Localization and Security Enhancement of Block-based Image Authentication

Abdelkader H. Ouda and Mahmoud R. El-Sakka, *Senior member, IEEE*
Computer Science Department
The University of Western Ontario
London, Ontario, Canada
{kader, elsakka}@csd.uwo.ca

***Abstract*— Most block-based image authentication techniques that are presented in the literature sacrifice localization accuracy in order to resist vector quantization (VQ) counterfeiting attacks. In this paper, we show that strong cryptography schemes, which produce a long signature, can be used to sign image blocks without regard to the size of these blocks. In addition, a new approach to generate overlapped watermark segments for image blocks is presented. These watermarks are generated using one-way function based on an NP-complete problem. Moreover, a block-based image authentication technique is proposed. This technique provides strong protection against the VQ attack, as well as a great enhancement in localization accuracy and system security.**

Keywords-component; image authentication; image signature; digital watermarking; cryptography; steganography

I. Introduction

The current digital revolution has caused a rapid development of multimedia technologies. While these technologies have provided unprecedented imaging services, opportunities for theft and misuse of intellectual property are also widespread. This issue highlights the demands for efficient image authentication techniques. The main function of image authentication techniques is to digitally sign the image data, or a condensed version of it. This signature forms all or part of the image watermark, which, at the end, is embedded into the image. Public-key cryptosystems [1] provide the capability to generate and verify signatures. Signature generation makes use of a private key to generate a digital signature. Signature verification makes use of a public key, which corresponds to the private key (they are not the same). Anyone can verify the signature by using the public key. Signature generation can be performed only by the possessor of the private key. The embedded watermark needs not only to authenticate the identity of the signatory and to detect unauthorized modifications to the image, but also to localize the location of these modifications.

Wong [2] presented a well known block-based image authentication technique. In this scheme, the image is divided into small blocks. Each block is cryptographically hashed and combined with a corresponding block in a selected binary image. This combination is signed and embedded into *the least significant bits* LSBs of the same image block.

Although the technique is able to verify the image integrity and to provide a proof of its identity, *the vector quantization VQ attack* by Holliman and Memon [3] was able to counterfeiting the embedded watermark. This is due to the independency of the image block signatures. This attack is based on the idea that, the attacker may build a large database of watermarked images that contain the same watermark which is formed using the same key. The watermarked image blocks in all images are sorted so that all blocks that correspond to the same binary image block will be classified in same class. The attacker might then use these blocks to counterfeit the target image. He/She can substitute specific blocks with other blocks from the database that belong to the same class. This substitution ensures that the extracted watermark has the same visual appearance as the extracted watermark from the original watermarked image.

This flaw leads to many other block-based image authentication techniques [4,5,6,7,8], which were proposed to defeat the VQ attack. These techniques either modify the original Wong scheme or produce new ideas that are also based on the original scheme. While these techniques have their advantages in defeating the VQ attack, they have sacrificed tamper localization accuracy of the original methods.

In this paper, we provide a new technique based on the sliding window idea to generate overlapped image block watermarks. This technique improves both the VQ attack resistance and the tamper localization accuracy.

The paper is organized as follows. Section II briefly discusses the drawbacks that might occur in the current solutions. Section III explains the proposed embedding and verification processes. The experimental results that support the proposed solution are shown in Section IV. Finally, the conclusion is offered in Section V.

II. Block-based Image Authentication Techniques

Wong and Memon [4] proposed three variations to the original Wong technique to solve the problem caused by VQ attack. The authors firstly recommended using a larger image block of 12x12 pixels to decrease codebook sizes. This is because smaller block sizes would increased the number of authentic blocks that can be obtained from an image, which in turn increase the chance to find a forgery image blocks. In their second proposal, they recommended adding the block index to

0-7803-9134-9/05/$20.00 ©2005 IEEE

the input of the hash function before hashing the image block. In this case, to forge one block of an image, the attacker is constrained to select only blocks from his/her database that has the same index. Yet, this may not be sufficient, considering that the attacker database is huge. Finally, the same authors proposed another idea to thwart this attack, where the hash function is applied to the entire image excluding the LSB, to produce an image-based index. This index is then added to the input of the hash function that hash the image block. While this approach might defeat the VQ attack completely, it also harms the localization accuracy. This is because any small modification in the most significant 7 bits of an image will always produce different image index that will lead to failed verification of each image block. Another factor that has a negative impact to the localization accuracy is illustrated by Ouda *et al.* [5]. They showed that in order to use a strong signature algorithm such as RSA [9] in Wong technique, the image blocks could not be less than 32x32 pixels. This is simply because the block signature (of length 1024 bits) needs to be completely embedded into the LSBs of the same block.

Fridrich [6] proposed an elegant and simple technique to resist the VQ attack. The idea is based on separating the authentication of the content and its origin. Instead of using a fixed binary image, as in Wong technique, a special symmetry structure in the logo is used to authenticate the block content. An 8x16 binary blocks are created and concatenated to form the binary image. These blocks are formed so that it is simply recognizable, and hold some information of the corresponding image block. This information could be block index, image index, time stamp, or author ID, etc. This approach makes it possible to identify swapped blocks between images (VQ attack), while being able to accurately localize tampered pixels. However, a block of size 8x16 pixels is not enough to hold the block signature. Assuming that a strong signature algorithm is used, the image block size should be at least 18x18, which decrease the localization accuracy.

An alternative method to defeat the VQ counterfeiting attack is proposed by Coppersmith *et al.* [7]. The authors proposed to divide the image into non-overlapped small blocks of size 24x24 pixels so that each block is a center of a 32x32 block that overlaps with its neighbors. RSA signature algorithm is used to sign each large block, however the signature will be embedded in the corresponding small block. This approach might decrease the risk of VQ attack, but at the same time it will flaw the localization accuracy. For instance, if a modification occurred in a pixel in a large block and outside the small one, the verification process will result in some ambiguity in tamper localization. Both the block belongs to the large block and the neighbor small block that contains the modified pixel will be declared unauthentic. I.e., the detection localization accuracy might become 48x48, which is a major drawback.

Celik *et al.* [8] have recently proposed another solution to the VQ attack. The image blocks are formed and signed in a hierarchical level. The produced signatures are embedded into the lowest level blocks of the hierarchy. The size of the lowest level blocks determines the accuracy of the alteration detection. At each successive level, image blocks are composed of 2x2 blocks at the preceding level of the hierarchy. The top level of the hierarchy forms the image itself. Since all the produced blocks are signed (including the whole image), the verification process will recognize any attempted of VQ attack. Note that, the LSBs of the lowest level blocks hold a portion of upper level signatures, together with its own signature. For example, consider dividing an image into blocks in three hierarchy level and the used signature algorithm produces ciphertexts of length *S* bits. In this case, each lowest level block will hold three bit-sequences of lengths *S*/16, *S*/4, and *S*. These correspond to one sixteenth, one fourth, and the entire top level block signature, respectively. For instance, if the RSA signature algorithm is used, the total number of embedded bits in each block is 21/16 × 1024 = 1344 bits, i.e., the smallest image block could not be less that 37x37. However, if the digital signature algorithm (DSA) [10] is used the total number of bits will be reduced to 21/16 × 320 = 420 bits. Note that, DSA produces signature of length 320 bits. Therefore, the smallest image block could be as small as 21x21. Yet, this performance distorts the localization accuracy of the Wong's technique. In addition, if one block of the lowest level is modified so that the alteration affects only the signature portion of the top level, this level will be declared unauthentic due to VQ attack, which is not the case.

We conclude from this discussion that most of the block-based image authentication techniques are forced to sacrifice tamper localization accuracy in order to build a strong defense against the VQ attack.

III. The Proposed Technique

In this paper we propose a new block-based image authentication technique. The block watermark generation and embedding processes are based on a secure one-way function, whereas the signing process utilizes the RSA public-key encryption scheme with 1024-bit modulus. The detail description of the proposed technique will be given in the following sections, but first we need to describe the used one-way function.

A. Knapsack One-way Function

A function $f(x)$ being one-way means that the transition from x to $f(x)$ is easy, whereas the inverse from $f(x)$ to x is intractable. The proposed technique utilizes a one-way function based on the *Knapsack problem* [11], where an *n*-tuple

$$A = (a_1, a_2, \ldots, a_n) \quad (1)$$

of distinct positive integers (items) , as well as another positive integer *K*, are given. The problem is to find a subset of items a_i that sums up to *K* exactly. E.g., consider A = (3,19,8,12,2,14) and K = 23. We observe that 23=3+8+12. Using vector dot product we may write $A.B$=23, where *B* is a vector (1,0,1,1,0,0). Hence, we define a function

$$f(B)=A.B \quad (2)$$

Note that, a solution can be found by checking through all subsets of A and check whether one of them sums up to *K*. In this example, i.e., 6 items, this means 2^6=64 subsets. Of course, this is manageable, but how about if the Knapsack contains 512 items? A search through 2^{512} different subsets is unmanageable. The Knapsack problem is known to be *NP*-complete [12]. Any *NP*-complete problem is considered intractable.

The proposed technique utilizes a knapsack-based function that contains 512 distinct positive integers. These integers are ranged from 1 to 2048.

B. Watermark Generation and Image Signature

The watermark generation and image signature processes are described in the following steps.

1) Reset the LSBs of the given image to zero.
2) Using a 36x36 sliding window, scan the entire image in a raster order from left to right and from top to bottom. The sliding step size is 12 pixels; see Fig. 1. Note that, at image boundaries, the window will move in a circular way so that, each pixel of the image will be scanned 9 times.

   Note also that, the size of the sliding window is chosen to be relatively large in order to resist the VQ attach. In addition, scanning the image in an overlapping way builds a strong connected chain of block watermarks to be broken by VQ attack.
3) At each window position, one watermark segment S_i will be produced as follows:
   a) Using the *secure hash function* SHA-512 [13], hash the corresponding image block to produce a block digest B_i of length 512 bits.
   b) Compute S_i using (3)

   $$S_i = f(B_i) = A.B_i \quad (3)$$

   where $f()$ is a one-way function; see Section III.A. Note that, the length of S_i will be at most 20 bits.
4) After scanning the entire image, concatenate the produced watermark segments into units U_j, such that each unit does not exceed 1023 bits (less than 51 segments per unit). E.g., in a 512x512 image, 1820 segments will be produced, and hence we will get 36 watermark units.
5) Sign each watermark unit U_j using the RSA-1024 encryption to produce the corresponding ciphertext C_j.

   $$C_j = E_{private}(U_j) \quad (4)$$

   where $E_{private}()$ is the RSA-1024 encryption using the image owner private key. Note that the C_j length will be at most 1024 bits.
6) Concatenate all C_j to produce the image signature Z.
7) Embed the image signature Z repeatedly into the LSB of the image. Note that, for an image of size 512x512 the length of Z will be at most 36864 bits. Therefore Z can be completely repeated in this image up to 7 times.

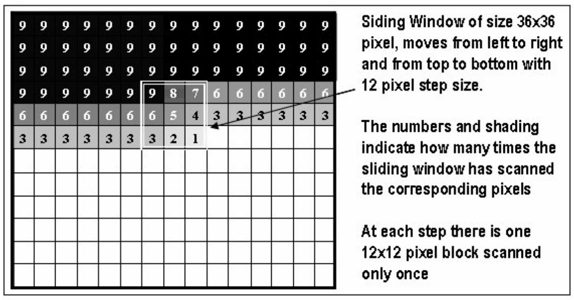

Siding Window of size 36x36 pixel, moves from left to right and from top to bottom with 12 pixel step size.
The numbers and shading indicate how many times the sliding window has scanned the corresponding pixels
At each step there is one 12x12 pixel block scanned only once

Fig. 1: The raster scan of the sliding window.

C. Watermark Extraction and Image Verification

The watermark extraction and image verification processes are described in the following steps.

1) Get a correct version of the image signature *Z,* by extracting all repetitions of the image signatures from the image LSBs, and applying a bit-wise majority function.
2) Partition Z to get the embedded cipertexts C_j.
3) Decrypt each C_j, by utilizing the same setting of the RSA that is used in the signing process:

   $$U_j = D_{public}(C_j) \quad (5)$$

   where $D_{public}()$ is the RSA-1024 decryption using the image owner public key.
4) Partition each U_j to get all watermark segments S_i.
5) Divide the image into 12x12 pixel block G_p, and mark each block by zero.
6) Use the 36x36 sliding window to scan the entire image exactly as in step 3 of the watermark generation process.
7) At each window position, generate the watermark segment S_j^*. Compare S_j^* with its corresponding S_i, if they are different, increase the marks of all current scanned G_p's by 1.
8) After scanning the entire image, for each block G_p that is marked by 9, produce a black block, and for all other G_p's blocks, produce white blocks.
9) Concatenate the produced white/black blocks to form a binary image of the same size as the original image. The black blocks (if any) indicate the unauthentic areas.

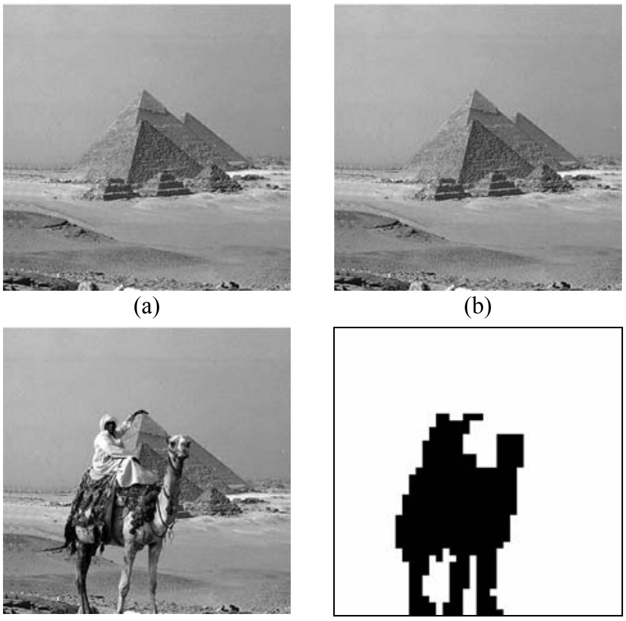
D. System Significant

The major features of the proposed technique are summarized in the following points:

- ***VQ attack Resistance***. Using a large overlapped sliding window kills any hope of the VQ attack. Note that, the window size could be even larger, with no side effect of the system performance. This is because the used hash function has no restriction on its input length, i.e., any window size will produce 512 bits hash.
- ***Good localization accuracy***. The proposed technique is able to detect the modifications in accuracy as small as 12x12 pixel blocks.
- ***Not forgeable***. The used one-way function blocked the way for the attacker to forge any part of the protected image. This is because it is computationally invisible to find an arbitrary block producing a specific watermark segment Si.
- ***Forward-compatible.*** The image block size does not depend on the length of the encryption key, therefore longer keys can be used, e.g., 2048 bits, in order to meet the feature security requirements.

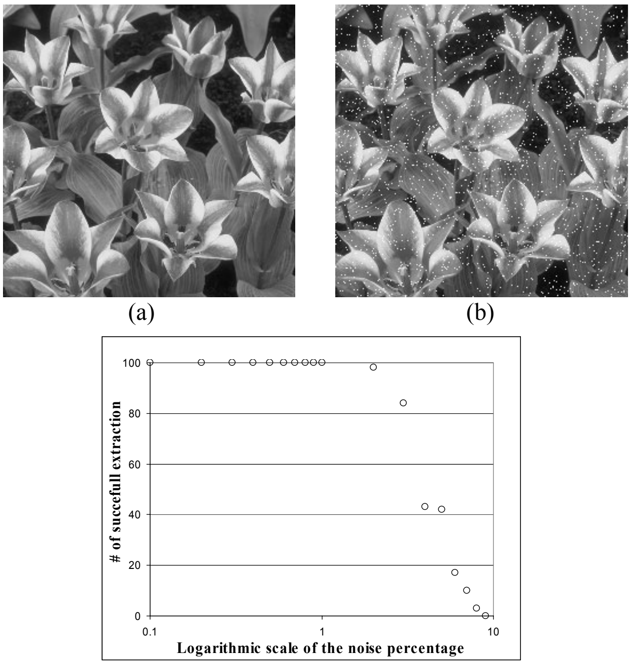
IV. Experimental results

To validate the proposed solution, two tests are made. The first test is conducted to show the localization and tamper detection ability of our technique. Fig. 2(a) shows the original image. Fig. 2(b) shows the watermarked image, after applying the proposed signing process. The watermarked image was modified by adding a man ridding a camel in the middle of the image; see Fig. 2(c). After applying the verification process, the extracted binary image is shown in Fig. 2(d), where the modified area (the man and the camel) can be detected with localization accuracy as small as 12x12 pixel block.


(a) (b)

(c) (d)

Fig. 2: (a) the original 512x512 image, (b) the watermarked image, (e) the modified watermarked image, and (d) The corresponding extracted binary image, with localization accuracy as small as 12x12 pixel block.



(a) (b)

(c)

Fig. 3: (a) A watermarked 512x512 gray scale image, (b) the watermarked image with 2% salt noises, (c) the number of successful extraction of the image signature

The second test measures the robustness of the image signature *Z* against image distortion. To simulate the distortion, different amounts of noise were spread across the watermarked image, by changing the value of randomly selected pixels to 255. The amount of noise is ranged from 0.1% to 10% of the image size. In this experiment, a 512x512 image is used; see Fig. 3(a). The experiment was repeated 100 times. The number of correct extraction of *Z* was counted; see Fig. 3(c). The result showed that *Z* can be successfully extracted even if almost 2% of the image is modified; see Fig. 3(b).

V. Conclusion

In this paper a new block-based image authentication technique is proposed. The watermark generation process utilizes a large overlapped sliding window, where one watermark segment is generated at each window position. The concatenation of these segments is signed using RSA-1024 encryption. This approach provides a strong line of defense against VQ attack. The image block watermark is generated based on a secure one-way function. The complexity of this function is based on Knapsack problem, which is known to be NP-complete. This function has a great impact in enhancing the localization accuracy, (i.e., 12x12 pixels). In addition, embedding the image block signature several times into the entire image increases the method's robustness characteristic.

References

[1] A. Menezes, P. Oorchot, and S. Vanstone, Handbook of Applied Cryptography, CRC Press, Florida, USA, 1997.

[2] P. Wong, "A Public Key Watermark for Image Verification and Authentication," IEEE International Conference on Image Processing (ICIP'98), vol. I, pp. 455–459, 1998.

[3] M. Holliman and N. Memon, "Counterfeiting attacks on oblivious block-wise independent invisible watermarking schemes," IEEE Transactions on Image Processing, vol. 9, no. 3, pp. 432–441, 2000.

[4] P. Wong and N. Memon, "Secret and Public Key Image Watermarking Schemes for Image Authentication and Ownership Verification," IEEE Transactions on Image Processing, vol. 10, no. 10, pp. 1593–1601, 2001.

[5] A. Ouda and M. El-Sakka, "A Practical Version Of Wong's Watermarking Technique", IEEE International Conference on Image Processing (ICIP'04), vol. 4, pp. 2615–2618, 2004.

[6] J. Fridrich, "Security of Fragile Authentication Watermarks with Localization," SPIE Security and Watermarking of Multimedia Contents, vol. 4675, pp. 691–700, 2002.

[7] D. Coppersmith, F. Mintzer, C. Tresser, C. Wu, and C. Yeung, "Fragile Imperceptible Digital Watermark with Privacy Control", SPIE, Security and Watermarking of Multimedia Contents, 79–84, 1999.

[8] M. Celik, G. Sharma, E. Saber and A. Tekalp, "Hierarchical watermarking for secure image authentication with localization," IEEE Transactions on Image Processing, vol. 11, no. 6 , pp. 585 –595, 2002.

[9] R. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," Communications of the ACM, vol. 21, pp. 120–126, 1978.

[10] National Institute of Standards and Technology, "Digital Signature Standard (DSS)", FIPS Publication no. 186, 1994.

[11] W. Brauer, G. Rozenberg and A. Salomaa, Public-key Cryptography, Monographs on Theoretical Computer Science, Springer-Verlag, 1990.

[12] U. Manber, Introduction to Algorithms, Addison-Wesley Publishing Inc., 1989.

[13] National Institute of Standards and Technology NIST, "Secure hash standard", Federal Information Processing Standards Publication FIPS, 180-1.