JOURNAL OF KOREA MULTIMEDIA SOCIETY, VOL. 8, NO. 12, DECEMBER 2005(pp. 1631-1637)

Parameter Estimations of ML Test Based Decoders for Perceptually Watermarked Images

Jingeol Lee[†]

ABSTRACT

Based on the generalized Gaussian pdf of DCT coefficients of images, Hernandez et al. propose the ML test applied watermark decoder. For images with watermarks shaped by the visibility thresholds of DCT coefficients and the luminance masking of human visual system, they conclude that the ML test with an appropriately chosen parameter associated with the pdf of DCT coefficients outperforms the correlation based decoder. In this paper, the parameter is estimated using various methods including a novel one for watermarks shaped by the visibility thresholds of DCT coefficients and the luminance masking as Hernandez et al. did and with the contrast masking added, and its effect on performance is compared.

Keywords: Image Watermarking, Human visual System

1. INTRODUCTION

As digital contents flood into internet, watermarkings have attracted attention as means for copy right protection. It has been well known that there are contradictory requirements on the watermarks, that is, robustness and imperceptibility. Recently, human perceptual models are applied to audio and image watermarks to compromise these requirements. In the case of image watermark, the Watson's visual model have been applied, which consists of three components, i.e., the visibility thresholds of DCT coefficients, the luminance masking, and the contrast masking [1]. Podilchuk and Zeng propose the still image watermarking adopting the Watson's visual model, which detects watermarks by normalized correlation coefficient on the assumption that the original images are available [2]. For the more practical situations under which the original images are not available, Hernandez et al. propose the ML test based decoder which relies on the fact that DCT coefficients of images are better modeled by generalized Gaussian than Gaussian pdf [3]. They apply the ML test to images with watermarks shaped by the visibility thresholds of DCT coefficients and the luminance masking of human visual system, and conclude that the ML test with an appropriately chosen parameter associated with the pdf outperforms the correlation based decoder. They also introduce two estimates of the parameter; one is derived by matching the sample mean and sample variance of the DCT coefficients, and the other is by applying ML estimate. However, the estimates are not proved to enhance the performance of the decoder. Besides, they do not adopt the contrast masking by which the signal strength of watermark can be increased such that the robustness of watermark is improved.

In this paper, the parameter is estimated using above stated methods and a novel one for watermarks shaped by the visibility thresholds of DCT coefficients and the luminance masking as Hernandez et al. did and with the contrast masking

※ Corresponding Author : Jingeol Lee, Address : (302-735) Department of Electronic Engineering, Paichai University, 439-6 Doma-dong, Seo-gu, Taejon, TEL : +82-42-520-5707, FAX : +82-42-520-5773
E-mail : jingeol@mail.paichai.ac.kr
Receipt date : Mar. 24, 2005, Approval date : July. 26, 2005
[†] Department of Electronic Engineering, Paichai University

added, and its effect on performance is compared.

Section 2 summarizes the Watson's visual model, and Section 3 reviews the watermark decoder proposed by Hernandez. Section 4 presents the existing and the novel parameter estimations. Section 5 presents the bit error rate (BER) measurements for the decoder under various conditions including adoption of specific masking model and the estimates of the parameter. Section 6 concludes the experimental results.

2. REVIEW OF THE VISUAL MODEL

The Watson's visual model consists of three components, i.e., the visibility thresholds of DCT coefficients, the luminance masking, and the contrast masking for block size of 8 by 8 pixels [1]. The visibility thresholds are maximum allowable variations of DCT coefficients by which visual artifacts are not perceptible, and are directly related to contrast sensitivity function (CSF) of human visual system. Peterson et al. propose experimental methods to measure the thresholds, and Ahumada et al. propose the formula which provides with the thresholds according to the display conditions [4, 5]. The luminance masking and the contrast masking allow higher thresholds to brighter regions of images and to higher DCT coefficients, respectively. In other words, the luminance masking models the masking effect in spatial domain, and the contrast masking in frequency domain. The luminance masking is modeled by

$$t_{ijk} = t_{ij}(c_{00k}/\bar{c}_{00})^{\alpha_T} \tag{1}$$

where t_{ij} is the visibility threshold for a coefficient of the DCT indexed by i and j, and c_{00k} is the DC coefficient for block k, and $\bar{c}_{00}$ is the DC coefficient corresponding to brightness of 128 in 8 bits image, and α_T is a parameter that controls the degree of the luminance masking. The typical value of α_T is 0.649. The contrast masking is modeled by

$$t'_{ijk} = \max(t_{ijk},\ \left|c_{ijk}\right|^{w_{ij}} t_{ijk}^{\ 1-w_{ij}}) \tag{2}$$

where w_{ij} is a parameter that controls the degree of the contrast masking, and it can take different value for each DCT coefficient. The typical value of w_{ij} is 0.7. The luminance masking affects all thresholds by the same amount according to the average brightness within a block, whereas the contrast masking raises the thresholds in case corresponding DCT coefficients are greater than the thresholds determined by the visibility thresholds and luminance masking, and holds the thresholds, otherwise. Thus, the adoption of the contrast masking permits large signal level of watermark, which leads to the improvement in robustness of watermark.

3. REVIEW OF ML TEST BASED DECODER

It is well known that the distribution of DCT coefficients of images fits in generalized Gaussian pdf given by [3]

$$f_x(x) = Ae^{-|\beta x|^c}. \tag{3}$$

A and β are functions of c and standard deviation σ as follows:

$$\beta = \frac{1}{\sigma}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}, \quad A = \frac{\beta c}{2\Gamma(1/c)} \tag{4}$$

where Γ is the gamma function. The generalized Gaussian distribution includes the Gaussian and the Laplacian distribution corresponding to $c=2$ and $c=1$, respectively.

The perceptually watermarked image and its ML test in the decoder based on the distribution are given by

$$Y[\mathbf{k}] = X[\mathbf{k}] + b_i t[\mathbf{k}] s[\mathbf{k}] \tag{5}$$

$$r_i = \sum_{\mathbf{k}\in S_i} \frac{\left|Y[\mathbf{k}] + t[\mathbf{k}]s[\mathbf{k}]\right|^{c[\mathbf{k}]} - \left|Y[\mathbf{k}] - t[\mathbf{k}]s[\mathbf{k}]\right|^{c[\mathbf{k}]}}{\sigma[\mathbf{k}]^{c[\mathbf{k}]}}, \tag{6}$$

where $\mathbf{k}$ represents 2- D indexes, and S_i is the set of DCT coefficients associated with an embedded bit $b_i \in \{-1,1\}$. $X[\mathbf{k}]$ is the DCT coefficients for an original image, $t[\mathbf{k}]$ is the masking thresholds, and $s[\mathbf{k}]$ is the pseudorandom sequences whose values depend on a secret key. The first and the second term in the numerator in (6) corresponds to the joint probability density corresponding to the hypothesis of $b_i = -1$ and $b_i = 1$, respectively. The decoded value of the bit b_i is determined by

$$\hat{b}_i = sign(r_i). \tag{7}$$

The ML test given by (6) includes the correlation normalized by the variance in case of $c[\mathbf{k}] = 2$.

4. ESTIMATES OF THE PARAMETER *C*

The value of $c[\mathbf{k}]$ in (6) can be adaptively estimated by matching the sample mean and the sample variance of DCT coefficients [3, 6]

$$c = F^{-1}\left(\frac{E[|X[\mathbf{k}]|]}{\sigma[\mathbf{k}]}\right), \tag{8}$$

where

$$F(c) = \frac{\Gamma(2/c)}{\sqrt{\Gamma(1/c)\Gamma(3/c)}}. \tag{9}$$

The other estimate of $c[\mathbf{k}]$ can be obtained by applying ML estimation, which results in finding a root of the following equation [7]:

$$\frac{\psi(1/c+1)+\log(c)}{c^2} + \frac{1}{c^2}\log\left(\frac{1}{N}\sum_{i=1}^{N}|X_i|^c\right) - \frac{\sum_{i=1}^{N}|X_i|^c \log|X_i|}{c\sum_{i=1}^{N}|X_i|^c} = 0, \tag{10}$$

where

$$\psi(\tau) = \frac{\partial \log\Gamma(\tau)}{\partial \tau}, \tag{11}$$

and N is the number of DCT coefficients. Hernandez et al. state that "note that even though the proposed procedures for estimating the c parameter are not equivalent to finding the value of c which optimizes the decoding or detection performance, they can be expected to lead to similar results" [3]. However, as will be shown in the experimental results, the existing estimates do not contribute to the performance improvements in terms of BER in the most of cases. Thus, the novel estimate, which produces the minimum BER, is proposed.

Assuming the outputs of the decoder as the Gaussian random variables, The SNR of the decoder is given by [3]

$$SNR = \frac{(E[r_i])^2}{Var[r_i]} = \frac{\left(\frac{1}{N}\sum_{\mathbf{k}}\frac{E[r[\mathbf{k}]]}{\sigma[\mathbf{k}]^{c[\mathbf{k}]}}\right)^2}{\frac{1}{N}\sum_{\mathbf{k}}\frac{Var(r[\mathbf{k}])}{\sigma[\mathbf{k}]^{2c[\mathbf{k}]}} + \frac{N-1}{N^2}\sum_{\mathbf{k}}\frac{E^2[r[\mathbf{k}]]}{\sigma[\mathbf{k}]^{2c[\mathbf{k}]}}} \tag{12}$$

where

$$E[r[\mathbf{k}]] = \frac{1}{2}\left[(|X[\mathbf{k}]|+2t[\mathbf{k}])^{c[\mathbf{k}]} + \left||X[\mathbf{k}]|-2t[\mathbf{k}]\right|^{c[\mathbf{k}]}\right] - |X[\mathbf{k}]|^{c[\mathbf{k}]} \tag{13}$$

$$Var(r[\mathbf{k}]) = \frac{1}{4}\left[(|X[\mathbf{k}]|+2t[\mathbf{k}])^{c[\mathbf{k}]} - \left||X[\mathbf{k}]|-2t[\mathbf{k}]\right|^{c[\mathbf{k}]}\right]^2, \tag{14}$$

and N is the number of bits. The BER can be expressed in terms of SNR for the bit by bit hard decoder as

$$\text{BER} = Q\left(\sqrt{SNR}\right) \tag{15}$$

where

$$Q(x) = \left(1/\sqrt{2\pi}\right)\int_x^{\infty} e^{-t^2/2}dt \tag{16}$$

It is certain with (15) and (16) that the maximum SNR leads to the minimum BER. Accordingly, an optimum estimate of c, which maximizes the SNR, can be derived by

$$c_{opt} = \max_c \; SNR \tag{17}$$

The proposed estimation relies on the decoder outputs such that its best performance is achieved whereas the existing ones are derived from the distribution of DCT coefficients. As Hernandez et al. state, the estimates derived from the distribution of DCT coefficients do not correspond to the optimum value of c leading to the best performance of the decoder. The proposed estimate formulated as an optimization is justified in the sense that it produces the maximum SNR at the decoder output, leading to the best performance in terms of BER.

5. EXPERIMENTAL RESULTS

The visibility thresholds of DCT coefficients for block size of 8 by 8 pixels are measured by the method proposed by Peterson et al. [4], and the masking thresholds are determined according to the Watson's visual model [1]. The pseudorandom sequences have two level discrete marginal distribution, $s[\mathbf{k}] \in \{-1, 1\}$. The 22 DCT coefficients where watermarks are embedded are shown as shaded in Fig. 1, which is the same condition as set by Hernandez et al. for the comparison [3]. The test images of 256 by 256 pixels, which have different features, are shown in Fig. 2.

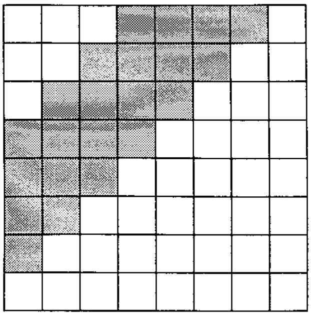

Fig. 1. DCT coefficients where watermarks are embedded.

The images are divided into blocks of 8 by 8 pixels, and the DCTs are taken for every block, by which 1024 coefficients are available for embedding a bit. A bit, $b_i \in \{-1, 1\}$ is multiplied to the masking thresholds and random sequence, and embedded in the same frequency component of the blocks.

The BER of the decoders are measured for $c = 1/2$, $c = 1$, $c = 2$, and estimates of c for the images with watermarks shaped by the visibility thresholds, the luminance masking and the contrast masking. The measurements shown in Table 1 are taken using 100 random sequences with $b_i = -1$ and

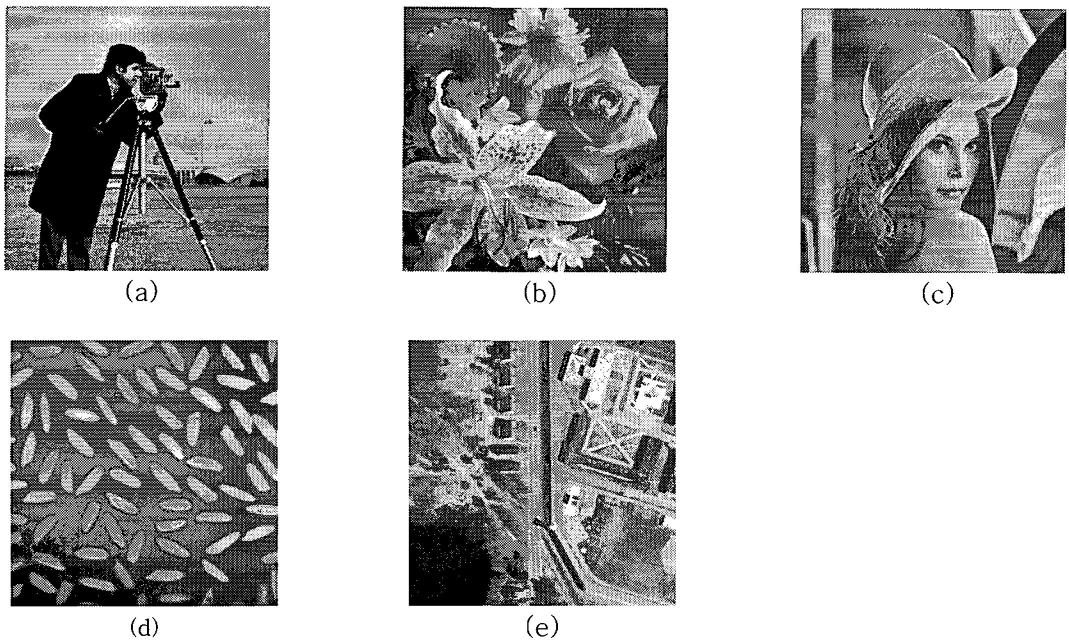

Fig. 2. Test images (a) Cameraman (b) Flowers (c) Lena (d) Rice (e) Aerial.

Table 1. BER measurements.

	$c=1/2$	$c=1$	$c=2$	Est_1	Est_2	Est_3
Cameraman	0.000(0.068)	0.000(0.037)	0.081(0.118)	0.000(0.086)	0.000(0.096)	0.000(0.031)
Flowers	0.022(0.528)	0.035(0.159)	0.100(0.037)	0.028(0.324)	0.025(0.370)	0.020(0.032)
Lena	0.001(0.118)	0.001(0.047)	0.065(0.060)	0.001(0.112)	0.001(0.131)	0.000(0.034)
Rice	0.000(0.009)	0.000(0.041)	0.053(0.013)	0.000(0.036)	0.000(0.005)	0.000(0.008)
Aerial	0.141(0.805)	0.073(0.228)	0.094(0.029)	0.088(0.316)	0.084(0.290)	0.064(0.028)

$b_i = 1$, respectively, which results in the BER as the number of error bits divided by 4400 bits. The Est_1, Est_2, Est_3 in Table 1 are estimates of c by matching the sample mean and the sample variance of DCT coefficients, by applying the ML estimation, and by applying the proposed estimation, respectively. The measurements for the cases that watermarks are shaped by the visibility thresholds and the luminance masking are compared with those in parenthesis for the cases that the contrast masking is added. From the measurements, it is certain that the performances with $c=1/2$ or $c=1$ are superior to those with $c=2$ for the cases that watermarks are shaped by the visibility thresholds and the luminance masking, which is the same as Hernandez et al. concluded. The performance comparison for the estimates shows that the proposed estimation provides the best performance for the most of cases except Rice with the watermark shaped with the contrast masking added, whereas existing ones do not contribute to performance improvement and produce even worse performances than those resulted from using constant values of c for the most of cases.

Fig. 3 shows the distributions of values of the parameter c derived by the proposed estimation for

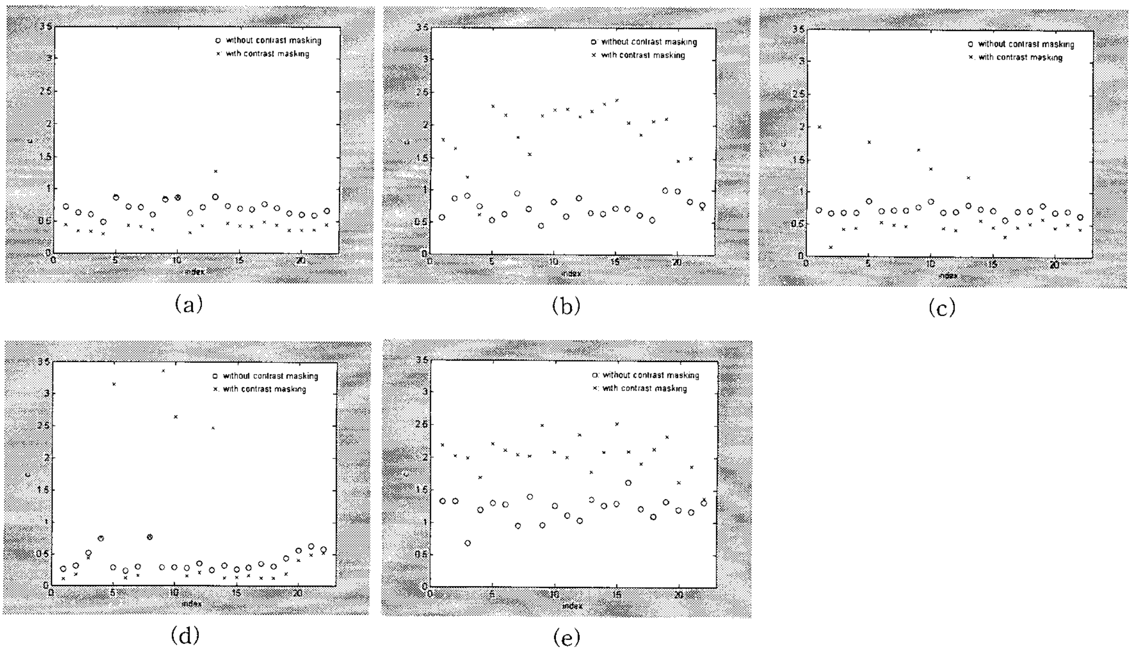

Fig. 3. Distributions of values of the parameter (a) Cameraman (b) Flowers (c) Lena (d) Rice (e) Aerial.

the test images with and without the contrast masking. For every test image, the values of the parameter without the contrast masking are different from 2 corresponding to the Gaussian pdf, which explains the Hernandez et al.'s conclusion. Besides, it is demonstrated that the optimum values of c without the contrast masking are different from ones with the contrast masking. Such the tendencies are observed in the experimental results for the constant values of c in Table 1. For example, in Flowers, without the contrast masking, $c=1/2$ produces the best performance among those resulted from the three constants whereas with the contrast masking, $c=2$ produces the best. However, the existing estimations relying on the distribution of DCT coefficients of original images do not depend on the adopted masking model. In addition to the BER measurements, this supports the justification of the proposed estimation.

6. CONCLUSIONS

The novel parameter estimation, which relies on the ML test based decoder outputs such that its best performance in terms of BER is achieved, is proposed, and its performance is experimentally demonstrated. The experimental results show that the proposed estimation provides the superior performances to those using the existing ones for nine out of ten cases given by the combination of the 5 test images with and without the contrast masking. Even the worst performance with the proposed estimation for Rice with the contrast masking is comparable to the minimum BER resulted from using the existing estimation. In contrast to the proposed estimation, it is shown that existing ones do not contribute to performance improvement for the most of cases. Also, it is demonstrated that the proposed estimation produces different values depending on whether the contrast masking is added, which is not the case in the existing ones.

7. REFERENCES

[1] Andrew B. Watson, "DCT Quantization Matrices Visually Optimized for Individual images," *Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display IV*, Vol. 1913, pp. 202-216, 1993.

[2] Christine I. Podilchuk, and Wenjun Zeng, "Image Adaptive Watermarking Using Visual Models," *IEEE Journal on Selected Areas in Communications*, Vol. 16, No. 4, pp. 525-539, May 1998.

[3] Juan R. Hernandez, Martin Amado, and Fernando Perez Gonzalez, "DCT-Domain Watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure," *IEEE Transactions on Image Processing*, Vol. 9, No. 1, pp. 55-68, Jan. 2000.

[4] H. A. Peterson, H. Peng J. H. Morgan, and W. B. Pennebaker, "Quantization of Color Image Components in the DCT Domain," *Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display II*, Vol. 1453, pp. 210-222, 1991.

[5] Albert J. Ahumada Jr., and Heidi A. Peterson, "Luminance Model-Based DCT Quantization for Color Image Compression," *Proc. SPIE Conf. Human Vision, Visual Processing, and Digital Display III*, Vol. 1666, pp. 365-374, 1992.

[6] Stephane G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 7, pp. 674-693, Jul. 1989.

[7] F. Muller, "Distribution Shape of Two Dimensional DCT Coefficients of Natural Images," *Electronics Letters*, Vol. 29, No. 22, pp. 1935-1936, Oct. 1993.

Jingeol Lee

He received the Bachelor of Engineering degree; the Master of Science degree from Korea University in 1981 and 1985, respectively; and the Doctor of Philosophy degree from University of Florida in 1994. From 1982 to 1990, he was with Agency for Defense Development, and from 1995 to 1996, he had worked for Samsung Electronics, where he was involved in research and development of military electronics throughout his job experiences. Since 1997 he has been with the Department of Electronic Engineering, Paichai University. His professional interests include watermarkings and acoustics.