hash function, MD5 [16]. Some fragile watermarking techniques [17-22] were usually based on the concept of checksum produced by secure hash functions (e.g. MD5) to verify the integrity of an image. The image digest for image H is generated by

$$d = \text{hash}(H) = (d_1, d_2, \ldots, d_{128}). \quad (2\text{-}5)$$

Since the MD5 hash function provides an image digest of 128 bits and the authentication codes are of length 16 bits, the image digest is needed to be converted from 128 to 16 bits. To achieve this goal, the 128-bit image digest is first divided into eight equal-sized sub-strings $\hat{d}$ (t=0,1,…7).

$$\hat{d}_t = d_{t\times 16+1} \| d_{t\times 16+2} \| \cdots \| d_{t\times 16+16}, t = 0,1,\ldots,7. \quad (2\text{-}6)$$

We are then able to generate a sequence of length 16 bits that characterizes the global features of image H with the following XOR operation

$$D = \hat{d}_0 \oplus \hat{d}_1 \oplus \cdots \oplus \hat{d}_7 \quad (2\text{-}7)$$

The authentication codes $A_{i,j}$ for block $H_{i,j}$ is finally constructed by XORing the two bit sequences $C_{i,j}$ and D derived from the local features of RGB blocks of image H and the global features of image H, respectively.

$$A_{i,j} = C_{i,j} \oplus D \quad (2\text{-}8)$$

b. The Recovery Codes Generation

The recovery codes embedded in block $H_{i,j}$ is composed of the RGB features of block $H_{m,n}$ whose index pair (m,n) is mapped to the pair (i,j) in

the block mapping sequence. The average intensity of each corresponding RGB feature within block $H_{m,n}$ to generate the color features, denoted by ($\overline{R}_{m,n}, \overline{G}_{m,n}, \overline{B}_{m,n}$), is first computed by

$$\overline{R}_{m,n} = \frac{1}{4}\sum_{l=1}^{4} R_{m,n}^{l}, \qquad (2\text{-}9)$$

$$\overline{G}_{m,n} = \frac{1}{4}\sum_{l=1}^{4} G_{m,n}^{l},$$

$$\overline{B}_{m,n} = \frac{1}{4}\sum_{l=1}^{4} B_{m,n}^{l}$$

Where $R_{m,n}^{l}$, $G_{m,n}^{l}, B_{m,n}^{l}$ are the color components of pixel l within block $H_{m,n}$.

It will take 24 bits to store the computed average color, but there are only 20 bits of storage that can be used for embedding the recovery data. Therefore, we need to transfer the color model from RGB model to HSV model. That is different with [9]. The reason why we use HSV model instead of YCbCr model which is proposed by [9] will be explained at the experimental result part of our proposed thesis.

We proceed with transforming the average color from RGB model into HSV model, denoted by ($h_{m,n}^{l}, s_{m,n}^{l}, v_{m,n}^{l}$). HSV model is a more intuitive method of

describing colors, and because the intensity is independent of the color information, this is very useful model for image processing.

The recovery data bits $E_{i,j}$ embedded in RGB features of block $H_{i,j}$ is created by concatenating all the bits of $h_{m,n}$ with the six MSBs of $s_{m,n}$ and $v_{m,n}$, that is, the two LSBs of the $s_{m,n}$ and $v_{m,n}$ are discarded.

$$E_{i,j} = h_{m,n}^{1} \| \cdots \| h_{m,n}^{8} \| s_{m,n}^{1} \| \cdots \| s_{m,n}^{6} \| v_{m,n}^{1} \| \cdots \| v_{m,n}^{6} \qquad (2\text{-}10)$$

3. Watermark Embedding

The generated authentication and recovery codes are embedded into the three LSBs of all the color channels of each pixel within the corresponding image block. The watermarked image H' is obtained after all the generated watermarks are embedded. How the 36-bit (a,e) are embedded in these four pixels of all RGB features of each block is shown in Fig. 2-2.

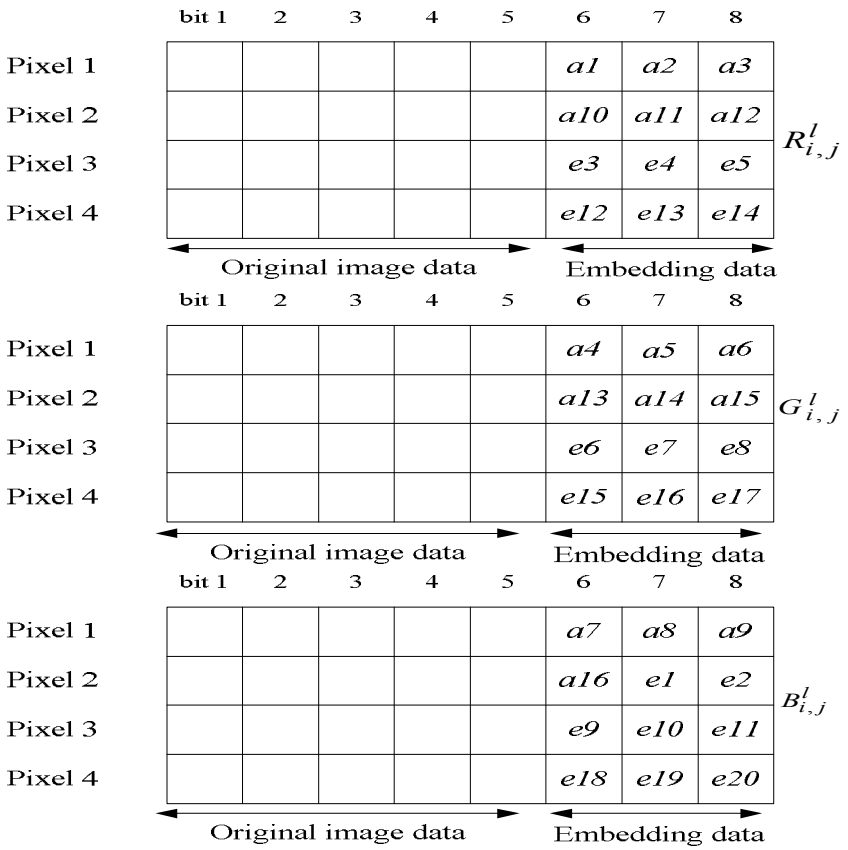


Figure 2- 2: The 36-bit watermark (a,e) embedded in pixel 1,2,3,4 of all RGB features of each block.

B. Block-based Tamper Detection

The flowchart of tamper detection is shown in Fig. 2-3. The detail steps about tamper detection are described below

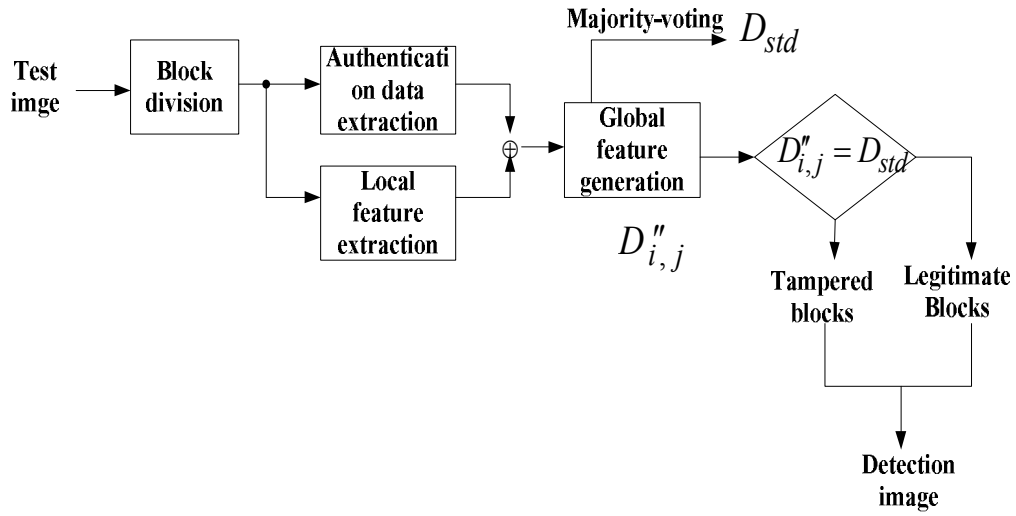


Figure 2- 3: The flowchart of block-based tamper detection.

Firstly, the test image H'' is divided into non-overlapping blocks of 2×2 pixels, as in the watermark embedding process. For each block, denoted as $H''_{i,j}$ $(1 \le i,j \le \frac{M}{2})$.

Secondly, the embedded authentication data bits $A''_{i,j}$ from the block $H''_{i,j}$ are extracted.

Thirdly, the color values and block indices of block $H''_{i,j}$, as in the watermark generation and embedding process, are then applied to generate the 16-bit sequence $C''_{i,j}$ that represents the local features of the block.

Lastly, the 16-bit sequence global features $D''_{i,j}$ of image H" is obtained by the following XOR operation.

$$D''_{i,j} = A''_{i,j} \oplus C''_{i,j} \quad (2\text{-}11)$$

If block $H''_{i,j}$ has been tampered, its corresponding sequence $D''_{i,j}$ would differ from those sequence derived from other blocks. Accordingly, we can identify the tampered blocks by employing a simple majority-selecting strategy, which means we choose the sequence with the maximum occurrence frequency from D'' as the standard sequence D_{std}, if $D''_{i,j} = D_{std}$, block $H''_{i,j}$ is legitimate; otherwise, it is regarded as a tampered block.

C. Block-based Tamper Image Recovery

After the detection stage, all the blocks are marked either valid or erroneous. There are 3 stages for the proposed recovery method.

We firstly list the following steps for recovery stage-1 and stage-2. The flowchart of tamper image recovery is shown in Fig.2-4 and detail steps will be described as follows:

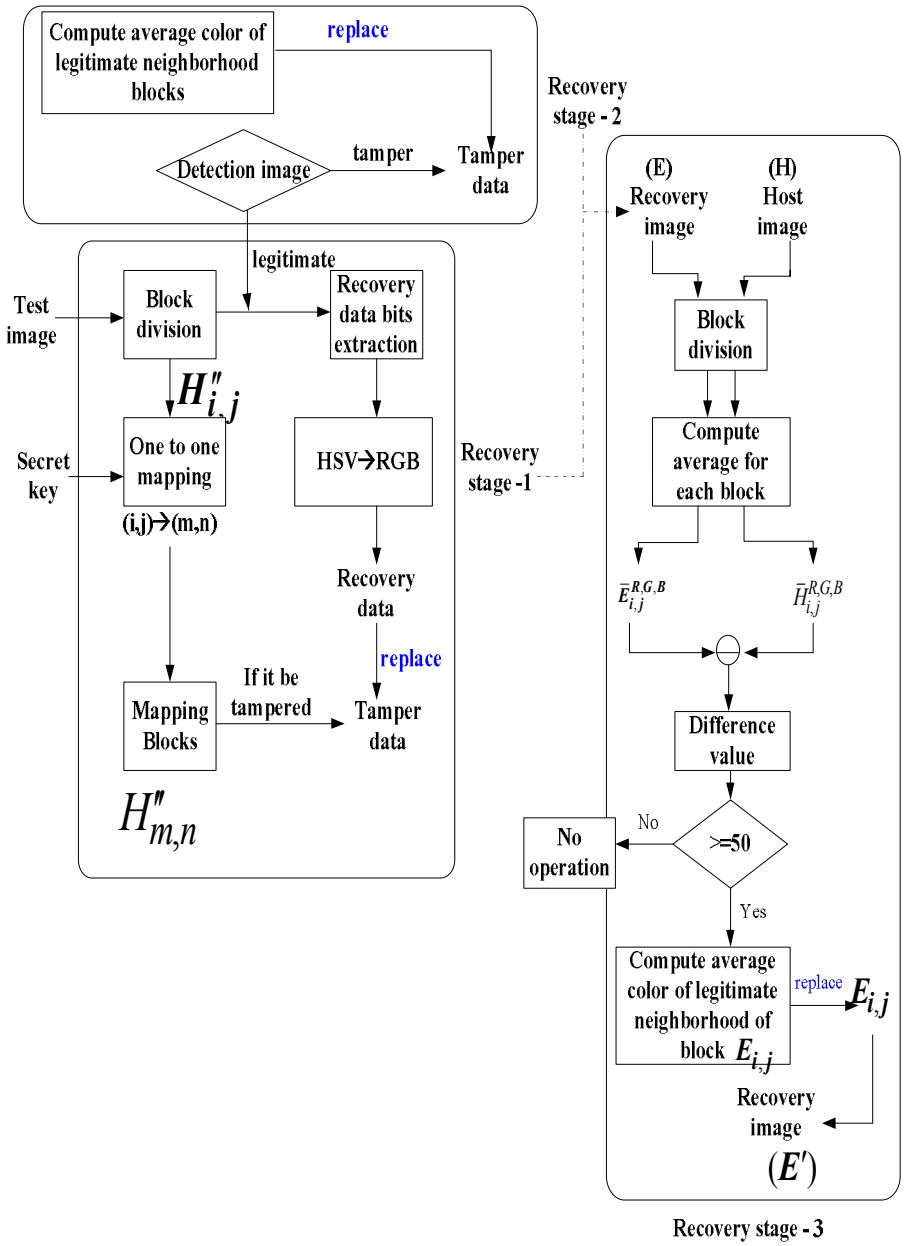


Figure 2- 4: The flowchart of block-based tamper recovery.

1. Stage-1 and Stage-2

Using block-mapping technique to recover each invalid block and mark the tampered block is valid for Stage-1 and using the average intensity of their 3×3 neighboring blocks values to recover the remaining invalid blocks.

- Use Eq. (1) and the secret key to locate block $H''_{m,n}$ in which the recovery data of block $H''_{i,j}$ is hidden and block-mapping detection image $D''_{i,j}$ to $D''_{m,n}$.
- If block $D''_{i,j}$ is legitimate, go to next step; otherwise go to step 6, which is also recovery stage2.
- Extract the recovery data $E''_{i,j}$ from block $H''_{i,j}$ and obtain the color features $h''_{i,j}$, $s''_{i,j}$, $v''_{i,j}$ of block $H''_{i,j}$.
- Pad the 6-bit $s''_{i,j}$ and $v''_{i,j}$, with two 0s to the end, respectively, and transform the color features from HSV model to RGB model to get $R''_{i,j}$, $G''_{i,j}$, $B''_{i,j}$.
- If block $D''_{m,n}$ is tampered, replace the color values of each pixel within $H''_{m,n}$ with the transformed color features $R''_{i,j}$, $G''_{i,j}$, $B''_{i,j}$ and mark block $D''_{m,n}$ valid. Go to Step 1 to recover the next tampered block.
- Compute the average color intensity of 3×3 legitimate neighboring blocks of each color block $H''_{i,j}$.

- Replace the color values of each pixel within block $H''_{i,j}$ with the average color intensity computed in Step 6, and go on Step 1 to recover the next tampered block.

2. Stage-3

We do the stage-3 for improving recovery abilities because of having some distortion of some extracted recovery codes.

Above-mentioned some distortion of some extracted recovery codes are due to block-mapping sequence. We know, if some pixels of original image positions are attacked, we need to extract the recovery data bits which are in mapping positions corresponding to original image positions and replace the tampered regions of original image positions. But if the tampered pixel of mapping position is also tampered, we can't get the right pixel to recover this tampered pixel. Therefore, we only use the stage-3 to solve such problem.

We make the difference value between the average intensity of recovered color blocks and the average intensity of original color blocks to find the positions of distortion pixels, if the difference value is more than threshold which is 50, we need to compute the average color intensity of legitimate neighboring blocks of block $\mathrm{H}''_{\mathrm{i,j}}$ again. At last, replace the color values of each pixel within block $\mathrm{H}''_{\mathrm{i,j}}$ with the average color values computed.

Finally, we accomplish all the recover image process to get the recovered color image.

III. Experimental Results

In this thesis, we have taken 24-bit color scale lena image and my photo as original images with size 256×256 is shown in Fig.5(a)(c). The watermarked image quality is measured through peak signal-to-noise ratio (PSNR). The watermarked images, which are given in Fig.5 (b)(d), are having PSNR values 42.958 dB and 42.5126 dB.

(a)

(b)

(c)

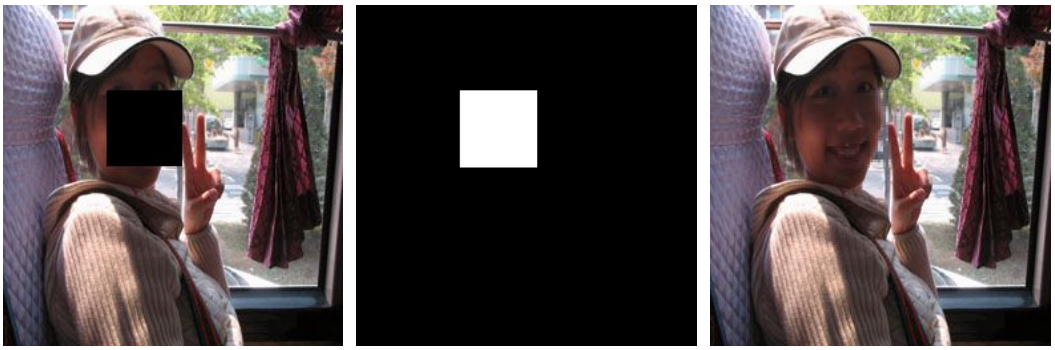
(d)

Figure 3- 1: (a) and (c) Original image, (b) and (d) Watermarked image.

The following few figures display the results of our proposed approach and we defined 4 kinds of attacks which are 5% tampered image, 10% tampered image, 20% tampered image and 30% tampered image at center and at upper left positions, respectively.

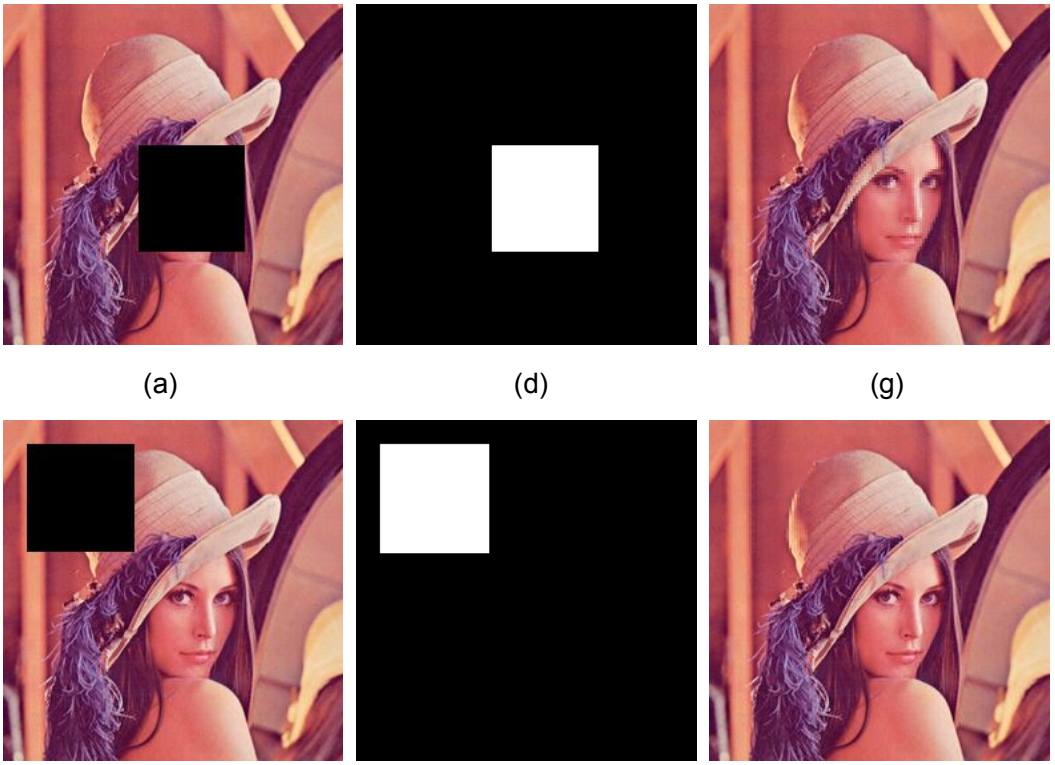
(a) (d) (g)

(b) (e) (h)



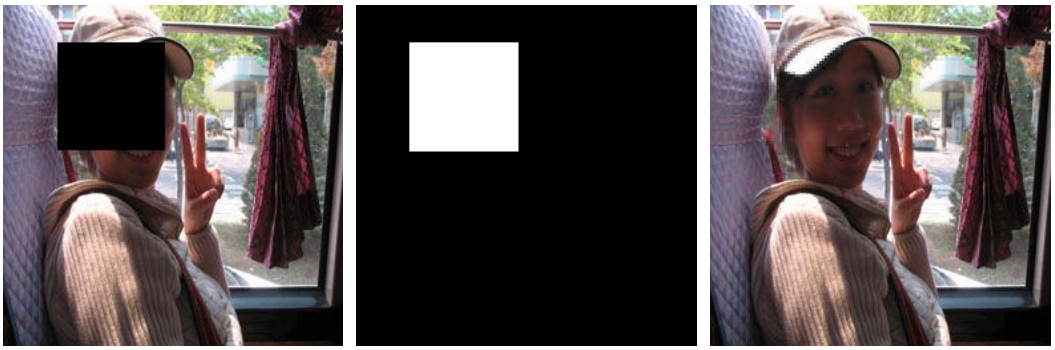
(c) (f) (i)

Figure 3- 2: (a)-(b) The 5% tampered lena image at center and at upper left, (c) The 5% tampered my image, (d)-(f) The detected tampered blocks of (a)-(c), (g)-(i) The recovered images of (a)-(c). (Note: white represents invalid, and black represents valid.).



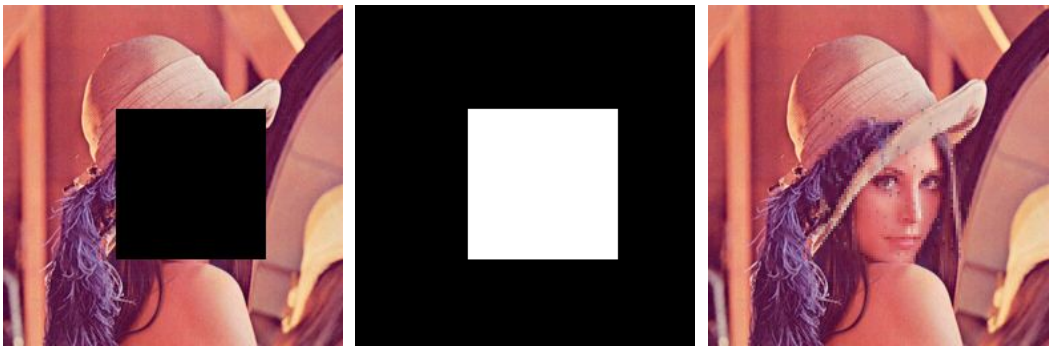
(a) (d) (g)

(b) (e) (h)



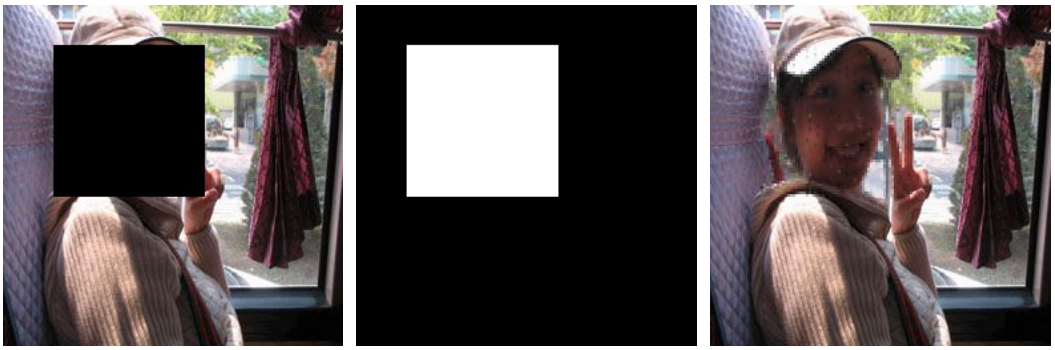
(c) (f) (i)

Figure 3- 3: (a)-(b) The 10% tampered lena image at center and at upper left, (c) The 10% tampered my image, (d)-(f) The detected tampered blocks of (a)-(c), (g)-(i) The recovered images of (a)-(c). (Note: white represents invalid, and black represents valid.).



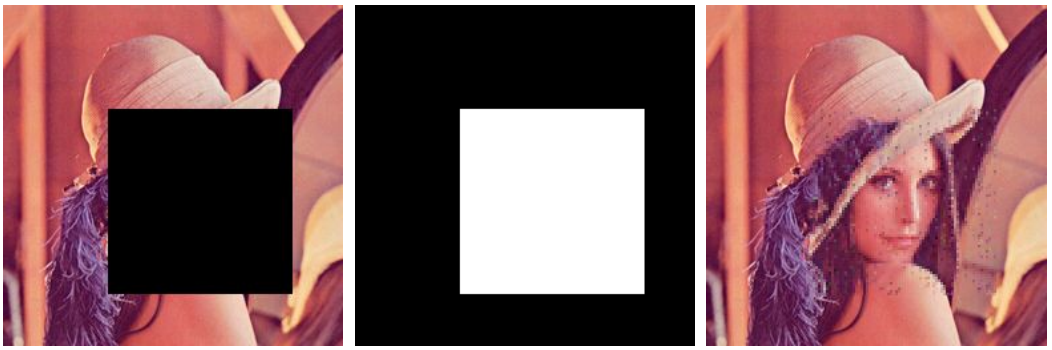
(a) (d) (g)

(b) (e) (h)

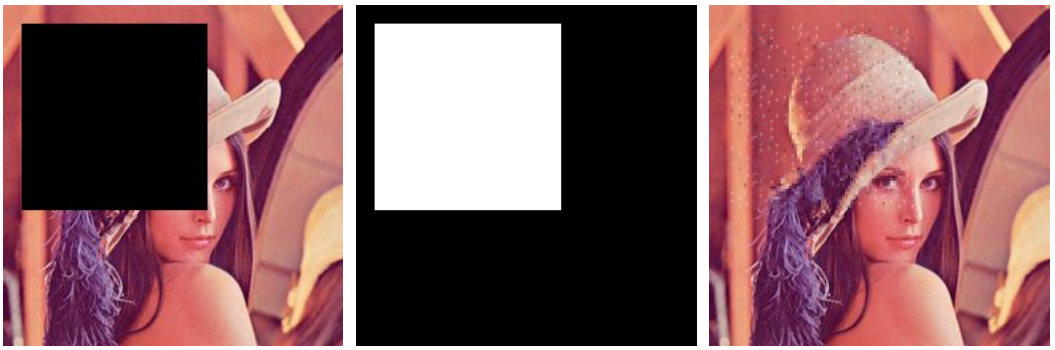


(c) (f) (i)

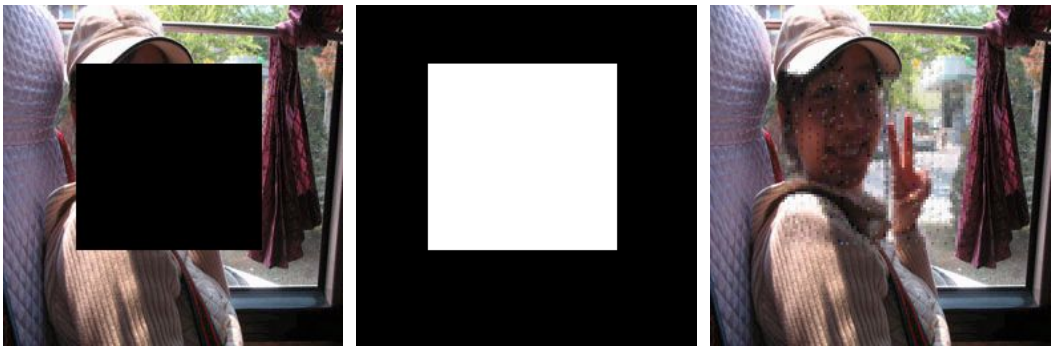
Figure 3- 4: (a)-(b) The 20% tampered lena image at center and at upper left, (c) The 20% tampered my image, (d)-(f) The detected tampered blocks of (a)-(c), (g)-(i) The recovered images of (a)-(c). (Note: white represents invalid, and black represents valid.).



(a) (d) (g)



(b) (e) (h)



(c) (f) (i)

Figure 3- 5: (a)-(b) The 30% tampered lena image at center and at upper left, (c) The 30% tampered my image, (d)-(f) The detected tampered blocks of (a)-(c), (g)-(i) The recovered images of (a)-(c). (Note: white represents invalid, and black represents valid.).

From above mentioned attacks for image to get, through using our proposed recovery method, we can also recover the image with good image visibility when the image is tampered by 30% image area. Irrespective of the tampered positions, they can be recovered well.

We need to explain the reason why we adopt HSV model instead of YCbCr model in here. We can use the contrast images which are shown in Fig.3-6 (a) and (b) to illustrate the reason.

(a) (b)

Figure 3- 6: (a) The recovery result of 10% tampered image using YCbCr model, (b) The recovery result of 10% tampered image using our proposed HSV model.

From the contrast result images to see, if we use YCbCr model method to recover, although the tampered image is recovered, some gray colors exist at the recovery part of tampered region. Using our proposed HSV model, we can overcome such problem.

In this thesis, we adopt correlation coefficient to show the similarity between the original watermarked image and recovered image. Equation (12) is about correlation coefficient.

$$\rho(\omega,\varpi)=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\omega(i,j)\varpi(i,j)}{\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\varpi^{2}(i,j)}\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\omega^{2}(i,j)}} \qquad (3\text{-}1)$$

Where ω is original watermarked image, ϖ is recovered image and m,n is the size of ω. The value of ρ lies between [-1,1]. If the value of ρ is equal to 1 then recovered image is just equal to original watermarked image. If the value of ρ is -1 then the difference is negative. In the following

TABLE I, we illustrate the recovered image quality through PSNR and list correlation coefficients of the recovered image.

At the same time, we also adopt the phrase peak signal-to-noise ratio, often abbreviated PSNR, which is very common in image processing. A sample use is in the comparison between an original image and a watermarked image. In this thesis, we use PSNR to compare the recovery image quality between a watermarked image and a recovery image.

Assume we are given a watermarked image f(i,j) that contains N by N pixels and a recovery image F(i,j). Error metrics are computed on the luminance signal only so the pixel values f(i,j) range between black (0) and white (255).

First we compute the mean squared error (MSE) of the recovery image as follows

$$\mathrm{MSE} = \frac{\sum [\mathrm{f}(\mathrm{i},\mathrm{j}) - \mathrm{F}(\mathrm{i},\mathrm{j})]^2}{\mathrm{N}^2} \qquad (3\text{-}2)$$

The summation is over all pixels. The root mean squared error (RMSE) is the square root of MSE. Some formulations use N rather than N^2 in the denominator for MSE.

PSNR in decibels (dB) is computed by using

$$\mathrm{PSNR} = 10\log_{10}(\frac{255^2}{\mathrm{RMSE}}) \qquad (3\text{-}3)$$

Because our original image and recovery image are color image, so in this thesis, we use the following algorithm to measure the PSNR between watermarked image and recovery image.

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{255^2 * 3}{\mathrm{RMSE}}\right) \quad (3\text{-}4)$$

Typical PSNR values range between 20 and 50. They are usually reported to two decimal points. The greater value, the better image quality can be shown. The actual value is not meaningful, but the comparison between two values for different recovery images gives one measure of quality.

Attacks \ **Positions**		***PSNR (dB)***	ρ
5% tampered	Center	44.362	0.99991
	Left	49.639	0.99997
	Mine	43.2560	0.9997
10% tampered	Center	41.348	0.99981
	Left	45.781	0.99994
	Mine	41.18	0.99956
20% tampered	Center	36.625	0.99939
	Left	39.263	0.99963
	Mine	36.07	0.99861
30% tampered	Center	33.922	0.99877
	Left	34.784	0.99927
	Mine	33.166	0.99728

Table 3- 1: PSNR of recovery image and the correlation coefficients of recovered image

Our proposed recovery method is obviously better than Ref.[6]. In [6], they have 68×68 tampered size of original image, which is about 7% tampered region, its PSNR value is 30.85 dB; for our proposed method at 10% tampered region, PSNR value is above 40 dB.

Because Ref.[9] use another method for measuring image quality, so we can't compare the results with ours directly. But we can see in Ref.[9], there is a vegetable image, we can see the tampered region of vegetable image is small, its PSNR value is 38.86 dB. We can estimate our proposed method is also better than [6].

IV. Conclusion and Future Work

In this paper, color image of tamper detection and recovery using block-based watermarking is presented. The proposed scheme embeds a 36-bit watermark consists of the authentication codes and the recovery codes into color features of each image block. In the tamper detection process, we use the generated global features of tampered watermarked image and employ a majority-selecting technique to examine the legitimacy of image blocks. Experimental results show our recovery images quality using proposed recovery method is better than others [6][9] and wherever the tampered positions are, it can sustain superior accuracy of tamper localization and can be successfully recovered with acceptable visual quality. From the result table, we also can find the recovery ability at left-tampered images is better than at center-tampered images.

For the thesis, we adopt the block-mapping technology to assure the main recovery function. However, from above-mentioned content we all know, it also has a problem. If the tampered pixel of mapping position is also tampered, we can't extract the right pixel to recovery this tampered pixel. So according to this problem, in the near future, we need to research how to mapping the pixel or block positions to enough distance where can not be also tampered.

Bibliography

[1] M. D.Swangson, B. B. Zhu and A. H. Tewfik, “Robust audio watermarking using perceptual masking,“ Signal processing Elesvier, vol.66, 1998, pp. 337-355.

[2] S. P. Monhanty, K. R. Ramakrishnan and M. Kankan-halli, “A Dual watermarking Technique for Images,” Proceedings of the seventh ACM international conference on Multimedia, Orlando, Florida,USA, 1999, pp.49-51.

[3] Y. Hu, S. Kwong and J. Huang, “Using Invisible watermarks to protect visibly Watermarked Images, “ Proceedings of the International Symposium on Circuits and Systems, vol. 5,2004, pp.584-587.

[4] K.-F. Li, T.-S. Chen, S.-C. Wu..: Image Tamper Detection and Recovery System Based on Discrete Wavelet Transform. In: IEEE Pacific Rim Conference on Communications, Computers and Signal processing, vol. 1, pp.164--167.(2001)

[5] P.-L. Lin, P.-W. Huang, A.-W. Peng.: A Fragile Watermarking Scheme for Image Authentication with Localization and Recovery. In: IEEE Sixth

International Symposium on Multimedia Software Engineering, 13--15, pp. 146--153. (2004)

[6] P.-L. Lin, C.-K. Hsieh, P.-W. Huang.: A Hierarchical Digital Watermarking Method for Image Tamper Detection and Recovery. In: Pattern Recognition 38 pp. 2519—2529. (2005)

[7] C.-L. Wang, R.-H. Hwang, T.-S. Chen, H.-Y. Lee.: Detecting and Restoring System of Tampered Image Based on Discrete Wavelet Transformation and Block Truncation coding. In: 19th International Conference on Advanced Information Networking and Applications. (2005).

[8] T.-Y. Lee, S.-D. Lin.: Dual Watermark for Image Tamper Detection and Recovery. In Pattern Recognition 41 pp. 3497--3506. (2008).

[9] M.-S. Wang, W.-C. Chen..: A Majority-voting Based Watermarking Scheme for Color Image Tamper Detection and Recovery. In: ScienceDirect Computer Standards & Interfaces 29 pp. 561—570. (2007).

[10] X. Xia, C. G. Boncelet and G. R. Arce, “A multi-resolution watermark for digital images,” Proceedings of IEEE Int. Conf. Image Processing, Santa Barbara, CA, vol.3, 1997, pp. 548-551.

[11] S. H. Wang and Y.P. Lin, “Wavelet tree quantization for copyright

protection watermarking," IEEE transactions on Image Processing, vol. 13, No. 2,2004, pp. 154-165.

[12] Z. Dawei, C. Guanrong and L. Wenbo, "A chaos-based robust wavelet-domain watermarking algorithm," J. Chaos Solitons Fractals, vol.22, 2004, pp. 47-54.

[13] P. Meerwald and A. Uhl, "A survey of Wavelet-Domain Watermarking Algorithms," Proceedings of SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents III, San Jose, CA, USA, vol. 4314, 2001.

[14] R. Liu and T. Tan, "An SVD-Based Watermarking Scheme for Protecting Rightful Ownership," IEEE Transactions on Multimedia, vol.4, no. 1,2002, pp. 121-128.

[15] E. Ganic and A. M. Eskicioglu, "Robust Embedding of Visual Watermarks Using DWT-SVD," Journal of Electronic Imaging, vol. 14, no. 4, 2005.

[16] W. Stallings, Cryptography and Network Security: Principles and Practice, Prentice-Hall, New Jersey, 2003.

[17] Cox, I. J., Kilian, J., Leighton, F. T., Shamoon, T., 1997. "Secure spread spectrum watermarking for multimedia." IEEE Trans. on Image Proc., vol. 6,

no. 12, pp.1673-1687.

[18] Chang, C.C., Hu, Y. S., Lu, T. C., 2006. “A watermarking-based image ownership and tampering authentication scheme.” Pattern Recognition Letter, 27, pp. 439-446.

[19] Walton, S., 1995. “Information authentication for a slippery new age.” Dr. Dobbs J.20 (4), pp.18-26.

[20] Schyndel, R. G., Tirkel, A.Z., Osbome, C.F., 1994, “A Digital Watermark”, Proceedings of the IEEE International Conference on Image Processing, Austin, Texas, vol.2, pp.86-90.

[21] Wolfgang, R. B., Delp, E. J., 1996, “ A Watermark for Digital Images“, Proceedings of IEEE International Conference on Image Processing, Lausanne, Switzerland, vol. 3, pp. 219-222.

[22] Ding, K., He, C., Jiang, L.G., Wang, H.X., 2005. “Wavelet-based semi-fragile watermarking with tamper detection.” IEICE Transactions on Fundamentals of Electronics, Communications and Computer Science, E88-A(3), pp. 787-790.