A Blind Authentication Purpose Discrete Wavelet Watermarking

Ahmet Şenol[1], Ersin Elbaşı[2], Kıvanç Dinçer[1], Hayri Sever[1]
[1] Computer Engineering Department, Hacettepe University, 06800 Çankaya Ankara
[2] Department of Computer Engineering, Çankaya University, 06790 Etimesgut Ankara

***Abstract*—Image watermarking is used for proving ownership of images, tracking advertisement broadcasts, testing the image's authenticity etc. For watermarking types that aim proving ownership, watermark must resist image operations and watermark must remain in the image after operations. Watermark must be fragile or semi-fragile for authentication watermarking types, where testing the image's being same as original is the main purpose. There are previous authentication-purpose studies in literature but not as much as proving-ownership types. In this study, a blind discrete wavelet transform based authentication purpose watermarking method is proposed. Although the study resembles the methods in literature, it is genuine in its method of embedding in DWT transform and detecting changes in images. The method is capable of detecting changes in images on 4-pixel block base.**

***Index Terms*— Blind, Authentication, Discrete Wavelet Transform, Image Watermarking, Semi-fragile.**

I. Introduction

The Internet made it possible to spread one's copyrighted property without owner's permission. People have to put or share their digital content for various reasons on the Internet by web sites, blog sites, e-mail, social networking etc. By the time, traditional copyright protecting methods such as sticking copyright labels on digital content packs, encrypting data while transferring became insufficient to satisfy needs. Encrypting data has its own problems such as sending keys to target person or content's being defenseless after decryption. Digital watermarking evolved as a new technology to solve digital content ownership proving. The digital content to be protected is called the host or cover data and a special type of data called watermark is embedded in the data itself. The watermark itself is relatively small amount compared to host cover data and becomes part of host data throughout the digital content's life. The watermark can be a visual company logo or a biometric face or sound of owner, a pseudo random number sequence etc. It is preferable for watermark data to have a normal distribution because natural image scenes have normal distribution and embedded watermark must not seem different from original by human eye. The type of watermark for proving ownership is called robust type of watermark and there are many studies on this subject [1]–[7].

A. Ş. Author (phone: +90-312-4175190/2641; e-mail: asenol@ kho.edu.tr).
E .E. Author (e-mail: eelbasi@cankaya.edu.tr)
K.D. Author (e-mail: kivanc.dincer@hacettepe.edu.tr)
H.S. Author (e-mail: sever@hacettepe.edu.tr) .

For any type of watermarking, fidelity is important. Fidelity is the extent the watermarked image looks like the original one i.e. the similarity between original image and watermarked image. Fidelity is measured by peak signal to noise ratio PSNR given in (1)

$$PSNR = 20 \log_{10}(255/RMSE) \quad (1)$$

where RMSE is square root of mean squared error between original and distorted images as in (2)

$$RMSE = sqrt((\sum_{i,j} (I^*_{ij} - I_{ij})^2) / (NxN)) \quad (2)$$

For some institutions or businesses such as military, aviation, satellite transmission, medical data, it is very important to ensure that downloaded content is same as original. For this purpose, authentication type of watermarking emerged. Authentication type of watermark is fragile or semi-fragile so that watermark disappears or become un-extractable when the watermarked content is modified to some extent. It is preferable for authentication type of watermark that semi-fragile watermark resist for not-ill-purpose image operations such as lossy compression, intensity adjustment, blur filter etc. Semi fragile watermark must deteriorate when a malign operation is applied to digital content such as changing face of a person in the image, putting a non-existent object in the image etc. In authentication type of watermarking, the algorithm should be able to decide whether the digital content in question is genuine without the original content at hand, i.e. algorithm should be blind type of watermarking.

Wolfgang and Delp developed a fragile and semi-fragile watermark for authentication purposes. In fragile watermark, they used the image hash and timestamp in watermarking phase and even a bit change in the watermarked content results in failing to authenticate the image. In their second variable-watermark two-dimensional algorithm (VW2D), the algorithm is able to categorize image as "unaltered", "slightly affected", "definitely altered but still originating from the watermarked image", "completely changed and not originating from watermarked image"[8].

Wong devised an authentication algorithm that is capable of detecting changes in pixel base [9]. He divided the image into blocks and used some blocks for calculating an MD5 hash using image size values M,N, the other block content that will hold the hash data, and private key of watermarking person. On the receiving side, image size, public key of sender, and MD5 holding block's content is used to authenticate the block. The MD5 holding block holds the MD5 hash of the other block in its least significant bits.

Chamlawi, Khan and Idris propose a method that embeds two watermarks for authentication and recovery purposes[10]. They use integer Wavelet transform (IWT) instead of discrete wavelet transform to reduce computational complexity.

II. Proposed Method

Watermark Embedding Algorithm:

1. Divide Original Image into 8x8 non overlapping parts

2. Take DWT of each block independently

3. Make pairs for 8x8 blocks

For the second block, for each (i,j) DWT value,
if the sum (HL_2(i,j)+ LH_2(i,j)+ HH_2(i,j)) <= Threshold
LL_1(i,j) = `floor(LL_1(i, j )/10)*10 + 2;`
else
LL_1(i,j) = `floor(LL_1(i, j )/10)*10 + 7;`

4. Take the inverse DWT transform and obtain the watermarked image.

The main idea is divide the image to non-overlapping blocks so that it is possible to localize image file changes. Block pairs are formed so that in one block some values are calculated, in the other block that calculated value is embedded as the watermark. The embedding procedure can be seen more clearly in fig.1. By dividing the LL value by 10 and taking the floor of this value, and then multiplying by 10, last digit of LL value is made zero. Then by adding 2, we make the last digit of LL value 2. For the other case, last digit of LL value is made 7. Two is the middle (also average) value for the interval {0,1,2,3,4}, seven is the middle value for {5,6,7,8,9}. By choosing these two medium values, some degree of resilience to value distortions is provided.

Authenticating Algorithm

1. Divide the Image to be authenticated in 8x8 non overlapping parts

2. Take DWT of each block independently

3. Make pairs for 8x8 blocks

a. For the second block, for each (i,j) DWT value,
calculate sum = (HL_2(i,j)+ LH_2(i,j)+ HH_2(i,j))
if (sum <= Threshold)
if (LL_1(i,j) >=0 and LL_1(i,j) <=4)
4 pixel values corresponding to this LLVal is genuine
else
4 pixel values corresponding to this LLVal not genuine
end if
else
if (LL_1(i,j) >=5 and LL_1(i,j) <=9)
4 pixel values corresponding to this LLVal is genuine
else
4 pixel values corresponding to this LLVal not genuine
end if
end if

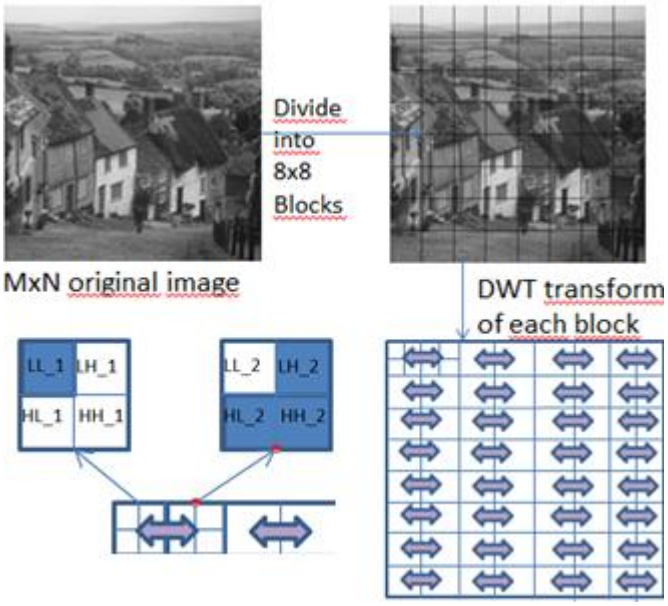

Fig.1. Watermark embedding algorithm

To decide the value to be used as threshold, the sum (LH,HL,HH) values are analyzed. Histogram of those values can be seen in Fig.2.

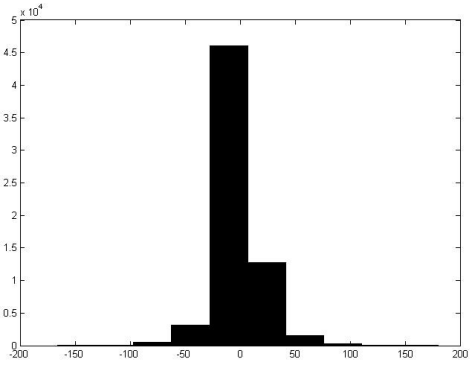

Fig.2. Histogram of sum(HL,LH,HH) values

By looking at the histogram, threshold value -8 is chosen.

III. Experiments and Results

The host image is a grayscale image. But the algorithm can be run on color images by taking the color image to YUV format or by watermarking one or all of the three color bands. The watermarked image is seen in fig.3. Watermarked image has PSNR value of 45.738338 which can be considered a good PSNR value. Algorithm can be run by dividing the image into 512x512(whole image), 64x64, 32x32, 16x16, 8x8 image blocks. One must take into consideration that change detection sensitivity decreases while the block size increases.

The watermarked image is modified in two places where two chimneys disappeared by block copy paste operations. Modified watermarked image is seen in fig.4. Since the aim is to detect changes in the image, altered image's non-professional looking modification is not bothered. It will not affect the algorithm's success whether the changes are made in a smooth way so that human eye cannot detect the change.

Fig.3.Watermarked image, PSNR : 45.788338

Fig.4 Modified watermarked image. Two chimneys are removed by copy paste.

The image authentication algorithm is run on modified image. The result is seen in fig.5. Two chimney regions where the image was modified is detected and marked by white (255) color value by the algorithm. The detected region is seen as blocky because the algorithm detects modified regions by 4-pixel base. By chance, for some values in modified region, sum(HL,LH,HH) are same as original resulting in a blocky appearance.

The authentication algorithm is tested for the attacks that are used for testing robust type of watermarking applications. The aim here is to test how the algorithm behaves in innocent type lossy compression or scale operations. The watermarked image is subjected to Jpeg compression by %75, %50, %25 image quality, blur filter 3x3, scale-rescale, gamma correction, Gaussian noise, histogram equalization, intensity adjustment.

Fig. 5 Image authentication result for modified image

The image authentication algorithm is run on images that are attacked by common image operations. Results of authentication can be seen in fig.8. For crop operation, authentication can detect the blocks that are original. For the other attacks, since almost all of the blocks are affected by the operation, almost all of the blocks are marked as modified. By intuition, it can be seen and decided that a common innocent operation is applied to the image.

The algorithm is tried on different images with different manipulations as seen on fig 6 and fig 7. The manipulated parts of the image are successfully detected by the algorithm.

IV. Conclusion

There are many studies for authentication purpose watermarking. Some of the studies are completely fragile that even a pixel value difference causes authentication to give negative result. Some studies propose semi-fragile type that tolerates some degree of innocent modifications to the image. Most of the studies do not fully describe the details of the algorithm so that the algorithm can be implemented and run.

In this study, a semi fragile blind DWT-based watermarking method is proposed. The algorithm is simple and well-presented so that it can be coded and run easily for testing

purposes. The algorithm can detect changes in the detail degree of 4 pixel blocks. Algorithm embeds the watermark in LL band of DWT transform values.

According to experiments done, when a simple operation is applied to whole of the image, the authentication result gives a hint to the observer or it can be decided automatically that the modification is ill-purposed or not.

The PSNR value for watermarked image can be fairly considered as high which is value 45.788338. The proposed study can be used as an alternative to previous authentication methods.

V. REFERENCES

[1] I. J. Cox, S. Member, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," in Image Processing, IEEE Transactions, 1997, vol. 6, no. 12, pp. 1673–1687.

[2] R. Dugad, K. Ratakonda, and N. Ahuja, "A new wavelet-based scheme for watermarking images," Proc. 1998 Int. Conf. Image Process. ICIP98 (Cat. No.98CB36269), vol. 2, pp. 1–5, 1998.

[3] A. M. ; G. E. Eskicioglu, "Robust DWT-SVD Domain Image Watermarking : Embedding Data in All Frequencies," 2004.

[4] E. Elbasi, A. M. Eskicioglu, and I. Science, "A DWT-based robust semi-blind image watermarking algorithm using two bands," vol. 6072, pp. 1–11, 2006.

[5] O. Jane and E. Elbasi, "A new approach of nonblind watermarking methods based on DWT and SVD via LU decomposition," Turkish J. Electr. Eng. Comput. Sci. doi10.3906/elk-1212-75, pp. 1–13, 2012.

[6] O. Jane, H. Gökhan, and E. Elbaşı, "A Secure and Robust Watermarking Algorithm Based on the Combination of DWT , SVD , and LU Decomposition with Arnold ' s Cat Map Approach," no. 3, pp. 306–310, 2014.

[7] H.-C. Huang and W.-C. Fang, "Metadata-based image watermarking for copyright protection," Simul. Model. Pract. Theory, vol. 18, no. 4, pp. 436–445, Apr. 2010.

[8] R. B. Wolfgang, E. J. Delp, R. B. Wolfgang, and E. J. Delp, "Fragile Watermarking Using the VW2D Watermark," Proc. SPIE/IS&T Int. Conf. Secur. Watermarking Multimed. Contents, vol. 3657, pp. 204–213, 1999.

[9] P. W. Wong and W. Road, "A Public Key Watermark for Image Verification and Authentication," Image Process. 1998. ICIP 98. Proceedings. 1998 Int. Conf., vol. 1, pp. 455–459, 1998.

[10] R. Chamlawi, A. Khan, A. Idris, and Z. Munir, "A Secure Semi-Fragile Watermarking Scheme for Authentication and Recovery of Images based on Wavelet Transform," vol. 2, no. 2, pp. 727–731, 2008.

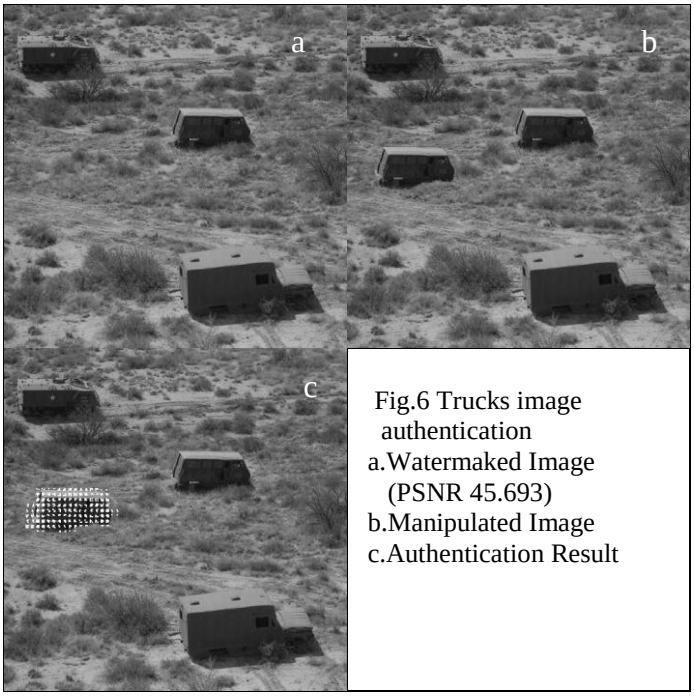

Fig.6 Trucks image authentication
a.Watermaked Image (PSNR 45.693)
b.Manipulated Image
c.Authentication Result

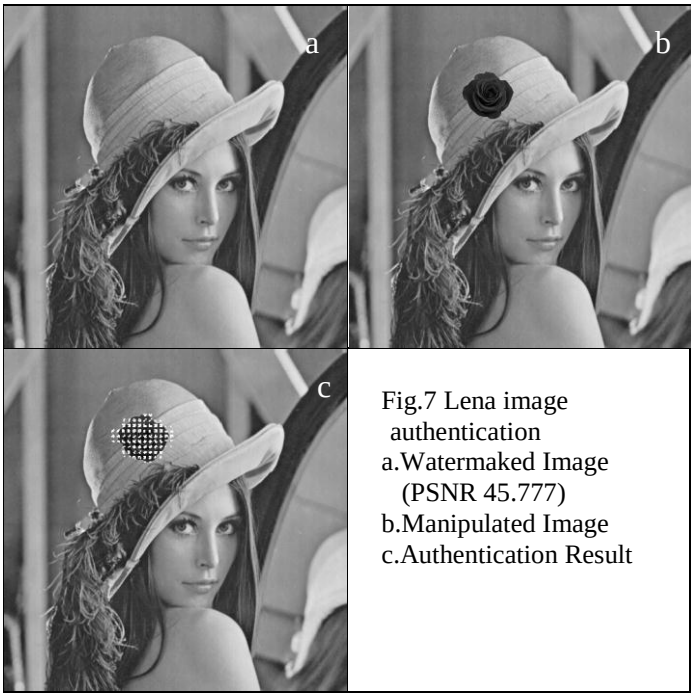

Fig.7 Lena image authentication
a.Watermaked Image (PSNR 45.777)
b.Manipulated Image
c.Authentication Result

Fig. 8. Image authentication applied to watermarked and modified images by various image operations