RELIABLE IMAGE/VIDEO WATERMARK RETRIEVAL IN THE PRESENCE OF LOSSY COMPRESSION

Qing Chen and Henri Maître

GET/Telecom Paris, LTCI UMR 5141 - CNRS
46 rue Barrault, 75013, Paris, France
phone: + 33 (0)1 45 81 76 55, fax: + (33) 1 45 81 37 94, email: chenq@iupui.edu, henri.maitre@telecom-paris.fr

ABSTRACT

Lossy compression is inherently contradictory with watermarking and makes the watermark retrieval unreliable. In this paper, we propose a watermarking technique capable to correctly retrieve watermark in both image and video, where watermarking is combined with JPEG/MPEG compression standards to adaptively obtain a proper trade-off among imperceptibility, achievable embedding capacity and robustness to lossy compression. Experimental evaluation demonstrates that algorithm achieve imperceptible watermark with error-free extraction performance and sufficient embedding capacity in the presence of lossy compression.

1. INTRODUCTION

In order to make a watermark-based application feasible, one of the fundamental issues is to find the proper trade-off among three important requirements: imperceptibility, capacity, and robustness. For most data hiding applications, a certain degree of compression is desired. However, the objective of lossy compression is fundamentally in conflict with watermarking. The retrieval of watermarks embedded in digital media is often affected if the marked contents undergo compression. The widespread use of hybrid DCT-based compression standards makes JPEG/MPEG-resisting watermarking an important aspect for algorithm design.

A great deal of research has proposed a variety of embedding strategies to trade-off those conflicting requirements. The algorithms operating on transform domains (e.g. DCT domain and wavelet domain) usually choose to insert watermark bits in the DCT coefficients at the middle frequencies (for block DCT-based algorithms) and medium frequency subbands with large level wavelet decomposition or coarsest subband with small level wavelet decomposition (for wavelet domain techniques) in order to have a good compromise between robustness and transparency. Others embed information by scaled addition according to the signal strength of the particular frequency components to provide a trade-off between imperceptibility and robustness [1]. Human visual system (HVS) models have also been incorporated to balance these opposing requirements [2][3]. The use of error-correcting codes (ECC) [4] either degrades the quality of marked signal or decreases the quantity of embedded data to achieve robustness. An information theoretic analysis of watermarking is presented in [5] for capacity problem. The framework shows the trade-off between achievable information hiding rates and allowed distortions. Overall, although the aforementioned papers have dealt with problems related to watermarking trade-off, some issues still remain. In addition to robustness and imperceptibility, there is a need for more comprehensive work. In some situations, Error-free watermark retrieval is critical to exploit the payload for further processing in the workflow. Thus, appropriate hiding capacity and accurate watermark retrieval are greatly desirable. These two requirements along with imperceptibility and robustness to compression are inter-dependent, mutually competitive, and need to be adapted dynamically according to media content and usage environment**.** This brings forward the important issue of adaptive trade-off among those conflicting features.

Our work is concentrated on the issue of error-free watermark retrieval in the presence of lossy compression. We are interested in encoder that adaptively adjusts data embedding to create a reliable watermark which tolerates lossy compression to a required level. Algorithm achieves lossless watermark retrieval in both image and video, where watermarking is integrated with JPEG/MPEG compression standards to adaptively obtain a proper trade-off among imperceptibility, achievable embedding capacity and robustness to block DCT-based lossy compression. By exploiting an HVS model, scheme estimates the image-/video-dependent embedding capacity resistant to the desired compression level in order to allocate information bits in a determinate way for consistent reliable retrieval performance.

2. THE PROPOSED ALGORITHM

The proposed strategy consists of A)To ensure the invisibility, we use a perceptual model, which provides an upper bound on the amount of available imperceptible modification even after subsequent compression; B)The watermark robustness to compression can be improved by using some kind of redundancy. Repetition, a simple and effective ECC resistant to fading-like attacks, is adjusted by quantization parameter to make the watermark survive a desirable quality level of JPEG/MPEG compression; C)To achieve reliable watermark recovery, the hidden watermark bits should be lower than the algorithm capacity. The estimate of embedding capacity is compromisingly attained under the constraint of an HVS model and at a desirable compression level, which is application-dependent and determined by JPEG/MPEG quantization parameter. Raise/lower the repetition factor will strengthen/weaken watermarking robustness and decrease/increase capacity.

2.1 Watermark insertion

Watson's HVS model describes JND (Just Noticeable Difference) based on three main concepts: frequency sensitivity, luminance sensitivity and contrast masking. The luminance sensitivity is:

$$t_k(i,j) = t(i,j)\left(\frac{c_k(0,0)}{\bar{c}(0,0)}\right)^{a_T} \qquad (1)$$

Where $t(i,j)$ is an original JND describing frequency sensitivity of HVS. Generally, JPEG quantization table, which can be scaled for a required level of original JND, is used as the frequency sensitivity. The $c_k(0,0)$ is the DC value of the block k, $\bar{c}(0,0)$ is the DC coefficient corresponding to the mean luminance of the display, and a_T, setting to 0.649, is a parameter that controls the degree of luminance sensitivity. Then the value of luminance sensitivity is adjusted by contrast masking:

$$m_k(i,j) = \max(t_k(i,j), |c_k(i,j)|^{w_{ij}} t_k(i,j)^{1 \;\; w_{ij}}) \qquad (2)$$

Where $c_k(i,j)$ is the (i,j)-*th* DCT coefficient of the block k. A typical empirical value for w_{ij} is 0.7.

The JND thresholds provide the location and maximum amplitudes of modifying signal that can be tolerated in every coefficient without causing the perceptual degradation. This allows us to insert the transparent watermark of maximal strength and to approach capacity that in turn, improves robustness to common image processing operations such as JPEG/MPEG compression.

Lin *et al.* [5] derived the private watermarking capacity by considering every coefficient as an independent random variable with its own noise distribution:

$$Cap_{total} = \sum_{i=1}^{n} Cap_i = \sum_{i=1}^{n} \frac{1}{2}\log_2(1+\frac{P_i}{N_i}) \qquad (3)$$

Where P_i and N_i are the power constraints of watermark and noise of the *i-th* coefficient, respectively. The n is the number of watermarked coefficients.

In the presence of JPEG compression, the major error source of watermark retrieval is due to the quantization of DCT coefficients. Denote C_i the *i-th* zigzag-scanned DCT coefficient in an 8x8 block, Q the quality factor of JPEG. When both embedding and perceptual coding are applied to image, the combined effect on any coefficient should not exceed JND threshold for watermark transparency. The gap between JND and the quantization error (determined by Q) is available room for watermarking. So, the maximal watermark strength of the *i-th* coefficient is

$$W_i = JND_i \quad E_i. \qquad (4)$$

The quantization error E_i which shows how much modification of DCT coefficient caused by the compression scheme is determined by:

$$E_i = C_i - Round(C_i / (Q \quad M_i)) \quad (Q \quad M_i) \qquad (5a)$$

Where, M_i is *i-th* value of quantizer matrix. Analogously, the quantization error of MPEG compression is:

$$E_i = C_i - Round(C_i / (q\_scale \quad M_i)) \quad (q\_scale \quad M_i) \qquad (5b)$$

Here, *q_scale* is MPEG quantizer scale. So the power constraints P_i of watermark in (3) is determined by:

$$P_i = W_i^2 = (JND_i \quad E_i)^2 \qquad (6)$$

In the case of blind watermarking, the host data is not available for extraction and should be considered as a source of noise. The noise N_i in (3) is

$$N_i = C_i^2 + E_i^2 \qquad (7)$$

To achieve reliability, we use repetition coding. The price paid for robustness is a reduction of the embedding data rate. For our blind watermarking system, every coefficient pair is used to embed one bit watermark. So repetition factor and 2 should divide Cap_{total} to obtain the embedding capacity of robust and blind watermarking system.

Watermark embedding is done by modifying a selected set of DCT coefficients at low to medium frequencies of luminance component of image or I-frame of video sequence. The practically used coefficients depend on the payload length and the energy profile of the candidates. Any candidate coefficient whose absolute magnitude is less than a threshold T, an empirical threshold for perceptual significance, is eliminated.

Before modifying the coefficients, we first compute the JND in block DCT domain and estimate embedding capacity C by exploiting Watson's model. The embedding bits should be a little less than C in order to ensure error-free retrieval. Then repetition factor *rept* is decided according to *Q/q_scale*. Every bit of input binary information a_j (j=1,...,N) will be repeated *rept* times and transformed to the watermark w_i $\{ \;\; 1\}$ ($i=j+N \;\; r$, r=0,1,...,*rept* 1). Allocate all the watermark bits equally to every DCT block.

In each DCT block, combine every two adjacent selected coefficients, in zigzag scanning order, into a pair (C_1, C_2). A pair with remarkable difference between its two members is unembeddable pair. Each coefficient pair is embedded 1-*bit* watermark w_i by modifying C_1, C_2 in the following way in order to gain the desired relations ($C_1' > C_2'$ *for* $w_i = 1$ or $C_1' < C_2'$ *for* $w_i = -1$):

$C_1' = C_1 \quad JND_1$ and $C_2' = C_2 \mp JND_2$ if $w_i = \quad 1$.

Obviously, there is no guarantee that the relations always hold. In order to achieve error-free watermark retrieval even after subsequent JPEG/MPEG compression, we adjust repetition factor accordingly to compensate the errors caused by the violations of the relations and the impact of compression in order to achieve trade-off among the abovementioned requirements. The repetition factor should be increased when there will be greater distortion in subsequent JPEG/MPEG compression (smaller *Q/q_scale*), and decreased when the compression distortion will be less (larger *Q/q_scale*). A watermark formed in this way is resistant to JPEG/MPEG compression using a *Q'/q_scale'* equal to or greater than the *Q/q_scale* used to embed the watermark.

2.2 Watermark extraction

The watermark retrieval doesn't use the original data. It is performed in the block DCT domain, like the embedding. In each 8 8 DCT block, combine two adjacent coeffi-

cients, in zigzag scan order, into one pair (C_1', C_2') to extract one watermark bit w_i by comparing the relation of C_1', C_2': $w_i = -1$ when $C_1' < C_2'$ and $w_i = +1$ when $C_1' > C_2'$.
Recover the input information a_j:

$$a_j = 1, \text{ if } \sum_i w_i > 0 \text{ and } a_j = 0, \text{ if } \sum_i w_i < 0$$

$$i = j + N \quad r; r = 0,...,rept \;\; 1$$

3. EXPERIMENTAL RESULTS

Our watermarking scheme has been implemented and tested on still images and video. We perform different tests in order to check transparency, reliability and capacity under different compression levels. For image watermarking, our results are given in two aspects: highest hiding data rate and fixed length of watermark on 256 256 grey images.

3.1 Highest hiding data rate for still images

Tests have been performed on "Flower" image. Figure 1 shows the original and watermarked images of "Flower" where the total bits embedded in the image are equal to the estimated capacity for different JPEG quality factors Q=100, 75, 50, and 25. The corresponding numbers of embedded data (equal to estimated data capacities) are 478, 102, 59, and 40 bits, respectively. Then the watermarked images are performed JPEG compression with those Qs used in embedding. The experimental results show the watermarks are below perceptual detection. The bit error rates (BER) of watermark retrieval are 0.21% (1 bit), 0.98% (1 bit), 0, and 0, respectively in these four cases. For reliable recovery, the embedding bits must be lower than the estimated capacity.

3.2 Fixed length of watermark (64 bits) for still images

We choose 64 bits to watermark the test images showed in Figure2. For all test images, we calculate the capacity C under different JPEG quality factors (from 90 down to 20). The image quality (measured as PSNR) and BER are calculated after watermarked image undergoes JPEG compression to the set level. The results are given in Table 1. Successful watermark retrieval is obtained when the number of bits embedded in an original image is less than the calculated capacity. Otherwise the BER may increase accordingly with the increase of the difference between estimated capacity and the number of practical embedded bits.

3.3 Video watermarking

For video watermarking, the implementation of video watermarking uses the software codec of MPEG Software Simulation Group as fundamental components to reduce overall video processing complexity. Watermark insertion/extraction incorporated with MPEG-2 encoder/decoder can run in real time. In most cases, the estimated capacities of single I-frame of TV CIF resolution are about hundreds of bits for accurate retrieval at high bit-rates and 64 up to 125 bits for low bit-rates. PSNR is above 40 dB down to around 35 dB for high bit-rates down to low bit-rates. Then the watermark data rate of error-free retrieval from moderate to high MPEG compression ratio may reach about 120 bit/s and up (for MPEG standard, there are at least 2 I-frames per second in video stream). No watermark errors occur when the watermark data rate is below the estimated capacity. However, some errors happen when payloads are higher than the estimated capacity.

We illustrate our watermarking scheme on the Susie video sequence. The spatial resolution of the sequences is 352x240 pixels (CIF resolution). Sequence is encoded by MPEG-2 using the 12-frame group of pictures (GOP) structure IBBPBBPBBPBB with a frame rate of 25fps. The proposed algorithm achieves 64 bits/I-frame rate for exact watermark retrieval at bit-rate as low as 1.152 Mbps. The comparison between the original and watermarked I-frames of Susie video sequence is showed in Figure 3. The frame has 78 bits of estimated capacity, and is embedded with 64-bit information. The PSNR is 41.25. The perceptual quality of watermarked video is almost the same quality as standard MPEG-2 encoder/decoder video without watermark. The embedded watermark results in slight increase of video file size (about 2%).

4. CONCLUSION

We have proposed an efficient compression-watermarking conjunction scheme which takes into account the effects of compression and HVS when embedding the watermark. Based on simulation results we draw the following conclusions. The algorithm achieves satisfactory adaptive trade-off among watermarking constraints: transparency, hiding capacity and robustness to JPEG/MPEG-2 compression with error-free watermark retrieval performance. Due to the repetition ECC, the watermark should have robustness to other fading-like attacks. For video watermarking, the watermark is only embedded in I-frame, so it is not sensitive to re-encoding with different GOP structure, P/B frame dropping and swapping. Watermarking in the DCT domain leads to better implementation compatibility with popular compression standards. The principle can also be applied to other hybrid DCT coding schemes. Furthermore, it provides a solution for watermark bit allocation among different media in a determinate way when multimedia incorporated watermarking is desirable.

REFERENCES

[1] I. J. Cox, J. Kilian, and T. Leighton, T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Proc.,* vol. 6, no. 12, pp. 1673-1687, Dec. 1997.

[2] C. I. Podilchuk and W. Zeng, "Image-adaptive watermarking using visual models," *IEEE Journal on Selected areas in Communications*, vol. 16, no. 4, pp. 525-539, May 1998.

[3] D. Kundur, "water-filling for watermarking?" in *Proc. IEEE ICME2000*, New York City, New York, pp.1287-1290, August 2000.

[4] J.R. Hernandez, F. Perez-Gonzales and J. M. Rodriguez "The impact of channel coding on the performance of spatial watermarking for copyright protection," in *Proc. IEEE ICASSP 98*, vol.5 May 1998, pp. 2973-2976

[5] P. Moulin and J. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Trans. Information Theory,* vol. 49, No. 3, pp 563-593, March 2003.

[6] C. Lin and S. Chang, "Watermark capacity of digital image based on domain-specific masking effects," in *Proc. IEEE Int. Conference on Information Technology: Coding and Computing*, April 2001, pp. 90-94.

(a) (b) (c) (d) (e)

Figure1: Original “Flower” image (a) and four watermarked “Flower” images that are embedded information bits equal to estimated data capacities corresponding to different JPEG compression *Q*s. (b): *Q*=100, (c): *Q*=75, (d): *Q*=50, (e): *Q*=25.

(a) Tank (b) Pentl (c) Lena (d) Elaine

Figure 2: Four test images for fixed length of watermark

Image	Quality Factor=90			Quality Factor=80			Quality Factor=70			Quality Factor=60		
	C(bits)	BER%	PSNR	C(bits)	BER%	PSNR	C(bits)	BER%	PSNR	C(bits)	BER%	PSNR
Tank	230	0	39.33	142	0	37.08	105	0	34.77	85	0	33.14
Pentl	226	0	37.57	139	0	35.73	103	0	33.88	83	0	32.40
Lena	192	0	37.92	120	0	36.62	90	0	34.68	72	0	33.19
Elaine	182	0	40.19	116	0	38.80	86	0	36.99	69	0	35.41

Image	Quality Factor=50			Quality Factor=40			Quality Factor=30			Quality Factor=20		
	C(bits)	BER%	PSNR	C(bits)	BER%	PSNR	C(bits)	BER%	PSNR	C(bits)	BER%	PSNR
Tank	71	0	31.47	61	0	30.07	53	1.56	28.97	48	15.62	27.85
Pentl	69	0	30.89	59	0	29.46	52	4.69	28.37	46	10.94	27.35
Lena	60	1.56	31.93	51	1.56	30.61	45	3.12	29.59	40	7.81	27.57
Elaine	58	0	34.31	50	4.69	33.03	44	17.19	31.93	39	15.62	31.09

Table 1. Estimated hiding capacity (C) of four images corresponding to different JPEG quality factors, and PSNR and bit error rate (BER%) of watermark extraction when they are embedded a 64-bit watermark and then compressed with the corresponding quality factors.

(a) (b) (c)

Figure 3. Comparison between the original and watermarked frames: (a) Original I-frame extracted from the Susie video sequence; (b) Corresponding watermarked frame; (c)Amplified difference between (a) and (b).