Data Hiding and Watermarking in JPEG Compressed Domain

by DC Coefficient Modification

Peter H. W. Wong*, Oscar C. Au**, Justy W. C. Wong***

Department of Electrical and Electronic Engineering
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong.

Email: eepeter@ee.ust.hk*, eeau@ee.ust.hk**, eejusty@ee.ust.hk***
Tel.: +852 2358-7053**

ABSTRACT

JPEG is a common image format in the WWW and can potentially be used to hide data for secure internet communication and watermark for copyright control. In this paper, we propose an algorithm to embed the secret or watermark information. The proposed algorithm, named Watermarking by DC Coefficients Modification (WDCM), assumes that the quality factor used in JPEG compression is known. We observe that it is perceptually undetectable if the DC coefficients in certain texture-rich blocks are modified by a small amount. We thus embed the secret information as a binary bit sequence in the quantized DC coefficients in those texture rich blocks. The watermark embedding process can be applied in compression domain without re-encode the data. The information bits are randomized by some pseudo-random noise (PN) sequences, the keys of which are needed for the decoding of the secret information. By embedding the information in the DC components, the proposed algorithm is robust to common JPEG compression if the quality factor is known.

1. INTRODUCTION

In recently years, many algorithms were proposed to embed robust watermark in digital images. Many of them focus on the robustness to common signal processing such as low pass filtering, rotation, scaling, cropping and compression [1-7] and some claimed that their algorithms are robust to JPEG compression[5-7]. However, those algorithms usually use normalized correlation as the measurement to detect the existence of the watermark information that is not suitable for data hiding. In [8], several data hiding techniques are proposed but they are not robust to JPEG compression. Embedding a bit sequence in the digital image is a difficult task since the bit sequence should be decoded correctly. In [5], an algorithm is proposed to embed a bit sequence in digital image by DCT coefficient removal but the modification of DCT coefficients in smooth regions may resulted in visual artifacts. In [2,6-7], spread spectrum technique is used to embed the watermark but the noise like watermark may be suppressed by JPEG compression significantly. In this paper, we propose a technique to embed a bit sequence in digital image. The bit sequence can be decoded in the compress domain with the secret key is given at the decoder. The secret key can be a public key or a private key.

2. WATERMARKING BY DC COEFFICIENTS MODIFICATION (WDCM)

Basically one bit of the secret or watermark information is embedded in an 8-by-8 block by modifying its quantized DC coefficients slightly. However, not all blocks are used. From our experiments, we observed that even a slight change in the quantized DC coefficients can cause visual artifacts in the smooth regions such as the shoulder of Lena shown in Fig. 1. However, in texture regions such as the hair and hat of Lena, a small change in the quantized DC coefficients can be perceptually undetectable. As a result, only the texture-rich blocks are used to embed the watermark information. To ensure the decoder selects the correct blocks to retrieve the watermark information, the quality factor should be given in the block selection process. We perform bandpass filtering on the JPEG decompressed image and examine the bandpassed blocks. Since only DC coefficients will be modified in the watermark embedding process, we used a 3-by-3 bandpass filter and the

boundaries of each block is neglected. If a bandpassed block is found to have significant energy, it is declared to be rich in texture and would be used to embed watermark information. This block information is not sent with the image because the same procedure can be performed in the decoder to find out the selected blocks. Note that our block selection algorithm is independent of the watermark being applied. It yields the same selected blocks regardless of the embedded watermark pattern. The procedures of the block selection process are as follows:

1. Compress and decompress the original image with the given quality factor. (If the watermark is to be embedded in JPEG compressed image, only decompression is needed.)
2. Apply bandpass filtering on the image. The boundaries of each block are treated as zero intensity.
3. Count the number of pixels that are greater than a threshold T_1 , if the number in a block is greater than T_2 , this block will be used to embed the watermark information.

Since the decompressed image is used and the bandpass filtering suppresses the DC energies, the decoder can extract the selected block without error if the AC coefficients remain unchanged. Here is how we embed one bit of secret or watermark information into the quantized DC coefficient of an 8x8 block. Let qDC_i be the quantized DC coefficient of block i , w_i be the 1-bit information to be embedded in block i , and n_i be the PN number which is uniformly distributed in $[R_1, R_2]$. The qDC_i is modified according to Eqn. 1 as follows:

$$qDC'_i = \begin{cases} \text{round}\left[n_i \cdot \text{round}\left(\frac{qDC_i}{n_i}\right)\right], & \text{if } \text{mod}\left[\text{round}\left(\frac{qDC_i}{n_i}\right), 2\right] = w_i \\ \text{round}\left\{n_i \cdot \left[\text{round}\left(\frac{qDC_i}{n_i}\right) + \beta n_i\right]\right\}, & \text{if } \text{mod}\left[\text{round}\left(\frac{qDC_i}{n_i}\right), 2\right] \neq w_i \end{cases} \quad (1)$$

The β is chosen to be either -1 or 1 such that the modification process minimizes the absolute difference between the original DC_i coefficient and the qDC'_i , as shown in Eqn. 2. Since the quantized DC coefficients can be easily extracted in the JPEG bits stream, there is no need to decompress and re-compress the image for the proposed algorithm. Moreover, the AC coefficients are unaffected. For a particular block, the amount of distortion of the quantized DC coefficient depends on the PN number n_i . The larger the n_i , the larger distortion is.

$$\beta = \begin{cases} 1, & \text{if } qDC_i \geq n_i \cdot \text{round}\left(\frac{qDC_i}{n_i}\right) \text{ or } \text{round}\left(\frac{qDC_i}{n_i}\right) = 0 \\ -1 & \text{if } qDC_i < n_i \cdot \text{round}\left(\frac{qDC_i}{n_i}\right) \end{cases} \quad (2)$$

In order to decode the watermark, the PN sequence is needed. The watermark w_i can be decoded according to the Eqn. 3:

$$w_i = \text{mod}\left[\text{round}\left(\frac{qDC'_i}{n_i}\right), 2\right] \quad (3)$$

3. SIMULATION AND RESULTS

In our simulation, the luminance component of two 512x512 testing images, 'Lena' and 'Pepper', were used. The original images were compressed to different qualities and then decompressed. The block selection based on the bandpass energy was carried out in the decompressed images. The watermark information was then embedded in the selected blocks according to the method described in section 2. The watermark information used was a random 1-bit sequence with the length depends on the number of selected blocks. It should be noted that the number of blocks used to embed the bits sequence depends on the nature of the image and the compression ratio. A PN sequence (n_i) which was uniformly distributed in $[3,4]$, was used to be the key for the decoding of the watermark information. T_1 and T_2 are chosen to be 8 and

18 respectively. The length of the PN sequence should be the same as the bit sequence. Several quality measures were recorded and listed in Table 1 and 2. The watermarked images were compressed using the JPEG compression function in MATLAB, with different quality factor corresponding to different quantization factor (Quant. factor) used in the quantization process of DC coefficients. The length of the bits sequence (no. of bits emb.) range from 879 to 1183 for Lena and from 665 to 1036 for 'pepper' respectively. From the tables, it can be observed that the bit per pixel (bpp) basically does not change after the watermarking information was embedded in the images. Fig. 2 and 3 shows the original 'Lena' and 'Pepper' respectively. Fig. 4 and 5 show the watermarked and decompressed 'Lena' with the quality factor of 60 and 30 corresponding to 0.74 bpp and 0.46 respectively. Fig. 6 and 7 show the watermarked and decompressed 'Pepper' with the same parameter of the images shown in Fig. 3 and 4. Fig. 8 and 9 show the 1-bit maps for 'Lena' and 'Pepper' respectively. The white blocks in the -but map indicate which blocks are selected to store the water information

4. CONCLUSION

In this paper, we proposed to modify the quantized DC coefficients in order to embed secret or watermark information. A pseudo noise (PN) sequence is used as the key for the decoding of the embedded information. Without the key, it is almost impossible to decode the watermark information. The embedded watermark is invisible because only blocks with rich texture information are used to embed the watermark.

5. ACKNOWLEDGEMENT

This work is supported in part by an HKTIIT grant.

REFERENCE

- [1] Mitchell D. Swanson, Bin Zhu and Ahmed H. Tewfik, "Multiresolution Scene-Based Video Watermarking Using Perceptual Models," *IEEE Journal on Selected Areas in Communication*, vol. 16, no. 4, pp. 540-550, May 1998.
- [2] Ingemar J. Cox, Joe Kilian, Tom Leighton and Talal Sharnoon, "Secure Spread Spectrum Watermarking for Multimedia," *NEC Research Institute, Technical Report 95-10*.
- [3] Xiang-Gen Xia, Charles G. Boncelet and Gonzalo R. Arce, "A Multiresolution Watermark for Digital Images," *Proc. of IEEE Int. Conf. Image Processing*, vol. 1, pp. 548-511, Oct. 97.
- [4] Joseph J.K. O Ruanidh and Thierry Pun, "Rotation, Scale and Translation Invariant Digital Image Watermarking," *Proc. of IEEE Int. Conf. Image Processing*, vol. 1, pp. 536-538, Oct. 97.
- [5] Gerrit C. Langelaar, Reginald L. Lagendijk and Jan Biemond, "Watermarking by DCT Coefficient Removal: A Statistical Approach to Optimal Parameter Settings," *Proc. of SPIE Sym. of Security and Watermarking of Multimedia Contents*, pp. 2-13, 1999.
- [6] F. Hartung, B. Girod, "Fast Public-Key Water-marking of Compressed Video", *Proc. of IEEE Int. Conf. Image Processing*, vol. 1, pp-528-531, Oct. 97.
- [7] P. H. W. Wong, O. C. Au and J. W. C. Wong, "Image Watermarking Using Spread Spectrum Technique in Log-2-Spatio Domain", to appear in *Proc. of 2000 IEEE Int. Sym. on Circuits & Systems (ISCAS)*.
- [8] W. Bender, D. Gruhl, N. Morimoto and A. Lu, "Techniques for Data Hiding," *IBM System Journal*, vol. 35, NOS 3&4, pp. 313-336, 1996.

Quality Factor	Quant. Factor	No.of bits emb.	PSNR w/o JPEG	PSNR w/ JPEG	Bpp. w/o water	Bpp. w/ water
100	1	1183	65.41	58.26	4.95	4.95
90	2	1231	59.16	40.72	1.81	1.81
80	4	1149	53.41	38.38	1.16	1.16
70	7	1104	48.77	37.01	0.89	0.89
60	9	1060	46.58	36.06	0.74	0.74
50	11	1015	45.08	35.31	0.64	0.64
40	16	1001	43.04	34.47	0.55	0.55
30	18	942	41.12	33.48	0.46	0.46
20	28	879	37.82	31.76	0.36	0.36

Table 1: Quality measures for 'Lena'

Quality Factor	Quant. Factor	No.of bits emb.	PSNR w/o JPEG	PSNR w/ JPEG	Bpp. w/o water	Bpp. w/ water
100	1	1036	65.96	58.31	5.30	5.30
90	2	1191	59.07	38.78	2.03	2.03
80	4	928	54.29	36.61	1.21	1.21
70	7	836	50.15	35.62	0.92	0.92
60	9	772	48.09	34.83	0.76	0.76
50	11	787	46.08	34.12	0.65	0.65
40	16	753	44.74	33.05	0.55	0.56
30	18	714	42.41	31.89	0.46	0.47
20	28	665	39.15	28.54	0.36	0.37

Table 1: Quality measures for 'Pepper'

Figure 2: Original 'Lena'

Figure 3: Original 'Pepper'

Figure 4: Watermarked and JPEG compressed 'Lena' with quality factor = 60. PSNR = 36.06dB, 1060 bits are embedded in the image. 0.74bpp.

Figure 5: Watermarked and JPEG compressed 'Lena' with the quality factor = 30. PSNR = 33.48dB, 942 bits are embedded in the image. 0.46bpp.

Figure 6: Watermarked and JPEG compressed 'Pepper' with the quality factor = 60. PSNR = 34.83dB, 772 bits are embedded in the image. 0.76bpp.

Figure 7: Watermarked and JPEG compressed 'Pepper' with the quality factor = 30. PSNR = 31.89dB, 714 bits are embedded in the image. 0.76bpp.

Figure 8: 1-bit map for 'Lena', quality factor is 60, 1160 blocks are selected to store the watermark information.

Figure 9: 1-bit map for 'Pepper', quality factor is 60, 772 blocks are selected to store the watermark information.