Hindawi Publishing Corporation
EURASIP Journal on Information Security
Volume 2011, Article ID 502748, 12 pages
doi:10.1155/2011/502748

Research Article

Research of Spatial Domain Image Digital Watermarking Payload

Mao Jia-Fa,[1,2] Zhang Ru,[1] Niu Xin-Xin,[1] Yang Yi-Xian,[1] and Zhou Lin-Na[1]

[1] *Key Laboratory of Network and Information Attack & Defense Technology of MOE, Beijing University of Posts and Telecommunications, Beijing 100876, China*
[2] *Department of Mathematics and Computer Science, Shangrao Normal University, Jiangxi 334001, China*

Correspondence should be addressed to Mao Jia-Fa, maojiafa@bupt.edu.cn

Received 26 July 2010; Accepted 21 February 2011

Academic Editor: Martin Steinebach

Watermarking payload is a topic in which the watermarking researchers have a great interest at present. Based on the constraint of "perceptual invisibility," this paper makes a study of the maximum watermarking payload of spatial domain image, which is related to not only embedding intensity, but also to factors such as the size of image, image roughness and visual sensitivity, and so forth. The correlation among the maximum payload and the embedding intensity and size of an image is theoretically deduced through the objective estimation indicator of the peak signal to the noise rate (PSNR) while the relationship model among watermarking payload and image roughness and visual sensitivity is deduced through effective experiments designed on the basis of subjective estimation indicators. Finally, taking all these relationship models into account, this paper proposes a watermarking payload estimation method and verifies its effectiveness through experiments.

1. Introduction

The research on technologies of information hiding and digital watermarking has developed for nearly twenty years. Information hiding is applied to covert communication, and digital watermarking is applied to copyright protection. They share one feature in common: When some data are embedded into the carrier data, no obvious damage is caused. Therefore, the key point of information hiding and digital watermarking is the same and that's what is called information hiding in a broad sense [1]. However, differences in their application environments result in different research emphases and requirements. Information hiding emphasizes on the resistance to steganalysis attacks while digital watermarking stresses the perceptual invisibility.

The existing research literature about information hiding capacity has established theoretical models for information hiding and drawn different capacity expressions for different models. Moulin and O'Sullivan [2] proposed an information hiding model by abstracting the process of information hiding and using the communication model to represent information hiding. The information hiding capacity is considered as the maximum of reliable transfer rate under the communication model. However, this abstract model is not suitable for the still image information hiding model and cannot be applied to estimate the spatial domain image steganographic capacity. Supposing that the carrier information is state traverse, Cohen and Lapidoth [3] provided the estimating range for information hiding capacity. But in reality, not all the image carriers are state traversed. Though the research of Somekh-Baruch and Merhav [4] is an advance for the Moulin model, it is still limited to the communication model. Reference [5] proposed a secure steganographic method based on the payload and analyzed the correlation between image complexity and payload, but this research is confined to the DCT domain and the payload of spatial domain format is not involved. References [6, 7] made an analysis of information hiding capacity by introducing the case theory, but this research can only be made when the carriers follow the Gaussian distribution.

This paper aims to make a research on the digital watermarking payload. Digital watermarking manages to embed secret information into the carrier data without affecting the use of carrier or arousing visual suspect. Once the watermarked carrier is suspected to have carried secret information, watermarking fails. The most direct constraint

for watermarking is "perceptual invisibility." When still images are used as the host image, the perceptibility is subject to subjective identification. The most direct reason for changes in perceptibility is the payload of the image. Given an image which has a certain size, if the watermarking algorithm is fixed, the maximum payload is also fixed. As a result, what the watermarking researchers are interested in recently is the maximum payload of still images under the constraint of "perceptual invisibility" [7].

Based on the constraint of "perceptual invisibility", this paper makes a study of the maximum digital watermarking payload of the spatial domain grayscale image. Factors restricting the maximum payload are not only internal but also external. The external factors are size of an image, embedding intensity, and so forth. while the internal factors are image roughness, visual sensitivity, and so forth. As is evident, just like a reservoir, the larger the image is, the larger the payload is. On the contrary, the greater the embedding intensity is, the smaller the payload. For instance is, to spatial domain embedding with the same embedding rate of 1 bpp (bits per pixel), higher bits embedding is more perceptible than lower bits embedding because changes in higher bits produce far more noise than those in lower bits embedding do. Different degrees in roughness result in different perceptibility because while it is difficult for naked eyes to identify the subtle changes in a highly rough image, it is easy to identify those changes in a smooth image [8–11]. The sensitivity of naked eyes to change in different images is varied, which is affected by brightness, image contrast, and so forth of the images. This paper carries on a research on the correlation between the payload and these factors, provides the payload estimation methods and verifies its applicability through experiments.

This paper is organized as follows. The external factors influencing the payload are discussed in Section 2. The internal factors influencing payload are elaborated on in Section 3. Section 4 introduces the subjective and objective estimation systems for perceptibility. In Section 5, the correlation between payload and the internal and external factors is discussed theoretically. Section 6 is devoted to the experiments and the testing results. The summary and future work are provided in Section 7.

2. External Factors Influencing Payload

Under the constraint of "perceptual invisibility," the external factors influencing the payload are mainly the size of an image and the embedding intensity. The size of the image is in direct proportion to the payload. It is like a reservoir; the larger the pool is, the larger the payload is. To study the influence of embedding intensity on watermarking payload, some knowledge about the digital watermarking embedding method should be introduced first.

The traditional image information hiding can be divided into two categories: spatial domain information hiding and transform domain (such as the DCT transform domain, the wavelet transform domain, etc.) information hiding [12]. Most watermarking methods in spatial domain embed the watermarking information directly into the original image information, such as embedding the watermarking information into the least significant bit (LSB) plane [13–15] of the original image.

Referring to references [4, 5, 16–19], a common spatial domain image watermarking formula can be summed up as follows:

$$f^s = f^c + \beta \cdot w \quad \text{or} \quad f^s = f^c(1 + \beta \cdot w). \tag{1}$$

Here, f^s and f^c refer to pixel values of the watermarked image and the clean image, respectively; w refers to the secret information embedded; β refers to the embedding intensity. In the LSB embedding, the value of $\beta \cdot w$ in the first part of formula (1) is $-1, 0$ or 1. When the value of β is very large, the embedded information causes great image distortion. Thus, the perceptibility is changed and embedding fails. To have a better understanding of the image payload, the definition of embedding intensity is introduced.

Definition 1 (Embedding intensity). Embedding intensity means to embed the secret information bit stream from a certain bit plane of the image, and if the secret information bit stream is not finished when this bit plane is full, it can be embedded into the higher bit plane until it is finished.

This paper grades the embedding intensity into eight levels, namely, $\beta = 1, 2, \ldots, 8$. When $\beta = 1$, the embedding begins from the first bit plane (also known as the least significant bits) line by line. If there is more secret information bit stream to be embedded, it can be embedded into the higher level until it is finished. While $\beta = 2$, the secret information bit stream is embedded from the 2nd bit plane. Similarly, it is embedded into the higher level until the secret information bit stream is finished. By inference, while $\beta = 8$, the watermark bit stream is embedded into the highest bit plane of the image.

The payload of an image is related to its embedding intensity. Under the constraint of "perceptual invisibility," it is obvious that when $\beta = 1$, the image has the largest payload. The reason why the payloads under different embedding intensities are researched is that when $\beta = 1$, it is actually an LSB watermarking method, for which the current watermarking analysis method is very effective. To avoid attacks on LSB, the watermarking researchers choose different embedding intensities to embed the secret information.

Generally speaking, improving the embedding intensity can increase the resistance capacity and robustness of smoothing, slightly recompression, Gaussian low-pass filtering, and LSB steganlysis attacks. However, the robustness of sharpen; geometric transform attacks cannot be strengthened. Therefore, robustness is not wholly decided by embedding intensity. This article is only an embedding payload reference to the watermarking researchers. According to this purpose, we do the research of the upper limit of embedding payload. The aim of this paper is to provide payload reference for watermarking researchers. To achieve this purpose, this paper studies the maximum payloads under different intensities.

3. Internal Factors Influencing Payload

Under the constraint of "perceptual invisibility," the internal factors influencing the payload are mainly the two factors of image roughness and visual sensitivity.

3.1. Image Roughness. Visual perceptibility of changes in image is not only related to variation but also to roughness of the image, just like when a smooth surface is stained, it is easy to identify but when the surface is rough, it is difficult to identify the stain. Therefore, the maximum payload of an image is closely related to the image itself.

3.1.1. 2D Histogram. A 2D histogram is based on the united probability distribution of a pixel pairs [20]. Take the two pixels $f(i, j)$ and $f(m, n)$ at (i, j) and (m, n) as an example, the distance between the two pixels is r and the line connecting the two points forms an angle of θ with the horizontal line. Suppose an image I with gray level L is given, then the 2D histogram can be expressed as follows:

$$M_{r,\theta}(a, b) = P\left(f(i, j) = a, f(m, n) = b \,\middle|\, \begin{array}{l} m = i + r\cos\theta \\ n = j + r\sin\theta \end{array} \right). \tag{2}$$

In (2), a and b refer to the gray values of pixel points of the image. The 2D histogram can be seen as an $M \times N$ matrix, which is called a gray subordinate matrix or empirical matrix (EM). If the pixel pairs are highly correlated, then the factors in $M_{r,\theta}(a, b)$ distribute close to the leading diagonal of the matrix. The approximate estimate of probability distribution of the EM is

$$P(a, b) \approx \frac{N(a, b)}{M}. \tag{3}$$

M in formula (3) is the number of image pixels. $N(a, b)$ refers to the number of times that $f(j, k) = a$ and $f(m, n) = b$ appear.

3.1.2. Measurement Indicators of Roughness. Yang et al. [21] have proposed many distribution indicators for texture roughness measurement, such as autocorrelation, covariance, moment of inertia, energy and entropy, and so forth. The histogram of the fine-grained texture than the histogram of the coarse texture is more evenly distributed in set $\{(r, \theta)\}$. The texture roughness can be measured through the distribution range of the units occupied by the histogram along the leading diagonal of the histogram. Therefore, this paper employs moment of inertia as the measurement indicator of texture roughness. The calculation formula for moment of inertia is as follows:

$$S(r, \theta) = \sum_{a=0}^{L} \sum_{b=0}^{L} (a - b)^2 P(a, b). \tag{4}$$

L refers to the highest gray value of the image. If the texture area has angular invariance, the moment of inertia $d(r)$ of various distances r can be worked out through the measurement angle of a single angle [22]

$$d(r) = \sum_{\theta} \frac{(S, \theta)}{N_\theta}. \tag{5}$$

The summation in formula (5) refers to that of the whole angle and the scale measurement area. N_θ refers to the number of angles. Since the distribution of image pixels is disperse to make the calculation easier, the parameter (r, θ) can be set as specific discrete value, such as $r = 1, 2, 3$, $\theta = 0$, $\pi/4$, $\pi/2$, $-\pi/4$. When the parameter (r, θ) is set as specific discrete values, the calculation formula for image roughness is

$$Rgh = E(d(r)) = \sum_{r} \alpha_r d(r). \tag{6}$$

$E(\cdot)$ refers to the mean operator, α_r refers to the weighting factor of moment of inertia $d(r)$, and $\sum_r \alpha_r = 1$. The image roughness represents the roughness of neighboring pixels of the image; thus, in this paper $r = 1, 2, 3$ and the weighting factors are 1/2, 1/3, 1/6. Figure 1 is 5 sample images and the second column of Figure 1 is the roughness value of sample images. In the sample images, the roughness of image (b) is the greatest and that of image (d) is the smallest.

3.2. Visual Sensitivity. Perceptibility change is directly related to the human visual system which is generally called visual sensitivity. The sensitivities of human visual system towards low brightness and high brightness are different. However, the human visual system is a complex biological system which has three stages of perception: encoding, representation, and comprehension [23]. There are many factors restricting the perceptibility of human's visual system. For instance, Mach bands phenomenon is a case in which a target is influenced by its surroundings and produces different perceptions. This phenomenon shows that brightness is not the monotonous function of visual sensitivity, that is, visual sensitivity is not only influenced by brightness but also by contrast of background. As a result, the perceptibility change caused by watermarking is closely related to the image contrast and brightness of the image.

3.2.1. Brightness and Contrast. The target brightness of illumination distribution $I(x, y, \lambda)$ is defined as [20, 24]

$$f(x, y) = \int_0^\infty I(x, y, \lambda) V(\lambda) d\nu, \tag{7}$$

where $V(\lambda)$ is the relative illumination efficiency function of the visual system. To human eyes, $V(\lambda)$ is a bell curve, whose features depend on whether it is scotopic or photopic vision. For a grayscale image, its brightness is the pixel value of the image. According to Weber's law, if a target's brightness f is perceived as different from its surroundings, their ratio is

$$\frac{|f_s - f|}{f} = \Delta c, \tag{8}$$

where f_s is the brightness value of the neighboring pixels. If $f_s = f + \Delta f$, Δf is very small, and only big enough to

(a) (b) (c) (d) (e)

Figure 1: Five sample images.

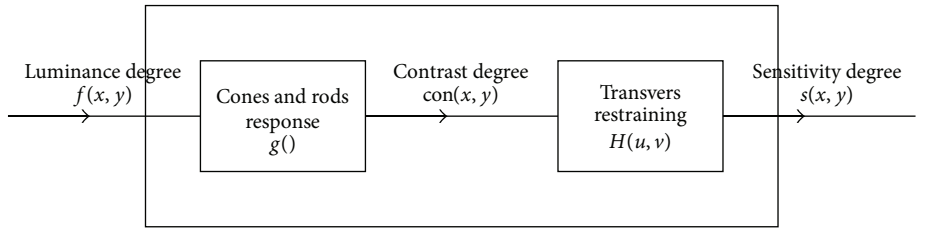

Figure 2: A simple homophony visual model.

distinguish different brightness. Then, (8) can be rewritten as

$$\frac{\Delta f}{f} = d(\log f) = \Delta c, \tag{9}$$

where $d(\cdot)$ is the differential operator. Formula (9) shows that the equal increment of the lightness logarithm can produce the feeling of equal difference, that is, $\Delta(\log f)$ is proportional to Δc, which is the change in contrast; thus, the following formula is obtained:

$$\mathrm{con} = a + b \log f. \tag{10}$$

Formula (10) is commonly called contrast, where a and b are constants. In researches on image coding, logarithm law contrast is the widest choice. Logarithm law contrast is defined as follows [20]:

$$\mathrm{con} = 106.3027 \log(f + 1). \tag{11}$$

3.2.2. Measurement Indicators for Visual Sensitivity. Visual sensitivity is also called perceptible brightness. The homophony visual model is introduced before visual sensitivity is defined. Figure 2 is a simplified homophony visual model. The light enters to the eyes, and the nonlinear response of the cone and rod is represented by the point nonlinear function $g(\cdot)$, producing contrast $\mathrm{con}(x, y)$. The lateral inhibition phenomenon is represented by a linear system which is spatially invariant and isotropous. Its response frequency is represented by filter $H(u, v)$

$$H(u, v) = H(\rho) = A\left(\alpha + \left(\frac{\rho}{\rho_0}\right)\right) \exp\left(-\left(\frac{\rho}{\rho_0}\right)^{\gamma}\right), \qquad \rho = \sqrt{u^2 + v^2}, \tag{12}$$

where A, α, γ, and ρ_0 all are constants. ρ_0 is the peak value frequency while $\alpha = 0$ and $\gamma = 1$. In image processing, it is suitable to choose $A = 2.6$, $\alpha = 0.0192$, $\rho_0 = 8.772$, and $\gamma = 1.1$. Figure 3 is the response curve of the linear system $H(\rho)$. From this curve, it can be seen that human eyes have inhibiting effect on low and high frequencies and are most sensitive to changes in medium frequency.

What are transmitted by linear system $H(\rho)$ are neural signals representing the perceptible lightness on the surface, that is, the sensitivity of human eyes to objects $s(x, y)$. From the visual model, the sensitivity $s(x, y)$ can be easily worked out:

$$s(x, y) = \Im^{-1}(\mathrm{con}(u, v)H(u, v)), \qquad \mathrm{con}(u, v) = \Im(\mathrm{con}(x, y)), \tag{13}$$

where $\Im$ refers to 2D Fourier transform and $\Im^{-1}$ refers to the 2D inverse Fourier transforms. To describe quantitatively the whole sensitivity of human eyes to the single $M \times N$ image, its average value is adopted to represent the image sensitivity ds

$$ds = E(s(x, y)) = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} |s(x, y)|. \tag{14}$$

The third column in Table 1 lists the sensitivity values of 5 sample images. Human eyes are most sensitive to changes in sample image (d) but least sensitive to changes in image (a).

3.3. Relationship between Image Roughness and Visual Sensitivity. It can be concluded from Sections 3.1 and 3.2 that image roughness is related to visual sensitivity. From the perspective of visual sensitivity, the visual sensitivity model (Figure 3) is not only related to the photo response of cone

Table 1: Image roughness and visual sensitivity of the sample images.

Images	Roughness	Sensitivity
(a)	119.1	79.348
(b)	906.51	83.581
(c)	377.23	85.204
(d)	106.78	95.899
(e)	666.28	90.293

and rod, but also related to the image contrast which is based on the image content. From the perspective of image roughness, its value is completely dependent on the image content. Thus, image roughness is related to visual sensitivity.

Three hundred cover images (for source of the images, see Section 6.1) are collected. Their image roughness Rgh_i and visual sensitivity ds_i can be worked out by making use of the image roughness and visual sensitivity, where $i = 1, 2, \ldots, 300$. Then, normalize them according to (15) and work out the image roughness and visual sensitivity after their normalization

$$N_{rgh_i} = \frac{Rgh_i - \min(Rgh_i)}{\max(Rgh_i) - \min(Rgh_i)},$$
$$N_{ds_i} = \frac{ds_i - \min(ds_i)}{\max(ds_i) - \min(ds_i)}. \quad (15)$$

The dotted line in diagram 4 is the connected line of dots (N_{rgh_i}, N_{ds_i}) of the image roughness and visual sensitivity of 300 clean images. When image roughness value is small, the sensitivity shock is high; when image roughness value is big, the sensitivity is stable, with its value around 0.4; when the roughness is in the middle (during the period of $[0.4, 0.5]$), the sensitivity reduces sharply. In general, the sensitivity reduces as the roughness increases. The line in Figure 4 shows this phenomenon.

4. The Estimation System of Image Visual Perceptibility

There are two estimation systems for image visual perceptibility: one is subjective, and the other is objective. According to the criteria of digital image processing [12], this paper adopts the perception rank for the subjective standard and PSNR for the objective standard to measure the distortion of the image.

4.1. Subjective Estimation. Ranks are based on the change rank when an image is compared with an originally clean image. Referring to a relevant image fidelity criterion [20], the image perception changes are rated into five ranks and each rank is quantized: unnoticeable (-2), not evident (-1), slightly evident (0), evident (1) and very evident (2). Different individuals have different perceptions and visual sensitivities. Therefore, subjective estimating is often

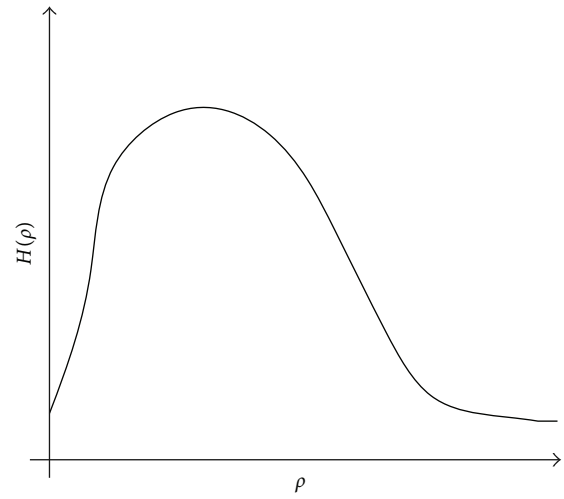

Figure 3: The curve of $H(\rho)$.

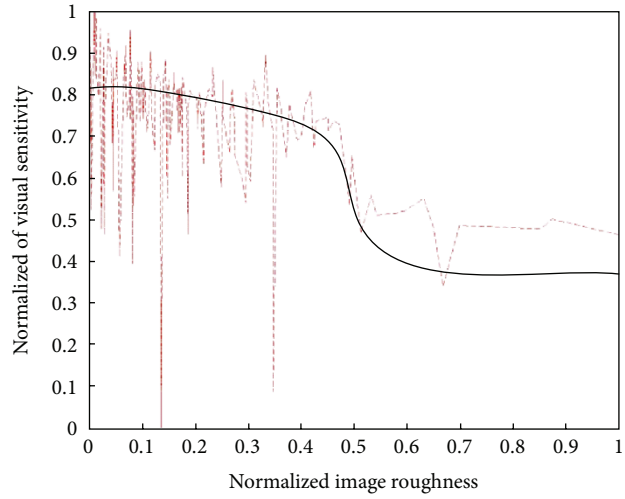

Figure 4: Relation between visual sensitivity and image roughness.

conducted by several image experts in watermarking field. The average ranks can be represented as (16)

$$R = \left(\frac{\sum_{i=1}^{n} s_i n_i}{\sum_{i=1}^{n} n_i} \right), \quad (16)$$

where s_i is the score of rank i, n_i refers to the number of the observers in this gradation, and n refers to the number of ranks. Figure 5 depicts a subjective decision device, the smaller the value of R is, the lower the perceptibility of the watermarked image is; the larger the value of R is, the easier the watermarked image is perceived. The change in image R which the observers cannot judge accurately should be less than –0.1. Suppose that the observer thinks of a tolerance range as 0.2, when the average rank R is between $[-0.1, 0.1]$, no judge is made. But when R is larger than 0.1, the observer can definitely judge that there's change in perceptibility. This means the embedding fails. Therefore, R of a watermarked image not to be perceived by the observers should be between $[-2, 0.1]$.

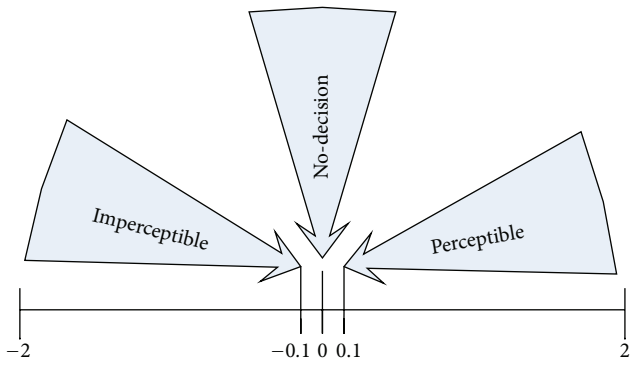

Figure 5: Discrimination classifier based on subjective estimating.

4.2. Objective Estimation. Objective estimation is a quantitative measurement, and PSNR is an effective visual fidelity indicator. Suppose that a watermarked image $f_s(x, y)$ is obtained after a clean image $f_c(x, y)$ is watermarked, then the mean square error (MSE) σ_e^2 of the watermarked image and the clean image is

$$\sigma_e^2 = E\Big[\,|\,f_s(x,y) - f_c(x,y)\,|^2\,\Big]. \tag{17}$$

Then, PSNR using dB as a unit is defined as

$$\text{PSNR} = \frac{10\log_{10}(255 * 255)}{\sigma_e^2}. \tag{18}$$

The amount of information of a arbitrary host image is defined as $255 * 255$, which is a fixed value. The variation of the image is only related to the MSE σ_e^2. The more data the watermarking researchers embed, the larger the MSE σ_e^2 is and the smaller the PSNR is. In such a situation, the watermarked image can be perceived more easily. On the contrary, when the data embedded is smaller, the MSE σ_e^2 is smaller, the PSNR is larger, and the watermarked image is less likely to be perceived.

Generally, the change in the image is imperceptible [20] when PSNR ≥ 40. But can it be concluded that when PSNR < 40, the images change is not perceptible? The answer is no because it is closely related to the internal factors (mainly the image roughness and visual sensitivity). For example, suppose that two images have different contents but the same variation, and then their PSNR values are the same. But it can happen that the change in one image is perceptible and the change in the other is not. However, for the same image, its imperceptible minimum PSNR is fixed. Hence, to the same image, PSNR is a criterion for both the visual perceptibility and the payload.

5. Analysis of Payload

The maximum payload of a given image under the constraint of perceptual invisibility is one of the main concerns for the watermarking researchers. From another perspective, what kind of image should be chosen as the carrier to hide a certain amount of watermarking is also the concern of the watermarking researchers. Both these two problems are related to payload.

Maximum payload refers to the maximum payload of the carrier under a certain constraint. Based on the constraint of "perceptual invisibility," maximum payload refers to the higher limit of watermarking data embedded into the image. If exceeding this limit, the watermarked image is perceived by the observer, that is, the observer discovers the change in the image quality, which is unbearable for the watermarking researchers because it means failure of the watermarking algorithm. But of course, the perception of this kind of change happens in the situation when the observer has the original host image.

From the analysis above, it can be concluded that the payload is not only related to embedding intensity but also to the factors such as the size of the image, roughness and visual sensitivity, and so forth.

5.1. Relation between Payload and Embedding Rate. Suppose that there is a spatial domain grayscale image $f(x, y)$ with a size of $M \times N$ to be embedded, under the constraint of perceptual invisibility, the arithmetic model for estimating its maximum payload is

$$C_f = \text{Re}(\text{size}, \beta, Rgh, ds), \quad \text{Condition: } R \leq 0.1. \tag{19}$$

Here, size $= M * N$, β refers to the embedding intensity, Rgh refers to image roughness, ds refers to visual sensitivity, and the constraint R refers to subjective estimating rank. Under most circumstances, the larger the size of the image is, the greater the payload is. It is like a reservoir. Therefore, formula (19) can be transferred as:

$$C_f = \text{size} * \text{Re}(\beta, Rgh, ds). \tag{20}$$

Suppose that the embedding rate is Re (bits per pixel, bpp). Then, the relation between the embedding rate and embedding intensity, roughness, and sensitivity is

$$\text{Re} = \frac{C_f}{\text{size}} = \text{Re}(\beta, Rgh, ds). \tag{21}$$

To obtain the maximum payload of the image, the relation between the embedding rate and embedding intensity, roughness and sensitivity should be obtained first. In the next section, the influence of embedding intensity on embedding rate is analyzed with the objective estimating system.

5.2. Relation between Embedding Rate and Embedding Intensity. To give a clear description of the relation between embedding rate and embedding intensity, the concept of embedding factor is introduced.

Definition 2 (Embedding factor). Embedding factor λ means to embed secret information bit stream only into a single bit plane of the image. For example, if the secret information bit stream is only embedded into the first bit plane, then $\lambda = 1$. If it is only embedded into the second bit plane, then $\lambda = 2, \ldots$.

When secret information bit steam is embedded into the ith bit plane, difference between the watermarked image and

the clean image only happens on the ith bit plane and the difference e is 0, −1, or 1. Its probability to appear is [25]

$$P(e) = \begin{cases} \frac{1}{4}, & e = -1, \\ \frac{1}{2}, & e = 0, \\ \frac{1}{4}, & e = 1. \end{cases} \quad (22)$$

If embed information with the embedding rate of $\mathrm{Re}_{\lambda=i}$ into the ith bit plane, its MSE is

$$E(\sigma_e^2 \mid \lambda = i) = (\mathrm{Re}_{\lambda=i}) * \sum_{e=-1}^{1} \left(e * 2^{i-1}\right)^2 * P(e). \quad (23)$$

If it is full imbedded on the ith bit plane, that is, when $\mathrm{Re}_{\lambda=i} = 1$, the mean square error on this level is

$$A_{\mathrm{er}_i} = \sum_{e=-1}^{1} \left(e * 2^{i-1}\right)^2 P(e). \quad (24)$$

Therefore, it is easy to obtain the relation between the embedding rate and the mean square error on the ith bit plane when it is not full imbedding

$$\mathrm{Re}_{\lambda=i} = \frac{\sigma_e^2}{A_{\mathrm{er}_i}}. \quad (25)$$

Suppose that the maximum MSE of an image when the images are visually imperceptible is σ_e^2, when embedding intensity $\beta = i$, the steps of method 1 to calculate the maximum embedding rate is as follows.

Step 1. Initialization: Re = 0.

Step 2. Making use of (24) to work out the full imbedding mean square error A_{er_i} on the ith bit plane.

Step 3. If $\sigma_e^2 \le A_{\mathrm{er}_i}$, then $\mathrm{Re} = \mathrm{Re} + (\sigma_e^2 / A_{\mathrm{er}_i})$. The end.

Step 4. Re = Re + 1, $\sigma_e^2 = \sigma_e^2 - A_{\mathrm{er}_i}$, $i = i + 1$, go to Step 2.

The method above is used when "the maximum MSE of visual imperceptibility" is clearly known. But in reality, the maximum mean square error of visual imperceptibility of an image is difficult to get beforehand because it is related to image roughness and visual sensitivity. Therefore, the relation between the maximum embedding rate and image roughness and visual sensitivity is deduced through subjective estimation (the MSE belongs to objective estimation) in the next section.

6. Experimental Derivation and Verification

To effectively estimate the payload of images under the constraint of perceptual invisibility in the carriers, this paper conducts an experiment to find the relation between payload and image roughness and visual sensitivity. Before the experiment, we make preparations as follows.

6.1. Data Preparation for Experiments. To deduce the relation between payload (or embedding rate) and image roughness and visual sensitivity, 300 various images have been collected, among which, 150 BMP images are downloaded from an image database [26] and 150 are classical images taken by the author with digital cameras which are often used in image treatment. To make the data models universal and reasonable, in the experimental images, there are simple images without any detail and images containing great details; there are images of mountains, rivers, people, animals, plants, and so forth. All the images are treated by using the ACDSee image treatment software, the colorful ones transferred into gray ones, non-BMP images transferred into BMP ones, and all of them are cut into sizes of 256 × 256. These images constitute clean image data. Figure 2 is a cover image of this specification.

6.2. Determination of Experiment Project. The aim of this paper is to estimate the maximum payload under the constraint of perceptual invisibility in the carriers. The payload is not only related to the image size and embedding intensity, but also to image roughness and visual sensitivity. We have discussed both the relation between the maximum payload and the image size and that between the embedding rate and the embedding intensity, and obtained the calculating formula (21) for maximum the embedding rate. Thus, the key to studying the payload is to study the relationship model of the embedding rate and the image roughness, visual sensitivity, for which the following two steps are of vital importance.

(1) The watermark method of increasing the payload dynamically. This watermark method means to increase the payload constantly in the process that the observer judges whether any visual perceptibility has happened. According to (1), we designed a watermark method which can change the embedding intensity β. Given a certain β, watermark information bit steam begins to be embedded from the βth bit plane. When β is embedded to full, there is no visual perceptibility happening in the image. Then, continue to embed from level $\beta + 1$, the watermark embedding will not stop until visual perceptibility happens in the image. Figure 6 is the watermark-image of different payload (or embedding rates) when the embedding intensity $\beta = 1$.

(2) Deciding the embedding intensity. When it is full embedding, the following can be worked out: $E(\mathrm{PSNR} \mid \lambda = 1) = 51.141$, $E(\mathrm{PSNR} \mid \lambda = 2) = 45.1205$, $E(\mathrm{PSNR} \mid \lambda = 3) = 39.099$. From these data we can see that when it is only embedded into the first bit plane (LSB embedding method), its PSNR is far higher than 40, which is to say naked eyes can hardly perceive the changes in the image. But from the 2nd bit plane, PSNR is lower than the secure value of 40, and the watermarked image may be perceived. When first, second, and third bit plane are all embedded with secret information, its PSNR is 37.9189, and then the possibility of

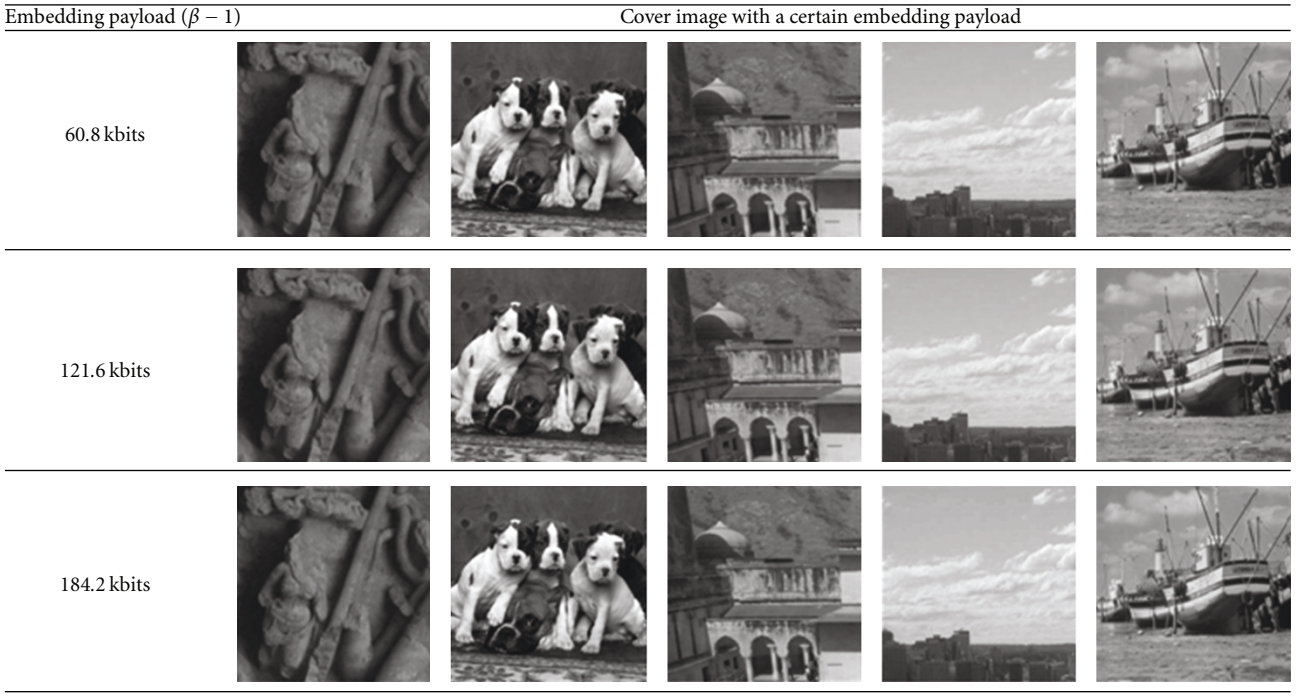

Figure 6: The sample images with payload of 60.8 k, 121.6 k, 184.2 k bits using our watermark method while $\beta = 1$.

its being perceived is greater. But this is only the possibility of being perceptible, whether it is really perceptible is closely related to image roughness and visual sensitivity. Since when only the first bit plane is embedded with information, its PSNR is far higher than the secure value of 40, it can be concluded that whether the first bit plane is embedded with information or not exerts little influence on visual perceptibility of the image. Therefore, this paper makes a research on the maximum payload (or maximum embedding rate) from level 2, that is, when $\beta = 2$.

After the embedding intensity is decided, the experiment plan can be designed as in Figure 7. Choose a host image at random. First, calculate its roughness and sensitivity. Then, without causing any visual perceptibility, embed watermarking information constantly until perceptual visibility happens to the image. The embedding rate value can be obtained by dividing the embedding amount until the last embedding by the size of the image.

6.3. Experiment of Estimating Maximum Payload. From the analysis above, it can be concluded that under the constraint of perceptual invisibility, the key to working out the maximum payload is to work out the maximum embedding rate. Seven postgraduates, all of whom have participated in image treatment for a long time, are invited to be the evaluation experts for perceptible changes, and 200 images are chosen from the image library of 300 at random to be the host images. In accordance with the experiment

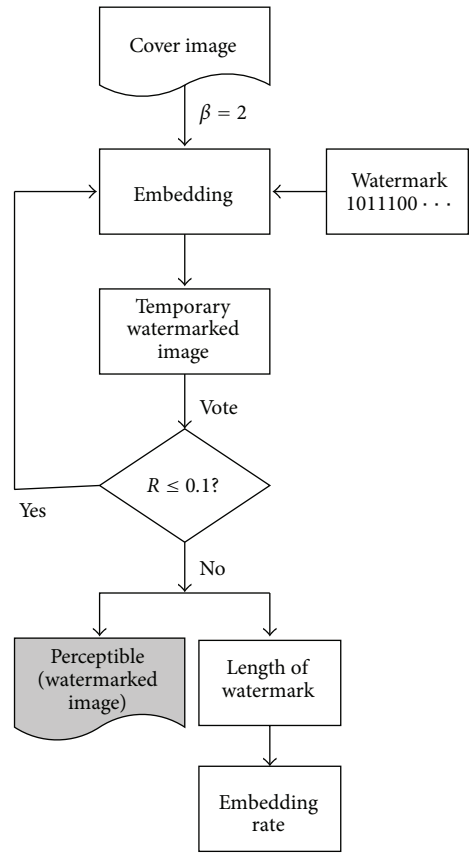

Figure 7: The block diagram of incremental embedding procedure to determine the embedding rate.

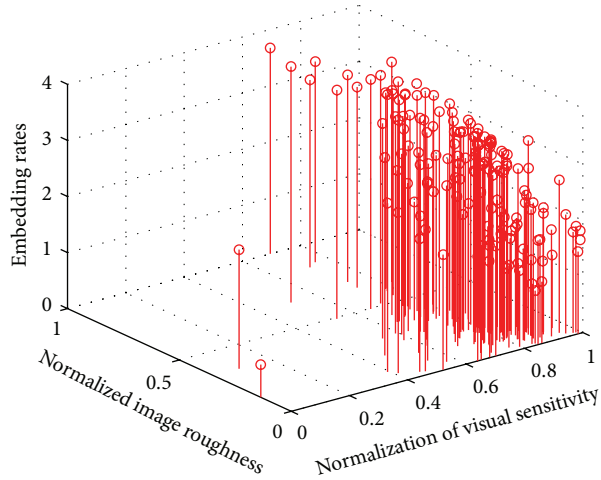

Figure 8: Relation among the embedding rate, roughness, and sensitivity of the two hundred images when $\beta = 2$.

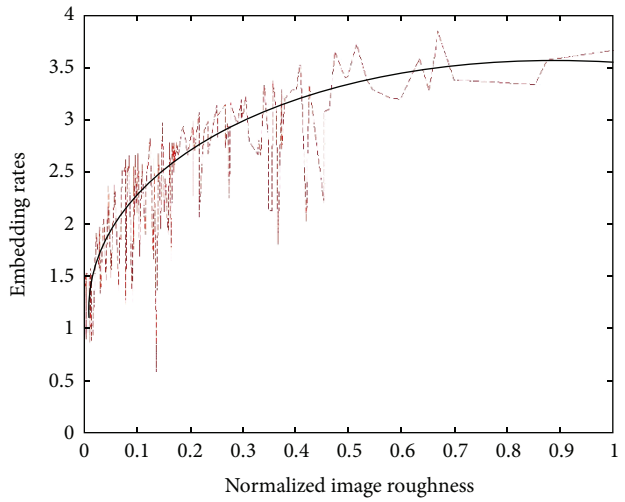

Figure 9: Binary relationship model between the embedding rate and roughness in Figure 8.

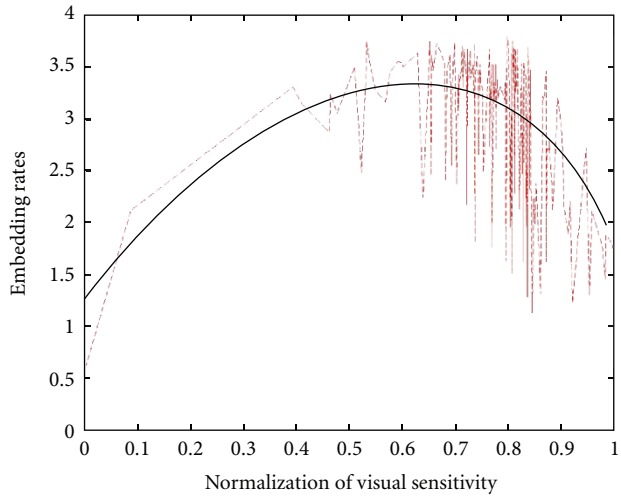

Figure 10: Binary relationship model between the embedding rate and sensitivity in Figure 8.

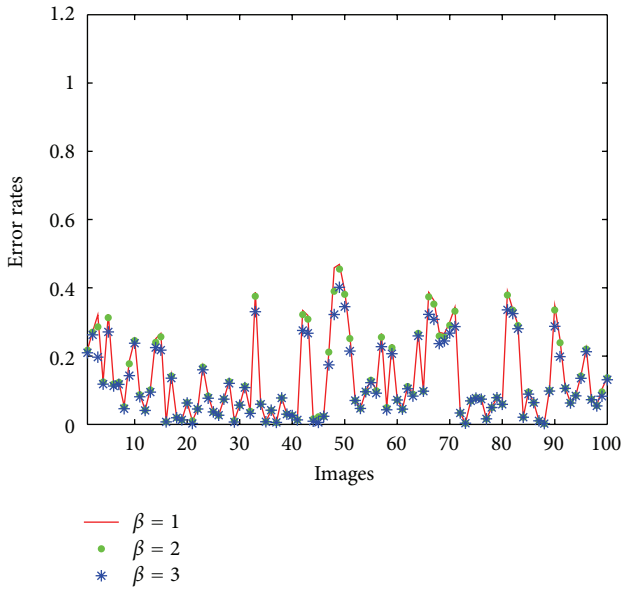

Figure 11: The error rate of 100 images.

plan, the information is embedded from the second bit plane. If the second bit plane is full, and the average level mark of the judging team is $R < 0.1$; the information is embedded into the next bit plane until $R \geq 0.1$, when this image cannot be embedded with information and the estimation is over. Then, the estimation of the next image begins. When $\beta = 2$, the maximum payload for this image is the information altogether until it is embedded into the last level. The relation between the embedding rate and image roughness and visual sensitivity of the 200 images is shown in Figure 8.

From Figure 8, the relationship model between the embedding rate and image roughness and visual sensitivity is hard to estimate. As a result, we divide the triadic relation of the embedding rate and image roughness and visual sensitivity into two binary relations between the embedding rate and image roughness and between the embedding rate and visual sensitivity. As to their relations, respectively, the readers can refer to the broken lines in Figures 9 and 10.

By observing Figure 9, one can find that though the embedding rate shocks greatly in the range of low roughness, the embedding rate, and the image roughness show the logarithm relation in general. We propose the relationship model between the embedding rate and image roughness as

$$\mathrm{Re}_3 = a + b\log_c(1 + d * Rgh) \quad (0 \leq Rgh \leq 1), \tag{26}$$

where a, b, c and d are unknown constants. Observing the trend in Figure 10, one can find that in general the embedding rate and visual sensitivity show a inverted U shape. As a result, we propose their relationship model as

$$\mathrm{Re}_4 = e - f * (\mathrm{Sens} - 0.6)^2 \quad (0 \leq \mathrm{Sens} \leq 1). \tag{27}$$

Table 2: The five images actual and estimated maximum payload in Figure 1.

The sample	Actual maximum payload (kilobits)			Estimated maximum payload (kilobits)		
Images	$\beta = 1$	$\beta = 2$	$\beta = 3$	$\beta = 1$	$\beta = 2$	$\beta = 3$
(a)	264.90	201.15	138.15	233.25	169.80	108.35
(b)	235.800	172.80	112.80	265.25	201.85	140.45
(c)	250.42	187.45	127.42	263.75	200.35	138.85
(d)	172	112	64	204.05	142.50	88.50
(e)	206.07	143.07	83.07	256.75	193.50	132.05

Similarly, e and f in formula (27) are unknown constants. Do formulas (26) and (27) by geometric mean, and the relation of the surface model between the maximum embedding rate and image roughness when $\beta = 2$.

$$\mathrm{Re}|_{\beta=2} = \sqrt{a + b\log_c(1 + d * Rgh)} * \sqrt{e - f(\mathrm{Sens} - 0.6)^2}. \tag{28}$$

using the actual embedding rate of the 200 images in experiments and the minimum mean square error of the embedding rate estimated in (27) as the constraint, we obtain that constants $a = 0.9430$, $b = 2.5613$, $c = 101.7520$, $d = 101.4355$, $e = 3.6930$, $f = 9.96500$. Making use (28), we can get the maximum embedding rate of the image when the embedding intensity $\beta = 2$. On the basis of this maximum embedding rate, how to work out the maximum embedding rate under various embedding intensities?

To work out the maximum embedding rate under various embedding intensities, the mean square error is still used as the transition. Suppose that the maximum embedding rate of an image when the embedding intensity $\beta = 2$ is obtained as $k \in R^+$ by using (28), then the maximum mean square error can be worked out according to method 2, which is described as follows.

Step 1. Initialization $i = \beta = 2$, $\sigma_e^2 = 0$.

Step 2. Making use of formula (24) to work out A_{er_i} on the ith bitplane.

Step 3. If $k \le 1$, then $\sigma_e^2 = \sigma_e^2 + k * A_{\mathrm{er}_i}$. The end.

Step 4. $\sigma_e^2 = \sigma_e^2 + A_{\mathrm{er}_i}$, $i = i + 1$, $k = k - 1$, go to Step 2.

When the maximum mean square error σ_e^2 is obtained, the method in Section 5.2 can be made use of to work out the maximum embedding rate $\mathrm{Re}(\beta, Rgh, \mathrm{Sens})$ under various embedding intensities. Then, the maximum payload can be worked out by using (20).

6.4. Testing Results. Under the constraint of perceptual invisibility, this paper proposes the estimation method for maximum embedding capacity. To evaluate the effectiveness of the estimation method, the error rate is introduced as the evaluation indicator. Suppose that the actual maximum capacity of an image is a, the estimated maximum capacity is b, then the error rate is

$$R_{\mathrm{error}} = \frac{|a - b|}{a}. \tag{29}$$

Table 3: The mean and standard deviation of 100 image's error rates.

Embedding intensity	Mean	Standard deviation
$\beta = 1$	0.139	0.12261
$\beta = 2$	0.13512	0.1171
$\beta = 3$	0.12309	0.10291
Average	0.1324	0.1142

Using the 100 images left as the test images, the experimental evaluation under three kinds of embedding intensities ($\beta = 1, 2, 3$) are carried out. First, according to the experiment plan in Section 6.2, the actual maximum payload of 300 images under $\beta = 1, 2, 3$ are worked out by the experts. Then, use our method to work out the estimated maximum payload, respectively. Table 2 shows the actual maximum payload and the estimated maximum payload of the five images in Figure 1 under three kinds of embedding intensities. From Table 2, one can easily conclude that the larger embedding intensity is, the smaller the payload is, and this is in accordance with what we discovered above.

The error rate in (29) reflects the deviation rate between the image actual payload and the estimated payload. Figure 11 shows the error rate of the payload of the 100 images tested. From Figure 11, it can be seen that the estimation of most images are highly accurate. There is little difference between the actual payload and the estimated payload. But difference between the actual payload and the estimated payload for few images is relatively great, with some approaching 50%. Table 3 shows the mean value and the standard deviation of the 100 images in the experiment. Generally, our estimation method is effective in that the average error rate of images tested under various intensities is less than 15% and the standard deviation is within 13%. From Table 3, it can be concluded that the larger the embedding intensity is, the smaller the difference between the estimated payload and the actual payload is and the higher the accuracy is. The reason is that when the embedding intensity is low, the payload is larger and it is easier for deviation to appear.

Table 4: Summarization for previous work and our proposed method.

Works	Domain	Number of factors	Constraint	Payload estimated method	Estimation error
[3]	Time-frequency domain	2	Imperceptible	$C = n^{-1} \max_{Q(X,U\|S,K)} (\min_{A(Y\|X)} J(Q,A))$	Unreported
[4]	Time-frequency domain	2	Imperceptible	$C = \max\left(\frac{1}{2}\log(1 + S(A; D_1, D_2, \sigma_s^2)\right)$	Unreported
[5]	Time-frequency domain	2	Imperceptible	$C = \frac{\max(\min(I(V;Y) - I(V;S)))}{n}$	Unreported
[6]	DCT domain	1	Secure steganography	Unreported	Unreported
[7]	Time-frequency domain	2	Minimum error rate	$C = \frac{1}{2}\log\left(1 + \frac{D_1}{D_2}\right)$	Unreported
This paper	Spatial domain	4	Visual imperceptible	$C = \mathrm{Re}(\mathrm{size}, \beta, Rgh, ds)$	13.24%

7. Conclusion

In the recent twenty years, the technology of information hiding has been widely applied to fields of copyright protection (digital watermarking), communication, and so forth. At present, most researches focus on how to embed information without visual distortion and there have been few researches on the maximum payload, that is, the maximum payload under the constraint of perceptual invisibility.

This paper proposes the estimation method for the maximum payload. The maximum payload is influenced not only by internal but also external factors. The external factors are mainly the image size, embedding intensity, and so forth while the internal factors are mainly the image roughness, visual sensitivity, and so forth. The size of image is in direct proportion to the payload while the embedding intensity is in inversely proportional to the payload because higher bits embedding generates more noise than lower bits embedding does and the noise is the normalized indicator of image distortion. Different degrees in roughness result in different perceptibility because while it is difficult for the human eyes to identify the subtle changes in a highly rough image, it is easy to identify such changes in a smooth image. The sensitivity of human eyes to changes in different images is varied, which is affected by image contrast and brightness. The correlation between the maximum payload and the embedding intensity and size of image is theoretically deduced through the objective estimation indicator of the peak signal to noise rate (PSNR) while the relationship model between watermarking payload and image roughness and visual sensitivity is deduced through effective experiments designed on the basis of subjective estimating indicators. Finally, taking into account of all these relationship models, this paper proposes the watermarking payload estimation method and verifies its effectiveness through experiments.

Table 4 summarizes both the estimation methods we have proposed before and the methods proposed in the previous literatures, which can be generalized as follows.

(1) Most references [3–5, 7] abstracted information hiding into a Communication Theory Model and draw different payload expressions from different models. However, this kind of abstraction of models can only act as a theoretical guide for hiding information capacity estimation of the real objective images and is not very much contributive to the accomplishment of the project. The estimation method proposed for hiding capacity estimation of the real objective images is more contributive to the Engineering Application.

(2) Reference [6] proposed a secure estimation method for steganographic capacity based on the DCT domain. It only proves the influence of image complexity on payload by doing some experiments but has not worked out the specific capacity estimation method.

(3) These references have not reported the deviation rate between the estimated value and the actual value. But this paper proves the effectiveness of our way of estimation through experimental tests.

There are still shortcomings in our method and further research is still needed to improve the estimating accuracy.

(1) The method is rough. This paper makes a study of the maximum watermarking payload of spatial domain image under the conditions of invisibility, in another word, the maximum embedding payload. Different area has the different payload capacity. For example, the payload of high roughness and perceptual invisibility areas is higher than the area of low roughness and visual sensitive areas. This article does not do further research of this aspect; it is the deficiency of this article and also further research directions of ours, which is closer to the practical applications.

(2) The experimental plan lacks novelty. Since evaluation of visual perceptibility in images is needed in the experiments, it costs much time of the experts. In the future work, better plans will be designed to save the experts time and to improve accuracy in estimating.

Acknowledgments

The authors thank the postgraduates in Information Security Center of Beijing University of Post and Telecommunication for their precious time devoted to the experimental evaluation in this paper. This work is supported by the National Basic Research Program of China (973 Program) (2007CB311203), the National Natural Science Foundation of China (no. 60821001), the Specialized Research Fund for the Doctoral Program of Higher Education (no. 20070013007) and the 111 Project (no. B08004), and the Shanghai Municipal Education Committee Scientific Research Innovation Project (no. 11YZ284).

References

[1] X. X. Niu and Y. X. Yang, "Study on the frame of information steganography and steganalysis," *Acta Electronica Sinica*, vol. 34, pp. 2421–2424, 2006.

[2] P. Moulin and J. A. O'Sullivan, "Information-theoretic analysis of information hiding," *IEEE Transactions on Information Theory*, vol. 49, no. 3, pp. 563–593, 2003.

[3] A. S. Cohen and A. Lapidoth, "The Gaussian watermarking game," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1639–1667, 2002.

[4] A. Somekh-Baruch and N. Merhav, "On the capacity game of public watermarking systems," *IEEE Transactions on Information Theory*, vol. 50, no. 3, pp. 511–524, 2004.

[5] H. Sajedi and M. Jamzad, "Secure steganography based on embedding capacity," *International Journal of Information Security*, vol. 8, no. 6, pp. 433–445, 2009.

[6] H. Y. Gao, *The theory and application of audio information hiding*, PH.D. dissertation, Beijing University of Posts and Telecommunications, Beijing, China, 2006.

[7] R. Chandramouli and N. D. Memon, "Steganography capacity: a steganalysis perspective," in *Security and Watermarking of Multimedia Contents*, vol. 5020 of *Proceedings of SPIE*, pp. 173–177, Springer, Santa Claru, Calif, USA, 2003.

[8] S. Ll, X. P. Zhang, and S. Z. Wang, "Digital image steganography based on tolerable error range," *Journal of Image and Graphics*, vol. 12, no. 2, pp. 212–217, 2007.

[9] H. Noda, J. Spaulding, M. N. Shirazi, and E. Kawaguchi, "Application of bit-plane decomposition steganography to JPEG2000 encoded images," *IEEE Signal Processing Letters*, vol. 9, no. 12, pp. 410–413, 2002.

[10] D. C. Wu and W. H. Tsai, "A steganographic method for images by pixel-value differencing," *Pattern Recognition Letters*, vol. 24, no. 9-10, pp. 1613–1626, 2003.

[11] X. Zhang and S. Wang, "Steganography using multiple-base notational system and human vision sensitivity," *IEEE Signal Processing Letters*, vol. 12, no. 1, pp. 67–70, 2005.

[12] S. Katzenbeisser and F. A. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House Press, Norwood, Mass, USA, 2000.

[13] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Systems Journal*, vol. 35, no. 3-4, pp. 313–335, 1996.

[14] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Processing*, vol. 66, no. 3, pp. 385–403, 1998.

[15] T. S. Chen, C. C. Chang, and M. S. Hwang, "A virtual image cryptosystem based upon vector quantization," *IEEE Transactions on Image Processing*, vol. 7, no. 10, pp. 1485–1488, 1998.

[16] A. S. Cohen and A. Lapidoth, "The capacity of the vector Gaussian watermarking game," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT '01)*, p. 5, June 2001.

[17] X. Zhang and S. Wang, "Efficient steganographic embedding by exploiting modification direction," *IEEE Communications Letters*, vol. 10, no. 11, pp. 781–783, 2006.

[18] Z. G. Qu, Y. Fu, X. Niu, Y. Yang, and R. Zhang, "Improved EMD steganography with great embedding rate and high embedding efficiency," in *Proceedings of the 5th International Conference on Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP '09)*, pp. 348–352, Tokyo, Japan, September 2009.

[19] W. N. Lie and G. S. Lin, "A feature-based classification technique for blind image steganalysis," *IEEE Transactions on Multimedia*, vol. 7, no. 6, pp. 1007–1020, 2005.

[20] A. K. Jain, *Fundamentals of Digital Image Processing*, Person Education, Inc., Publish as Prentice Hall, 1989.

[21] C. H. Yang, C. Y. Weng, S. J. Wang, and H. M. Sun, "Adaptive data hiding in edge areas of images with spatial LSB domain systems," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 488–497, 2008.

[22] E. L. Hall, "Survey of preprocessing and feature extraction techniques for radiographic images," *IEEE Transactions on Computers*, vol. 20, no. 9, pp. 1032–1044, 1971.

[23] J. F. Delaigle, C. Devleeschouwer, B. Macq et al., "Human visual system features enabling watermarking," in *Proceedings of IEEE International Conference on Multimedia and Expo*, pp. 489–492, Lusanne, Switzerland, 2002.

[24] C. K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," *Pattern Recognition*, vol. 37, no. 3, pp. 469–474, 2004.

[25] A. D. Ker, "A capacity result for batch steganography," *IEEE Signal Processing Letters*, vol. 14, no. 8, pp. 525–528, 2007.

[26] http://sipi.usc.edu/database/database.cgi?volume=textures.