J. Vis. Commun. Image R. 38 (2016) 775–789

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Effective self-embedding watermarking scheme for image tampered detection and localization with recovery capability ☆

Durgesh Singh *, Sanjay K. Singh

Department of Computer Science and Engineering, Indian Institute of Technology (BHU) Varanasi, Varanasi, Uttar Pradesh 221005, India

ARTICLE INFO

Article history:
Received 16 March 2016
Revised 23 April 2016
Accepted 25 April 2016
Available online 27 April 2016

Keywords:
Fragile watermarking
Image authentication
Tamper localization
Image recovery
Discrete Cosine Transformation

ABSTRACT

This paper presents a Discrete Cosine Transformation (DCT) based effective self-recoverable fragile watermarking scheme. For each 2×2 non-overlapping block, two authentication bits, and ten recovery bits are generated from the five most significant bits (MSBs) of pixels. Authentication bits are embedded in the three least significant bits (LSBs) of the block itself while recovery bits are embedded in the three LSBs of the corresponding mapped block. The proposed watermarking scheme is also effective because the authentication of each block is based on two levels hierarchical tampered detection mechanisms. So the detection of tampered block can be ensured with high probability. The experimental results demonstrate that the proposed scheme not only outperforms high-quality restoration effectively, but also removes the blocking artifacts and improves the accuracy of tamper localization due to the use of very small size blocks, smoothing function and two levels tampering detection mechanisms.

1. Introduction

With the powerful image and signal processing technologies and the wide availability of multimedia editing tools, digital image content is prone to illegal manipulations and alterations. The content of a digital image can easily be replaced with fake content. In the critical situations, such as for evidence in a court of law, a small amount of content modification in digital imaging can change the judgment [1]. Therefore, in these days the tampered image detection and localization are important issues. This problem can be overcome by using digital watermarking techniques. Digital watermarking is a technique of imperceptibly altering an image to embed secret message or information about that image. The image, in which secret message is embedded is known as cover or host image and the secret message or information is known as watermark [2–4]. In addition to the content authentication bits which are used for authentication purposes, recovery (i.e. restoration or reference) bits are used to recover the corrupted content of the host image. The recovery bits are basically the compressed form of the host image called image digest and describe the content of the host image. The watermark scheme in which watermark generated from the cover image is known as self-embedding watermarking scheme.

Digital watermarking techniques can be broadly classified as robust watermarking, semi-fragile watermarking and fragile watermarking. In the robust watermarking technique, the watermark is capable of resisting intentional or unintentional manipulations. It can be employed for copyright protection and ownership verification [2,4–8]. The fragile watermarking is intended to be watermark destroyed any modification (unintentional or intentional) in the image [2–4,8–14]. In the semi-fragile watermarking, watermark is designed to resist non-intentional manipulations caused by the common image processing operations like JPEG compression, scaling, sharpening and is fragile against intentional manipulations like content modified by fake information [2,13,15–17]. The fragile and semi-fragile watermarking schemes are mainly employed for image and video content authentication.

Due to sensitivity towards the modification, fragile watermarking schemes are used (area like court of evidence) where the exact authentication required. Fragile watermarking schemes for authentication, can be further categorized into two main categories as pixel-wise fragile watermarking schemes [12,18–20] and block-wise fragile watermarking schemes [1,3,4,10,11,17,21–23]. In pixel-wise fragile watermarking schemes, the watermark information generated from the gray values of host pixels and is embedded into the host pixels themselves. In block-wise fragile watermarking schemes, the cover image is divided into sub-blocks. Each sub-block has watermark information and is authenticated by the successful retrieval of the watermark embedded in it. If the watermarked image is modified (malicious or non-malicious), the watermark of a particular sub-block is not retrieved

☆ This paper has been recommended for acceptance by Prof. M.T. Sun.

* Corresponding author.
E-mail addresses: durgesh.rs.cse13@iitbhu.ac.in, durgeshcse@gmail.com (D. Singh), sks.cse@iitbhu.ac.in (S.K. Singh).

http://dx.doi.org/10.1016/j.jvcir.2016.04.023
1047-3203/

successfully, then that sub-block is identified as tampered or invalid block. The least significant bit (LSB) method in data hiding has the benefit of simplicity, easily detection and high embedding capacity [24,25]. So, the least significant bit (LSB) method is generally preferred for watermark embedding in the fragile image watermarking.

The rest part of this paper is organized as follows. Section 2 gives a review of related work. Section 3 describes the watermark embedding and extraction procedures of the proposed scheme. The experimental results and analysis are given in Section 4. Finally, the concluding remarks are given in Section 5.

2. Related work

The performance of self-embedding schemes are generally given in terms of recovered image quality, and recovery (or restoration) conditions. The recovered image quality can be decided by a peak signal to noise ratio (PSNR) and the restoration conditions are typically the tampering rate (i.e. amount of modifications). These recovered image quality and tampering rate are strongly interrelated, and the trade-off between them needs to be properly balanced. Due to this trade-off, there is no general model for the content recovery problem despite the existence of the variety of schemes that are capable to recover the tampered content [4,21,22].

In many self-embedding schemes, the recovery bits of a particular sub-block are always hidden in another block (or blocks) of the image. This recovery may fail when the block (or blocks) containing its recovery bits is (are) tampered. This is called tampering coincidence problem. The schemes [3,10,11,26–28] are not able to handle this problem. Although, in the schemes [3,27], two copies of recovery bits were embedded into the image and the second chance for block recovery was provided. In [28], authors proposed a hierarchical watermarking scheme. In this scheme, four levels tampered detection procedures were used. It was based on 2 bits authentication watermark and performed in 2×2 sized sub-blocks of a block of size 4×4. This scheme is not robust against the collage and vector quantization(VQ) attacks [29] because the authentication bits of a block were independent of the other blocks. The recovery bits are just an average value of 6 MSB-layers of sub-blocks. So, the quality of recovered image is also not very good in this scheme.

The authors of schemes [21,30] proposed self -embedding watermarking scheme using a reference sharing mechanism to resolve the tampering coincidence problem. However, the quality of the recovered tampered image in both schemes is improve by embedding the redundant information in the cover image, but both schemes are not able to resist the VQ attack due to the independence between blocks. The localization accuracy is also reduced because of using a large block size of 8×8.

Zhang et al. [31] also proposed two self-embedding watermarking schemes by utilization of reference sharing mechanism. The watermark was derived from 5 MSB-layers of the cover image. In the second technique, the cover image was decomposed into three levels based on the hierarchical self-embedding scheme. The first scheme, the original data in five layers of original watermarked image can be recovered when tampering rate is no more than 24%. In the second scheme, the reference sharing methods with different restoration capabilities are employed to protect the data at different levels. So that a better recovered image can be obtained in second scheme, from a tampered version with less fake content. The second scheme is capable to recover up to 66% tampering rate.

To avoid tampering coincidence problem, another self-embedding watermarking scheme was proposed by Zhang et al. [32], in which watermark of a block was scattered in the whole image. In this scheme, the image pixels were permuted using a secret key and then divided them into a series of pixel-pairs. The recovery bits were generated by the XOR operation within the 5 MSB-layers of pixel pairs and the authentication bits were derived from 5 MSB -layers and recovery bits. Finally, watermark data were embedded into the three LSB-layers. In this scheme, the five MSBs of a pixel pair can be either fully or partially recovered by using reference data. The remaining uncertainty is resolved by manipulating local pixel correlations.This scheme is capable of restoring up to 54% tampering rate.

In [33], Qian et al. proposed a watermarking scheme based on Discrete Cosines Transform (DCT) to avoid tampering coincidence problem. The DCT coefficients of 8×8 blocks were encoded into different numbers of bits and the authentication-bits and restoration-bits were embedded into the three LSBs planes of the cover image. However the accuracy of tampered localization decreases because of using a large block size.

Zhang et al. [34] proposed a self-embedding fragile watermarking scheme based on DCT and fractal compression coding. In this scheme, the watermark was generated by using interleaved and overlapped block structure. Three versions of recovery bits of each block were embedded into different quadrants, which provide two chances for block recovery in case of one is destroyed. However, in this scheme the accuracy of tamper localization also decreases because of using a large block size.

To improve the quality of recovered image, an effective self -embedding watermarking scheme is proposed in this paper. Here, a block-wise mechanism is used for tampered region detection, localization and recovery. In this scheme authentication bits are embedded in the block itself instead of the mapped block, so the false tampered detection error can be minimized. This scheme is also effective because it is using two level tampered detection procedures. So the accuracy of tamper detection and localization is enhanced. During tampered detection procedures, if tampered areas are not detected at level-1 examination, it will be detected at level-2. So, the tampered regions are detected with the high probability. This scheme is using very small size of the block (i.e. 2×2) and is also secured by using predefined user keys. There are some strong points in the scheme: (1) Very less sensitive to error pixels; (2) higher localization accuracy; (3) The high quality of recovered image because DCT coefficients are used for recovery; (4) negligible blocking artifacts in recovered image. If the recovery bits are unable to extract from mapped block, then recovery in this scheme is based on its 3×3 neighborhood blocks, Hence, the tampering coincidence problem with this scheme is also avoided. Experimental results show that the proposed scheme allows high quality recovery up to 50% tampering rate.

3. Proposed algorithm

The proposed watermarking scheme can be explained in three steps: watermark generation and embedding, tamper detection and localization, and tamper image recovery procedures. Let the cover image I has M_1 rows and M_2 columns where both M_1 and M_2 are multiples of 2. Let M represent the total number of pixels ($M = M_1 \times M_2$).

3.1. Watermark generation and embedding procedure

Watermark is generated from the five MSB layers of the cover image itself. The three LSB layers of the cover image are substituted with the watermark bits. Two types of the watermark are generated for each block of size 2×2, where the first type of watermark is called authentication bits (2 bits) which are used to locate and detect the tampered blocks and the second type of watermark is