Robustness Enhancement of Content-Based Watermarks Using Entropy Masking Effect

Amir Houman Sadr[1] and Shahrokh Ghaemmaghami[2]

[1] Electrical Engineering Department,
Sharif University of Technology, Azadi st., Tehran, Iran
houmansadr@mehr.sharif.edu
[2] Electronic Research Center, Sharif University of Technology, Azadi st., Tehran, Iran
ghaemmag@sharif.edu

Abstract. Image-Adaptive watermarking systems exploit visual models to adapt the watermark to local properties of the host image. This leads to a watermark power enhancement, hence an improved resilience against different attacks, while keeping the mark imperceptible. Visual models consider different properties of the human visual system, such as frequency sensitivity, luminance sensitivity and contrast masking. Entropy masking is another human visual system's characteristic, which rarely has been addressed in visual models. In this paper we have utilized this masking effect to improve the robustness of Image-Adaptive watermarks while keeping their transparency. Experimental results show a significant amount of enhancement to the power of watermark. The work has been expanded to video watermarking, considering special properties of the entropy masking effect.

1 Introduction

The extreme development of Internet has made the transmission, distribution and access to digital media very convenient. As a result, media producers are more frequently dealing with illegal and unauthorized usage of their productions. Amongst all digital media, video files could be the most valuable products that are being used vastly, while violating copyright laws that could impose huge damage to filmmaking industry.

Over the last two decades digital watermarking has been addressed as an effective solution to safeguard copyright laws, and an extensive research activity has been done on the area. Generally speaking, a digital watermark is an invisible mark that is inserted into a digital media such as audio, image or video and could be utilized to identify illegal distributions of copyright protected digital media and also lawbreaking customers. A digital watermark must have special features to make its desired functionalities. The embedded mark should be robust enough against various watermarking attacks while keeping the perceived quality of the host signal unchanged (the imperceptibility requirement). Watermarking attacks consist of deliberate attacks made maliciously to remove or change the mark sequence by lawbreakers and unintentional attacks caused as a result of different kinds of coding and compression made to the digital media prior to transmission and/or storage and

M. Barni et al. (Eds.): IWDW 2005, LNCS 3710, pp. 444 – 458, 2005.

also errors occurred during the transmission of the media through the transmission networks. In addition, inserted watermark must prohibit the malicious insertion of additional marks by embedding the maximum possible watermark power, exploiting the maximum affordable capacity of the digital media.

Fig. 1. Comparison of *Cox* scheme (Image-Independent) and *IA-DCT* scheme (Image-Adaptive): (a) top-left: Original image, (b) top-middle: watermarked image by *Cox* scheme, (c) top-right: watermarked image by *IA-DCT* scheme, (d) bottom-middle: Cox's watermark and (e) bottom-right: IA-DCT's watermark (watermarks are scaled to be visible)

Watermarking algorithms, as far as digital images are concerned, can be classified into two different categories: Image-Independent watermarking schemes and Image-Adaptive watermarking schemes. Watermarking schemes of the first class embed the mark discarding Human Visual System (HVS) characteristics and without using any visual model. In other words, watermark insertion algorithm performs the same for every host image without considering its special characteristics. *Cox* watermarking scheme [1] is a well-known scheme, which belongs to this category and is based on the spread spectrum communications. In this scheme, a DCT (Discrete Cosine Transform) transform is performed on the whole image and then the watermarking sequence is embedded in the predetermined perceptually significant DCT coefficients (low frequency components except DC coefficient). On the other hand, Image-Adaptive watermarking schemes exploit visual models to adapt the mark sequence to the local properties of the host image providing an optimal transparency and robustness. *Podilchuk et al* proposed two watermarking algorithms, which are classified in this category of image watermarking scheme [2]. Their *IA-DCT* scheme is the modified version of the Cox scheme using *Watson visual model* [3], in which DCT transform is performed on non-overlapping 8*8 blocks of the host image and then the mark sequence is inserted in each block considering its visual characteristics according to Watson's visual model. Figure 1 shows the watermarked images of *Lena* using Cox algorithm and IA-DCT algorithm in addition to their corresponding

watermarks. As shown in the figure, the second class of image watermarking schemes allocates watermark positions according to local characteristics of the host image with respect to a visual model. As a result, for images with fairly uniform perceptual characteristics, image-adaptive watermarks provide a watermark power close to what the non-image-adaptive algorithms insert, while for more complex host images, taking advantage of local properties of the images leads to insertion of higher power and more watermark bits.

In this paper, we will use a rarely used masking effect to improve watermark power in Image-Adaptive watermarking schemes. In section 2, the concept of *entropy masking* is described, following a general introduction to visual models. In section 3, the mentioned masking effect is exploited in the watermark insertion process. Some simulation results are given in section 4 and the paper is concluded in section 5.

2 Visual Models and Entropy Masking

Over recent years, there has been tremendous effort in order to understand and model the Human Visual System and applying it to different image processing applications [2]. Such effort has been examined for solving various problems and has resulted in different levels of success. Recently, visual models have been developed as a result of the efforts taken place in the field of image and video compression, which desire to improve the quality of the compression exploiting HVS characteristics. Basically, both image watermarking and image compression are concerned of the image redundancy, which is to be reduced in the case of compression, while is employed to insert the mark in the case of watermarking. As a result, visual models devised in the area of image compression can also be suited to the watermarking problem.

In compression applications, a common approach to perceptual coding is to derive an image dependent mask containing the *JND*'s (Just Noticeable Difference) to gain perceptual based quantizers and to perceptually allocate the bit positions. The same approach can be utilized in the case of watermarking problem to find upper bounds of the watermark intensity levels in different regions of the image to assure the watermark's transparency, while providing the maximum affordable robustness of the mark. Also, JND's can be used to determine an upper bound of the number of watermarks that can be inserted in a particular image with a low error probability, which can be referred to as the *watermark capacity*.

In contrast with compression applications, watermarking algorithms can fully utilize the local information extracted from the visual models since the original image is available at the receiver [2]. As an example of visual models, we consider two perceptual models which have been applied to the baseline of the JPEG coder. *Watson* model [3] and *Safranek-Johnston* model [4] are based on the same image independent component, utilizing *frequency sensitivity,* which is the human's eye sensitivity to sine wave grating at various frequencies. This component is based on the work done in [5] with a minimum viewing distance of four picture heights and a D65 monitor white point. As a result, a frequency threshold value is obtained for each DCT basis function that we mention it by $t_{u,v}^{F}$ where u and v are the indexes of the block and in the case of the work done for JPEG compression, this results in an 8*8 matrix of threshold values. Frequency sensitivity provides a static JND, which only depends on

the viewing conditions and is independent from the image. Generally speaking, HVS decreases its sensitivity to very low and very high frequencies. Furthermore, Watson refined his model by adding a *luminance sensitivity* and *contrast masking* component [3]. Luminance sensitivity states the detecting threshold of noise on a constant background. This is a nonlinear parameter, which depends on the local characteristics of the image. Watson estimated luminance sensitivity in [3] by the formula

$$t^{L}_{u,v,b} = t^{F}_{u,v}\left(\frac{X_{0,0,b}}{X_{0,0}}\right)^{a}, \tag{1}$$

where $X_{0,0,b}$ is the DC coefficient of the DCT transform over the b's block, $X_{0,0}$ is the DC coefficient of the DCT transform over the whole image and a is a parameter for controlling the degree of luminance sensitivity which was set to 0.649 by the authors in [5]. *Contrast masking* is the third component that Watson used in his model, which states that a signal can be masked in the presence of another signal especially when the both signals have the same spatial frequency, orientation and location. This allows for a more dynamic JND threshold allocation. Considering $X_{u,v,b}$ as the DCT coefficients of block b and $t_{u,v}{}^{L}$ as the threshold values derived from the viewing conditions and also the luminance masking, a contrast masking threshold is derived as

$$t^{C}_{u,v,b} = Max\left[t^{L}_{u,v,b}, \left|X_{u,v,b}\right|^{\omega_{u,v}}\left(t^{L}_{u,v,b}\right)^{1-\omega_{u,v}}\right], \tag{2}$$

where $\omega_{u,v}$ assumes a different value for each DCT basis function, where a value of 0.7 was derived for it in [3].

Here, we address another masking effect that was first introduced by *Watson et al* [6]. Watson classified all the previous studies on visual masking into two classes: *contrast masking* and *noise masking* depending primarily on whether the mask is deterministic or random. Generally, contrast masking refers to a decrease in the effective gain of the early visual system. On the other hand, noise masking is explained by an increased variance in some internal decision variable [6]. Clearly, contrast masking is the one which plays the role in image quality models. Watson et al. introduced a third type masking effect, which is *deterministic* but *unfamiliar* and called it *entropy masking*. Watson performed a number of experiments to determine the detectability threshold of an especial Gabor function added to various backgrounds by some observers. Each trial of the experiment consists of two intervals, which in one of them the single background is shown to the observer and in the other the background containing the target (Gabor function) is shown to the observer, from which he/she must detect the background containing the target. The mean results of the Watson's experiments are shown in table 1, where *none* represents the condition without any background, *cos* represents the situation with a cosine background having the same frequency and orientation as the Gabor function, *random* represents the situation that a new sample of a bandpass noise (with a bandwidth equal to Gabor function's frequency) was used in each interval of each trial, *twin* is a especial case of *random* experiment, in which the same background is used in two intervals of each trial and the *fixed* experiment is the case where the bandpass noise background is fixed for all the trials.

Table 1. Mean threshold and threshold elevation for different experimental conditions. Elevations are relative to no mask condition. (Table from [6]).

	None	Cos	Random	Twin	Fixed
Threshold(dB)	-27.59	-24.92	-14.08	-12.80	-19.34
Elevation(dB)	0	2.67	13.51	14.79	8.24

The small threshold elevation in the *cos* case is as a result of contrast masking. As stated, all the experiments are under the same contrast masking effect, so the 10.84 dB threshold elevation of *random* experiment in comparison to *cos* experiment must be due to *noise masking*. Surprisingly, threshold elevation in the *twin* condition is approximately the same as that in the random condition while we have no more *noise masking* because of utilizing the same background for the both intervals of every trial. Watson et al stated that this threshold elevation is due to unfamiliarity of the observer to background which they called it entropy masking to reflect the notion that the masking is a function of the degree to which the mask is unknown. Entropy is a measure of information in a signal which is by definition that which we do not know.

An interesting phenomenon was observed dealing with entropy masking. In the *fixed* condition of the experiments (where the bandpass noise is fixed for all the trials), as we raise the number of observations, the detectability threshold decreases (see Table 1). This is called *learning* characteristic upon it entropy masking decreases if the complexity of the background is *learned* by the observers. Other experiments utilizing *white noise*, *fixed white noise* and *fixed image* as a background showed that learning ability and speed of learning of the background is a function of its *simplicity* [6]. In other words, an image is learned very rapidly while a fixed white noise either not learned or learned very slowly.

In the next sections of this paper, we consider the entropy masking effect in the Watson's visual model in order to improve watermarking power of the IA-DCT scheme [2] (which is an Image-Adaptive watermarking scheme), and then implement this power enhancement on video streams with respect to learning characteristic of the entropy masking effect.

3 Watermark Insertion and Detection

As mentioned earlier, Watson in [6] introduced a new masking effect and called it entropy masking. The goal of this paper is to exploit this masking effect to increase watermark power in content based (image adaptive) watermarking schemes, hence to improve their resilience against various intentional and unintentional attacks.

For the aim of representing and measuring the amount of complexity, we define a *spatial complexity* parameter. To avoid massive computational complexity (especially when dealing with video host signals) we define this parameter for *n*n* blocks of pixels (rather than for each pixel). On the other hand, spatial complexity for each pixel of an *n*n* block *b* is defined as the entropy of the block

$$En_b = -\sum_x p(x).\log p(x) \ , \tag{3}$$

where $p(x)$ is the ratio of the number of x'th symbol occurrence in block b over the block size. Before computing the spatial complexity as above, pixel values must be quantized. Figure 2 shows the complexity parameter for pixels of an 8bit image for different number of quantization levels (block size is set to 4). With no quantization (256 levels of quantization), the complexity parameter is uniform (high) for almost all the pixels (figure2b). On the other hand, a 2-level quantization yields the complexity function representing the edges in the image (figure 2d). Our simulation results show that for an 8bit image, 32 levels of quantization make the defined complexity parameter to represent high entropy regions of the image, which is affected by the entropy masking effect (figure 2b).

Fig. 2. Sketching the defined complexity function for different number of quantizing levels: (a) top-left: original image, (b) top-right: no quantization, (c) bottom-left: 32 levels of quantization and (d) bottom-right: 2 levels of quantization.

Figure 3 shows our complexity function on *Lena* image for different values of size of the block n. It can be seen that for large n (figure 3a) our complexity parameter returns a high value for some non-complex pixels in the neighborhood of the edges and this will lead some impairment in this regions after applying watermark utilizing entropy masking effect. On the other hand, figure 3c returns a more realistic value for all the pixels but leading to more computational complexity. Our simulations showed that choosing n equal to 4, will make an acceptable tradeoff between precision and complexity (figure 3b). An alternative approach in defining a complexity parameter

will be similar to what *Kim et al* [7] done in the DCT domain. In this manner, the entropy of eight neighbors of a pixel is defined as a measurement for its complexity. Figure 3d shows this measure for the same image. Clearly, here we have a more precise measure of complexity for all the pixels but this will impose about 8 times more complexity to our watermarking algorithm than our simple complexity parameter. We do not need such accuracy, so we do not suffer such a high computational complexity.

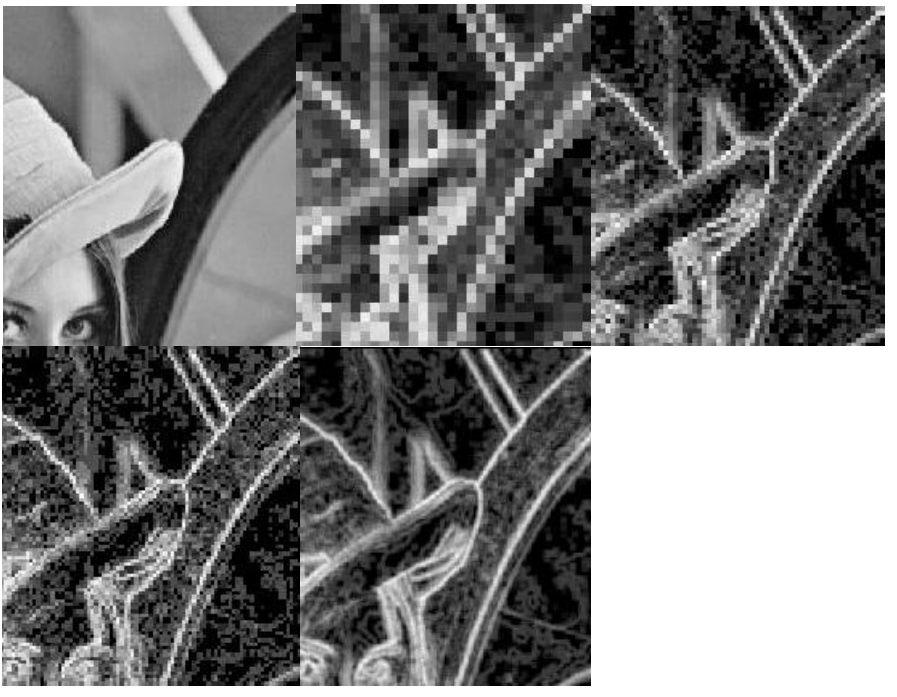

Fig. 3. Sketching the defined complexity parameter for different values of block size: (a) top-left: original image, (b) top-middle: block size equals 8, (c) top-right: block size equals 4, (d) bottom-left: block size equals 3 and (e) bottom-middle: a different complexity parameter

As an example of content based watermarking schemes, we first introduce *IA-DCT* watermarking scheme [2] and then apply our power improvement idea on it using complexity function which we defined above. In IA-DCT method, the host image is first divided into 8*8 non-overlapping blocks and 8*8 DCT transforms are performed on these blocks. Then watermark will be inserted into each DCT coefficient as below

$$\begin{aligned} X'_{u,v,b} &= X_{u,v,b} + jnd_{u,v,b} * \omega_{u,v,b} && \text{if } X_{u,v,b} \geq jnd_{u,v,b} \\ X'_{u,v,b} &= X_{u,v,b}\ , && \text{if } X_{u,v,b} < jnd_{u,v,b} \end{aligned} \tag{4}$$

where $X_{u,v,b}$ is the (u,v)-th coefficient of DCT transform over b'th block, *jnd* is the corresponding Just Noticeable Difference, $\omega_{u,v,b}$ is the corresponding watermark bit

and $X'_{u,v,b}$ is the resulted watermarked DCT coefficient. *jnd* is the contrast masking that was defined by Watson and mentioned in the previous section. Considering En_b as the complexity parameter of block b, we modify this JND considering entropy masking effect as below

$$\begin{aligned} jnd^*_{u,v,b} &= jnd_{u,v,b} * f(En_b) && \text{if } En_b > k \\ jnd^*_{u,v,b} &= jnd_{u,v,b}\ , && \text{if } En_b < k \end{aligned} \tag{5}$$

where $jnd^*_{u,v,b}$ is the modified JND, f is a function of defined complexity parameter and k is a threshold discussed in the following. We keep the JND of the three low-frequency components of each block unchanged to avoid changes to image luminance.

Through a meticulous inspection of a set of images of various degrees of complexity, we found out that choosing the threshold, *k,* equal to 2 would lead our complexity function revealing complex regions affected by the entropy masking phenomenon. Figure 4 illustrates regions of Lena image whose entropy (as defined by the complexity parameter) is greater than *k,* for different values for *k*. It is observed that regions having entropy more than 1 covers a vast area of the image (fig 4b). On the other hand, regions with entropy more than 3 are restricted to neighborhoods of the edges (fig 4d), while a reasonable result is obtained with *k=2* (figure 4c).

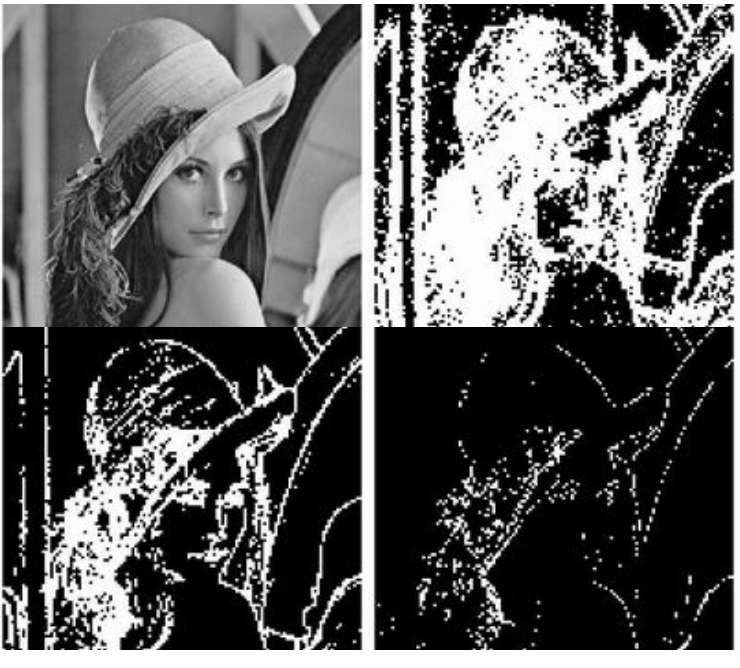

Fig. 4. Regions having entropy more than *k* threshold: (a) top-left: original image, (b) top-right: k=1, (c) bottom-left: k=2 and (d) bottom-right: k=3

According to our subjective experiments on images with various degrees of complexity, it is concluded that a simple linear function, constructed by a line connecting highest affordable entropy (which is six) and the minimum entropy masking capability (which is two), yields a reliable imperceptibility margin in the entropy space. This is expressed as:

$$f(x) = 0.2 * x + 0.8 \ . \tag{6}$$

This means that making the mark's intensity twice in regions having the maximum entropy keeps the mark still imperceptible due to the entropy masking effect, where we can have just a 20% intensity increment for regions of minimum entropy, as two extremes over the margin. Selection of a linear function for *f(x)* minimizes the computational complexity, as compared to possibly more accurate non-linear alternatives.

The detection process is the same as that introduced in [2] for a typical IA-DCT scheme. The original image is subtracted from the received possibly distorted image that results in a difference image. The correlation between this difference and the watermark sequence is then evaluated, based on a certain threshold, to check if the image contains a watermark. The detection scheme based on normalized correlation can be expressed as [1]:

$$\omega^*_{s,u,v,b} = X^*_{u,v,b} - X_{u,v,b} \ , \tag{7}$$

$$\omega^*_{u,v,b} = \frac{\omega^*_{s,u,v,b}}{JND_{u,v,b}} \ , \tag{8}$$

$$\rho_{\omega\omega^*} = \frac{\omega \ .\omega^*}{\sqrt{E_\omega . E_{\omega^*}}} \ , \tag{9}$$

where $\omega^*.\omega$ denotes the dot product, $JND_{u,v,b}$ is the corresponding JND threshold, $\omega_{u,v,b}{}^*$ is the received possibly distorted watermark, E_ω represents the $\omega.\omega$ and $\rho_{\omega\omega*}$ is the normalized correlation coefficient between the two signals ω^* and ω. If ω is normally distributed and identical to ω^*, the correlation coefficient approaches to one. Independence of ω^* and ω leads a normally distributed correlation coefficient with zero mean. A blind detection also can be performed by estimating the JND thresholds from the received image. The estimated JND form the watermarked image highly resembles the JND used for the mark insertion; however, the blind detection scheme is less robust to various watermarking attacks.

At this point, we want to perform our power enhancement method to video frames. Because of the *learning* effect mentioned in section 2, we should act a little conservatively when applying entropy masking JND elevation to still images. Similarly, for video frames, we consider that the JND elevation of a complex scene resulted from entropy masking will vanish if the scene is being repeated on consecutive frames. So, we have to define a *motion* parameter to reflect how much the components of a video frame have been changed in respect to another frame. This parameter could be defined as the absolute difference or square difference of the frame components and could be performed pixel-wise or block-wise. By spending the

price of more complexity, a BMA block search yields to more accurate results. For convenience, we defined the *motion* parameter to be the pixel-wise square difference between the selected video frames. Figure 5 shows the result of our motion parameter on two video frames. Our simulations on various video sequences showed that regions having motion parameter lower than 0.1 represent the *static* regions. Regions which remain static over a number of frames will be *learned* and its increased JND due to entropy masking will be retreated.

Fig. 5. Our motion parameter on two video frames. (Left): first video frame, (Middle): seventh video frame of the *momdaughter* sequence and (Right): the corresponding motion parameter.

Motion parameter can be used to repair the JND thresholds in different manners. One can evaluate the motion parameter between the interested frame and a number of its followers to determine the static regions of that frame. But, this will yield a huge computational complexity in the watermark insertion and detection processes of a video sequence. Our experiments showed that a motion parameter evaluation between the interested frame and only two frames located 6 and 12 frames later will give a similar measurement. So, in the case of video watermarking, after evaluating entropy of the pixels (as defined earlier) the following algorithm refines them before being performed on the JND threshold of the frame pixels:

```
For every pixel p of the i'th frame
If motion_p(i,i+6)<0.1 then En_p=En_p-Th
If motion_p(i,i+12)<0.1 then En_p=En_p-Th
```

where $motion_p(i,j)$ is our motion parameter for pixel *p* between i'th and j'th frames and En_p is the defined complexity parameter for pixel *p*. *Th* is the amount of entropy decrement which we set it to be 0.3 after a number of experiments.

4 Results

As described in the previous section, we exploit entropy masking effect to improve the watermark power in IA-DCT scheme [2] which is one paradigm of Image-Adaptive watermarking schemes. We used different host images with various degrees of complexity to simulate the proposed power-improved watermarking scheme. Figure 6 comprises our power improved watermarking method with the normal IA-DCT scheme for two types of host images (a high-entropy image and a medium-entropy image). As shown in the figure, we have strengthened the watermark power in high entropy regions of the host image by relying on the entropy masking effect.

As mentioned in the previous section, we have an extra power of the mark in regions that are being masked by the entropy masking effect. Expectedly we expect an improved resilience against various intentional and unintentional watermarking attacks due to this power enhancement. Table 2 shows the amount of improved mark power for two kinds of the images. Clearly, more power enhancement can be done for high-entropy host images.

Table 2. Power enhancement percentage for medium-entropy and high-entropy images

Image Type	*Lena* (a medium-entropy image)	*Baboon* (a high-entropy image)
Power Enhancement	35%	84%

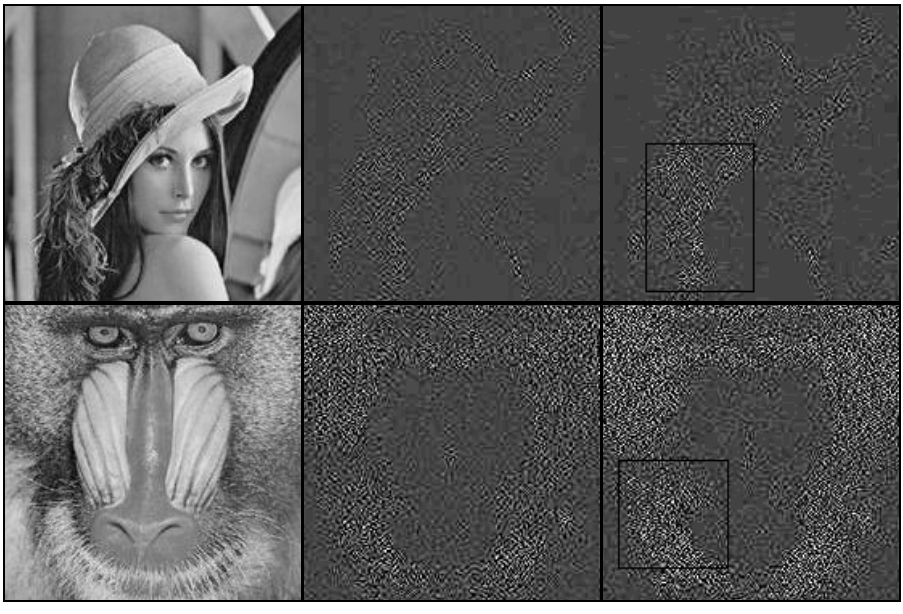

Fig. 6. Comparison between the typical IA-DCT method and its power improved version using our method for a medium-entropy (Lena) and a high-entropy image (Baboon). (left): original image, (middle): IA-DCT watermark and (right): the corresponding power improved watermark. (Watermarks are scaled to be visible.)

There are three kinds of attacks to the IA-DCT scheme considered in [2]: JPEG Compression, Cropping, and Scaling. We investigated the robustness of the improved IA-DCT scheme to the same attacks to compare the proposed method to the typical IA-DCT scheme.

JPEG Compression acts as a low-pass filtering, which zeros out frequency components. We marked different images using the typical IA-DCT scheme and its improved version, and then performed JPEG compression with various quality factors

(Q) on the marked images. Subsequently, we verified the existence of the watermark in the compressed images, evaluating the correlation coefficient as in (9). In the power-improved method, the correlation coefficient showed different amounts of elevation depending on the complexity of the host image, while the false correlation value (correlation by a non-relevant watermark) remained the same. Table 3 depicts the ratio of the correlation coefficient of the improved IA-DCT scheme over the typical IA-DCT scheme's correlation coefficient. As shown, there is not a significant elevation in the correlation coefficient, especially for low-entropy images. This is due to the fact that the improved-power mark has been inserted in high-frequency (complex) regions of the image, which are highly degraded by the JPEG compression.

Table 3. Enhancement in the detection value of the power-improved method to JPEG compression

Q factor	80	60	40	20	10	5
Baboon	1.22	1.22	1.22	1.20	1.37	1.15
Lena	1.06	1.05	1.08	1.09	1.03	1.10

Cropping can be taken as the dual of the JPEG Compression, which zeros out spatial components of the image. Because of its structure, the typical IA-DCT scheme is quite robust to only-cropping attack. We cropped several images to one-sixteenth of their original size (keeping the central part of the image) and performed JPEG compression with various quality factors. Table 4 shows the ratio of the correlation coefficient of the improved method to the typical IA-DCT's correlation coefficient, stating higher robustness of the power-improved method to the cropping attack.

Table 4. Enhancement in the detection value of the power-improved method to Cropping (one-sixteenth) followed by the JPEG compression

Q factor	80	60	40	20	10	5
Baboon	1.49	1.43	1.34	1.48	1.09	1.29
Lena	1.97	1.46	1.55	1.18	1.13	0.98

Table 5. Enhancement in the detection value of the power-improved method to Scaling

Scaling factor	2	4
Baboon	2.38	1.41
Lena	1.53	1.42

Pudlichuk et al. also investigated the robustness of the IA-DCT method to the *Scaling* of a watermarked image. Similar to the procedure given in [2], we lowpass filtered the watermarked image using four-tap filter prior to downsampling by 2 in each direction. The resulting image is upsampled prior to calculation of the correlation coefficient. Table 5 shows the ratio of the correlation coefficient of the

improved method to the typical IA-DCT's correlation coefficient. A high amount of improvement in the correlation coefficient is achieved, while the false correlation exhibits no elevation.

To prove the watermarking imperceptibility, we performed a number of subjective tests employing three independent observers. Each trial was composed of a typical IA-DCT watermarked image and its power-improved version in a random order. In every trial, each observer had to distinguish the image with more impairment (due to the watermark insertion) or remark his ambiguity. By changing the time each observer has to make his decision, three different kinds of experiments conducted: fast-decision experiment, medium-decision experiment and slow-decision experiment with two, five and ten seconds permitted for each trial respectively. Table 6 shows the success percentage of different kinds of experiments for two kinds of high-entropy and medium entropy images. By definition, an experiment is done successfully if the observer chooses the typical IA-DCT watermarked image as the image with more impairment or reveals his ambiguity.

As the results show, while enhancing a significant amount of watermark power, the mark remains truly imperceptible for fast-decision and medium-decision experiments as a result of entropy masking phenomenon. Logically, awarding a habit-time to observers gives them more chance to distinguish the power-improved mark because of *learning* property. On the other words, *learning* the complex background leads its entropy masking effect to decline as stated in section 2. As a result, we have to do our treatment more conservatively when exploiting the entropy masking effect for still images in respect to their functionalities.

We also implemented our method to improve the watermark power in video sequences. As explained in section 3, we have considered the *Learning* property of entropy masking in the process of mark insertion in video frames. Figure 7 shows the mark's power enhancement for a medium-entropy video frame. Obviously, our method strengthens the mark power in regions having a large amount of spatial entropy and/or motion entropy with respect to the temporal activity.

Table 6. Success percentage for three kinds of experiments performed on medium-entropy and high-entropy images

Image Type	*Lena* (a medium-entropy image)	*Baboon* (a high-entropy image)
Fast-decision Experiment	100%	100%
Medium-decision Exper.	95%	97%
Slow-decision Experiment	87%	91%

Table 7. Power enhancement percentage for medium-entropy and high-entropy video sequences

Video Sequence	*Momdaughter* (medium-entropy)	*Coastguard* (high-entropy)
Power Enhancement	28%	57%

Table 7 shows the average power enhancement per frame for two kinds of the video streams. Again, more power enhancement can be achieved for high-entropy sequences.

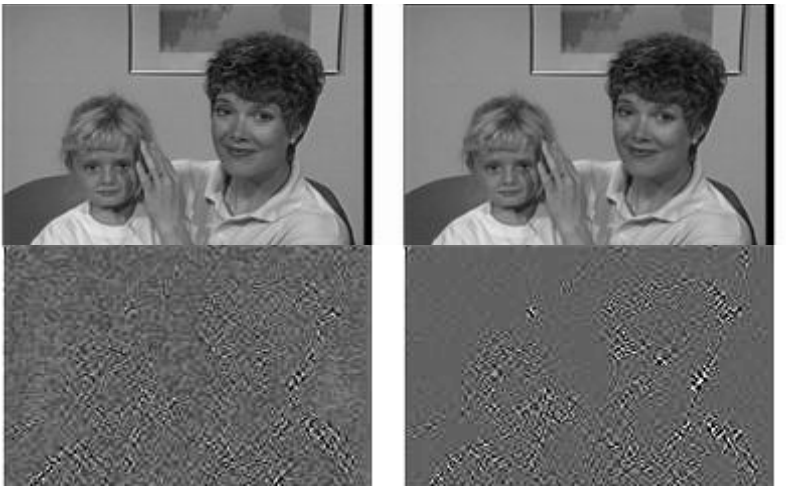

Fig. 7. Comparison between the typical IA-DCT watermarking scheme on video sequences and its power improved version using our method for a medium-entropy video sequence (Momdaughter). (left): video frame watermarked with the IA-DCT method and its corresponding watermark and (right): our power improved version. (Watermarks are scaled to be visible)

Similar to what was done for still images a number of subjective experiments were done to assure the imperceptibility of the power increased mark. Again, each trial of the experiments consisted of an IA-DCT frame by frame marked video sequence and its power improved twin (according to our method) in an unknown ordering, and the observers had to distinguish the sequence having more impairment or claim their ambiguity. Table 8 shows the success percentage of the experiments for two kinds of the sequences (high-entropy and medium-entropy sequences). Results promise a high assurance of imperceptibility while achieving the valuable mark's power improvement. Because of considering the *learning* property, there is no more need to perform conservatively similar to what concluded for the still images.

Table 8. Success percentage of our experiments performed on medium-entropy and high-entropy video sequences.

Video Sequence	*Momdaughter* (medium-entropy)	*Coastguard* (high-entropy)
Experiment's Success Percentage	94%	97%

5 Conclusions

In this paper, we have discussed the entropy masking effect, which was initially introduced by Watson et al. [3], and utilized it to improve the watermark power in content-based watermarking schemes. The proposed method leads to enhancement of the watermark's robustness against various intentional and unintentional attacks. As an example of IA watermarking scheme, we have implemented our method on the IA-DCT scheme [2] and have shown that a significant amount of power enhancement can be achieved exploiting the mentioned masking effect. In fact, we have increased the power of the mark in the regions that are being masked by the entropy masking effect. Because of the learning effect of the entropy masking effect, we must act a little conservatively (not using all the power enhancement of the method) when applying our method on still images, depending on how the marked image is used. However, exploiting this masking effect will result in a more assurance of the mark's imperceptibility with the same watermark power, if there is an upper bound limitation on the mark's power.

We also applied our method to video sequences. The same power enhancement achieved while keeping the marked sequence imperceptible. As cited, we considered the learning property in the watermark embedding process of the video frames, so there is no more need for working conservatively. We tried our method to be as simple as possible, because of the computational constraints in video watermarking. So, different approaches can be implemented with varying degrees of complexity.

Acknowledgements. The authors would like to thank Professor F. Marvasti, director of the Advanced Communications Research Institute for his cooperation and M. Sohizadeh Abianeh for his reviews.

References

1. I. J. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," NEC Research Institute Tech. Rep. 95-10, 1995.
2. C. I. Podilchuk and W. Zeng, "Image-Adaptive Watermarking using visual models", *IEEE Journal on selected areas in communications,* vol. 16, no. 4, May 1998, pp. 525-539
3. A. B. Watson, "DCT quantization matrices visually optimized for individual images," in *Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display IV,* Feb. 1993, vol. 1913, pp. 202–216.
4. R. J. Safranek and J. D. Johnston, "Perceptually based prequantization for image compression," in *Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display V, 1994.*
5. H. A. Peterson, A. J. Ahumada, Jr., and A. B. Watson, "Improved detection model for DCT coefficient quantization," in *Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display IV,* Feb. 1993, vol. 1913, pp. 191–201.
6. A. B. Watson, R. Borthwick and M. Taylor, "Image quality and entropy masking", in *Proc. SPIE Conf.,* vol. 3016, 1997
7. S. W. Kim and S. Suthaharan, "An entropy masking model for multimedia content watermarking", in *Proc. 37th Hawaii International Conference on System Sciences, 2004*