# DIGITAL WATERMARKING FOR IMAGE AUTHENTICATION WITH LOCALIZATION

*Huajian Liu, Martin Steinebach*

Fraunhofer IPSI - Integrated Publication and Information Systems,
Darmstadt, 64293, Germany
Email: {liu, steinebach}@ipsi.fraunhofer.de

## ABSTRACT

In this paper we propose a novel watermarking scheme for image authentication. A high localization of tampering detection is achieved by applying a random permutation process where every embedded watermark bit verifies random image positions instead of a local image block. Thereby the resolution of tampering detection is significantly improved in comparison to existing solutions while keeping the payload low. Furthermore, the proposed scheme doesn't embed the watermark locally but distributes it into the suitable embedding wavelet coefficients, avoiding embedding in smooth regions. Therefore, the scheme is intrinsically secure to block-based local attacks and retains high fidelity of the watermarked image. Scalable sensitivity of tampering detection is also enabled in the authentication process. Experimental results demonstrate the performance and effectiveness of the scheme for image authentication.

***Index Terms***— Image authentication, digital watermark, tampering localization

## 1. INTRODUCTION

In recent years digital watermarking has become a very active research field and been widely accepted as a promising technique for multimedia security. Image authentication is one of the application fields of digital watermarking, which allows us to recognize manipulations in images.

Unlike the classical message authentication, in image authentication, not only the integrity of the image content needs to be verified, but also tampering localization is very useful in practical applications, which identifies the positions where the tampering occurred. With the help of localization information, other parts of the image can still remain trustworthy and useful when the original data is not available. It can also help to infer the attacker's motives in applications such as forensic evidences.

In order to achieve the capability of localizing tampered regions, many existing watermarking schemes embed the watermark in a block-based way [1-4]. The image is divided into blocks and the watermark information is embedded into every block. The block content authentication is done by verifying whether the watermark can be successfully extracted from the block.

Therefore the maximum detection resolution is based on the block size. In [2,3], the block size of 8x8 is used and then the maximum detection resolution is only 8x8 block. In order to increase the detection resolution, the smaller block size is required but this will lead to high watermark payload. Subsequently, higher watermark payload will cause more artifacts. In [4] the detection accuracy is improved to 2x2 block, but the watermark payload is also increased to 1 bit per 2x2 block. So the challenge is how to increase the detection resolution with embedding the same or less watermark information.

Furthermore, in order to protect the whole image by the block-based schemes, the authentication data, i.e. the watermark, must be embedded into every block over the whole image. However, it is very difficult to embed the data in smooth regions without causing noticeable artifacts. It becomes even worse when embedding in smaller blocks. In [5], the random shuffling is used to handle the uneven distribution of embedding capacity in order to use the watermark capacity more efficiently.

In addition, another problem of block-based schemes is the security to local attacks. Because the block-based schemes embed the watermark locally, they show their weakness in local attacks, like copy and paste, vector quantization attack, which swap blocks in the same image or from different images [6,7].

In this paper we propose a novel watermarking scheme for image authentication to detect and localize the tampered regions. We apply a random permutation process to reduce the necessary watermark payload instead of trying to utilize the maximal watermark capacity as in [5], while keeping high tampering detection resolution. The embedded watermark is distributed in the wavelet coefficients suitable for embedding, which causes less perceptual artifacts. Also the random permutation enhances the security of the whole system against local attacks.

The paper is organized as follows. In section 2, we introduce the proposed scheme, including the watermark

1-4244-0481-9/06/$20.00 ©2006 IEEE

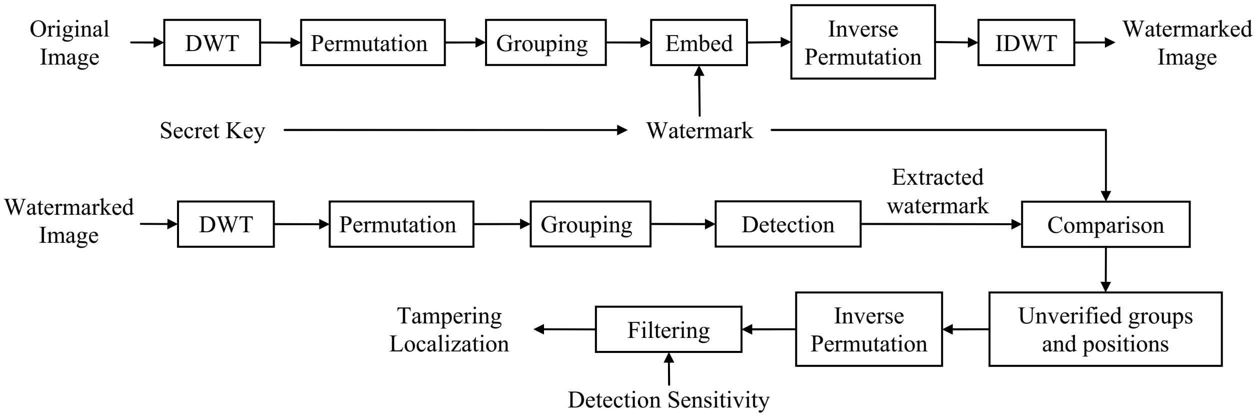

Figure 1: Block diagram of the proposed scheme

embedding, detection and image authentication. In section 3, the experimental results are presented. We conclude the paper in section 4.

## 2. PROPOSED SCHEME

The block diagram of the proposed scheme is shown in Figure 1, consisting of the watermark embedding process, detection process and image authentication.

### 2.1. Random permutation

The proposed scheme performs the watermark embedding in the Discrete Wavelet Transform (DWT) domain. The image is decomposed by *R*-level wavelet transform. The watermark is embedded into the three subbands of the selected level *r*.

We firstly concatenate all the wavelet coefficients of the three subbands into a single string. Three coefficients in the three subbands, which correspond to the same spatial location, are continuously adjacent in the new sequence. Let $f_{HL}$ , $f_{LH}$ , $f_{HH}$ denote respectively the coefficients of the different subbands. The coefficients are rearranged in the following way:

$\{ f_{HL}(0,0)$ , $f_{LH}(0,0)$ , $f_{HH}(0,0)$ , $f_{HL}(0,1)$ , $f_{LH}(0,1)$ , $f_{HH}(0,1)$ , $\cdots\cdots$ , $f_{HL}(m-1,n-1)$ , $f_{LH}(m-1,n-1)$ , $f_{HH}(m-1,n-1) \}$.

Then the concatenated coefficients are randomly permutated, controlled by a secret key. A minimal distance between the original adjacent coefficients is required in order to ensure the coefficients are enough randomly distributed.

After the random permutation, the string is divided into groups with a fixed group size *G*. In every group, one watermark bit is embedded. The embedded bit will monitor all the members of this group. The random permutation process distributes the coefficients suitable for watermark embedding evenly over all the groups. This property ensures the embedding process to make no modification in the smooth image regions and therefore improves the fidelity of the watermarked image.

The group size *G* will affect the watermarked image's quality and the maximal localizable tampered area. With a larger *G*, fewer watermark bits will be embedded and a higher fidelity of the watermarked image will be achieved, but this will not decrease the detection resolution.

The selected wavelet level *r* decides the maximum detection resolution and will affect the watermark robustness. Embedding in higher wavelet level will render higher robustness against the common image processing, e.g. JPEG compression, while it decreases the accuracy of tampering localization.

### 2.2. Watermark embedding

The watermark consists of a binary random sequence, generated by the secret key. The random sequence serves as an authentication code. This code is compared with the retrieved watermark in the authentication process, similar to a Look Up Table (LUT) structured algorithm.

In every group of the random string, all the wavelet coefficients are summed up. The summation is quantized by a quantization step *Q* as shown in the following formula.

$$s_j = \sum_{i=0}^{G-1} f_j(i) = \lfloor s_j / Q \rfloor \cdot Q + \Delta_j \qquad (1)$$

where $f_j(i)$ is the *i*th coefficients in the *j*th group and $s_j$ is the summation of the coefficients in that group. $\Delta_j$ is the quantization residue.

To embed the watermark bit, the summation $s_j$ is quantized based on the watermark bit. The quantization process is shown in Figure 2. The $s_j$ is modified to the nearest 0 bin or 1 bin according to the watermark bit.

In order to modify the summation $s_j$, we propose two methods to update the coefficients. The simple way is to

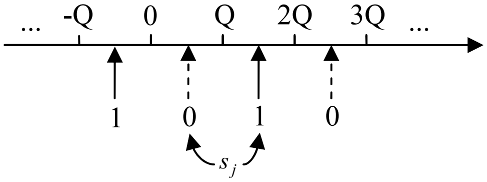

Figure 2: Quantization process

modify the coefficient with the maximal magnitude in the group, because the modification of such coefficients with large magnitude causes less noticeable artifacts than other small coefficients. As mentioned in section 2.1, the random permutation process ensures the large coefficients are distributed evenly over all the groups. The coefficient with maximal magnitude is updated as follows.

$$\delta_j = s_j^* - s_j \qquad (2)$$

$$f_{j,\max}^*(i) = f_{j,\max}(i) + \delta_j \qquad (3)$$

where $s_j^*$ is the expected summation value of the *j*th group and $\delta_j$ is its difference from the original summation value. $f_{j,\max}(i)$ is the coefficient with the maximal magnitude in the *j*th group.

The second method is to update every coefficient in the group proportionally. The amount of modification of every coefficient is decided by the proportion of the magnitude of every coefficient as follows.

$$f_j^*(i) = f_j(i) + \frac{\left|f_j(i)\right|}{\sum_{i\in G}\left|f_j(i)\right|}\delta_j \qquad (4)$$

In either formula (3) or (4), a perceptual model can be used to decide the most suitable coefficient to modify instead of the coefficient magnitude. The perceptual model in [8], which takes into account the local brightness, frequency and texture, can be applied to calculate the suitable embedding strength of all the wavelet coefficients of the three subbands at level *r*. Larger embedding strength indicates the corresponding coefficient is more suitable to be modified, because its modification will cause less noticeable artifacts.

After all the watermark bits are embedded, the wavelet coefficients are put back to their original positions in the subbands from the string. The inverse wavelet transform is then performed to obtain the watermarked image.

## 2.3. Watermark detection

The watermarked image is firstly decomposed by wavelet transform. Then the coefficients in the three subbands of the level *r* are concatenated in the same way as the embedding process and then randomly permutated by the correct secret key.

The permuted coefficients are then divided into groups with the same group size *G* as the embedding process. The watermark bit is extracted by quantizing the summation of all the coefficients in every group.

$$r_j = \lfloor s_j^* / Q \rfloor \qquad (5)$$

where $r_j$ is the quantization result in the *j*th group.

## 2.4. Image authentication

Every extracted watermark bit is compared with the embedded one generated by the secret key. For every group, if the extracted bit does not match the embedded one, the whole group is considered unverified and every group member is marked as an unverified coefficient.

All the coefficients are then mapped back to their original positions in the wavelet subbands by the inverse permutation. The unverified coefficients will randomly scatter over the subbands. If there is a tampered region in the watermarked image, in every subband there will be a region with much higher density of unverified coefficients at the location corresponding to the tampered region, because all unverified groups contain one or more coefficients from the tampered region. All the other isolated unverified coefficients come from the same groups which the tampered region belongs to. Due to the random permutation, they are distributed over the subbands sparsely and are considered as noises.

Then we construct a matrix of the same size as the subband at the *r*th level wavelet transform, i.e. $1/4^r$ of the image size, in which every position corresponds to a $2^r \times 2^r$ pixels block of the image. We consider a position in the matrix as unverified when there is an unverified coefficient at the corresponding position in any subband. In this way, the isolated unverified coefficients in the subbands still randomly scatter over this matrix as noises, while in the tampered region the density of the unverified coefficients becomes higher than in any subband.

A noise filter, e.g. a median filter, is used to filter out the noise coefficients. Then the tampered region will be easily picked out. A properly designed noise filter can not only remove the noises, but also can compensate for an insufficient random permutation or watermark detection errors. When the coefficients in one group are changed more than $1.5Q$, a watermark extraction of missing to detect the tampering may occur. After filtering out the noises, the remaining unverified positions indicate the tampered region in the image. Since the matrix size is $1/4^r$ of the image size, it provides a maximum detection resolution of $2^r \times 2^r$ blocks in the image.

Furthermore, the sensitivity of tampering detection can be adjusted by choosing different filtering sizes. Based on different application requirements, by presetting the filter dimension, the scheme can identify tampering of various sizes, bypassing the smaller alterations but detect the bigger ones.

Increasing the group size G will not decrease the tampering detection resolution, but it decides the maximum tampered area that can be localized. A larger G will cause more unverified coefficients outside the tampered region. Therefore, when a very large area is tampered, too many unverified coefficients will make it difficult to filter out the correct tampered region.

## 3. EXPERIMENTAL RESULTS

An example image of 720x576 is shown in Figure 4. In the following experiments, the first coefficient update method is applied in the watermark embedding process.

The PSNR of the watermarked image with different *G* and *Q* is shown in Figure 3 (with *r*=1). A larger *G* decreases the embedded watermark payload and renders better image quality. On the contrary, a larger quantization step *Q* will degrade the watermarked image quality. With *r*=1 the watermark can only survive JPEG compression with quality factor 100. With *r*=2 or higher *r*, the watermark can resist JPEG compression with quality factor 70 or lower by applying various *Q* and *G*.

In the following experiments, we let *r*=1, *Q*=6 and *G*=12. The PSNR of the watermarked image is 49.06dB and the embedded watermark is completely imperceptible. Figure 4 (a) shows the original image and (b) is a tampered version, in which the man in the image is removed and one window of the house is deleted. The Figure 4 (c) is the image authentication result using the proposed scheme. The localized tampered regions are depicted in white color. A median filter with size 5×5 is used as the noise filter for localization.

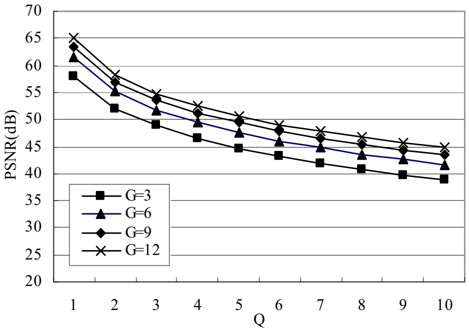

Figure 3: Watermarked image quality (PSNR)

## 4. CONCLUSION

In this paper we propose a watermarking scheme to detect and localize the tampered regions in the image. By applying a random permutation process, the proposed scheme significantly improves the resolution of tampering detection with low watermark payload. The scheme is intrinsically secure to local attacks, because the watermark is randomly distributed into the suitable wavelet coefficients instead of being embedded locally.

## 5. REFERENCES

[1] Ö. Ekici, B. Sankur and M. Akçay, "Comparative evaluation of semifragile watermarking algorithms", Journal of Electronic Imaging, 13(1), 209-216, January 2004.

[2] C.Y. Lin and S.F. Chang, "Semi-fragile watermarking for authentication JPEG visual content", Proc. of SPIE, San Jose, CA, vol. 3971, pp. 140-151, 2000.

[3] M. Wu and B. Liu, "Watermarking for image authentication", Proc. of IEEE International Conference on Image Processing (ICIP'98), Chicago, IL, vol.2, pp.437-441, Oct. 1998.

[4] D.A. Winne, H.D. Knowles, D.R. Bull, and C.N. Canagarajah, "Digital watermarking in wavelet domain with predistortion for authenticity verification and localization", Proc. of SPIE, San Jose, CA, vol. 4675, pp. 349-356, 2002.

[5] M. Wu and B. Liu, "Digital watermarking using shuffling", Proc. of IEEE International Conference on Image Processing (ICIP'99), Kobe, Japan, vol.1, pp.291-295, Oct. 1999.

[6] J. Fridrich, "Security of fragile authentication watermarks with localization", Proc. of SPIE, San Jose, CA, vol. 4675, pp. 691-700, January, 2002.

[7] A.H. Ouda, M.R. El-Sakka, "Localization and security enhancement of block-based image authentication", Proc. of IEEE International Conference on Image Processing (ICIP 2005) vol.1, pp.673-676, Sept. 2005.

[8] A.S. Lewis, G. Knowles, "Image compression using the 2-D wavelet transform", IEEE Trans. Image Processing, Vol. 1, pp. 244-250, Apr. 1992.

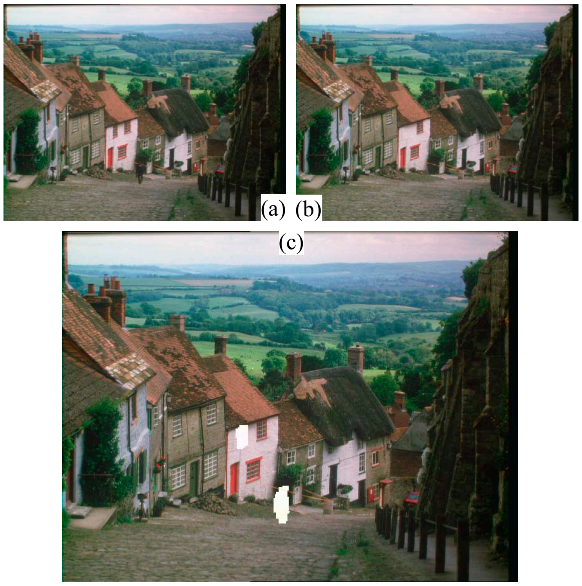

Figure 4: (a) Original image (b) Tampered image (c) Image authentication result