A CONTENT-BASED IMAGE AUTHENTICATION SYSTEM WITH LOSSLESS DATA HIDING

Dekun Zou[1], Chai Wah Wu[2], Guorong Xuan[3], Yun Q. Shi[1]

[1]Dept. of ECE, New Jersey Institute of Technology, Newark, NJ 07102
[2]IBM Thomas J. Watson Research Center, Yorktown Heights, NY 10598
[3] Dept. of Computer Science, Tong Ji University, Shanghai, P.R.China

ABSTRACT

In this paper, we present a novel content-based image authentication framework which embeds the authentication information into the host image using a lossless data hiding approach. In this framework the features of a target image are first extracted and signed using the Digital Signature Algorithm (DSA). The authentication information is generated from the signature and the features are then inserted into the target image using a lossless data hiding algorithm. In this way, the unperturbed version of the original image can be obtained after the embedded data are extracted. An important advantage of our approach is that it can tolerate JPEG compression to a certain extent while rejecting common tampering to the image. The experimental results show that our framework works well with JPEG quality factors greater than or equal to 80 which are acceptable for most authentication applications.

1. INTRODUCTION

Recently, digital multimedia content such as images and sound data are proliferating at a rapid pace, enabling new ways of communications in our daily life. However, with the wide availability of editing software, it is very easy to manipulate the content of such multimedia data. The question of how to protect the data integrity of the multimedia content is becoming an important issue. Authentication of multimedia data provides an answer to this question. This paper focuses on authentication of images.

Early attempts at image authentication are mainly based on fragile watermarking approaches where the multimedia data is treated as digital bits and signed using traditional cryptographic techniques. The signature information can be inserted into spatial domain [1], DCT domain [2] or wavelet domain [3]. In these algorithms any change to the image, however minor, will render the image inauthentic. As a result, even the slightest use of JPEG lossy compression will result in an inauthentic image.

More recently, many content-based multimedia authentication schemes have been proposed which can tolerate minor modifications such as JPEG compression which do not alter the content of the multimedia data [4-6]. The goal in content-based image authentication is to verify the content of an image, and not its representation. This is a useful idea, but how to characterize the content of an image is still a difficult unsolved problem. Two problems need to be addressed in content-based authentication schemes. First, how tolerant is it to minor modifications? Second, how hard is it to forge an authentic image? In many proposed schemes, the focus is on the first question. In [7] a distortion bounded approach is proposed to address the second question. In this scheme, the target image is quantized and used as the "original" image. The quantized image is then hashed and encrypted to generate a signature. The same quantization is performed prior to verification. Consequently, distortion less than half of the quantization step size will not affect the verification. The use of cryptographic techniques prevents forgery to some extent. The main drawback of this algorithm is that the image quality of the quantized "original" is degraded when compared with the unmodified original image. In [8-9], an algorithm is proposed to address this problem while preserving the cryptographic solution to counter forgery. By introducing a quantization function index vector, the original image can be preserved. In [8-9], the main focus is on label-based techniques where the authentication information is stored as a label attached to the image. Some early image formats do not support labels and labels can be disconnected easily from the image. We propose to extend the approach in [8-9] to watermarking or data hiding solutions by combining it with a lossless data hiding technique and demonstrate its robustness and effectiveness.

The rest of this paper is organized as follows. Section 2 presents the procedure for generating the authentication information. In Section 3, we present the lossless data hiding method used in our framework. In

Section 4, the verification procedure is described. Next, we offer some experimental results in Section 5. Finally, conclusions and problems for future research are discussed in Section 6.

2. GENERATION OF INFORMATION FOR AUTHENTICATION

The generation of the information needed for authentication (for simplicity, the word "tag" is used from now on to represent this authentication information) is adopted from [8-9].

Properly scaled DCT coefficients of image blocks are useful representations of the image content. The original image (I) is first divided into N by N blocks. Then a 2-dimensional DCT is applied to each block. Next, all the (scaled) coefficients are arranged into a feature vector (V). Following that, each element in vector V is quantized by a quantization function chosen among two different quantization functions. The quantization function chosen is the one which has the least quantization error for this element. Thus, a quantized feature vector V' is formed. The information about which quantization function is selected is stored in an index vector X. Appending X to V' forms W which is then signed by the DSA digital signature algorithm resulting in the signature S. Since many of the high frequency coefficients are zero after quantization and the same quantization function is chosen, the vector X can be compressed effectively into X' by a lossless compression algorithm. Finally, the verification tag to be embedded into I is formed by appending X' to S. The procedure is illustrated in Figure 1.

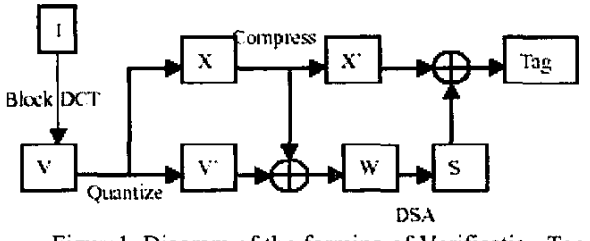

Figure1. Diagram of the forming of Verification Tag.

As we have discussed in Section 1, the approach in [7] requires the original image to be quantized in order to make it authenticable. Specifically, the feature value is shifted to the middle of quantization step. This avoids the situation where the feature value is near the boundary of a quantization region and where minor distortion results in false rejection. This quantization results in a degraded original image. The introduction of an index vector X in our approach avoids this problem and the original image does not need to be quantized to be authenticable.

3. TAG EMBEDDING

After the tag has been generated, it is embedded into the image itself by a data hiding method. To maintain the benefit of not having to degrade the original image to make it authenticable, a lossless data hiding technique is used. On the other hand, the tag should survive minor distortion such as JPEG compression. The data hiding method we use will be based on the circular histogram algorithm proposed in [10].

First, the original image I is divided into 8 by 8 blocks. One bit of information is embedded into each block. The 64 pixels in each block are pseudo-randomly divided into two groups A and B, each of which has 32 pixels. For each group, the histogram of the gray scale values is mapped onto a circle. Specifically, the gray level of pixels is quantized into Q ranges with step size of P=256/Q (assuming the gray levels are represented by 8 bits). The circle is equally divided into Q segments each of which represents a range of gray value. If the gray level of a pixel falls into the range representing it, a unit weight will be put in the middle of the segment. That is, each pixel generates one weight. After all the pixels of a pixel group in one block are processed, there will be 32 weights put on the circle. Then the center of mass of the circle can be obtained by calculating the barycenter of all the weights. Then the same steps are applied to the other pixel group of the same block.

Statistically, the centers of mass of the two pixel groups should be close since the pixels of the two groups are randomly selected. To embed one bit of information into this block, the gray level of pixels in this block are modified according to following rule:

To embed "1"

C'=C+P for pixels in group A
C'=C-P for pixels in group B

To embed "0"

C'=C-P for pixels in group A
C'=C+P for pixels in group B

where C is the pixel value of the original image, C' is the pixel value of marked image and P is the step size P=256/Q.

We can see that if "1" is embedded, all the gray levels of pixels of group A will be increased by P. If we map the new values onto a circle, the weight representing a pixel will be put into the segment next to the original one in the clockwise direction. As a result, the center of mass of group A will turn clockwise by 360/Q degrees. Similarly, the center of mass of group B will turn counterclockwise by the same degree. To extracting this bit, we simply compare the angle difference between the centers of mass of pixels of group A and group B. If the angle of A is greater than that of B, bit "1" is extracted. Otherwise, bit "0" is extracted. The original image can be obtained by applying the operation inverse to the embedding rules.

For some blocks, the centers of mass of group A and group B may be so far away that errors will occur when extracting. Specifically, if the angle of the center of mass of group A is over 360/Q degrees greater than that of group B, and the bit "0" is embedded here, then after the embedding operation is performed, the angle of group A will still be greater than that of B and thus an erroneous bit "1" is extracted. To avoid this problem, such blocks are skipped when embedding information. However, to ensure the correct recognition of such blocks during extraction, the angle difference will be expanded to be larger than what would occur when information is embedded.

If the authenticable image is subject to JPEG compression, the gray values of some pixels will shift slightly and the center of mass of a pixel group will not shift significantly. As a result, the embedded bit can still extracted reliably.

After the tag is inserted, the resulting image is ready for authentication. We refer to this image as the authenticable image.

4. VERIFICATION PROCEDURE

To verify the content of an authenticable image I^a, the embedded information is first extracted and the original image restored. The embedded information contains the signature S and vector X'. Next, X' is decompressed to obtain the index vector X. The same procedure as Section 2 is applied to the restored authenticable image to generate quantized features except that instead of using the quantization errors as a criterion to selecting the quantization functions used, the extracted index vector X is used to select the quantization functions. The quantized features along with X are then verified using the extracted signature S. If they are verified then the content of I^a is authentic (Figure 2).

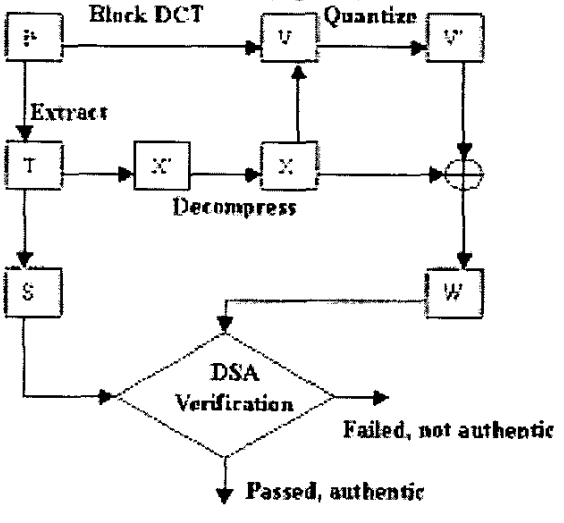

Figure 2. Diagram of the Verification Procedure.

Several situations will result in the failure of the verification. The first reason is if the image is degraded too much so that the embedded tag could not be extracted correctly. The second reason is when the tag extraction is correct, but the image is corrupted such that the quantized features do not verify with the extracted signature S. It is desirable to choose system parameters so that these two scenarios occur with similar levels of distortion to the image.

5. EXPERIMENT RESULTS

To test the effectiveness of our framework, we apply our method to various images. In our experiments, the block size for feature extraction is 32 by 32. For robustness considerations, error correction coding is applied to the authentication tag before embedding. Two kinds of attack are considered: JPEG compression and slight tampering of the image content. Here, we do not apply JPEG to the tampered image. Table 1 shows the experiment results of the "Lena" image. (PSNR is calculated by comparing the original image with the image after the hidden data has been extracted.) JPEG90 refers to JPEG compression with quality factor 90. Figure 3 shows the original and the authenticable "Lena". Figure 4 shows the tampered "Lena" image.

Table 1. Experiment Results of "Lena"

Lena Image	Authentication	PSNR (dB)
No JPEG	PASSED	Infinity
JPEG90	PASSED	38.72
JPEG80	PASSED	34.66
JPEG75	PASSED	33.14
JPEG70	PASSED	31.49
JPEG65	FAILED	29.86
Tampered	FAILED	31.86

Table 2 shows the results for the "Baboon" image. Figure 5 shows the original and the authenticable image. Figure 6 shows the tampered image.

Table 2. Experiment Results of "Baboon"

Baboon Image	Authentication	PSNR (dB)
No JPEG	PASSED	Infinity
JPEG90	PASSED	36.68
JPEG85	PASSED	33.57
JPEG80	PASSED	31.74
JPEG75	FAILED	30.18
Tampered	FAILED	33.59

We can see from the experimental results that changing the content renders the image inauthentic, while

JPEG compression with quality factor higher than 80 will be verified as authentic.

6. CONCLUSIONS AND FUTURE WORK

In this paper, a new watermarking framework is proposed for image content authentication such that the original image can be restored, is robust to JPEG compression and is signed with cryptographic signature algorithms. According to our experiment results, we claim that our system can survive JPEG compression with quality factor 80.

In our experiments, we tested over 1000 images and find that our approach is not suitable for some images. Approximately 5 percent of all the images we tested have problems using our algorithm. Future work includes refining our method to be applicable to more images and to increase the robustness of the system to tolerate lower JPEG compression quality factors.

7. ACKNOWLEDGEMENTS

This work is partially supported by New Jersey Commission of Science and Technology via NJWINS, New Jersey Commission of Higher Education via NJ-I-TOWER, and NSF via IUCRC.

8. REFERENCE

[1] S. Walton, "*Information Authentication for a Slippery New Age*", Dr. Dobbs Journal, 20(4), pp18-26, April 1995.

[2] M. Wu, B. Liu, "*Watermarking for Image Authentication*", Proceedings ICIP-98, vol. 2, pp. 437-441, 1998.

[3] G.R. Arce, Liehua Xie, "*Joint wavelet compression and authentication watermarking*" Proceedings ICIP-98, vol.2 pp. 427-431, 1998.

[4] M. Schneider, Shih-Fu Chang, "*A robust content based digital signature for image authentication*", Proceedings ICIP-96, vol. 3, pp. 16-19, 1996.

[5] M. P Queluz, "*Towards robust, content based techniques for image authentication*", Proceedings of IEEE Workshop on Multimedia Signal Processing, pp. 297 –302, 1998.

[6] W. Wolf, Xiangyang Kong, H Yu, "*Techniques for content-based graph authentication*", IEEE Transaction on Multimedia, vol. 8, no. 4, pp 38 –45, 2001.

[7] N. Memon, P. Vora, B. Yeo, M. Yeung, "*Distortion Bounded Authentication Techniques*", Proceedings of SPIE, vol. 3971, pp. 164-174, 2000.

[8] C. W. Wu, "*Limitations and Requirements of Content-based Multimedia Authentication Systems*", Proceedings of SPIE, vol. 4314, pp. 241-252, 2001.

[9] C. W. Wu, "*On the Design of Content-based Multimedia Authentication Systems*", IEEE Transaction on Multimedia, vol. 4, no. 3, pp. 385-393, Sept. 2002.

[10] C. De Vleeschouwer, J.F. Delaigle, B. Macq, "*Circular Interpretation of Histogram for Reversible Watermarking*", Proceedings of IEEE Workshop on Multimedia Signal Processing, Cannes, France, pp. 345 –350, 2001.

(a) Original (b) Authenticable

Figure 3. Lena Image.

Figure 4. Tampered Lena image.

(a) Original (b) Authenticable

Figure 5. Baboon Image.

Figure 6. Tampered Baboon image.