RESEARCH ARTICLE

An Efficient Self-Embedding Watermarking Scheme for Colour Image Tamper Detection and Recovery

Shruthy V C¹, Saira Varghese²

¹Computer Science Department & CUSAT, India
²Computer Science Department & CUSAT, India
¹shruthyvc91@gmail.com; ²saira.manoj@gmail.com

Abstract—Digital watermarking is a method for inserting the watermark information into an image, which can be later extracted for various purposes including copyright identification and content authentication. In this study, the watermarking issue for colour image authentication is presented to resist malicious tampering. The main aim of this paper is to preserve the colour moments such that the tampered region of the colour image can be recovered with high quality. Here our input is a colour image in YCbCr colour space. Each colour component of the image is first partitioned into non-overlapping blocks of size 8x8 for watermark embedding. Since we divide each colour component into non-overlapping blocks of size 8x8 we can achieve a better tamper localization capability. The watermark information embedded into each block consists of the authentication data and the feature information. The authentication data is generated with the help of two stage dual parity check method and feature information is generated through bi level moment preserving technique. In this system the use of an optimization technique known as Artificial Bee Colonyization will help us to locate the watermark bits at best location. So the attackers feel difficulty to create attacks on the watermark bits. In detection process, the authors propose a two-stage dual parity-check method and use morphological operations to prove the validity of the image blocks. The first method is proposed to check two set dual authentication data for obtaining better results of tamper detection and the latter is used to further improve the neighborhood connectivity of the results. The use of morphological operations like erosion, dilation, open and close will yield better results because it can filter single and multiple noises in the image. In the recovery process, the feature information of each block is used for reconstruction. The simulation result shows that the proposed self-embedding watermarking scheme is able to effectively detect the tampered region with high detection rate and recover the tampered region with high quality. This scheme outperforms most of the existing watermarking methods. This method has got a better PSNR value and can decrease bit error rate and increase accuracy of detection.

Keywords—Self-embedding, Bi level moment preserving technique, two stage dual parity check method, Artificial Bee Colony Optimization, Morphological operations
I. INTRODUCTION

Image authentication techniques have recently gained great attention due to its importance for a large number of multimedia applications. Digital images are increasingly transmitted over non-secure channels such as the Internet. Therefore, military, medical and quality control images must be protected against attempts to manipulate them; such manipulations could tamper the decisions based on these images [1]. To protect the authenticity of multimedia images, we have a lot of methods. These methods include conventional cryptography, watermarking methods and digital signatures. A good image authentication system has the properties like sensitivity, robustness, localization, recovery, security, portability, and also it should be neither complex nor slow. Now a day’s all of us have internet connectivity and all multimedia documents can be conveniently accessed through the internet. We know that there are a lot of image processing tools and most of them are available free of cost. So anyone can download images and then perform attacks on it easily. In order to protect our images we can make use of watermarking techniques.

The aim of watermarking is to include subliminal information in a multimedia document to ensure a security service or simply a labelling application. This can be possible to recover the embedded message at any time, even if the document was altered by one or more non-destructive attacks, whether malicious or not [2]. We have different types of watermarking schemes. First one is fragile watermarking scheme where we insert a specific watermark so that any attempt to alter the content of an image will also alter the watermark itself. It yields better tamper detection. But we cannot recover the original image when it is tampered. Second one is semi fragile watermarks it is less sensitive to classical user modifications. Third one is robust watermarking scheme, where a robust watermark is embedded into the image. Fourth one is block based; here we are dividing the original image into blocks and embedding a robust watermark on each block and check the integrity of the block. Fifth one is feature based watermark, in which the features of the image are extracted and then stored in pixels. During tamper detection extracted and computed features are compared, if they match then valid else invalid. The last one is self-embedding watermarking, where we embed an approximation image into itself.

Now a day’s there has been an explosive growth in the digital media processing. We have a lot of techniques to achieve image authentication. But the problem is some of them are concentrating only on gray scale image authentication and recovery some others for colour image authentication, and some for recovery. We have a lot of techniques that achieves tamper detection and recovery of colour images, but the recovery quality of the images are compromised in most of the schemes.

A. Problem Statement

KC Liu’s self-embedding watermarking scheme [3] concentrates on improving recovery quality and it achieves a better PSNR value of around 42dB. Other than lower PSNR value we have other problems like 1) insecure mapping, 2) compromised quality, 3) problems related to watermark embedding position and 4) exploration of strong image processing tools.

1) Use of insecure block mappings
There are some watermarking methods that divide the original image into blocks and use a block mapping sequence to embed watermark information of one block to its mapped block. We know that one to one mapping is insecure mapping. Because of the limited number of freedom, linear transforms are easily recovered from a few sample images, and are weak from security point of view. Insecure mapping lead the system vulnerable to four scanning attack and synchronous counterfeiting attacks.

2) Quality of recovered image is compromised
Most of the watermarking schemes have lower PSNR value (below 40dB) which will make the detection of difference between original & recovered image by human eyes easier.

3) Embedding watermark bits in LSBs
Most of the watermarking schemes embed watermark bits in their LSBs, so they can be recovered easily and are vulnerable to lossy compression.

4) Exploration of Strong image processing tools
Due to the exploration of strong image processing tools images are vulnerable to security attacks.

We know that if the recovered image has a PSNR value of 40dB or above then the human eyes cannot distinguish between original image and recovered image. So our aim is to improve the PSNR value. So we can adopt the Liu’s scheme and then improve the scheme by using optimization technique, and key based pseudo random permutation.

The rest of the paper is organized as follows. Related works will be discussed in section 2. In Section 3, the technique known as bi level moment preserving technique is described. The proposed watermarking scheme is
II. RELATED WORKS

The watermarking techniques are categorized as 1) Watermarking methods for gray scale image authentication [4 - 6], 2) Watermarking methods for gray scale image authentication and recovery [7 - 11], 3) Watermarking methods for colour image authentication and recovery [12 - 16] and 4) Watermarking methods concentrating on improving recovery quality [3, 17&18]. That is some methods are concentrating only on authentication, some others are for recovery and if both met then the quality will be compromised. We are interested in improving recovery quality. In Wang’s scheme [37] automatic image authentication and recovery is proposed. Here they make use fractal encoding and image inpainting methods and achieves a recovery quality of around 38dB. Luis [18] use halftone of an image as its watermark and embed it into LL sub band of IWT. Then recovery of tampered image is done through previously trained MLP and so is complex. Liu’s self embedding watermarking scheme [3] achieves a PSNR value of around 42dB. But it has got some disadvantages 1) insecure watermark embedding position, 2) problems with watermark embedding position.

III. BI LEVEL MOMENT PRESERVING TECHNIQUE

Moment-preserving technique is an image thresholding method. It classifies the pixels of a given image into many groups and the pixels in each group are assigned to a certain grey value such that the moment of the thresholded image and moment of original image are same. Bi level thresholding means to classify the pixels of a given image into two groups one including those pixels with their gray values above a certain threshold, and the other including those with gray values equal to and below the threshold. If we use more than one threshold, that is if we have t thresholds then we have (t+1) representative grey values. This is known as multilevel thresholding. In this paper, we are using the bi-level method because of its simplicity.

Given an image I with n pixels whose gray value at pixel (x, y) is denoted by I(x, y), we want to threshold I into two pixel classes, the below-threshold pixels and the above-threshold ones. As depicted in [19], the moments of the input image are computed before thresholding in order to determine the threshold and two representative grey values. The jth moment of the input image I is defined as follows

\[ m_j = \sum_{i=1}^{N} (I(i))^j \]  

(1)

Where j is the order of the moment, I(i) is the grey value of the pixel i and N is the total number of pixels in the image I. By using the histogram of the image I, (1) can be rewritten as

\[ m_j = \sum_{k=1}^{N} n_k (c_k)^j \]  

(2)

Where n_k and b_k are the number and the fraction, respectively, of the pixel with grey value c_k in the image. The same formula can be applied to the bi-level image I’, that is, the thresholded image maintaining the same moment of the image I. That is, I’ is obtained by replacing each pixel in the image I with one of the two representative grey values according to the classification result. In this case, we assume that the pixels below the threshold are replaced by c_l and the pixels above the threshold are replaced by c_h. c_l and c_h are the two representative grey values in the bi-level moment-preserving technique. Thus, (2) can be extended to give the following equations for the image H while the first three-level moments are preserved.

\[ b_l c_l + b_h c_h = m_l \]  

(3)

\[ b_l c_l^2 + b_h c_h^2 = m_l \]  

(4)

\[ b_l c_l^3 + b_h c_h^3 = m_l \]  

(5)

\[ b_l c_l^4 + b_h c_h^4 = m_l \]  

(6)

Where b_l and b_h are the fractions of the pixels with the two representative grey values c_l and c_h, respectively, in the thresholded image I’. The reason why we require the above equations of the first three-level moments is to solve c_l, c_h, b_l, and b_h. From (3) to (6), the two representative grey values can be solved as follows

\[ c_h = \left(-z - \sqrt{z^2 - 4y}/2\right) \]  

(7)

\[ c_l = \left(-z + \sqrt{z^2 - 4y}/2\right) \]  

(8)

For

\[ y = (m_3m_3 - m_2^2)/(m_0m_2 - m_1^2) \]  

(9)

\[ z = (m_3m_3 - m_2m_3)/(m_0m_2 - m_1^2) \]  

(10)

Then, the class fractions p_l and p_h can be solved and given by

\[ b_l = (c_l - m_l)/(c_h - c_l) \]  

(11)

\[ b_h = 1 - b_l \]  

(12)
as the bi level thresholding method described in [19], the threshold $t$ can be determined while the following equivalent equation is achieved.

$$b_x = \frac{\sum_{k\in\mathbb{N}} n_k}{N}$$  \hspace{1cm} (13)

Actually, the accumulation result in (13) is not exactly equivalent to $b_x$ for discrete grey values; the threshold could thus be selected such that the accumulation result is closest to $b_x$. As described above, we use this threshold to classify the pixels of the image $I$ into two groups where the grey values of the pixels in each group are, respectively, assigned to $c_x$ and $c_y$ such that the moment of $I$ is maintained.

In the prior watermarking works presented in [9–11], the image is first divided into non-overlapping blocks. The block mapping process is used to embed the feature information of each block to its mapping block. The feature information of each block is essentially the statistics of the block in the image for recovery of the tampered region. Although the watermarked image is tampered by malicious attacks, the block in the tampered region can be restored by the corresponding feature information. In these works, the feature information of each block is regarded as the reduced content of the block and is actually its average intensity. An example is given in Fig. 1 to show the comparison of a block of $8 \times 8$ pixels and its recovered blocks as the block is tampered. Figs. 1a and b show the original block pattern and its recovered block pattern after recovering with the average intensity, respectively. It is obvious that the recovery quality is unsatisfied, especially when the image is treated as the evidence for the court judgments. On the contrary, the bi-level moment-preserving technique can be adequately utilized to give the feature information of the block consisting of a threshold and two representative values for better recovery quality. Fig. 1c shows the recovered result of applying the moment-preserving technique to Fig. 1a. It clearly keeps the pattern of Fig. 1a and illustrates better performance than Fig. 1b. If we use three levels or any other higher level we can achieve more recovery quality. Here we use bi level moment preserving because of its simplicity.

**IV. PROPOSED SYSTEM**

Here we are discussing about the proposed self embedding watermarking scheme. This scheme is applicable for colour images in $YCbCr$ colour space. We know that human eyes are more sensitive to changes in luminance rather than chrominance. In watermarking we hide watermark data in colour components. If we use the classical RGB components, all of them contribute to the illumination and each have a distinct colour. So while this is a good “additive colour” representation for displays (TVs), it’s not so nice to represent/store/transmit the image in that form. This scheme includes two different phases. First phase is watermark embedding and the second phase is tamper proofing and tamper recovery phase. In watermark embedding phase we have block division and block mapping and then watermark generation and embedding. In tamper proofing and recovery phase, we have block division and mapping, watermark extraction, and recovery phases. Figure 2 shows an overview of the system.

A. Watermark Embedding Phase

The Watermark embedding phase occurs in the sender side. In watermark embedding phase, first of all divide the colour image into non overlapping blocks of size $8 \times 8$. Then use a key based pseudorandom
permutation to obtain the block mapping sequence. Then generate authentication data and then generate the feature information. The authentication data and feature information are together called watermark data. The watermark of block B1 is embedded in block B2, which is obtained from the secure block mapping information. Then embed the watermark in a position which is obtained from the ABC optimization technique. Now we got the watermarked image. Figure 2 shows the functional block diagram for watermark embedding.

1) **Block Division and Block Mapping**: The host colour image in YCbCr colour space is first divided into non-overlapping blocks of size 8x8 and is done for each colour component. We can achieve a better tamper localization through block division of 8x8. We know that insecure mapping has two problems. First it is vulnerable to four scanning attack and synchronous counterfeiting attack; second it can be recovered from few sample images. Then a secure mapping sequence is generated with the help of a key based pseudorandom permutation. A block mapping sequence S is computed from a key based pseudo random permutation \[S (1) \ldots S (N)\] of the integer interval \[1 \ldots N\].

2) **Watermark Generation**: The watermark consists of authentication data and feature information. So in order to generate a watermark we have to generate authentication data and feature information. Authentication data can be generated using two stage dual parity check method. For that first of all we have to set two LSBs of each block to zero and then compute the average intensity of the block. This average intensity is represented as K here. This K is converted into binary and we get \[K_7 \ K_6 \ K_5 \ K_4 \ K_3 \ K_2 \ K_1\]. Authentication data is of four bits. The dual parity check codes \((r_1, s_1)\) and \((r_2, s_2)\) can be generated from the following equations.

\[
\begin{align*}
r_1 &= \begin{cases} 
    k_7 \oplus k_5 \oplus k_3, & \text{if } q \mod 2 = 0 \\
    k_6 \oplus k_4 \oplus k_2, & \text{if } q \mod 2 = 1 
\end{cases} \\
s_1 &= \begin{cases} 
    0, & \text{if } r_1 = 1 \\
    1, & \text{if } r_2 = 0 
\end{cases} \\
r_2 &= r_1 \\
s_2 &= s_1
\end{align*}
\]

Here \([q_1]\) is a pseudorandom sequence generated from a 2-bit seed. The authentication data of 4 bits \((r_1, s_1)\) and \((r_2, s_2)\) is obtained.

Now we have to generate feature information. Feature information can be generated from bi-level moment preserving technique. In order to generate feature information of the block first we have to compute threshold value, and two representative gray values in such a way that the moments of the block B1 and the moments of thresholded block B1 are equal or nearly equal. Here lies the colour moment preserving. By using threshold value generate bitmap for block B1. In the bitmap one representative gray value represents 0 and another one represents 1. The feature information of the block B1 include bitmap of 64 bits and the two representative values with its two LSBs truncated. High recovery quality is achieved through preserving colour moments.

3) **Watermark Embedding**: In this method watermark of a block is embedded into its mapped block. The mapped block is obtained from the block mapping sequence generated by using key based pseudorandom permutation. The position in which the watermark bits to be embedded in mapped block is obtained by using
artificial bee colony optimization algorithm. Most of the watermarking schemes embed their watermark bits in LSBs of mapped block. These are vulnerable to lossy compression and also an attacker can easily extract watermark bits from LSBs.

B. Tamper Detection and Recovery

Sender sends a watermarked image and receiver receives it. Then it undergoes block division, block mapping, tamper detection and recovery. Figure 3 shows the functional diagram for tamper detection and recovery.

4) Block Division And Block Mapping: Each colour component of the watermarked image in YCbCr colour space is divided into blocks of size 8X8. Then generate a block mapping sequence using the same method as earlier. Here also we are using key based pseudorandom permutation for generating secure block mapping sequence. This key can be exchanged between sender and receiver using any secure key exchange method (eg Diffie - Hellman key exchange method, key exchange method using Elliptic Curve Cryptography).

5) Authentication Data Extraction: We know that the watermark information of block B_1 is embedded into block B_2 and now we got the watermarked image block B_k1. Perform artificial bee colony optimization technique and find out the watermark bit positions. Out of that positions first four bits are authentication data then extract the authentication data. Now we got two set dual parity check bits (r_k1, s_k1) and (r_k2, s_k2).

6) Authentication Data Generation: In order to generate authentication data we have to use the same two stage dual parity check method. Take the block B_1 in the watermarked image, that is B_k1. Set the two LSBs of each pixel in the block to zero. Then compute the average intensity of the block and then denote it as K^c. Then compute the two set dual parity check bits according to the pseudo random sequence \{q_i\} generated by using the key information.

7) Tamper Detection: During tamper detection first of all extract the authentication data from watermarked block and then generate the authentication data. Lets denote authentication data extracted by A_E and authentication data generated by A_G. Then check whether A_E equals A_G then mark the block valid and if A_E not equals A_G then mark the block invalid. Valid means no tampering and invalid means tampering occurred. At last generate the binary image. In that binary image zero indicates valid and one indicates invalid.

8) Tamper Recovery: If blocks are tampered then we have to recover it. During tamper proofing, the binary image obtained in the previous step undergoes some morphological operations. Then check whether the mapped block is valid or invalid. If the mapped block is invalid then recovery is not possible and if the mapped block is valid then recovery is possible. In order to recover the original image the feature information is extracted from the mapped block. The feature information includes bit sequence of the bitmap, 12 bit representative values. Pad zeros at the end of representative values. Replace the pixel value within block B_k1 with x_{(10)}(10) and y_{(10)}(10) according to c_{k1}^{i,j} = 1,2 ...16.
V. COMPARISON WITH EXISTING METHODS

Here we are comparing watermarking methods concentrating on improving recovery quality with proposed scheme.

Table I  
Comparison of different watermarking methods

<table>
<thead>
<tr>
<th>Title</th>
<th>Advantages</th>
<th>PSNR value (Approximately)</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed System</td>
<td>• Secure mapping.</td>
<td>42dB</td>
<td>• Manual selection of ROI is a time consuming process.</td>
</tr>
<tr>
<td></td>
<td>• Immune to Synchronous counterfeiting attack &amp; 4 scanning attack.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Immune to lossy compression.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Better tamper localization.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Better tamper recovery.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Automatic image authentication and recovery using fractal code embedding and image inpainting. [17]</td>
<td>• Effective.</td>
<td>38dB</td>
<td>• Manual selection of ROI is a time consuming process.</td>
</tr>
<tr>
<td></td>
<td>• Automatically localize tampered region.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Watermarking-based image authentication with recovery capability using Halftoning and IWT [18]</td>
<td>• Robust to JPEG compression attack.</td>
<td>38dB</td>
<td>• Complex for implementing.</td>
</tr>
<tr>
<td>Self-embedding watermarking scheme for colour images by bi-level moment-preserving technique. [3]</td>
<td>• Better PSNR value.</td>
<td>42dB</td>
<td>• Insecure mapping.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Vulnerable to watermarking attacks.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Vulnerable to lossy compression.</td>
</tr>
</tbody>
</table>

This method can be further improved by using any high level moment preserving technique. This method can be extended to video authentication and recovery also.

VI. CONCLUSION

The proposed watermarking method is applied to colour images for tamper detection and recovery. We have incorporated the bi-level moment-preserving technique, two-stage dual parity-check method, ABC optimization technique and morphological operations into the proposed watermarking scheme to perform the tamper proofing with high detection rates and to achieve the high quality of the restored colour image. The majority of edges and textures in the tampered colour image is successfully recovered and preserved. The proposed scheme also provides an effective and low complexity tamper proofing method. Since we use ABC optimization technique to find locations for embedding watermark bits so attacker cannot easily tamper watermark and is not vulnerable to lossy compression attack. This method uses a secure mapping so it is not vulnerable to security attacks. The proposed scheme outperforms the relevant existing schemes in deriving higher quality of the recovered colour image whereas the watermarked image is nearly lossless with the original image.
ACKNOWLEDGEMENT

First and foremost of all, I express my heartfelt gratitude to God Almighty for giving me an opportunity to excel in my efforts to complete this work on time. I cordially thank and acknowledge all my friends, parents and all my well wishers for supporting me directly or indirectly in my hard times.

REFERENCES