Primal Sketch Based Adaptive Perceptual JND Model for Digital Watermarking

Yana Zhang[1,2], Cheng Yang[1], Qi Zhang[1], Pamela Cosman[2]

[1]Communication University of China, Beijing, China
[2]Department of Electrical and Computer Engineering, University of California, San Diego, USA

***Abstract*—Watermarking algorithms based on the Just Noticeable Distortion (JND) model show great superiority over other methods, in terms of the watermark imperceptibility and robustness. However, the existing wavelet-based JND models are based on global coefficients, without consideration of image content characteristics. Following the pixel-wise masking idea, a primal sketch based adaptive perceptual JND model (PSAPM) is proposed in this paper, in which an improved watermarking algorithm is designed. It better describes the behavior of the watermark embedder. Experiments show that our algorithm is robust against attacks, including cropping, noise addition and JPEG compression.**

Keywords—JND model; primal sketch; digital watermarking

I. INTRODUCTION

The most important aspects of digital image watermarking algorithms are imperceptibility and robustness. It is required in practice that the embedded watermark is invisible to normal observers and difficult to be removed. But robustness and perceptual transparency are two conflicting factors.

In digital media, the core content or region-of-interest remains basically the same when it is processed or attacked. For example, with a portrait of a person that has been JPEG compressed and cropped, the person (core content) is still recognizable after compression, but a tree in the background may be cropped out. If the watermark were embedded in the core content region, it would typically be robust. However, this kind of strategy would impair visual quality and also it may be difficult to define and locate core content. The Just Noticeable Distortion (JND) model can be utilized in this kind of strategy to reach a balance between imperceptibility and robustness. A typical JND model [1] is a function of local luminance, frequency and contrast. Zhang [2] gave another formula of luminance adaptation adjustment and a contrast masking incorporating block classification. Wei [3] considered gamma correction which was applied to compensate for the luminance adaptation effect for a more accurate result. Niu [4] exploited a combined JND model, using the visual saliency to modulate JND values. Barni's typical JND model in the DWT domain [5] is more consistent with the perceptibility of the human visual system (HVS) and is used widely in watermarking systems. Xie [6] and Zolghadrasli [7] optimized Barni's model in terms of accurate luminance sensitivity and texture masking. Baaziz [8] proposed a function of watermarking embedding strength based on luminance sensitivity and texture masking in stationary wavelet transforms. Due to its strong similarity to the way the HVS processes images, JND models in the DWT domain have demonstrated great advantages over the JND models in other domains. But the existing JND models in the DWT domain have not exploited content analysis. So our plan is to build a JND model based on both media content analysis and perceptual decomposition.

In this paper, a brief introduction of primal sketch theory is presented, and we propose a novel JND model based on a primal sketch for watermarking embedding. We conclude from the experiments that it results in better watermarking performance.

II. PRIMAL SKETCH BASED JND MODEL

Generally speaking, an image could be decomposed or expressed by pixels, as well as objects and scenes because of hierarchical perception. Following Marr's insight, Guo [9] proposed a generative image representation called primal sketch, which integrates two modeling parts. The first one explains the structural part of an image, such as object boundaries. The second one models the remaining textural part without distinguishable elements. They adopted an artist's terminology, calling those two "sketchable" and "non-sketchable" respectively.

However, Guo's model has always been used only in image synthesis and compression. Both the "sketchable" and "non-sketchable" portions are reconstructed by the image base and MRF (Markov Random Field) model. The reconstructed image has large distortion, as shown in Fig. 1(b) and (d) (referenced from [9] [10]). The similarity between the synthesis image and the original image depends on the completeness of the image base and that of the MRF model. When we turn to digital watermarking systems, it is not wise to introduce distortion in pre-processing the original image. Therefore our proposed model PSAPM is explained as shown in Fig. 2. Starting with an original image (Fig. 3a), we first model the structure template (Fig. 3b) by Gabor bases. The structural image is a binary image, and points out the sketch location. Second, we multiply the structural image by the original image to obtain the sketchable image (Fig. 3c). We introduce the term "un-sketchable image" which is defined as the difference between the original image

and the sketchable image. The un-sketchable image (Fig. 3d) is part of the original image, while the non-sketchable image in Guo's terminology is a synthesis image reconstructed by an MRF model. We consider that the original image is divided into the two portions: sketchable image and un-sketchable image, so that it can be reconstructed (Fig. 3f) seamlessly and losslessly. In addition, the computing time is much less than that of Guo's theory.

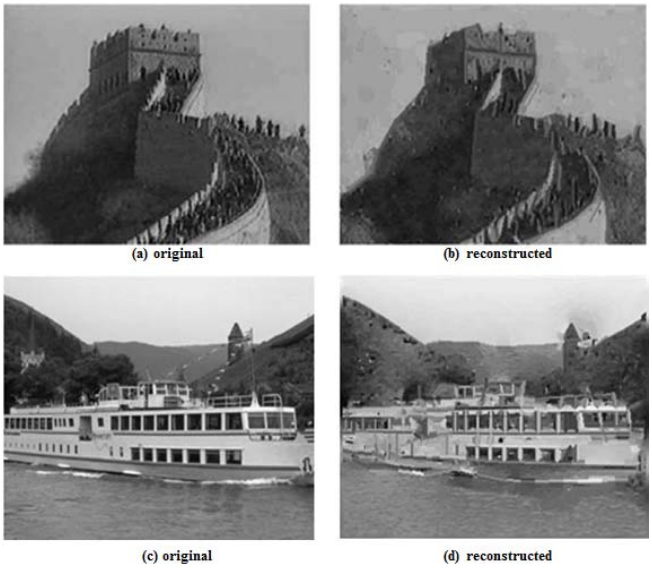


Fig. 1. Image expression based on primal sketch [9] [10]

Then we could analyze the masking effect of the un-sketchable image and that of the sketchable image, and get the JND values of the whole image. In Fig. 3e, brighter pixels indicate larger JND values. The basic characteristics of the HVS include the frequency sensitivity, luminance masking, contrast masking and texture masking. Some modifications of Barni's computing procedure are introduced here, in order to better describe the behavior of the HVS for the watermarking system. The 3-level Haar DWT is used as a perceptual decomposition.

$\mathrm{JND}_{s_l}^{\theta}(i,j)$ denotes the JND values of the sketchable image, and it is computed as the product of two items, depending on the orientation θ, resolution level l and location (i,j):

$$\mathrm{JND}_{s_l}^{\theta}(i,j) = \alpha \cdot \mathrm{Frq}_s(l,\theta) \cdot \mathrm{Lum}_s(l,i,j) \qquad (1)$$

As proposed in Barni's JND model [5], $\alpha = 0.5$. $\mathrm{Frq}_s(l,\theta)$ denotes the frequency sensitivity of the sketchable image. Considering that the sketchable image is full of high frequency and human eyes are less sensitive to high frequency, we modify the values for different orientations compared with those of Barni's model. Our experiments (same procedure as discussed in Sec. 4) showed that the values 1.5 if $\theta = \mathrm{HH}$ and 1.2 if $\theta = \mathrm{HL}$ or LH provide better results than $\sqrt{2}$ if $\theta = \mathrm{HH}$ and 1 if $\theta = \mathrm{HL}$ or LH as proposed by Barni.

$$\mathrm{Frq}_s(l,\theta) = \begin{Bmatrix} 1.5 & \text{if } \theta = \mathrm{HH} \\ 1.2 & \text{if } \theta = \mathrm{HL \text{ or } LH} \\ 1 & \text{if } \theta = \mathrm{LL} \end{Bmatrix} \cdot \begin{Bmatrix} 1.00 & \text{if } l = 1 \\ 0.32 & \text{if } l = 2 \\ 0.16 & \text{if } l = 3 \end{Bmatrix} \qquad (2)$$

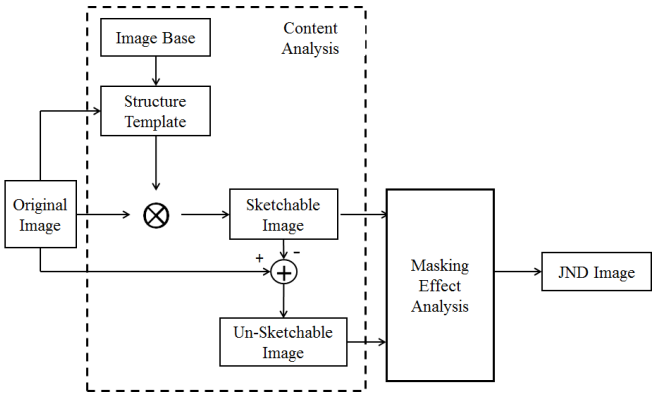


Fig. 2. Primal sketch based JND model

$\mathrm{Lum}_s(l,i,j)$ denotes the luminance masking of the sketchable image, and it is computed as in [5]:

$$\mathrm{Lum}_s(l,i,j) = 1 + \frac{1}{256} I_l^{LL}(i,j) \qquad (3)$$

$\mathrm{JND}_{u_l}^{\theta}(i,j)$ denotes the JND values of the un-sketchable image, and it is computed as the product of three terms:

$$\mathrm{JND}_{u_l}^{\theta}(i,j) = \alpha \cdot \mathrm{Frq}_u(l,\theta) \cdot \mathrm{Lum}_u(l,i,j) \cdot \mathrm{Tex}_u(l,i,j)^{\beta} \qquad (4)$$

As proposed in Barni's JND model, $\beta = 0.2$. We use $\mathrm{Frq}_u(l,\theta)$ and $\mathrm{Lum}_u(l,i,j)$ to denote the frequency sensitivity of the un-sketchable image and the luminance masking of the un-sketchable image respectively, and take the values from Barni's model:

$$\mathrm{Frq}_u(l,\theta) = \begin{Bmatrix} \sqrt{2} & \text{if } \theta = \mathrm{HH} \\ 1 & \text{others} \end{Bmatrix} \cdot \begin{Bmatrix} 1.00 & \text{if } l = 1 \\ 0.32 & \text{if } l = 2 \\ 0.16 & \text{if } l = 3 \end{Bmatrix} \qquad (5)$$

$$\mathrm{Lum}_u(l,i,j) = \begin{cases} 2 - L_u(l,i,j) & \text{if } L(l,i,j) < 0.5 \\ 1 + L_u(l,i,j) & \text{others} \end{cases} \qquad (6)$$

$$L_u(l,i,j) = \frac{1}{256} I_l^{LL}(i,j) \qquad (7)$$

The third term $\mathrm{Tex}_u(l,i,j)$, which measures the texture activity in the neighborhood of a coefficient, is computed as follows:

$$\mathrm{Tex_u}(l,i,j) = \sum_{k=0}^{3-l} \frac{1}{16^k} \sum_{\theta}^{HL,LH,HH} \sum_{x=-2}^{2} \sum_{y=-2}^{2} \left[I_{k+l}^{\theta}\left(y+\frac{i}{2^k}, x+\frac{j}{2^k}\right)\right]^2 \times \mathrm{Var}\{I_l^{LL}(i,j)\} \quad (8)$$

Here $\mathrm{Var}\{I_l^{LL}(i,j)\}$ is the local variance of the sub-band in a neighborhood (5×5 window in our model) corresponding to the location (i,j). The wider window shows more accurate variance, along with higher computation complexity, compared to Barni's model.

Finally, we combine the two functions together to get the JND value of every coefficient.

$$\mathrm{JND}_l^{\theta}(i,j) = m \times \mathrm{JND}_{u_l}^{\theta}(i,j) + n \times \mathrm{JND}_{s_l}^{\theta}(i,j) \quad (9)$$

We expect to embed a relatively large number of watermark bits, and need to put more bits in the un-sketchable area. When level l=3, we tried values of $m \in [0.9,1], n \in [0.01,0.1]$, with step sizes of 0.1 and 0.01 for m and n respectively. The experiments showed $m = 0.9$ and $n = 0.01$ provide the best results.

So we can see that an original image could be divided into a sketchable image (made up by textons) and an un-sketchable image (made up by textures). The JND image combines features of these two images, and points out the highest imperceptible threshold, so that if the watermark is inserted with that embedding intensity it would be transparent. Also, the JND image reflects the main image characteristics locally, so that if the watermark is embedded in these areas it would be robust.

Fig. 3. (a) original image (b) structure template (c) sketchable image (d) un-sketchable image (e) JND image (f) reconstructed image

III. WATERMARKING ALGORITHM

The digital watermarking algorithm will be designed based on the above JND model, including the following steps, shown in Fig. 4.

1) In the watermark generator, we obtain the random watermarking sequence (see Fig. 5) through binarization and chaotic scrambling based on the logistic function [11]:

$$X_{n+1} = \mu X_n (1 - X_n) \quad (10)$$

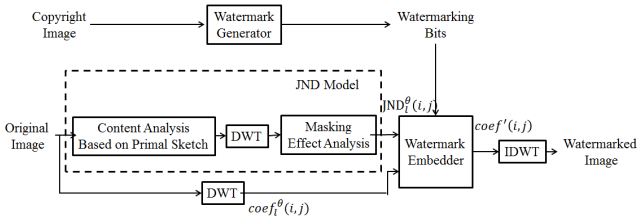


Fig. 4. Watermarking Embedding Procedure

When $3.5699456 < \mu < 4$, the logistic works in chaotic status. n denotes the index of the watermarking bits. In our experiments, $X_1 = 0.5$. After we obtain the chaotic sequence $\vec{X} = \{X_1, X_2, X_3, \cdots, X_n\}$, we need to sort it to get the reordered bits of the watermarking sequence.

(a) (b) (c)

Fig. 5. (a) copyright image (b) watermarking image (c) chaotic scrambling image

2) Use the primal sketch model to find the sketchable part I_s of an original image I. Leave the residual part as the un-sketchable portion I_u.

3) Transform the un-sketchable portion I_u and the sketchable portion I_s respectively into the DWT transform domain $\tilde{I}_u$ and $\tilde{I}_s$, in which the coefficients could be expressed as $\mathrm{dwt}_l^s(i,j)$, where l is level and s is orientation.

4) Use the proposed JND model to obtain the JND values.

5) The watermarking embedding strength should be modulated as:

$$s(i,j) = \alpha \times \mathrm{JND}(i,j) \quad (11)$$

α is a modulating factor, which is defined through experiments according to the application requirement.

6) The embedding strategy is expressed as:

$$\mathrm{coef}'(i,j) = \begin{cases} \mathrm{coef}(i,j) - \mathrm{mod}(\mathrm{coef}(i,j),\ s(i,j)) + \frac{s(i,j)}{4} & \text{if } W = 0 \ \&\ P = 1 \\ \mathrm{coef}(i,j) - \mathrm{mod}(\mathrm{coef}(i,j),\ s(i,j)) + \frac{3\times s(i,j)}{4} & \text{if } W = 1 \ \&\ P = 1 \\ \mathrm{coef}(i,j) & \text{if } P = 0 \end{cases} \quad (12)$$

7) After IDWT, we achieve the final watermarked image.

The watermark extraction scheme is non-blind. We use the original image and follow the similar steps for achieving the same intensity. If the remainder $R(i,j)$ lies in $(0, \frac{s(i,j)}{2}]$, the embedded watermark is 0. Otherwise if the remainder lies in $(\frac{s(i,j)}{2}, s(i,j)]$, the embedded watermark is 1.

$$R(i,j) = \mathrm{mod}(\mathrm{coef}(i,j), s(i,j)) \quad (13)$$

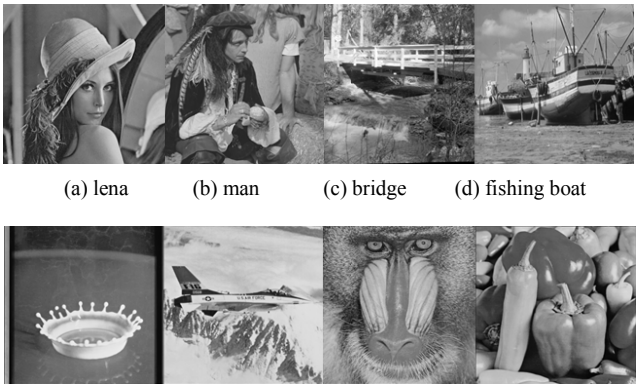
$$w = \begin{cases} 0 & \text{if } 0 < R(i,j) \le s(i,j)/2 \\ 1 & \text{if } s(i,j)/2 < R(i,j) \le s(i,j) \end{cases} \quad (14)$$

After anti-scrambling and undoing the sorting operation, we obtain the watermark.

IV. EXPERIMENTS AND RESULTS

To evaluate objectively the performance of our new JND model, we conduct a comparative test between the PSAPM model and Zolghadrasli's JND model [7] (which is the latest improvement of Barni's JND model in watermarking systems, and denoted Z.A. model here) in the same watermarking system given in Sec. 3. The conditions for them are kept as similar as possible. The watermark is embedded in the HL2 sub-band after 3-level DWT. The watermarking performance evaluation contains three aspects: watermark capacity, imperceptibility and robustness which are in conflict with each other. In our experiments, we compare BER (Bit Error Rate) under different kinds of attacks with different attacking intensities, while the two systems being compared have the same number of embedding bits and similar imperceptibility. Then we could analyze the effect of the different JND models. In this case, we set $m = 0.9, n = 0.01$ in the PSAPM scheme. When $\alpha = 1.15$ for PSAPM and $\alpha = 1.7$ for Z.A. measured by experiments, the two are similarly imperceptible, as will be discussed in the next subsection.

The eight original images with different characteristics and complexities are all JPEG format, size 512 x 512, 8-bit grayscale (Fig. 6). The copyright images are shown in Fig. 7 and their numbers of embedding bits are 128 x 128, 64 x 64 and 32 x 32.



(a) lena (b) man (c) bridge (d) fishing boat

(e) milk drop (f) plane (g) baboon (h) peppers

Fig. 6. Original Images

(a) chm.bmp (b) cuc.jpg (c) dell.jpg

Fig. 7. Copyright Images

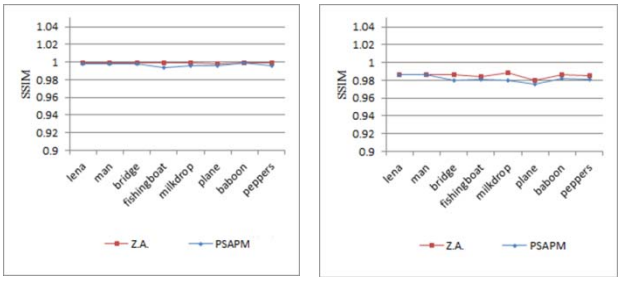
A. Imperceptibility Assessment

We use DSCQS (Double Stimulus Continuous Quality Scale) for subjective perception assessment. Fifteen observers rated the quality of watermarked images in the same environment (according to Rec. ITU-R BT.500-11T). As shown in Fig. 8 and Table I, the two watermarked images have similar imperceptibility.

(a) original image (b) PSAPM model (c) Z.A. model

Fig. 8. Watermarked Images

We use SSIM (Structural Similarity Index Measurement) for objective perception assessment. When the difference in SSIM values is less than 0.02, we conclude the images have similar perceptual quality. As shown in Fig. 9 and Table 1, the two watermarked images have similar imperceptibility. So our imperceptibility assessment showed that the modulating factors $\alpha = 1.15$ for PSAPM and $\alpha = 1.7$ for Z.A. are appropriate. Under this circumstance we can compare robustness fairly.



(a) embedding 32 x 32 bits (b) embedding 128 x 128 bits

Fig. 9. SSIM Comparison

TABLE I. IMPERCEPTIBILITY ASSESSMENT RESULTS

(EMBEDDING 64 X 64 BITS)

Images	SSIM value		subjective perception assessment value	
	PSAPM	Z.A.	PSAPM	Z.A.
lena	0.994	0.996	4.7	4.6
man	0.994	0.996	4.8	4.8
bridge	0.996	0.996	4.9	4.8
fishingboat	0.989	0.995	4.8	4.6
milkdrop	0.990	0.996	4.5	4.6
plane	0.990	0.993	4.4	4.5
baboon	0.996	0.996	5.0	5.0
peppers	0.990	0.995	4.7	4.8

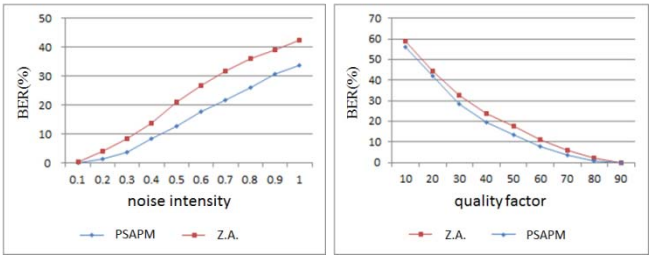
B. *Robustness Assessment*

We use Stirmark 4.0 for the robustness assessment, choosing 4 different kinds of attacks, each with 10 different levels of attacking intensity (shown in Table 2). To analyze robustness, we compare the BERs between the embedded watermark and extracted watermark. We also analyze the adaptability of the PSAPM model for different types of images.

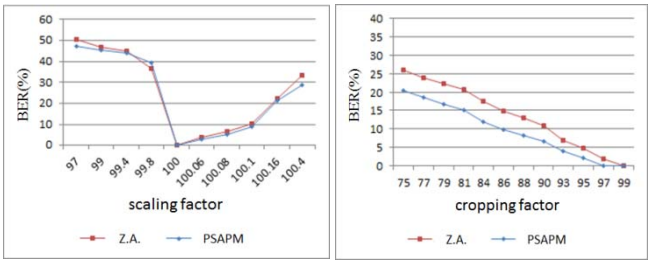
TABLE II. ATTACKING PARAMETERS IN ROBUSTNESS ASSESSMENT

Attack type	Attacking Intensity	Comment
additive noise	start=0.1,end=1,step=0.1	dB
JPEG Compression	start=10,end=90,step=10	Smaller quality factor means compression ratio is larger.
scaling	list=97,99,99.4, 99.8,100,100.06, 100.08,100.1, 100.16,100.4	When this factor is greater than 100, it is enlargement.
cropping	list=75,77,79,81, 84,86,88,90,93,95,97,99	Smaller cropping factor means more image cut.

Here we only present the experiments with embedding 64 x 64 bits. Fig. 10 (a) and (c) are the BER results of the image "milk drop", and (b) and (d) are for "peppers" and "fishing boat" respectively. Our algorithm with PSAPM is more robust than the watermarking scheme with the Z.A. model at surviving under additive noise, JPEG compression and cropping attacks. PSAPM is equal to the Z.A. model in defending against scaling attacks. Fig. 11 and Fig. 12 show the BER results for all 8 images. In Fig. 11(a) and (b), "milk drop" has larger BER values than other images do, because simple images with less fine detail are more noise sensitive. In Fig. 12(a), "baboon" has larger BER values than other images do, because complicated images are more sensitive to distortions of location. In Fig. 12(b), "plane" has lower BER values than the other images, because the main subject is salient in the central portion of the image, so it is robust to a cropping attack which cuts the surrounding portion. As a result, our proposed JND model works with complicated images in additive noise and JPEG compression attacks. It is fit for simple images in scaling attacks. For the images with the core content in the central area, the PSAPM model can defend against cropping attacks when the cropping intensity is lower than 25%.

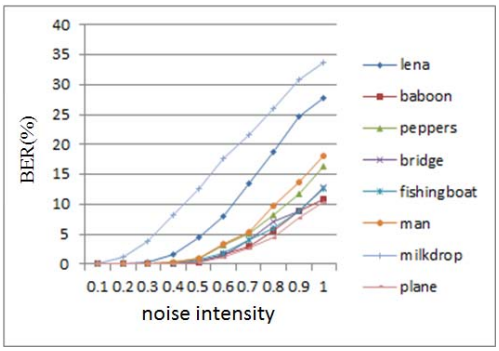


(a) additive noise (b) JPEG compression

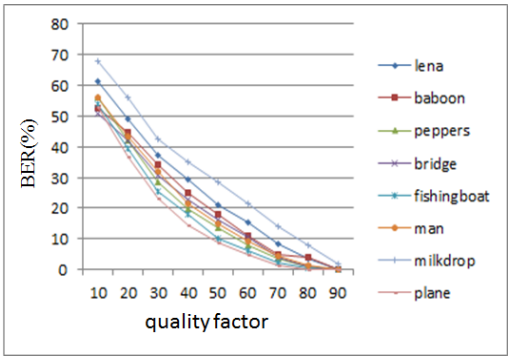


(c) scaling (d) cropping

Fig. 10. BER for different attacks (with embedding 64 x 64 bits)

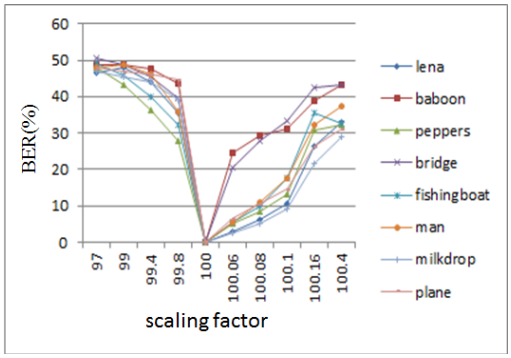


(a) additive noise

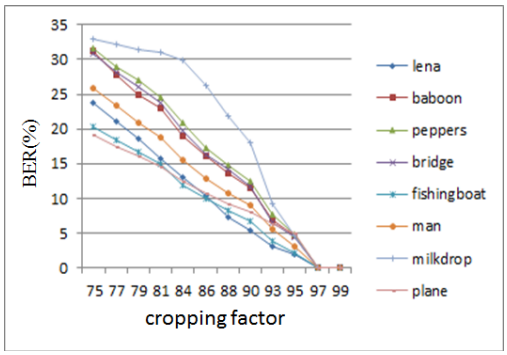


(b) JPEG compression

Fig. 11. BER of PSAPM for different types of images (with embedding 64 x 64 bits) for additive noise and JPEG attacks



(a) scaling



(b) cropping

Fig. 12. BER of PSAPM for different types of images (with embedding 64 x 64 bits) with scaling and cropping attacks

V. CONCLUSIONS

In this paper we introduce an image decomposition method based on a primal sketch into a watermarking system, so that an original image is divided into a sketchable image and an un-sketchable image. The JND image reflects the main image characteristics locally. We demonstrate through experiments that this method is suitable for watermarking systems, and it outperforms the JND model of [7]. Through the Haar wavelet transform, we discuss perceptual masking in different levels of content regions. Experimental results show that digital watermarking algorithms based on our PSAPM model would survive in cropping, additive noise and JPEG compression attacks.

VI. ACKNOWLEDGEMENT

The work on this paper was supported by the National Nature Science Foundation of China (60902061), the National Key Technology R&D Program (2012BAH17F01、2012BAH37F03、2012BAH02B03), the National Culture S&T Promotion Program (WHB1002), the China 863 Project (2012AA011702) and the SARFT Scientific and Technological Project (2012-20, 2011-34, 2012-27), the New Century Excellent Talents in University.

REFERENCES

[1] A. B. Watson. "DCT quantization matrices visually optimized for individual images." in SPIE Proceedings (1993): 202-216.

[2] X. H. Zhang, W. S. Lin, and P. Xue. "Improved estimation for just-noticeable visual distortion." Signal Processing 85.4 (2005): 795-808.

[3] Z. Y. Wei and K. N. Ngan. "Spatio-temporal just noticeable distortion profile for grey scale image/video in DCT domain." IEEE Trans. on Circuits and Systems for Video Technology, 19.3 (2009): 337-346.

[4] Y. Q. Niu, M. Kyan, S. Krishnan, and Q. Zhang. "A combined just noticeable distortion model–guided image watermarking." Signal, Image and Video Processing 5.4 (2011): 517-526.

[5] M. Barni, F. Bartolini, and A. Piva. "Improved wavelet-based watermarking through pixel-wise masking." IEEE Transactions on Image Processing, 10.5 (2001): 783-791.

[6] G. Xie and H. Shen. "Toward improved wavelet-based watermarking using the pixel-wise masking model." IEEE Intl. Conf. on Image Processing (ICIP), 2005. Vol. 1. (2005): I-689.

[7] A. Zolghadrasli, and S. Rezazadeh. "Evaluation of spread spectrum watermarking schemes in the wavelet domain using HVS characteristics." Signal Processing and Its Applications, 2007. 9th International Symposium on. IEEE, (2007): 1-4.

[8] N. Baaziz, D. Zheng, and D. Wang. "Image quality assessment based on multiple watermarking approach." IEEE 13th International Workshop on. Multimedia Signal Processing (MMSP), (2011): 1-5.

[9] C. E. Guo, S. C. Zhu, and Y. N. Wu. "Towards a mathematical theory of primal sketch and sketchability." Proc. 9th IEEE Intl. Conference on Computer Vision, (2003): 1228 - 1235.

[10] C. E. Guo, S. C. Zhu, and Y. N. Wu. "Primal sketch: Integrating structure and texture." Computer Vision and Image Understanding 106.1 (2007): 5-19.

[11] T. Xie and H. Xing. "New color image encryption scheme based on chaos." Application Research of Computers, vol. 30 No. 1, (2013): 318-320.