

WATERMARK IMPLEMENTATION IN DIGITAL PHOTOGRAPHY

Sammy H.M. Kwok, Edmund Y. Lam

Department of Electrical and Electronic Engineering,
The University of Hong Kong,
Pokfulam Road, Hong Kong SAR

ABSTRACT

Digital watermarking can be used for tamper proofing of images taken by digital cameras. Such tamper proofing can be applied in authentications for courtroom evidence, insurance claims, copyright claims and journalistic photography. In this paper, we discuss the implementation of digital watermarking in the digital camera environment using a fragile watermark embedded in the discrete wavelet domain of the images. We propose an architecture to implement the scheme under the constraints of the digital camera environment, such as with a limited computation power. We also evaluate the versatility of the scheme to detect and characterize signal modifications from a number of distortions, such as substitution of data, filtering, and lossy compression, through software simulation. Results show that the proposed scheme is feasible and effective for authentication of images taken by digital cameras.

Keywords – *digital watermarking implementation, authentication, digital photography, wavelet transform.*

1. INTRODUCTION

Digital photography is rapidly gaining popularity, and so the authentication of images taken by digital cameras has become a great concern. These images are easily modified or forged using widely available editing software. Such problems have hindered the application of digital images for courtroom evidence, insurance claims, copyright claims and journalistic photography.

By using digital watermarking techniques, information such as the author's credentials and time stamps can be embedded into digital images for tamper-proofing or authentication. Such an embedded watermark is transparent to general readers of the images, but it can be extracted and verified for detecting any tampering with the images using a relevant detection tool. Over the past decade, many invisible watermarking schemes have been proposed in the literature [1,2].

For example, Kundur and Hatzinakos proposed a digital watermarking scheme for Telltale tamper proofing and authentication [3]. In the proposed scheme, a fragile digital watermark can be embedded in the discrete wavelet

transform (DWT) domain of an image by quantizing the corresponding coefficients. Since the watermark is placed in the DWT domain, it allows the detection of changes in localized spatial and frequency domain region of the images. Besides, embedding the watermark in the DWT domain instead of spatial domain also improves the imperceptibility of the watermark. Yet, in the paper, the relationships between the parameters in the algorithm and the implementation costs were not discussed. We thus could not directly apply their scheme for digital camera without further study on its feasibility given the constraints in the digital camera environment, such as a limited computational power and the availability of processing power in the in-camera image compression [4].

In Chen and Chang, a digital watermarking scheme using SHA and RC6 cryptographic algorithms was proposed [5]. The scheme promises robustness against the JPEG compression, the EZW compression, noise corruption, cropping and sharpening, implying that the watermark can be used for image authentication. However, this paper also did not give a detailed consideration on the implementation perspectives of the watermarking scheme. In addition, the relationship between the robustness of the watermark and the parameters in the algorithm has not been discussed.

The objective of our work is to develop a feasible and effective digital watermarking scheme for digital cameras so as to provide a cost-effective and practical tool for authentication and tamper-proofing of digital images taken by digital cameras. We pay particular attention to the implementation details of the scheme in the context of digital cameras. The proposed framework, together with discussions on the choice of parameters, are given in Section 2. They are tested through extension simulation given in Section 3. We conclude our findings in Section 4.

2. PROPOSED TECHNIQUE

2.1. The proposed digital watermarking scheme

As with traditional color processing, we first convert an image from an RGB color space to the YCbCr color space [6]. Then the Y component of the image is down-sampled to form a grayscale image of resolution of 1M pixels (assuming that the original is between 2M and 8M pixels, true for most digital cameras today). Afterwards, a watermark is embedded in the image by quantizing the coefficients of the 5th level

DWT of the image. Finally, the image is converted back to spatial domain by IDWT and a watermarked image is formed by up-sampling the image and adding it with the original Cb and Cr color components.

The watermark extraction process is similar to the watermark embedding process. The whole watermark embedding process is shown in Figure 1 whereas the watermark extraction process is shown in Figure 2.

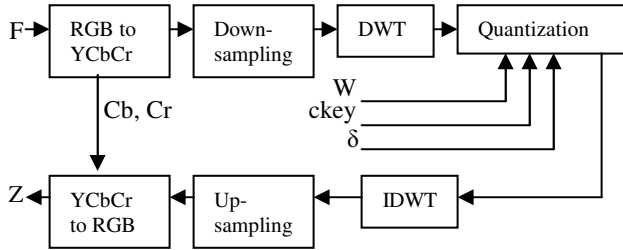


Fig. 1. Proposed watermark embedding process.

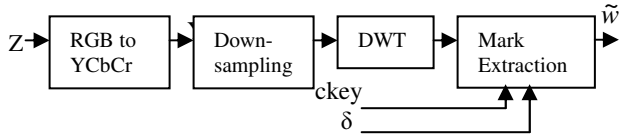


Fig. 2. Proposed watermark extraction process.

In the quantization process, the parameters for the quantization, namely the watermark (w), the coefficient selection key ($ckey$) and the quantization magnitude (δ), form a user-selectable key for the watermarking process. Since only the user of the digital camera knows the key for the watermarking, security against forgery is guaranteed.

To assess the artifacts created in the watermarked image by the watermarking scheme, both qualitative observations and the peak signal-to-noise ratio (PSNR) of the watermarked image are proposed to use. PSNR is defined as

$$PSNR(F, Z) = 10 \log_{10} \left[\frac{\left(\max_{m,n} F(m, n) \right)^2}{\frac{1}{N_F} \sum_{m,n} (Z(m, n) - F(m, n))^2} \right] \quad (1)$$

where $F(m, n)$ = the host image
 $Z(m, n)$ = the watermarked image.

A fragile watermarking scheme implies that the extracted watermark will not be identical to the original one whenever there is any distortion in the watermarked image, resulting in the failure of the authentication procedure. To measure the credibility of the modified image, the mean square error (MSE) between the original watermark and the extracted watermark can be used. MSE is defined as

$$MSE(w, \tilde{w}) = \frac{1}{N_w} \sum_{i=1}^{N_w} (w(i) - \tilde{w}(i))^2 \quad (2)$$

where w = the original watermark
 \tilde{w} = the extracted watermark
 N_w = the length of the watermark

A small value of the MSE (For example, a value less than 0.02) indicates the extracted watermark is still credible. When the MSE value is near 0.5, we can consider that the extracted watermark is totally uncorrelated to the original one.

2.2. Algorithmic features of the scheme

The algorithmic features of the scheme are carefully chosen and fine-tuned after considering the implementation costs and constraints, and scope of application in the digital camera environment. The following discusses the optimal choice of all the algorithmic features:

1. Size of the image to be watermarked: The size of images taken by digital cameras ranges from 2M pixels to 8M pixels. As the processing time for watermark embedding is proportional to the image size to be watermarked, the performance of the algorithm can be more predictable and consistent if the image size can be fixed to a certain small and reasonable value. Thus, we propose to down-sample the host image to a fixed size of 1M pixels before the actual watermark embedding process.

2. Maximum level of wavelet decomposition (L): Usually, images taken by digital cameras are stored in JPEG format with a mild compression ratio. This still tampers with the watermark embedded in the image especially in its high-frequency components. In order to maintain the credibility of the watermark after JPEG compression during image storage, we choose the maximum level of wavelet decomposition, L to be 5.

3. Coefficient selection key ($ckey$): “ $ckey$ ” is used to select which DWT coefficients of the host image are to be used for the watermark embedding. The key is generated by randomly selecting a coefficient from either one of the three detailed wavelet coefficients (horizontal, vertical and diagonal) of the image at the 5th level of decomposition for each spatial location. To generate $ckey$, the following procedures are used:

- a) The user of the digital camera selects a user password.
- b) The password is then hashed by using MD5 and the hash result acts as the key for generating $ckey$ by random using RC4 [7,8].

Since only the user of the digital camera knows the password, it is unlikely for any one to find or generate the same $ckey$ easily for forging the user’s watermarks.

4. Generating the watermark (w): As the size of the detailed coefficients at the 5th level of DWT is 32 x 32, the size of w can be set to 32 x 32 bits = 1024 bits = 128 bytes. To generate w , the first 112 bytes of w are assigned with the author’s credentials such as the author’s name and a time

stamp showing the time when the image is taken (If the length of the author's credentials is less than 112 bytes, they can be repeated multiple times to increase the length). Then, the last 16 bytes of w are assigned with a keyed hash value of the author's credentials where the key for hashing is selected by the user of the camera. With this hash value, the authenticity of the author's credentials can be verified.

5. Quantization magnitude (δ): During the quantization process, δ can be used for establishing an appropriate sensitivity of the watermark to changes in the image. A smaller value of δ will make the quantization process of the detailed wavelet coefficients of the image finer and hence makes minor changes in the image easier to detect. δ can be selected by the user of the camera ranging from 1 to 3 so that the user can choose his preference on the sensitivity of the watermark to the changes in the watermarked image.

3. SIMULATION AND DISCUSSION

In the software simulation, a digital image (shown in Figure 3a) with resolution of 1600 x 1200 pixels was used and the following string of 112 characters long acts as the author's credentials: "Sammy Kwok 2005-03-1917:07 Sammy Kwok 2005-03-19 17:07Sammy Kwok 2005-03-19 17:07Sammy Kwok 2005-03-19 17:07".

To test the robustness of the watermarking scheme, different kinds of tampering and distortions were introduced into the watermarked image using a commercial image-editing software, and the resultant images were examined.

3.1. Maximum level of wavelet decomposition, L

We embedded watermarks with $\delta = 2$ and L ranging from 1 to 5, then stored the watermarked images in JPEG format with a mild compression ratio, and finally measured the MSE value of the extracted watermarks. The watermarked image with $L = 5$ and $\delta = 2$ is shown in Figure 3b and the MSE of the extracted watermarks are shown in Figure 4.

In the case of $L = 5$, the extracted author's credential is "Seemy(Kwok 2005-03-19 ◀wz0| Sammy Kwok 2005-03-19 17:07Salmy Kwok :005/0;-1y 17:07Sammy0Kwo{ 2005-03-19 17:07". Even though there are differences with the watermark we use, they are small enough that we can extract an entire occurrence of "Sammy Kwok 2005-03-19 17:07" unaltered. Such is not the case for other values of L . Furthermore, the experiment results show that the MSE value of the extracted watermark in images stored in JPEG format can only be less than 0.02 if L is at least 5. To tradeoff the consideration on the processing time requirement for calculating the wavelet transform, the optimal level of wavelet decomposition L is seen to be 5 in our scheme.

3.2. Perceptibility of the watermark

δ has a direct effect on the perceptibility of the watermark embedded in the image. To measure the effect, we embedded the watermark into the image with δ ranging from 1 to 5. Then, the perceptibility of the watermark in the watermarked

images was subjectively inspected and the PSNR of the watermarked images was calculated. The results are shown in Figure 5. We find that, for a sufficiently imperceptible watermark, $\delta \leq 3$ is necessary. Thus, in our proposed scheme, users are allowed to choose the value of δ ranging from 1 to 3 so as to select the robustness of the watermark for different applications.



Fig. 3a. Original image



Fig. 3b. Watermarked image

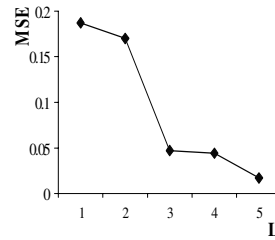


Fig. 4. MSE of extracted watermark vs L

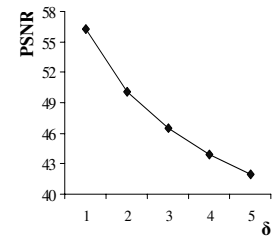


Fig. 5. PSNR of watermarked image vs δ

3.3. The robustness of the watermark

To test the tamper proofing performance of the watermarking scheme, the watermarked image was modified by the following six different ways. The simulation results are summarized in Table 1.

Tampers	MSE
Mild JPEG compression	0.0176
Mild JPEG compression + Removal of the beetle	0.0273
Mild JPEG compression + Increase in Red by 50%	0.2891
Mild JPEG compression + Increase in color saturation by 50%	0.2607
Mild JPEG compression + Rotation by 0.1 degree	0.2275
Mild JPEG compression + Soften filtering	0.4316
Mild JPEG compression + Sharpen filtering	0.1299

Table 1. MSE of the extracted watermark under different kind of tampers.

These results show that the proposed watermarking scheme can be used to detect the distortion and tampering with the image such as pixel modification, color modification, and filtering. The hash value of the extracted watermark can be used to detect any tampering with the watermarked image

and the MSE and the content of the extracted watermark can be used to assess the seriousness of the tampering.

3.4. Implementation costs consideration

We need to be concerned with the implementation cost in five particular areas:

1. Color space conversion: Most digital cameras in the market support storing images in JPEG format. Because JPEG compression involves RGB to YCbCr color space conversion, the color space conversion required in the proposed scheme does not impose any extra implementation cost.

2. Discrete wavelet transform (DWT): Since L is set to 5 and the size of the host image is down-sampled to 1M pixels, the total computations for the DWT and IDWT are about 2.8 million operations.

3. Generation of the coefficient selection key (ckey): "ckey" is generated once only when the user of the camera changes his password. Thus, no extra implementation cost is required during each watermark embedding process.

4. Generation and embedding of the watermark (w): The amount of operations for hashing and watermark embedding is 17.5K and 1K respectively. Thus, the total computations required for the generation and embedding of the watermark is about 18.5K operations.

5. Extraction of watermark and tamper assessment: Since the extraction of watermark and tamper assessment are done by software tools running in computers, no extra implementation cost in digital cameras is required.

Overall, we can conclude that the proposed scheme does not consume a lot of extra computational power, and is feasible to be implemented on a digital camera with only a minimal load on its performance.

4. CONCLUSION

Digital watermarking can be used for the tamper proofing of multi-media data such as still images. In many previously proposed digital watermarking schemes, the suitability and implementation perspectives of the watermarking schemes in the digital camera environment have not been sufficiently addressed. In this paper, we propose a feasible and cost-effective digital watermarking scheme for providing means to authenticate digital images taken by digital cameras.

By software simulations, all the algorithmic features of the scheme were tested. Their effects on the effectiveness and robustness of the watermark and the relationships between the features and the implementation costs were studied. In addition, the robustness of the watermark against different kinds of distortions and tampering which are common to digital images was verified.

The simulation results indicate that the proposed scheme can be implemented in digital cameras without imposing a significant extra implementation cost. Further research can concentrate on the optimization of the architecture for hardware implementation of the proposed scheme using

ASIC or FPGA [9], or software approach such as assembly programs in DSP or dedicated image processors.

5. REFERENCES

- [1] C. T. Hsu and J. L. Wu, "Hidden Digital Watermarks in Images", *IEEE Transactions on Image Processing*, vol. 8, no. 1, pp. 58-68, January 1999.
- [2] C. I. Podilchuk and E. J. Delp, "Digital Watermarking Algorithms and Applications", *IEEE Signal Processing Magazine*, pp. 33-48, July 2001.
- [3] D. Kundur and D. Hatzinakos, "Digital Watermarking for Telltale Tamper Proofing and Authentication", *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1167-1180, July 1999.
- [4] E. Lam, "Image Restoration in Digital Photography", *IEEE Transactions on Consumer Electronics*, vol. 49, no. 2, pp. 269-274, May 2003.
- [5] Y. C. Chen and L. W. Chang, "A Secure and Robust Digital Watermarking Technique by the block cipher RC6 and Secure Hash Algorithm", *Proc. IEEE Int. Conf. Image Processing, 2001*, vol. 2, pp. 518-521, October 2001.
- [6] R. Gonzalez and R. Woods, *Digital Image Processing*. Second edition. Prentice Hall, 2002.
- [7] A. J. Menezes, P. V. Oorschot and S. A. Vanstone, *Handbook of Applied Cryptography*, CRC Press, Boca Raton, 1996.
- [8] B. Schneier, *Applied Cryptography: Protocols, Algorithms, and Source Code in C, Second Edition*, John Wiley & Sons, New York, 1996.
- [9] N. J. Matha, D. Kundur and A. Sheikholeslami, "Hardware Implementation Perspectives of Digital Video Watermarking Algorithms", *IEEE Transactions on Signal Processing*, vol. 51, no. 4, pp. 925-938, April 2003.