WATERMARKING FOR IMAGE AUTHENTICATION

Min Wu *Bede Liu*

Department of Electrical Engineering
Princeton University, Princeton, NJ 08544, USA
Fax: +1-609-258-3745 {minwu, liu}@ee.princeton.edu

ABSTRACT

A data embedding method is proposed for image authentication based on table look-up in frequency domain. A visually meaningful watermark and a set of simple features are embedded invisibly in the marked image, which can be stored in the compressed form. The scheme can detect and localize alterations of the original image, such as the tempering of images exported from a digital camera.

Keywords: digital watermark, image authentication.

1. INTRODUCTION

Digital information revolution has brought about many advantages and new issues. With the ease of editing and perfect reproduction, the protection of ownership and the prevention of unauthorized manipulation of digital audio, image, and video materials become important concerns. Digital watermarking, a scheme to embed special labels in digital sources, has made considerable progress in recent years. There are several categories of watermarking schemes [1]. Among them, *fragile watermarking* is a technique to insert a signature for image authentication. The signature will be altered when the host image is manipulated. In this paper, we focus on watermarking for digital image authentication.

An effective authentication scheme should have the following desirable features:

1. to be able to determine whether an image has been altered or not;
2. to be able to locate any alteration made on the image;
3. to be able to integrate authentication data with host image rather than as a separate data file;
4. the embedded authentication data be invisible under normal viewing conditions;
5. to allow the watermarked image be stored in lossy-compression format.

Previously published methods for image authentication do not satisfy all the requirements. The digital signature proposed in[2], as well as the content based signature reported in [3] and [4] do not satisfy requirements 2 and 3. The pixel-domain scheme [5] can not be stored in lossy compression format. In addition, two frequency domain data hiding schemes [6][7] may be used for authentication, but they cannot always locate alteration and the distortion introduced by embedding is larger than the proposed scheme.

This paper presents a new authentication scheme by embedding a visually meaningful watermark and a set of simple features in the frequency domain of an image via table look-up. This scheme can be applied to compressed image using JPEG or other techniques such as Wavelet compression, and the marked image can be kept in the compressed format. Any alteration made on the marked image can be localized, making it suitable for "trustworthy" digital camera. Furthermore, this scheme is computationally efficient, and can be applied to video.

2. TWO ASPECTS OF PROPOSED AUTHENTICATION APPROACH

We shall present our approach in the context of grey scale JPEG compressed images. The extension to images compressed using other means and to color images are reasonably straightforward and discussed in Section 7.

A block diagram for the embedding of data is given in Fig. 1 which, aside from the block labeled "embed",

This research is supported by a New Jersey State R&D Excellence Award and NSF grant MIP-9408462.

is identical to that of the JPEG compression [8]. Watermark is inserted into the quantized DCT coefficients via a look-up table.

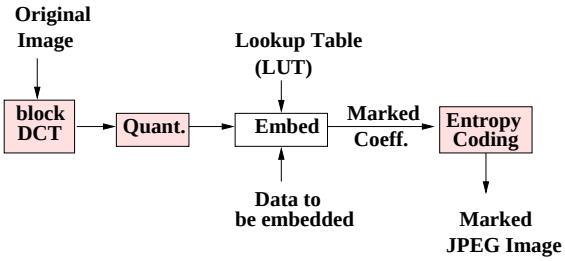

Figure 1: Block Diagram of Embedding Process

In the next section, we present how to embed data. The question of what data are used for authentication is discussed in Section 4.

3. PROPOSED FREQUENCY-DOMAIN DATA EMBEDDING SCHEME

3.1. Embedding Binary Bits Via Table Look-up

A proprietary look-up table (LUT) is generated beforehand by the owner of the image or the digital camera manufacturers. The table maps every possible value of JPEG coefficient randomly to "1" or "0" with the constraint that runs of "1" and "0" are limited in length. To embed a "1" in a coefficient, the coefficient is unchanged if the entry of the table corresponding to that coefficient is also a "1". If the entry of the table is a "0", then the coefficient is changed to its nearest neighboring values for which the entry is "1". The embedding of a "0" is similar. This process can be abstracted into the following formula where v_i is the original coefficient, v_i' is the marked one, b_i is the bit to be embedded in, and $LUT(\cdot)$ is the mapping by look-up table:

$$v_i' = \begin{cases} v_i & \text{if } LUT(v_i) = b_i \\ v_i + \delta & \text{if } LUT(v_i) \neq b_i, \text{ and} \\ & \delta = \min_{|d|}\{d \in \mathbf{Z} : LUT(v_i + d) = b_i\} \end{cases}$$

The extraction of the signature is simply by looking up the table. That is,

$$\hat{b_i} = LUT(v_i')$$

where $\hat{b_i}$ is the extracted bit.

The basic idea of the embedding process is also illustrated by the example in Fig. 2. In this example, zeros are to be embedded in two AC coefficients with values "-73" and "-24" of an 8x8 image block. The entry in the table for coefficient value "-73" is "1". So to embed a "0", we go to its closest neighbor for which the entry is "0". In this case, "-73" is changed to "-74". Since the entry for coefficient value "24" is "0", it is unchanged.

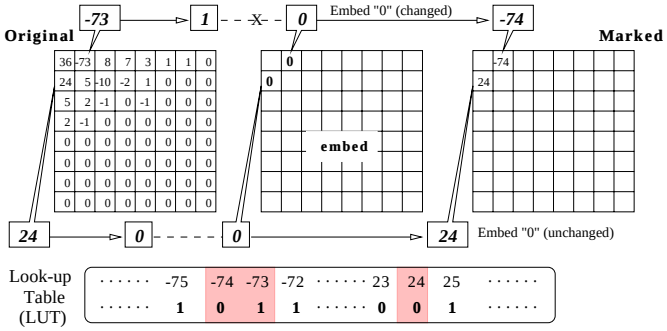

Figure 2: Frequency-domain Embedding Scheme

3.2. Considerations For Minimizing Distortion

Several steps are taken to ensure that the markings are invisible:

- The run of "1" and "0" entries in the LUT is constrained to avoid excessive modification on the coefficients;
- DC coefficient in each block is not changed to avoid blocky effect unless the quantization step is very small;
- Small valued coefficients are not modified to avoid high frequency distortion.

The above constraints make the number of bits that can be embedded varies significantly from block to block. Also, extraction error may occur due to image format conversion and other causes. For these reasons, we choose to embed one bit per 8x8 block by embedding the same bit to all embeddable coefficients in the block and detecting that bit by majority voting. The detection of an error in the embedded bit signals alteration of that block, which will be discussed in next section. Better encoding schemes will be discussed in Section 7.

The constraint of leaving DC coefficients and small valued AC coefficients unchanged makes it impossible

to embed data in smooth blocks where all AC coefficients are very small hence unembeddable. We shall discuss this problem in Section 5.

4. CHOICE OF EMBEDDED DATA FOR AUTHENTICATION

The authentication data we embed in an image consists of a visually meaningful binary pattern and some content features. The visually meaningful pattern serves as a quick check for alteration and location of possible alteration. A simple example of content features is the most significant bit of macroblock mean intensity. Another example is the sign of intensity difference between macroblocks. The combination of these two types data is suitable for such applications as image authentication for "trustworthy" digital cameras. The embedding and verifying procedures are summarized in Fig. 3.

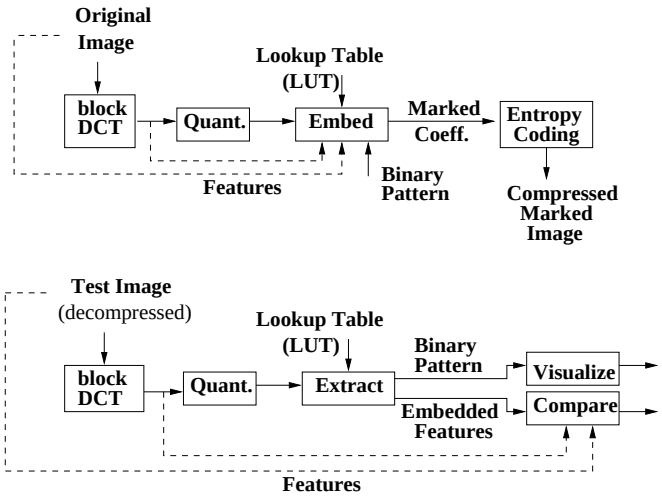

Figure 3: Embedding and Extraction Procedures

5. SMOOTH REGION EMBEDDING

Authentication data cannot be embedded in smooth blocks because the AC coefficients are small. In a typical natural image such as the one shown in Fig. 4, about 20% of the blocks are smooth, which is illustrated in Fig. 5.

We propose a solution to embed data in smooth region by embedding data in blocks that are not smooth but whose location bears a fixed relationship to the smooth block in question. A simple implementation, called *backup embedding*, is illustrated in Fig. 6. Instead of embedding one bit independently in each block,

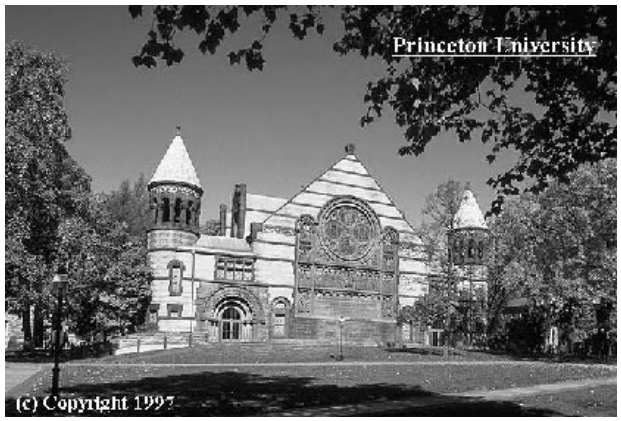

Figure 4: Original Image (75% JPEG Compression)

we take the macroblock as a unit. We use two bits to embed data corresponding to the current macroblock, and use the other two bits to embed a backup copy of data corresponding to the companion macroblock. In the illustration of Fig. 6, the companion blocks are half picture height apart.

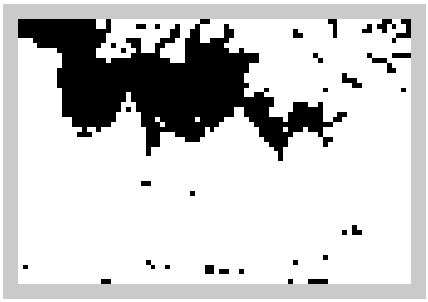

Figure 5: Smooth Blocks of Fig. 4 (shown in black)

6. EXPERIMENTAL RESULTS

A JPEG compressed image of size 640x432 is shown in Fig. 4. Shown in Fig. 7(a) is the same image but with a 40x27 binary pattern "PUEE" of Fig. 7(b) and the MSB of macroblock mean intensity embedded in. This image is indistinguishable from the original. The smooth regions of the image is shown in Fig. 5. For these blocks, the backup embedding is used.

The marked image is modified by changing "Princeton University" to "Alexander Hall" and "(c) Copyright 1997" to "January 1998", shown in Fig. 8(a). This

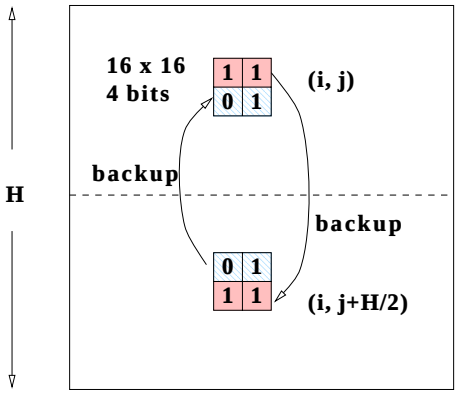

Figure 6: Backup Embedding For Smooth Region

image is again stored in the JPEG format. For ease of visualization, we shall denote the block that embeds "0" as a black dot, the block that embeds "1" as a white dot, and the block with no majority as a gray dot. With these notation, Fig. 8(b) and Fig. 8(c) show the extracted binary pattern and content feature matching result of the edited image. The random pattern corresponding to the altered blocks are clearly identifiable. Also, very few unembeddable data are shown as isolated gray dots by the backup scheme.

Figure 7: $\frac{a}{b}$ (a) Marked Image; (b) Embedded Binary Pattern.

Figure 8: $\frac{a}{b|c}$ (a) Marked Image After Editing (stored in 75% JPEG); (b) Extracted Binary Pattern from Edited Image; (c) Feature Matching Result.

7. IMPROVEMENT AND EXTENSIONS

7.1. Double Watermarking For Authentication and Ownership Verification

Previous work on double watermarking mainly emphasized on embedding multiple labels (such as ownership label, recipient label, etc.) using the same embedding approach. No published work discussed to verify ownership as well as signal and locate alteration. A key issue for this kind of double watermarking is that no original "true" image can be used since the validity of this original image has not been verified yet.

The proposed new scheme can be combined with a previous robust watermarking and detection scheme on ownership verification [9] for double watermarking. This double watermarking has been tested successfully.

7.2. Embedding More Data

We have described previously a method to embed 1 bit per 8x8 block. The amount of data embedded can be increased by more sophisticated backup and coding.

7.3. Multi-level Data Embedding and Unequal Error Protection

We mentioned in Section 4 that two sets of data, namely, a meaningful visual pattern and a set of low-level con-

tent features, are embedded in the image for authentication purpose. *Multi-level data embedding and unequal error protection* are possible in order to embed several sets of data with different levels of error protection.

7.4. Other Compression Format

Besides JPEG compressed image, we also tested Wavelet compression and found our approach effective.

7.5. Color Image and Video

For color images, we may work in YCrCb coordinates and use the new approach to mark luminance components while leaving chrominance components unchanged; or we may apply the approach to chrominance components as well to embed more data. We may also work in other color coordinates, such as RGB.

The proposed scheme can be applied to MPEG compressed digital video. A simple approach would be to mark I-frames of video streams using our new scheme. In addition, I-frame serial number can be used as part of embedded data to detect modification such as frame reordering and frame dropping.

8. CONCLUSION

In this paper, we have presented a frequency domain technique for image authentication in which the signature is embedded in the image and the marked image can be kept in the compressed form. The scheme can detect tempering of the marked image and can locate where the tempering has occurred. Potential applications include trust-worthy digital cameras and camcorders. Image samples and other information are available at http://www.ee.princeton.edu/~minwu/rsch_authwmark.html . Refer to the files attached in the CD-ROM for full-size high-resolution images of Fig. 4, 7(a), and 8(a).

9. REFERENCES

[1] F. Mintzer, G. W. Braudaway, M. M. Yeung: "Effective and Ineffective Digital Watermarks", *ICIP*, 1997

[2] G. L. Friedman: "The Trustworthy Digital Camera: Restoring Credibility to the Photographic Image", *IEEE Trans. on Consumer Electronics*, Nov. 1993.

[3] M. Schneider, S-F. Chang: "A Robust Content Based Digital Signature for Image Authentication", *ICIP*, 1996.

[4] D. Storck: "A New Approach to Integrity of Digital Images", *IFIP Conf. on Mobile Communication*, 1996.

[5] M. M. Yeung, F. Mintzer: "An Invisible Watermarking Technique for Image Verification", *ICIP*, 1997.

[6] M. D. Swanson, B. Zhu, A. H. Tewfik: "Robust Data Hiding for Images", *IEEE DSP Workshop*, 1996.

[7] E. Koch, J. Zhao: "Towards Robust and Hidden Image Copyright Labeling", *IEEE Workshop on Nonlinear Signal and Image Processing*, 1995.

[8] G. K. Wallace: "The JPEG Still Picture Compression Standard",*IEEE Trans. on Consumer Electronics*, vol.38, no.1, pp18-34, 1992

[9] W. Zeng, B. Liu: "On Resolving Rightful Ownerships of Digital Images by Invisible Watermarks", *ICIP*, 1997.