

CONTENT BASED COLOR IMAGE ADAPTIVE WATERMARKING SCHEME

Huajian Liu, Xiangwei Kong, Xiangdong Kong, Yu Liu

Department of Electronic Engineering,
Dalian University of Technology,
Dalian 116023, China

ABSTRACT

A color image adaptive watermarking scheme is proposed. To increase the robustness and perceptual invisibility, a model called ICPM (image complexity and perceptibility model) is proposed to analyze the different sub-blocks' properties. A novel technique is proposed for the watermark casting and retrieval by utilizing the properties of the image itself. Experimental results demonstrate the robustness of the algorithm to many attacks, such as JPEG, JPEG2000, and A/D conversion.

1. INTRODUCTION

With an explosive growth in digital imaging technology and applications, digital images and video are now widely distributed on the Internet and via CD-ROM. The copyright protection of digital images and video has a great need against pirating. As a solution to this problem, a digital watermark technology is now drawing the attention as a new method of protecting copyright for digital data.

In general, a digital watermark technique must satisfy the following two properties.

1. The embedded watermark does not distort visually the image and should be perceptually invisible.
2. The watermark must be difficult to remove. It should also be robust to common signal processing and geometric distortion, such as A/D and D/A conversion, compression and rescaling.

In recent years, different schemes have been proposed. W. Bender[1] proposes a watermarking method called 'Patchwork'. It encodes a watermark by modifying a statistical property of the image. Kutter et. al [2] propose a new method based on amplitude modulation. Single watermark bit is multiply embedded by modifying pixel values in the blue channel, since the human eye is less sensitive to changes in this band. These modifications are proportional to the luminance. Cox et. al [3] propose a DCT based spread spectrum watermarking technique. A pseudo-random sequence is embedded into the significant DCT coefficients and is retrieved by calculating the similarity function of the original watermark and extracted watermark. H.J. Mike Wang [4] proposes a wavelet-based watermark algorithm. Based on the principle of multi-threshold wavelet codec (MTWC), the method searches the significant wavelet coefficients to embed the watermark in order to increase the robustness. The embedding strength in each subband is determined by the threshold of the subband. Polilchuk C. I. and Zeng W. [5] propose two kinds of adaptive watermarking

methods. One is based on discrete cosine transform (IA-DCT), the other is based on discrete wavelet transform (IA-W). The watermark is embedded according to the JND threshold. N. Kaewkammerd et. al [6] propose a wavelet based adaptive watermarking scheme. The human visual system (HVS) is employed to determine the weighting function $T(x,y)$ to control the watermark casting process.

These different watermarking schemes usually use the pixels or transform coefficients to embed the watermark information in the whole image [1-6]. But when the watermarked image is attacked and the pixel values or coefficients have relatively large changes, the methods will fail to retrieve the watermark. And most of such schemes take the whole image as an interesting region, but in fact the different regions of an image have different perceptual invisibility and robustness for watermarking. In this paper, we propose a novel watermarking method for color images, which takes into account the different properties of different parts of the image. The proposed method inserts a watermark not into a whole image region but only into interesting regions. To extract the interesting regions, we propose an image complexity and perceptibility model (ICPM) to analyze each sub-block's properties. Also the proposed scheme is not only to change individual coefficient's value to embed each watermark bit as usual methods do [1-4], but to change the coefficients' statistical distribution of some regions to get more robust performance. The experimental results demonstrate the proposed algorithm is not only robust to the general signal processing attacks but also robust to JPEG2000 compression and A/D conversion such as print and scan processing.

2. THE DIAGRAMS OF WATERMARK CASTING AND RETRIEVAL

The diagrams of watermark casting and retrieval are shown in Fig.1 and Fig.2.

3. INTERESTING REGIONS SEARCHING

The proposed methods in recent years embed the watermark into the whole image or the sub-blocks according to the human visual system. But the maximal embedding strength in one sub-block may not be the maximal strength in the whole image, and even may be the relatively small one. Therefore, the perceptual invisibility and robustness can not be satisfied simultaneously.

In an image, the different parts have the various perceptual invisibility and robustness for the watermark embedding. To

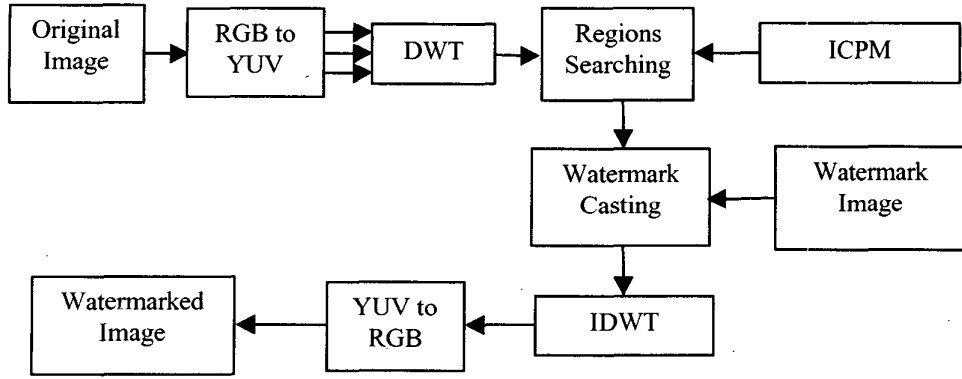


Figure 1. The block diagram of the watermark casting

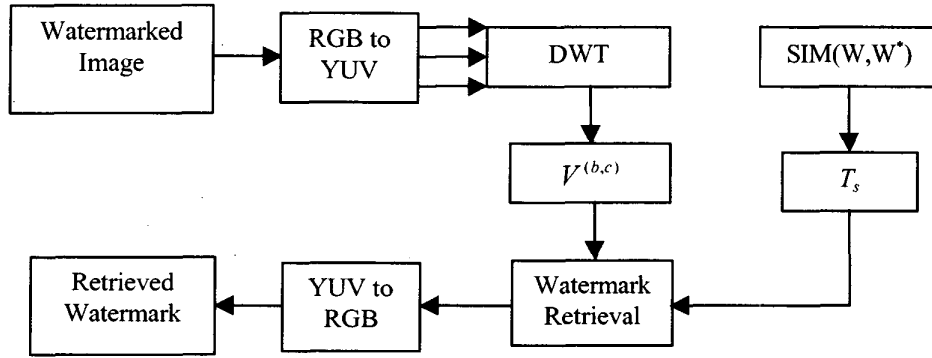


Figure 2. The block diagram of watermark retrieval

obtain the better performance, we analyze the different parts' properties of an image and find the interesting regions for watermarking.

Before the region searching and watermark casting, the color images with RGB channels are converted into YUV channels by using the CCIR 601 standard. Then the original image is divided into $n \times n$ blocks and a proper r -level wavelet transform is applied to Y, U and V channels separately in every block. To search the interesting regions, an image complexity and perceptibility model (ICPM) is defined as follows:

$$M^{(b,c)} = \frac{m^{(b,c)} \sum_{(x,y) \in S_i} |E(x,y)|}{n^{(b,c)} \sum_{i=0}^{n/2^r} \sum_{j=0}^{n/2^r} \|E(x+i, y+j) - |E(x,y)|\|}$$

where $E(x,y)$ is the coefficient value of the decomposition at position (x,y) in block b and channel c . $m^{(b,c)}$ and $n^{(b,c)}$ are the scaling factors. S_D and S_i are the non-overlapped subsets of subbands. S_D is usually selected in low frequency subbands, and S_i is selected in high frequency subbands. According to the $M^{(b,c)}$ value, we can choose N_w blocks as interesting regions.

Some assumptions are made for the proposed scheme: The attacker does not have the original image as well as source codes of watermark casting and retrieval. The information available to the public is only the protected image. Even though the attacker knows that the wavelet transform is used and understands the algorithm of ICPM, the following information are kept confidential to protect the location of interesting regions:

1. Wavelet transform structure and the number of the transform level r ;
2. Wavelet filter (e.g. Harr, Daubechies, etc.)

By different wavelet structure and filter, our scheme will embed the watermark into different interesting regions. Therefore, to keep these information confidential can avoid attackers from using inverse watermarking operations to remove the watermark.

4. WATERMARK CASTING SCHEME

The watermark, $W(i,j)$, is a binary bitmap or logo which can be either the private fingerprint or the company trademark to claim the ownership. In most of proposed watermarking methods, the watermark bit is usually embedded into the individual pixel or coefficient. However, when the watermarked image is attacked, the individual pixel's value or coefficient maybe changes a lot

and the watermark can not be successfully retrieved. If the watermark is presented by the image oneself information, the better robustness can be obtained. Simultaneously, the watermark's perceptual invisibility can be satisfied better.

The watermark casting process is performed as follows.

If $W(i,j)=0$,

$$E_{S_A}^{(b,c)}(x,y) = \text{sign} \times C^{(b,c)} \times A^{(b,c)}(x,y) \in S_A$$

where sign is the sign value (-1 for negative and 1 for positive) of $E_{S_A}^{(b,c)}(x,y)$. $C^{(b,c)}$ is the embedding strength factor. And $A^{(b,c)}$ is defined as

$$A^{(b,c)} = \frac{1}{M_s} \sum_{(x,y) \in S_A} |E^{(b,c)}(x,y)|$$

where M_s is the number of the coefficients in S_A subsets.

If $W(i,j)=1$,

$$E_{S_A}^{(b,c)}(x,y) = \text{sign} \times C^{(b,c)} \times Z^{(b,c)}(x,y)$$

where $Z^{(b,c)}$ is the constant that is determined by coefficients in S_A subsets of block b and channel c . In our experiments, it is set to

$$Z^{(b,c)}(x,y) = \begin{cases} \text{MAX}^{(b,c)}, & |E_{S_A}^{(b,c)}(x,y)| > A^{(b,c)} \\ \text{MIN}^{(b,c)}, & \text{else} \end{cases}$$

where $\text{MAX}^{(b,c)}$ and $\text{MIN}^{(b,c)}$ is the maximum and minimum of the absolute values of coefficients in S_A subsets of block b and channel c .

To avoid severe visual quality degradation in the watermarking casting process, the Y, U and V channels must be carefully selected according to the human visual system.

5. WATERMARK RETRIEVAL SCHEME

Unlike the usual methods, in the retrieval process we do not only consider the individual coefficient, but take into account the difference of the individual coefficient and the property of whole region. By utilizing the correlation of coefficients, we obtain the better robustness.

The watermark retrieval process is performed as follows.

To retrieve the watermark, the variance of the coefficients of S_A subsets in every selected block is calculated:

$$V^{(b,c)} = \frac{1}{|S|} \sum_{(x,y) \in S} \left| E^*(b,c)(x,y) - A^*(b,c) \right|^2$$

where E^* and A^* are both obtained from the attacked watermark image.

Then the watermark can be decoded via

$$W^*(i,j) = \begin{cases} 0, & V^{(b,c)} < T_s \\ 1, & V^{(b,c)} > T_s \end{cases},$$

where W^* is the retrieved watermark, and T_s is a threshold, which can be automatically adjusted by calculating the similarity of the original watermark and the retrieved watermark.

6. EXPERIMENTAL RESULTS

In the experiment, we evaluated the proposed watermarking scheme by testing with several 512x512 still color images. We choose a 40x20 bitmap as the watermark. After the embedding process, the PSNR of watermarked image without any attack is above 40 dB and the watermark can be correctly retrieved. As shown in Fig. 1, there is no visible degradation in the watermarked image, i.e., the watermark is completely perceptual invisible.

Noise is one of common distorts in the image processing and transmission. In the experiment, we add 10% uniform noise into the watermarked image and the watermark can still be retrieved successfully as shown in Fig.3(a). The error ratio is only 1.0%.

JPEG is a widely used compression format and the default quality factor is 75%. In the case of JPEG compression with quality factor 65% and 50%, the watermark can survive successfully as shown in Fig.3(b-c) and the error ratio is 0.75% and 3.125% respectively.

JPEG2000 is the new generation compression standard, which is based on wavelet transform. In our experiments, we test the watermarked image with JPEG2000 compression using LuraWave SmartCompress. When the quality is 95 and 90, the reconstructed watermarks are shown in Fig.3(d-e) with the error ratio only 0.25% and 3.5%.

Rescaling is very easy to perform in the edition of digital images. So the watermarking technique must be robust to the attacks of rescaling. We test our algorithm in the case of rescaling the watermarked image by 0.5x0.5, 0.5x0.6, 2x2 and 2x3 respectively using StirMark 3.1. The experiment results show the watermark can still be retrieved as shown in Fig.3 (f-i) with the error ratio only 2.875%, 2.875%, 0.25% and 0.375%.

A/D and D/A conversion is a very serious attack to the watermarked image. After A/D conversion, the image's fidelity decreases and some usual methods fail to retrieve the watermark. In our experiments, the watermarked image was printed with an Epson Stylus Photo 710 color printer, with a resolution of 360dpi on standard paper. The image was then scanned using an Epson Perfection 1200 Photo scanner with a resolution of 300dpi. The experimental results demonstrate our algorithm is very robust to the A/D conversion and the error ratio is only 2.875%. The reconstructed watermark after print and scan processing is shown in Fig.3(j). To confirm the performance of the proposed scheme, we compare our result with that obtained using the algorithm proposed in [4]. The result obtained using algorithm proposed in [4] has a error ratio of 19.3% after the same print and scan processing.

Some test results are shown in Fig.4 and Table 1.

7. CONCLUSION

We propose a novel color image adaptive watermarking algorithm. By using a model called ICPM, the proposed scheme selects the interesting regions instead of the whole image to embed the watermark to make it completely perceptual invisible and more robust to attacks. A novel technique for watermark casting and retrieval is also proposed. The experimental results demonstrate the proposed algorithm is robust to many attacks, such as noise, JPEG, JPEG2000, rescaling and the A/D and D/A conversion.

8. REFERENCE

- [1] W. Bender, D. Gruhl, 'Technique for Data Hiding', MIT Media Lab, Cambridge MA Tech. Rep, 1994
- [2] Kutter, F. Jordan, and F. Bossen, 'Digital signature of color images using amplitude modulation', in Storage and Retrieval for Image and Video Databases V, SPIE Vol. 3022, pp. 518-526
- [3] COX,I.J. , KILIAN,J. , LEIGHTON,T. and SHAMOON,T. 'Secure spread spectrum watermarking for multimedia', *IEEE Trans. on Image Processing*, 1997, 6(12), pp. 1673-1687
- [4] SU,P. , KUO,C.J. , and WANG,H.M. 'Blind digital watermarking for cartoon and map images', IS&T/SPIE Conference on Security and Watermarking of Multimedia Contents, San Jose, California, January 1999, pp. 296-305
- [5] PODICHUK, C.I., and ZENG, W. 'Image adaptive watermarking using visual models', *IEEE J. Sel. Areas Commun.*, 1998,16, pp. 525-538
- [6] N. Kaewkamnerd, K.R. Rao, 'Wavelet based image adaptive watermarking scheme', *Electronics Letters*, 2000,2, vol. 36, pp. 312-313

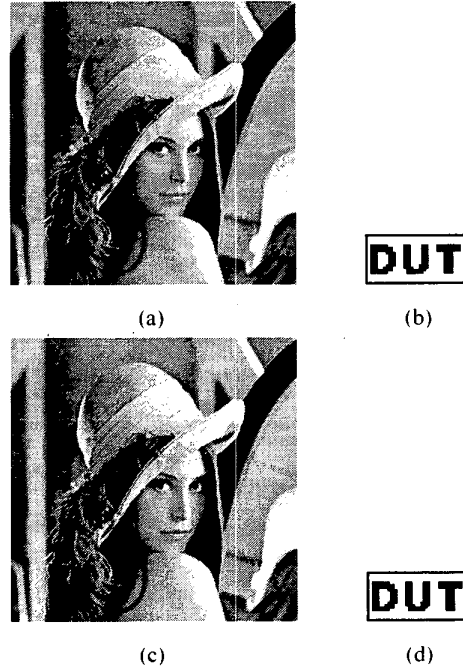


Figure 3. (a)original color Lena image; (b)original watermark; (c)watermarked image(PSNR=40.44; (d)retrieved watermark (no error).

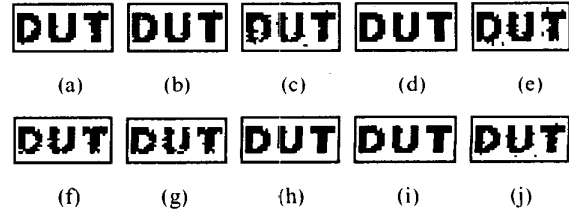


Figure 4. Retrieved watermarks after attacks (a)10% uniform noise; (b)JPEG(65%); (c)JPEG(50%); (d)JPEG2000(95%); (e)JPEG2000(90%); (f)0.5x0.5 rescaling; (g)0.5x0.6 rescaling; (h)2x2 rescaling; (i)2x3 rescaling; (j)print and scan.

Table 1 Experimental results after attacks

Attacks	10% uniform noise	JPEG (65%)	JPEG (50%)	JPEG 2000 (95%)	JPEG 2000 (90%)
Error ratio(%)	1.000	0.750	3.125	0.250	3.500
Attacks	Rescaling (0.5x0.5)	Rescaling (0.5x0.6)	Rescaling (2x2)	Rescaling (2x3)	Print and scan
Error ratio(%)	2.875	2.875	0.250	0.375	2.875