SIViP (2011) 5:517–526
DOI 10.1007/s11760-010-0164-x

ORIGINAL PAPER

A combined just noticeable distortion model–guided image watermarking

Yaqing Niu · Matthew Kyan · Sridhar Krishnan · Qin Zhang

Received: 1 September 2009 / Revised: 19 April 2010 / Accepted: 20 April 2010 / Published online: 8 May 2010
© Springer-Verlag London Limited 2010

Abstract Perceptual watermarking should take full advantage of the results from human visual system (HVS) studies. Just noticeable distortion (JND), which refers to the maximum distortion that the HVS does not perceive, gives us a way to model the HVS accurately. In this paper, we exploit a combined JND model, which represents an additional, accurate, perceptual visibility threshold profile to guide watermarking for digital images. The combined JND model-guided watermarking scheme, where visual models are fully used to determine image-dependent upper bounds on watermark insertion, allows us to provide the maximum strength transparent watermark. Experimental results confirm the improved performance of our combined JND model. Our combined JND model is capable of yielding higher injected-watermark energy without introducing noticeable distortion to the original image and outperforms the relevant existing visual models. Simulation results show that the proposed JND model-guided image watermarking scheme is more robust than other algorithms based on the relevant existing perceptual models while retaining the watermark transparency. At the same time, the proposed combined JND model has much lower computational complexity compared with the relevant existing perceptual models.

Keywords Human visual system · Just noticeable distortion model · Perceptual watermarking · Contrast sensitivity function

Y. Niu (✉) · Q. Zhang
Information Engineering School, Communication University of China, 202#, 1 Ding Fu Zhang Dong Jie, 100024 Chao Yang District, Beijing, China
e-mail: niuyaqing@cuc.edu.cn

Q. Zhang
e-mail: zhangqin@cuc.edu.cn

Y. Niu · M. Kyan · S. Krishnan
Department of Electrical and Computer Engineering, Ryerson University, 350 Victoria Street, Toronto, ON, Canada

M. Kyan
e-mail: mkyan@ee.ryerson.ca

S. Krishnan
e-mail: krishnan@ee.ryerson.ca

1 Introduction

The rapid growth of the Internet has created a need for techniques that can be used for copyright protection of digital images and videos. One approach is to introduce a digital watermark into images or video sequences. For a well-designed watermark, there are many requirements including imperceptibility, robustness and capacity. However, in order to maintain the image quality and at the same time increase the probability of watermark detection, it is necessary to take the human visual system (HVS) into consideration when engaging in watermarking research [3,13,19].

HVS makes final evaluations on the quality of images that are processed and displayed. Just noticeable distortion (JND), which refers to the maximum distortion that the HVS does not perceive gives us a way to model the HVS accurately and can serve as a perceptual visibility threshold to guide image watermarking. An early perceptual threshold estimation in DCT domain was proposed by Ahumada [1], which gives the threshold for each DCT component by incorporating the spatial Contrast Sensitivity Function (CSF). This scheme was improved by Watson [17] after the luminance adaptation effect had been added to the base threshold, and contrast masking [9] had been calculated as the elevation factor. In [20], an additional block classification-based contrast masking and

luminance adaptation was considered by Zhang for digital images. A spatial JND model proposed by Wei [18] incorporates new spatial CSF, luminance adaptation and contrast masking.

Previous watermarking researches have only partially used the results of the HVS studies [4,6,11,13,19]. Many previous image watermarking algorithms utilize visual models to increase the robustness and transparency. The perceptual adjustment of the watermark is mainly based on Watson's spatial JND model [11,13,14,19]. An image-adaptive watermarking procedure based on Watson's spatial JND model was proposed in [13]. In [19], the DCT-based watermarking approach uses Watson's spatial JND model in which the threshold consists of spatial frequency sensitivity, luminance sensitivity and contrast masking. An Energy Modulated Watermarking Algorithm Based on Watson's Spatial JND Model was proposed in [11]. During the modulation, Watson's perceptual model is used to restrict the modified magnitude of DCT coefficients. In [14], Watson's perceptual model and a modified Watson's perceptual model were used to adaptively select the quantization step size for modifications to dither-modulated Quantization Index Modulation (QIM). In the modified Watson model, the luminance masking is simply modified to scale linearly with valumetric scaling. The main drawback of utilizing Watson's visual models for images watermarking is that it does not satisfactorily provide the maximum strength transparent watermark. The obtained watermark is not optimal in terms of imperceptibility and robustness.

In this study, we propose an effective combined JND model-guided image watermarking scheme in the DCT domain. The combined JND model is more accurate since the spatial CSF is designed based on psychophysical experiment and a more precise luminance adaptation function is used. The watermarking scheme is based on the design of an additional accurate JND visual model, which is fully used to determine scene-adaptive upper bounds on watermark insertion. Thus, performance in terms of robustness and transparency is obtained by embedding the maximum strength watermark while maintaining the perceptually lossless quality of the watermarked images. This paper is organized as follows. In the next section, the details of the proposed combined JND model is presented, so visibility threshold in DCT sub-bands is derived. And Sect. 3 gives the details of the proposed combined JND model-guided image watermarking scheme. The proposed scheme is tested against common attacks such as additive Gaussian noise, valumetric scaling, geometric distortion and JPEG compression in Sect. 4. A detailed comparison of the proposed combined JND model-guided watermarking scheme with three relevant existing perceptual model-guided watermarking schemes is also presented in this section. Finally, the paper is concluded in Sect. 5.

2 Combined JND model

Just noticeable distortion (JND) refers to the maximum distortion threshold that the HVS cannot perceive; therefore, it is an efficient model to represent perceptual redundancies. Here, we compute a combined JND model that incorporates Spatial CSF, luminance adaptation and contrast masking all together.

2.1 Spatial CSF

Human eyes show a band-pass property in the spatial frequency domain. The Spatial CSF model is the reciprocal of the base distortion threshold, which can be tolerated for each DCT coefficient. Literature [18] shows that the base threshold for DCT domain T_{BASEs}corresponding to Spatial CSF model can be expressed by (1)

$$T_{\mathrm{BASE_S}}(n,i,j) = s \cdot \frac{1}{\phi_i \phi_j} \cdot \frac{\exp\left(c\omega_{ij}\right) / \left(a + b\omega_{ij}\right)}{r + (1-r)\cdot \cos^2 \varphi_{ij}}. \tag{1}$$

where T_{BASEs} is the base threshold corresponding to Spatial CSF model and n is the index of a block in the image, i and j are the DCT coefficients' indices. $a = 1.33$, $b = 0.11$, $c = 0.18$, $s = 0.25$. φ_i and φ_j are DCT normalization factors by (2), where m represents i or j; ω_{ij} is the spatial frequency that can be calculated by (3); N is the dimension of the DCT block; θ_x and θ_y are the horizontal and vertical visual angles of a pixel by (4), Where l is the viewing distance and Λ stands for the display width/length of a pixel on the monitor. r is set to 0.6, and φ_{ij} stands for the directional angle of the corresponding DCT component by (5).

$$\phi_m = \begin{cases} \sqrt{1/N}, & m = 0 \\ \sqrt{2/N}, & m > 0 \end{cases} \tag{2}$$

$$\omega_{ij} = \frac{1}{2N}\sqrt{(i/\theta_x)^2 + \left(j/\theta_y\right)^2} \tag{3}$$

$$\theta_\hbar = 2 \cdot \arctan\left(\frac{\Lambda_\hbar}{2 \cdot l}\right) \quad (\hbar = x, y) \tag{4}$$

$$\varphi_{ij} = \arcsin\left(\frac{2\omega_{i,0}\omega_{0,j}}{\omega_{ij}^2}\right) \tag{5}$$

2.2 Luminance adaptation

The HVS is more sensitive to the noise in medium gray regions, so the visibility threshold is higher in very dark or very light regions. Because our base threshold is detected at the 128 intensity value, for other intensity values, a modification factor needs to be included. This effect is called the luminance adaptation effect.

The curve of the luminance adaptation factor is a U-shape, which means the factor at the lower- and higher-intensity regions is larger than the middle-intensity region. An empirical formula for the luminance adaptation factor a_{Lum} in [18] is shown as (6), where $I(n)$ is the average intensity value of the n_{th} block in the image.

$$a_{\text{Lum}}(n) = \begin{cases} (60 - I(n))/150 + 1 & I(n) \le 60 \\ 1 & 60 < I(n) < 170 \\ (I(n) - 170)/425 + 1 & I(n) \ge 170 \end{cases} \tag{6}$$

2.3 Contrast masking

Contrast masking refers to the reduction in the visibility of one visual component in the presence of another one. The masking is strongest when both components are of the same spatial frequency, orientation and location. To incorporate contrast masking effect, we employ contrast masking a_{contrast} [9] measured as (7), where $C(n, i, j)$ is the (i, j)th DCT coefficient in the nth block in the image.

$$a_{\text{contrast}}(n, i, j) = \max\left(1, \left(\frac{C(n, i, j)}{T_{\text{BASE}}(n, i, j) \cdot a_{\text{Lum}}(n)}\right)^{\varepsilon}\right) \tag{7}$$

To find a suitable value for ε, Zhang's perceptual model [20] proposed a much lower value 0.36. But Zhang's model tends to underestimate JND in edgy blocks. We exploit ε fixed to 0.7 as in Watson's perceptual model [9].

2.4 Complete JND estimator

The overall JND (8) can be determined by the base threshold T_{BASEs}, the luminance adaptation factor a_{Lum} and the contrast masking factor a_{Contrast}.

$$T_{\text{JND}}(n, i, j) = T_{\text{BASEs}}(n, i, j) \cdot a_{\text{Lum}}(n) \cdot a_{\text{contrast}}(n, i, j) \tag{8}$$

$T_{\text{JND}}(n, i, j)$ is the complete image-driven JND estimator, which represents the additional accurate perceptual visibility threshold profile to guide watermarking for digital images.

3 The JND model-guided watermarking scheme

We exploit the combined JND model-guided watermarking scheme to embed and extract watermarking. The scheme first constructs a set of approximate energy sub-regions using the Improved Longest Processing Time (ILPT) algorithm [5,21] and then enforces an energy difference between every two sub-regions to embed watermarking bits [8,10] under the control of our combined JND model.

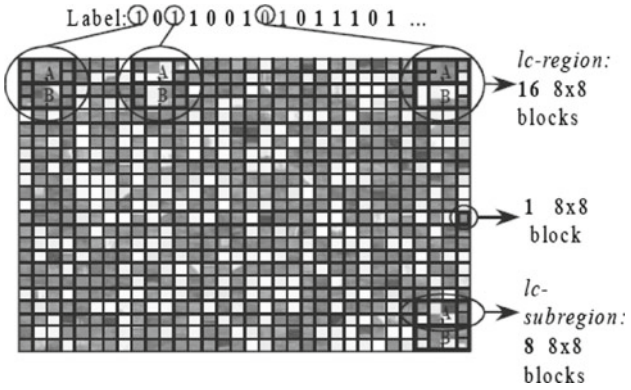

Fig. 1 Watermark bit corresponding to approximate energy sub-regions constructed by ILPT

3.1 The construction of approximate energy sub-regions

The watermark bit string is embedded bit-by-bit in a set of regions (each region is composed of $2n$ 8×8 DCT blocks) of the original image. Each region is divided into two sub-regions (each sub-region is composed of n 8×8 DCT blocks). A single bit is embedded by modifying the energy of two sub-regions separately. However, for better imperceptibility, approximate energy sub-regions need to be constructed using ILPT, so that the original energy of each sub-region in one region are approximate.

Each bit of the watermark bit string is embedded in its constructed bit-carrying region. For instance, in Fig. 1, each bit is embedded in a region of $2n = 16$ 8×8 DCT blocks. The value of the bit is encoded by introducing an energy difference between the low-frequency DCT coefficients of the top half of the region (denoted by sub-region A) containing in this case $n = 8$ 8×8 DCT blocks, and the bottom half (denoted by sub-region B) also containing $n = 8$ 8×8 DCT blocks. The number of watermark bits that can be embedded is determined by the number of blocks in a region.

3.2 The embedding procedure

Diagram of combined JND model-guided watermark embedding is shown in Fig. 2. The embedding procedure of the scheme is described as the following steps:

a) Decompose the original image into non-overlapping 8×8 blocks and compute the energy of the low-frequency DCT coefficients in the zigzag sequence.
b) Obtain approximate energy sub-regions by ILPT algorithm.
c) Map the index of the DCT blocks in a sub-region according to ILPT.
d) Use our combined JND model described in Sect. 2 to calculate the perceptual visibility threshold profile for

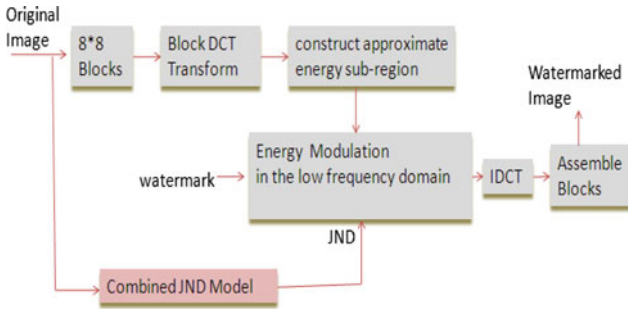

Fig. 2 Diagram of combined JND model-guided watermark embedding

an image such that the watermark remains imperceptible and robust.

e) If the watermark to be embedded is 1, the energy of sub-region A should be increased and the energy of sub-region B should be decreased. If the watermark to be embedded is 0, the energy of sub-region A should be decreased and the energy of sub-region B should be increased. The energy of each sub-region is modified by adjusting their low-frequency DCT coefficients accordingly under the control of our combined JND model as (9).

$$C(n,i,j)^m = \begin{cases} C(n,i,j) + Sign(C(n,i,j)) \\ \quad \cdot f(C(n,i,j), T_{\mathrm{JND}}(n,i,j)), PM \\ C(n,i,j) - Sign(C(n,i,j)) \\ \quad \cdot f(C(n,i,j), T_{\mathrm{JND}}(n,i,j)), NM \end{cases} \tag{9}$$

where $C(n,i,j)^m$ is the modified DCT coefficient, Sign(.) is the sign function, PM is the positive modulation, which means increase the energy and NM is the negative modulation, which means decrease the energy, $T(n,i,j)$ is the perceptual visibility threshold by our combined JND model and $f(.)$ can be expressed by (10)

$$f(C(n,i,j), T_{\mathrm{JND}}(n,i,j)) = \begin{cases} 0, & \text{if } C(n,i,j) < T_{\mathrm{JND}}(n,i,j) \\ T_{\mathrm{JND}}(n,i,j), & \text{if } C(n,i,j) \geq T_{\mathrm{JND}}(n,i,j) \end{cases} \tag{10}$$

f) Conduct IDCT to the energy modified result to obtain the watermark embedded image.

3.3 The extraction procedure

Diagram of combined JND model-guided watermark extraction is shown in Fig. 3. The extraction procedure is described as the following steps:

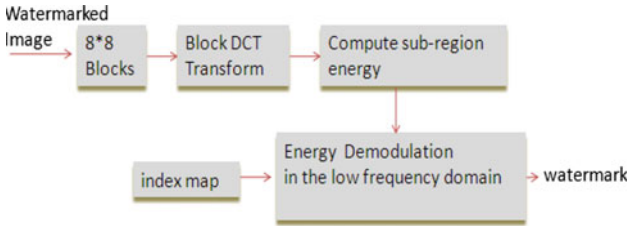

Fig. 3 Diagram of combined JND model-guided watermark extraction

a) Decompose the watermark embedded image into non-overlapping 8×8 blocks and compute the energy of the low-frequency DCT coefficients in the zigzag sequence.
b) Energy of each sub-region is calculated according to the index map.
c) Compare the energy of sub-region A with sub-region B. If the energy of sub-region A is greater than the energy of sub-region B, the watermark embedded is 1. If the energy of sub-region A is smaller than the energy of sub-region B, the watermark embedded is 0. The watermark is extracted accordingly.

4 Experimental results and performance analysis

The proposed algorithm is composed of two parts. The first part of the proposed algorithm is in Sect. 2, which derived the combined JND model. In Sect. 3, the second part of the proposed algorithm gives the details of how to utilize the combined JND model to guide image watermarking. Experimental results are then given according to these two parts. For the first part, we perform experiments in Sect. 4.1 to evaluate the performance of our combined JND model. For the second part, we perform experiments in Sect. 4.2 to evaluate the performance of our JND model-guided watermarking scheme focusing on the watermark's visual quality, capacity, robustness and computational complexity.

4.1 Evaluation on combined JND profile for still images

In this experiment, the generated JND profile can be used to guide noise shaping in images to evaluate the performance of different JND models. Watson's spatial JND model [17] (referred to as Model 1 hereinafter), Zhang's spatial JND model [20] (referred to as Model 2 hereinafter) and Wei's spatial JND model [18] (referred to as Model 3 hereinafter) were implemented and compared with our JND estimator.

Eight images are chosen as test images for this experiment (bridge, baboon, Lena, woman, man, lake, crowd, couple, shown in Fig. 4 with different visual contents).

Noise is added to each DCT coefficients of the images as (11). Where f takes $+1$ or -1 randomly, $T_{\mathrm{JND}}(n,i,j)$

Fig. 4 Images for the experiments **a** bridge, **b** baboon, **c** Lena, **d** woman, **e** man, **f** lake, **g** crowd, **h** couple

represents the JND obtained via each model, $C'(n, i, j)$ is the noise-injected DCT coefficient.

$$C'(n, i, j) = C(n, i, j) + f \times T_{\mathrm{JND}}(n, i, j) \tag{11}$$

For a convincing evaluation of our spatial JND estimator for still images, we tested in two aspects. One is to measure how much the visual content variations are after each JND model-guided noise injection, and the second is to assess the quality of each noise-injected image. The PSNR tests were used to measure the variations introduced by noise injection. On the other hand, the subjective viewing tests were used to assess the quality of the resultant visual content. A better JND model allows higher injected-noise energy (corresponding to lower PSNR) without jeopardizing picture quality (measured by the subjective viewing tests).

4.1.1 PSNR tests

Figure 5 shows the result of the Bridge image after the JND-guided noise injection as (18) by Model 1–3, and our model. We find the noise hardly noticeable in four resultant images, which showed that each of the four models yield values no larger than the actual HVS thresholds. Our model yields the lowest PSNR (model 1 PSNR = 27.6 dB; model 2 PSNR = 26.8 dB; model 3 PSNR = 28.3 dB; ours PSNR = 26.6 dB) which reflects that our JND model allowed for higher injected-noise energy without introducing noticeable visual distortions.

As aforementioned, at the same level of perceived picture quality, a better JND model yields more aggressive JNDs (resulting in lower PSNR). Figure 6 shows the PSNR of the noise-injected images with four JND profiles. As can be seen, our model yields more aggressive JNDs than models 1–3 toward the actual HVS thresholds, with the evidence of an average PSNR reduction of 0.9 dB from model 1, 0.5 dB from model 2 and 1.3 dB from model 3.

4.1.2 Subjective viewing tests

The subjective viewing tests were conducted with a ThinkPad X61 12.1 TFT LCD display. In each session of the experiment, two images of the same scene (the original as left image and noise- injected version as right image) were juxtaposed on the screen. Ten subjects were asked to give scores for all the image pairs, using the continuous quality comparison scale shown in Table 1 and a viewing distance of four times the image height.

Table 2 shows the mean subjective scores of eight image pairs by models 1–3 and our model. As can be seen, the mean subjective scores resulting from three models are very similar; hence, our model leads to similar visual quality on average as models 1–3 in the noise-injected images.

From the experimental results in Sect. 4.1, we can see our JND model correlates with the HVS the best with the evidence of being capable of yielding higher injected-noise energy (we can see lowest PSNR in Fig. 6) without introducing noticeable distortion to the original image (we can see same perceptual quality performance in subjective viewing tests in Table 2 and Fig. 5).

4.2 Evaluation on performance of combined spatial JND model-guided image watermarking

We construct a series of tests to observe the performance of our combined spatial JND model-guided image watermarking in terms of watermark's visual quality, capacity and robustness. The 512 × 512 Lena image is used for experiments. The original watermark is a binary image of the logo of Communication University of China with size 32 × 64.

4.2.1 Estimation of visual quality

As discussed in Sect. 4.1, our combined spatial JND model correlates with the human visual system very well. We can

Fig. 5 Resultant bridge images **a** with model 1, **b** with model 2, **c** with model 3 and **d** with our model

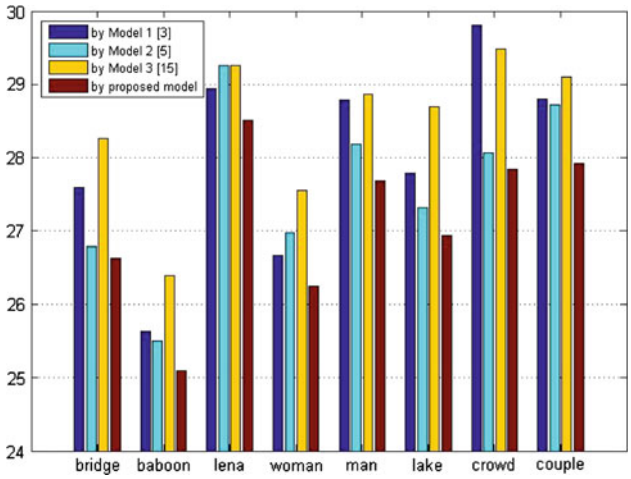

Fig. 6 PSNR with four JND profiles

use our model to guide noise shaping in each DCT coefficients of the images yet the difference is hardly noticeable. To embed watermark with the guiding of our model is similar to the process of noise shaping. As described in Sect. 3, not all the coefficients are modified with the consideration of lossy compression so the visual quality is even better than the test result in Sect. 4.1. Fig. 7a, c–f show the Lena image before and after watermark insertion guided by four different JND models. We can see no obvious degradation in

Table 1 Quantitative score of an image pair

Score	Description
−3	The right one is much worse than the left one
−2	The right one is worse than the left one
−1	The right one is slightly worse than the left one
0	The right one has same quality as the left one
1	The right one is slightly better than the left one
2	The right one is better than the left one
3	The right one is much better than the left one

Table 2 The mean subjective score

Three JND model	The mean subjective score
Model 1 [17]	−0.29
Model 2 [20]	−0.38
Model 3 [18]	−0.32
proposed model	−0.34

Fig. 7 **a** Original image, **b** watermark, **c** watermarked image by model 1 [17], **d** watermarked image by model 2 [20], **e** watermarked image by model 3 [18], **f** watermarked image by proposed model

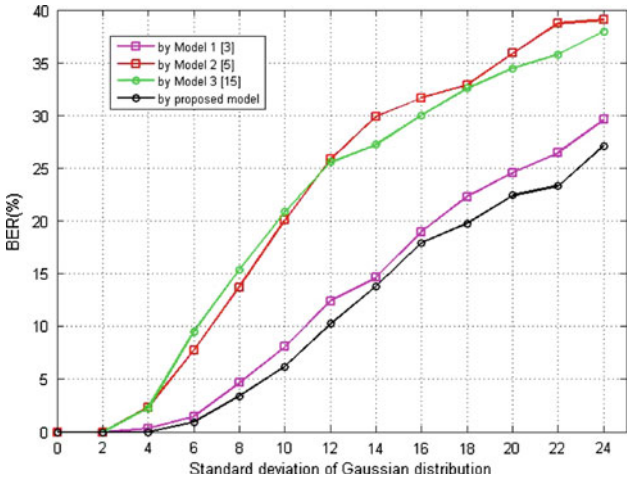

Fig. 8 Robustness versus Gaussian noise

Fig. 7c–f whose PSNR are 36.7, 46.7, 43.0 and 35.8 dB; SWR are 25.4, 35.5, 31.1 and 24.5 dB, respectively. That reflects our JND model allowing higher injected-noise energy without introducing noticeable visual distortions.

4.2.2 Estimation of capacity

The watermark bit string is embedded bit-by-bit in a set of regions of the original image. Each region is divided into two sub-regions. A single bit is embedded by modifying the energy of two sub-regions separately. For each 512×512 image, the number of bits can be embedded is determined by the number of 8×8 DCT blocks in a region. We set the number of blocks at 2 in each region in the following experiments. That means each 512×512 image would embed 2,048 bits.

In Sect. 4.2, we focused on testing that for the same number of embedded bits (2,048 bits) which watermarking scheme is more robust based on the relevant JND model while retaining the watermark transparency.

4.2.3 Estimation of robustness

In practice, watermarked content has to face a variety of distortions before reaching the detector. We present robustness results with different attacks like Gaussian noise, JPEG compression, volumetric scaling and geometric distortion. Robustness results of algorithms based on models 1–3 were compared with results of the algorithm based on our JND model shown in Figs. 8, 9, 10 and Table 3. For each category of distortion, the watermarked images were modified with a varying magnitude of distortion and the bit error rate (BER) of the extracted watermark was then computed.

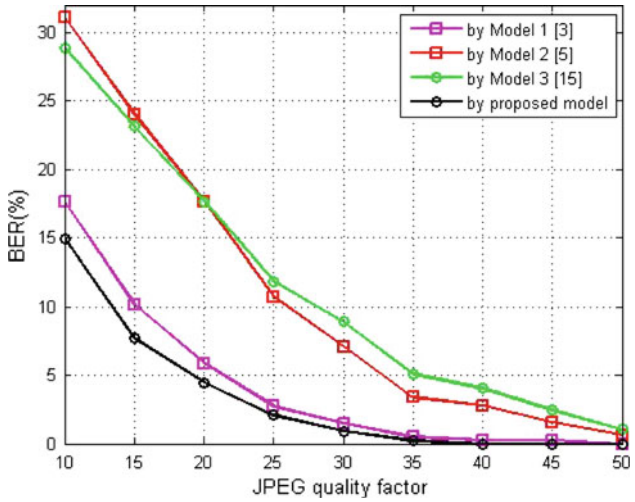

Fig. 9 Robustness versus JPEG compression

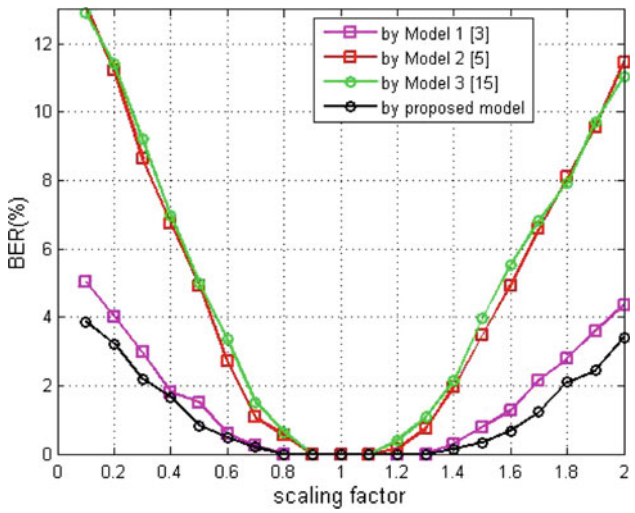

Fig. 10 Robustness versus volumetric scaling

The normal distributed noise with mean 0 and standard deviation σ is added to the watermarked images, where σ varies from 0 to 25. The experimental results are presented in Fig. 8. From the robustness results with Gaussian noise in Fig. 8, the watermarking scheme based on our combined spatial JND model performs better than algorithms based on models 1–3. Figure 8 demonstrates that our scheme's performance against Gaussian noise has an average bit error rate value 1.4% lower than algorithm based on model 1, 10.2% lower than algorithm based on model 2, and 9.7% lower than algorithm based on model 3. Following the practice, we considered the watermarking scheme to be robust if at least 80% of the watermarks were correctly retrieved, i.e. the BER is below 20%. If we establish a threshold at a bit error rate of 20%, then the algorithms based on models 1–3 are robust against Gaussian noise with standard deviation, σ, of only 16.5, 10 and 9.8 or less. In contrast, the watermarking scheme based on our combined spatial JND model is robust to standard deviations up to 18.

The JPEG compression is implemented to the watermarked images, where JPEG quality factor varies from 10 to 50. The experimental results are presented in Fig. 9. From the robustness results with JPEG compression in Fig. 9, the watermarking scheme based on our combined spatial JND model performs better than algorithms based on models 1–3. Figure 9 demonstrates that our scheme's performance against JPEG compression has an average BER value 0.9% lower than algorithm based on model 1, 7.6% lower than algorithm based on model 2, and 8.1% lower than algorithm based on model 3. We can see that if the JPEG quality factor is decreased to 10, there will be about 17.7% bit error rate introduced by model 1, while only 15% is based on our model. As the JPEG quality factor decreases, the gap between the bit errors of model 1 compared to our model increases.

From the robustness results with volumetric scaling in Fig. 10 and geometric distortion in Table 3, the watermarking scheme based on our combined spatial JND model performs better than algorithms based on models 1–3. The experiment shown in Fig. 10 reduced the image intensities as scaling factor varied from 1 to 0.1 and increased the image intensities as scaling factor varied from 1 to 2. Figure 10 demonstrates that our scheme's performance against valumetric scaling has an average bit error rate value 0.4% lower than algorithm based on model 1, 3.7% lower than algorithm based on model 2 and 3.8% lower than algorithm based on model 3. Table 3 demonstrates that our scheme performances are slightly

Table 3 Bit error rate after some geometric distortions

Attack	Bit error rate (by model 1 [17]) (%)	Bit error rate (by model 2 [20])(%)	Bit error rate (by model 3 [18])(%)	Bit error rate (by proposed model)(%)
1_Row_1_Col_removed	2.10	2.44	2.34	1.90
1_Row_5_Col_removed	1.95	2.34	2.25	1.71
5_Row_1_Col_removed	1.90	2.25	2.10	1.71
5_Row_17_Col_removed	2.25	2.78	2.59	2.10
17_Row_5_Col_removed	1.95	2.20	2.15	1.81

Table 4 Time expended for computing four spatial JND models

Images	Time expended (for model 1 [17]) (ms)	Time expended (for model 2 [20]) (ms)	Time expended (for model3 [18]) (ms)	Time expended (for proposed model) (ms)
Lena	20	156	922	15
Lake	22	168	953	16
Crowd	21	160	943	16
Baboon	24	173	968	17
Couple	20	153	916	15

better than algorithms based on models 1–3 in all cases where some geometric distortions were applied.

4.2.4 Estimation of complexity

For each 512 × 512 image, we tested the expended time for computing four spatial JND models listed in Table 4. It shows that the computation times for models 1–3 are much longer than the computation time of our JND model. This is mainly because model 2 and model 3 need more steps to find the contrast masking value, while our model has much simpler spatial CSF computational method compared with model 1.

We have presented the results of our investigation on the performance of our JND model-guided watermarking scheme in terms of watermark visual quality, capacity, robustness and computational complexity. The improvement evident above is due to our combined spatial JND model for still images, which correlates with the human visual system better than models 1–3, allowing higher injected-watermark energy and obtain better robustness in image watermarking while retaining the watermark transparency. At the same time, our JND model has much lower computational complexity than models 1–3.

5 Conclusions

Perceptual watermarking should take full advantage of the results from human visual system (HVS) studies. Just noticeable distortion (JND), which refers to the maximum distortion that the HVS does not perceive, thus gives us a way to model the HVS accurately. An effective combined JND model-guided image watermarking scheme in DCT domain is proposed in this paper. The watermarking scheme is based on the design of an additional accurate JND visual model, which is fully used to determine scene-adaptive upper bounds on watermark insertion. Thus, performance in terms of robustness and transparency is obtained by embedding the maximum strength watermark while maintaining the perceptually lossless quality of the watermarked images. Simulation results show that the proposed scheme is more robust than other algorithms based on the relevant existing perceptual models while retaining the watermark transparency. At the same time, the proposed combined JND model has much lower computational complexity compared with the relevant existing perceptual models.

Acknowledgments We acknowledge the funding provided by the National Natural Science Foundation of China (Grant No. 60832004) and Key Construction Program of the Communication University of China "211" Project. Also, I would like to thank Yunfeng Wu, Xiaoli Li and Beibei Jiao for their support.

References

1. Ahumada, A.J. Jr., Peterson, H.A.: Luminance-model-based DCT quantization for color image compression. In: Proc. SPIE **1666**, 365–374 (1992)
2. Daly, S.: Engineering observations from spatio velocity and spatiotemporal visual models. Proc. SPIE **3299**, 180–191 (1998)
3. Doerr, G., Dugelay, J.L.: A guide tour of video watermarking. Signal Process. **18**(4), 263–282 (2003)
4. Huang, J., Shi, Y.Q.: Adaptive image watermarking scheme based on visual masking. Electron. Lett. **34**, 748–750 (1998)
5. Jungwattanaki, J., Reodecha, M., Chaovalitwongse, P., Werner, F.: An evaluation of sequencing heuristics for flexible flowshop scheduling problems with unrelated parallel machines and dual criteria. Otto-von-Guericke-Universitat Magdeburg, Preprint 28/05, pp. 1–23 (2005)
6. Kankanhalli, M.S., Ramakrishnan, K.R.: Content based watermarking of images. In: Proceedings of the 6th International Conference on Multimedia, pp. 61–70 (1998)
7. Laird, J., Rosen, M., Pelz, J., Montag, E., Daly, S.: Spatio-velocity CSF as a function of retinal velocity using unstabilized stimuli. Proc. SPIE **6057**, 32–43 (2006)
8. Langelaar, G.C., Lagendijk, R.L.: Optimal differential energy watermarking of DCT encoded images and video. IEEE Trans. Image Process. **10**(1), 148–158 (2001)
9. Legge, G.E.: A power law for contrast discrimination. Vision Res. **21**, 457–467 (1981)
10. Ling, H.F., Lu, Z.D., Zou, F.H.: Improved differential energy watermarking (IDEW) algorithm for DCT-encoded imaged and video. In: Proceedings of the 6th International Conference on Signal Processing, pp. 2326–2329 (2004)
11. Ling, H.F., Lu, Z.D., Zou, F.H., Li, R.X.: An energy modulated watermarking algorithm based on Watson perceptual model. J. Softw. **17**(5), 1124–1132 (2006)

12. Methodology for the subjective assessment of the quality of television pictures. ITU-R BT.500-11 (2002)
13. Podilchuk, C.I., Zeng, W.: Image-adaptive watermarking using visual models. Proc. IEEE **16**, 525–539 (1998)
14. Qiao, L., Cox, I.J.: Using perceptual models to improve fidelity and provide invariance to valumetric scaling for quantization index modulation watermarking. IEEE Trans. Inf. Forensics Secur. **2**(2), June (2007)
15. Schütz, A.C., Delipetkos, E., Braun, D.I., Kerzel, D., Gegenfurtner, K.R.: Temporal contrast sensitivity during smooth pursuit eye movements. J. Vis. **7**(13), 1–15 (2007)
16. Tourancheau, S., Callet, P.L., Barba, D.: Influence of motion on contrast perception: supra-threshold spatio-velocity measurements. Proc. SPIE, **6492** (2007)
17. Watson, A.B.: A technique for visual optimization of DCT quantization matrices for individual images. Soc. Information Display Dig. Tech. Papers XXIV, pp. 946–949 (1993)
18. Wei, Z., Ngan, K.N.: Spatial Just Noticeable Distortion Profile for Image in DCT Domain. In: Proceedings of IEEE International Conference on Multimedia and Expo, pp. 925–928 (2008)
19. Wolfgang, R.B., Podilchuk, C.I., Delp, E.J.: Perceptual watermarks for digital images and video. Proceedings of the IEEE, Special Issue on Identification and Protection of Multimedia Information 87, pp. 1108–1126 (1999)
20. Zhang, X.K., Lin, W.S., Xue, P.: Improved estimation for just-noticeable visual distortion. Signal Process. **85**(4), 795–808 (2005)
21. Zou, F.H.: Research of robust video watermarking algorithms and related techniques. A Dissertation Submitted for the Degree of Doctor of Philosophy in Engineering, Hua zhong University of Science & Technology (2006)