WAVELET DOMAIN WATERMARKING CAPACITY ANALYSIS

Zhang Fan, Zhang Hongbin

Computer institute, Beijing University of Technology
Pingleyuan 100#, Chaoyang District, Beijing, China (Asia)
phone: +86 10 6739 2717, fax: +86 10 6739 2878, email: justzf@sina.com

ABSTRACT

Capacity is an important character of digital watermarking. Research of image watermarking capacity is to study how much information can be hidden in an image. Recently, wavelet transform has been applied widely in watermarking research as its excellent multi-resolution analysis property. According to watermarking methods in wavelet domain, watermark's embedding and extracting realized in wavelet domain, we think that watermarking capacity should be analyzed in wavelet domain also. This paper presents a watermarking capacity analysis method based on content of wavelet subbands by using Watson quantization matrix and Noise Visibility Function (NVF), and discusses watermarking capacity of blind watermarking and non-blind watermarking in this scenario.

1. INTRODUCTION

Image watermarking capacity is an evaluation of how much information can be hidden with in a digital image. Watermarking capacity is determined by the statistical model used for the host image, by the distortion constraints on the data hider and the attacker, and by the information available to the data hider, to the attacker, and to the decoder. The purpose of watermarking capacity research is to analyze the limit of watermark information while satisfying the invisibility and robustness of watermarking. The research of watermarking capacity has an important meaning for more effective watermarking algorithms to be designed.

Several works on watermarking capacity have been presented in recent years. Servetto considered each pixel as an independent channel and calculated the capacity based on the theory of Parallel Gaussian Channels (PGC) [1]. Barni's research focused on the image watermarking capacity in DCT and the DFT domain [2]. Moulin's work studied a kind of watermarking capacity problem under attacks [3,4]. Lin presented zero-error information hiding capacity analysis method in JPEG compressed domain using adjacency-reducing mapping technique [5, 6].

Image watermarking capacity is a complex problem. It may be influenced by many factors. The content of image has great influence to watermarking capacity. This influence has two aspects, on the one hand, we hide information using the content of image, on the other hand, in blind watermarking, the content of image becomes an obstacle when we detect watermark. We think that watermarking capacity should be associated with the content of image; different image has different watermarking capacity. Obviously, more watermark information can be transmitted in a complex image compare with a flat image such as a pure white color image. However, some previous works about watermarking capacity do not utilized the content of images fully, watermarking capacity is calculated using a given power signal-to-noise ratio.

Watermarking capacity can also be influenced by the watermark's strength, but a high strength watermark not always means a high watermarking capacity. For example, assume a watermarking algorithm is to add one to each pixel's amplitude value and then watermark information is one bit. If we embed a stronger watermark, to add ten to each pixel, and then watermark information is also one bit, only watermark's delectability is improved. Some previous works calculate each image pixel's watermarking capacity respectively or treat each image pixel as a separate channel. Obviously, these methods cannot explain above problem. When we calculate an image pixel's watermarking capacity, we cannot control the change of other pixels. These methods ignore the relationship among image pixels. But according to information theory, the diversity of amplitude value is very important for watermarking to transmit information.

Almost all previous works on watermarking capacity are realized in spatial domain. Recently, wavelet transform has been applied widely in watermarking research as its excellent multi-resolution analysis property. Watermarking algorithms based on wavelet become the major research direction. According to those watermarking methods, watermark's embedding and extracting realized in wavelet domain, we think that watermarking capacity should be analyzed in wavelet domain.

According to above analysis, we think that the watermark power should be constrained according to the content of wavelet subbands. In this paper, we present a content adaptive watermarking algorithm using Watson quantization matrix and Noise Visibility Function (NVF), and we discuss watermarking capacity problem of blind watermarking and non-blind watermarking in this scenario.

2. ADAPTIVE WATERMARKING

In watermarking schemes, image is considered as a communication channel to transmit messages, its power constrains on watermark are determined by the Human Vision System (HVS) model. HVS models have been studies for many

years. These models were explored to describe human vision mechanism such as spatial frequency orientation, sensitivity on local contrast, adaptation and masking etc.

Watson proposed a mathematical model for the Discrete Wavelet Transform (DWT) noise detection thresholds that is a function of level, orientation, and display visual resolution [7]. This allows calculation of a perceptually lossless quantization matrix for which all errors are in theory below the visual threshold. The quantization matrix is composed of quantization factor for each level and orientation. The quantization factor can be written as the equation (1):

$$Q_{\lambda,\theta} = \frac{2}{A_{\lambda,\theta}} a 10^{k\left(\log\frac{2^{\lambda} f_0 g_\theta}{r}\right)^2} \quad (1)$$

Where λ and θ denote wavelet level and orientation respectively, γ is display resolution, $A_{\lambda,\theta}$ denote the basis function amplitudes, and a, k, f_0, g_θ are constant.

Watson's perceptual model based on experience, satisfy the requirement of HVS. But this model is independent of the content of image. In image compression of wavelet domain, the quantization is uniform, wavelet coefficient be quantized using same quantization factor in same level. According to watermarking capacity, if we embed same watermark strength in a wavelet subband, watermarking capacity is the minimum, almost no information been transmitted. So for estimating the maximum information of watermark, we should design watermark strength based on the content of wavelet subbands.

According to analysis above, we can compute the perceptually lossless threshold using Watson's perceptual model, but we should associate it with other masking function to found a model to constrain watermarking embedding.

Recently, Voloshynovsky proposed a texture masking method, Noise Visibility Function (NVF)[8,9]. The NVF is the function that characterizes local image properties, identifying textured and edge regions where the watermark should be more strongly embedded. The NVF can be applied in either spatial domain or wavelet domain.

NVF can be written as the equation (2):

$$NVF(i,j) = \frac{w(i,j)\sigma_n^2}{w(i,j)\sigma_n^2 + \sigma_x^2(i,j)} \quad (2)$$

Where σ_n^2 denote noise variance, $\sigma_x^2(i,j)$ denote the local variance of the image in a window centered on the pixel with coordinates $(i,j), 1 \le i,j \le N$, $w(i,j)$ is a weighting function depends on shape parameter γ. $w(i,j)$ can be written as the equation (3):

$$w(i,j) = \gamma[\eta(\gamma)]^{\gamma} \frac{1}{\|r(i,j)\|^{2-\gamma}} \quad (3)$$

Where

$$r(i,j) = \frac{x(i,j) - \bar{x}(i,j)}{\sigma_x}, \quad \eta(\gamma) = \sqrt{\frac{\Gamma(\frac{3}{\gamma})}{\Gamma(\frac{1}{\gamma})}},$$

$\Gamma(t) = \int_0^{\infty} e^{-u} u^{t-1} du$, $\bar{x}(i,j)$ denote the mean of the image.

There are two kinds of NVF, based on either a non-stationary Gaussian model of the image, or a stationary Generalized Gaussian model. In the non-stationary Gaussian model, the data is assumed to be locally i.i.d. (Independent identically distribute) Gaussian process, while in the stationary Generalized Gaussian model the data is assumed to be globally i.i.d.. Because watermarking problem has a close relationship with the local properties of image, we think that the NVF based on non-stationary Gaussian model more suits for watermarking problem. Assume the host image is a Gaussian process. In the case of non-stationary Gaussian model, NVF can be written as the equation (4):

$$NVF(i,j) = \frac{1}{1 + \sigma_x^2(i,j)} \quad (4)$$

Once we have computed the NVF, we can obtain the allowable distortions of each wavelet coefficient by computing:

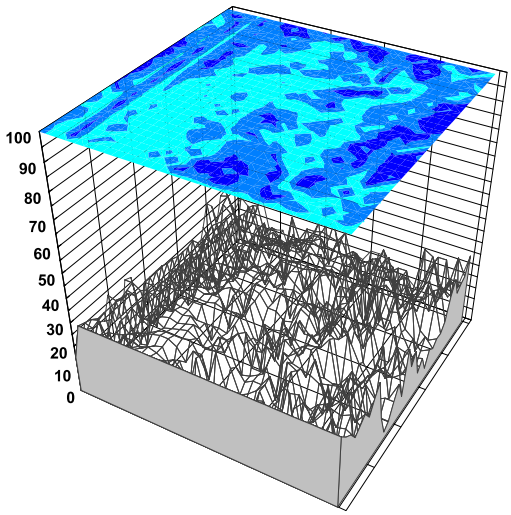
$$\Delta(i,j) = (1 - NVF(i,j)) \cdot S_0 + NVF(i,j) \cdot S_1 \quad (5)$$

Where S_0 and S_1 are the maximum allowable wavelet coefficient distortions in textured and flat region respectively. Typically S_0 may be as high as 30 while S_1 are usually about 3. In flat regions the NVF tends to 1 so the first term of equation (5) tends to 0, and consequently the allowable wavelet coefficient distortion dependent on S_1 that is small. Intuitively this makes sense since we expect that the watermark distortions will be visible in flat regions and less visible in textured regions. According to above equation, the watermark embedded in the texture or the edge regions is stronger than in the flat regions. If we embed maximum allowable watermark in each wavelet coefficient, the robustness of watermarking will have a good performance. By this way, we can achieve the best trade-off between robustness and invisibility.

According to equation (5) we can calculate the maximum allowable watermark amplitude value of each wavelet coefficient while keeping watermark's invisibility. All watermark amplitude value can build an image; we name the image Maximum Watermark Image (MWI). Σ_w^2 denote the variance of MWI. Figure (1) shows Lena and its Maximum Watermark Image.

In Figure (1), bright parts denote the regions where allow bigger watermark amplitude value; dark parts denote the regions where allow smaller watermark amplitude value. From the contour map on the top, an approximate outline of Lena can be identified. In the complex texture regions or in the edge regions, for example, the regions of Lena's hair, allow bigger watermark amplitude value. This case is coincides with our estimate.

(a)



(b)

Fig. 1. Lena (a) and its Maximum Watermark Image (b) with a contour map on the top.

3. WATERMARKING CAPACITY

Watermarking can be viewed as a form of communication and image can be considered as a communication channel to transmit messages. So, watermarking capacity problem can be solved using traditional information theory. Figure (2) is a model of watermarking communication in wavelet domain. In this model, we design a switch, which control if the host image can be used in watermarking extracting. The switch on denotes non-blind watermarking and the switch off denotes blind watermarking.

Some previous works on image watermarking capacity consider each pixel as an independent AWGN channel, and watermarking capacity is summation capacity of all channels. According to the analysis of section (2), we don't think so. We think that the image should be treated as an AWGN channel with power constraint P_S.

Many previous works using a given power constrain to calculate watermarking capacity. Obviously, these works did not utilize the content of images fully. Their result cannot express the influence of image's content and the difference between one image to another. We think that the power constraint P_S should be calculated based on the MWI, which is the strongest watermark while cannot be apperceived. Assume σ_w^2 denote the variance of the MWI and σ_n^2 denote the variance of noise, then, in non-blind watermarking scenario, image watermarking capacity can be written as:

$$C = W \log_2\left(1 + \frac{\sigma_w^2}{\sigma_n^2}\right) \tag{6}$$

Where W denote the bandwidth of channel. What is the bandwidth of an image? Assume the size of image is $N \times N$, the number of pixels is $M = N \times N$. According to Nyquist sampling theory, if we want to express all the pixels correctly, sampling points should be 2 W at least. So the bandwidth of an image is $W=M/2$.

Costa has studied the channel capacity problem of so-called dirty paper communication in 1983 [10]. The capacity problem of blind watermarking is the same as the problem described by Costa. According to Costa's work, the capacity of blind watermarking was the same as the non-blind watermarking cases. So, in both cases, watermarking capacity is same.

4. EXPERIMENTS

In experiment, the 256×256 standard test image Fishingboat is used. A biorthogonal 9/7 DWT was used to decompose the host image into four levels. Using the NVF, we can calculate the maximum allowable watermark amplitude value of each wavelet coefficient while keeping watermark's invisibility. A window of size 3×3 is used for NVF computing. We calculate the variance of the MWI and finally we can calculate image's wavelet domain watermarking capacity in condition of different noise power constrains. We also calculate the spatial domain watermarking capacity in same noise power constrains as contrast. The experiment result of watermarking capacity is shown in table (1) and figure (3).

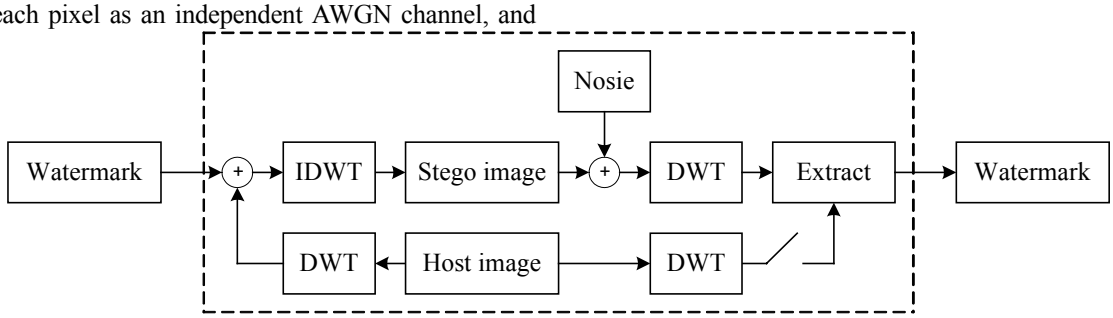


Fig. 2. Watermarking communication model in wavelet domain

Table 1. Watermarking capacity of Fishingboat in spatial domain and wavelet domain

σ_n^2	Spatial domain (bit)	Wavelet domain (bit)
1	128032	84813
2	110410	68364
3	98309	57914
4	89214	50543
5	82005	45006
6	76088	40664
7	71106	37151
8	66833	34239
9	63113	31783
10	59835	29677

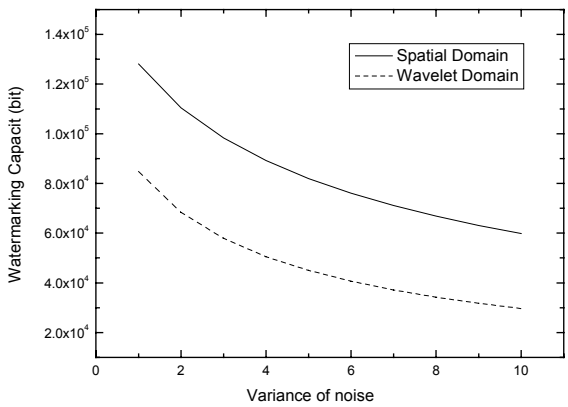


Fig. 3. Watermarking capacity of Fishingboat in spatial domain and wavelet domain

The result shows that watermarking capacity in wavelet domain is smaller than in spatial domain. We think the reason is that energy of image in wavelet domain is concentrated on low frequency subbands while other subband's energy are small.

Fig. 4. Fishingboat's stego image and noised stego image

Figure (4) is Fishingboat's stego image according to the watermarking algorithm in section (2), PSNR is 29.08 dB; and noised stego image (σ_n^2 =4), PSNR is 28.16 dB.

5. CONCLUSIONS

According to watermarking methods in wavelet domain, watermark's embedding and extracting realized in wavelet domain, we think that watermarking capacity should be analyzed in wavelet domain. In watermarking schemes, image is considered as a communication channel to transmit messages. But because of the requirements of robustness and visibility, watermarking has some characteristics different from traditional communication. Watermarking capacity should be associated with the content of wavelet subbands. In this paper, we present an adaptive watermarking capacity analysis method based on content of wavelet subbands, and discusses watermarking capacity of blind watermarking and non-blind watermarking in this scenario.

REFERENCES

[1] S D Servetto, C I Podilchuk, K Ramchandran. Capacity Issues in Digital Image Watermarking [A]. *IEEE Intl. Conf. on Image Processing* [C], 1998, volume 1:445~449.

[2] M Barni, F Bartolini, A De Rosa, A Piva. Capacity of the watermarking-channel: how many bits can be hidden within a digital image [A]. *Security and Watermarking of Multimedia Contents, Proceedings of SPIE* [C], 1999, Vol. 3657:437~448.

[3] P Moulin, M K Mihcak. A Framework for Evaluating the Data-Hiding Capacity of Image Sources [J]. *IEEE Trans. on Image Processing*, 2002,Vol. 11, No. 9:1029~1042.

[4] P Moulin. The Role of Information Theory in Watermarking and Its Application to Image Watermarking [J]. *Signal Processing*, 2001,Vol. 81, No. 6:1121~1139.

[5] C Y Lin, S F Chang. Zero-error Information Hiding Capacity of Digital Images [A]. *IEEE Intl. Conf. on Image Processing* [C], 2001,Vol. 3:1007~1010.

[6] C Y Lin, Watermarking and Digital Signature Techniques for Multimedia Authentication and Copyright Protection. Ph.D. Thesis [D], Columbia University, 2000.

[7] Watson A B, Yang G Y. Visibility of wavelet quantization noise [J]. *IEEE Trans image Processing*. 1997. 6(8): 1164~1174.

[8] S Voloshynovskiy, S Pereira, V Iquise, T Pun. Attack modeling: Towards a second-generation benchmark [J]. *Signal Processing*, 2001, Vol.81, No. 6:1177~1214.

[9] S Pereira, S Voloshynovskiy, T Pun. Optimal transform domain watermark embedding via linear programming. Signal Processing [J], 2001, Vol. 81, No. 6: 1117~1335

[10] M Costa. Writing on Dirty Paper. *IEEE Transactions on Information Theory* [J], 1983, 29(3): 439~441.