A Fast and Effective Steganalytic Technique against JSteg-like Algorithms

Tao Zhang
Department of Information Science,
University of Information Engineering
P.O.Box1001, No.306, Zhengzhou,
450002, P.R.China
86-371-3531563
brunda@sina.com

Xijian Ping
Department of Information Science,
University of Information Engineering
P.O.Box1001, No.37, Zhengzhou,
450002, P.R.China
86-371-3531563
rbfn@sina.com

ABSTRACT
Detection of hidden messages in images, also known as image steganalysis, is of great significance to network information security. In this paper, we propose a fast and effective steganalytic technique based on statistical distributions of DCT coefficients which is aimed at two kinds of popular JSteg-like steganographic systems, sequential JSteg and random JSteg for JPEG images. Our approach can not only determine the existence of hidden messages in JPEG images reliably, but also estimate the amount of hidden messages exactly. Its advantages also include simplicity, computational efficiency and easy implementation of real-time detection. Experiment results show the superiority of our approach over other steganalytic techniques.

Keywords
Information Hiding, Steganography, Steganalysis, JPEG, JSteg

1. INTRODUCTION
Information hiding is a recently developed technique in the information security field and has received significant attention from both industry and academia. There are two main branches – steganography and digital watermarking. As a new way of covert communication, the main purpose of steganography is to convey messages secretly by concealing the very existence of messages, while digital watermarking is mainly used for copyright protection of electronic products [1].

In this paper, we focus on the detection of hidden messages embedded in image using steganographic algorithms, also known as image steganalsis. Steganalysis is of great significance to network information security.

Detection of hidden messages can be viewed as a passive attack against steganography. The aim of steganalysis is to determine the[1] existence of hidden messages in the given image without access to original carrier-image. In steganalysis we don't care which bits carry what information, which is we are not trying to read the hidden messages, just trying to determine the existence. If one can show that an image conceals a message, any message, whether it can be read or not, then the stego system has failed and correspondingly the steganalysis system has gained its ends.

N.F. Johnson and S. Jajodia made a careful analysis of signatures introduced by current steganographic software [2]. A good survey on steganalytic techniques is given by Jiri Fridrich in [3]. For JPEG image file format is most frequently used one through internet, people pay more attention to steganalytic techniques specifically designed to defeat those steganographic methods using JPEG file as carrier-images. In this paper, we focus on steganalysis against two kinds of JSteg-like steganographic systems. Pfitzman and Westfeld proposed an effective Chi-Square steganalytic technique that can reliably detect images with secret messages embedded in consecutive pixels [4,5]. This method provides very reliable results when the messages are embedded sequentially in image. Provos pointed out that the method could still be used for detection of randomly scattered messages by applying the same idea to smaller portions of the image [6]. However, he hasn't given any further details for this generalized approach.

So-called universal blind steganalysis including Memon's approach based on image quality measures and Farid's approach based on higher order statistics, is a meta-detection method in the sense that after training on original and stego-images database it can detect the existence of hidden messages embedded using any steganographic method regardless of the embedding domain [7,8]. Such steganalytic algorithms usually find an appropriate set of sensitive statistical quantities with "distinguishing" capabilities. Clustering algorithms or regression models can then be used to construct a classifier for carrier-images and stego-images from the collected experimental data. Universal blind steganalysis algorithms are more flexible because they can be quickly adjusted to new or completely unknown steganalytic methods. However, from the experimental results given in [7] and [8], we know that they are generally less accurate and reliable than those algorithms specifically targeted to a specific steganographic method.

In this paper, we propose a fast and effective steganalytic technique based on statistical distributions of DCT coefficients which is aimed at two kinds of popular JSteg-like steganographic systems, sequential JSteg and random JSteg for JPEG images. Our approach can not only determine the existence of hidden messages in JPEG images reliably, but also estimate the amount of hidden

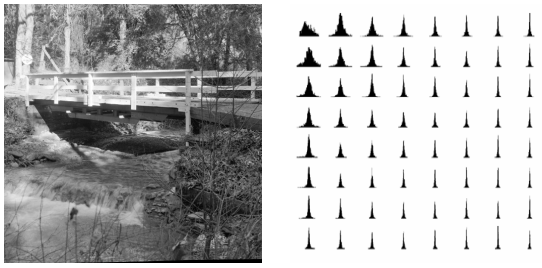
SAC 2003, Melbourne, Florida, USA

messages exactly. Its advantages also include simplicity, computational efficiency and easy implementation of real-time detection.

2. STATISTICAL DISTRIBUTIONS OF DCT COEFFICIENTS

The JPEG (Joint photographic expert group) standard is a widely used image data compression standard. First, the encoder divides an image into 8×8 blocks of pixels in the YCbCr colorspace. Then they are run through a DCT (discrete cosine transform) and the resulting frequency coefficients are quantized according to a predefined quantization table to remove the high frequency components. Finally, an entropy encoder is used to further compress the quantized DCT coefficients.

Over the past two decades, there have been various studies on the statistical distributions of the DCT coefficients for images [9-11]. Figure 1(b) shows a typical plot of the histograms of the DCT coefficients. The image used here is the "bridge" image shown in Figure 1(a) from the standard image processing library. The upper left coefficient is called the dc coefficient while the rest are ac coefficients.



(a) **(b)**

Figure 1. (a)Standard image "Bridge"; (b) DCT coefficients distributions of "Bridge".

Early on, it was conjectured that the AC(Alternating Current) coefficients have Gaussian distributions. However, soon experimental results like Figure 1(b) indicated that they resemble Laplacian distributions when the Kolmogorov–Smirnov goodness-of-fit test is used. The probability density function of a Laplacian distribution can be written as

$$\boldsymbol{p}(\boldsymbol{x}) = \frac{\boldsymbol{m}}{2}\exp\{-\boldsymbol{m}|\boldsymbol{x}|\} \tag{1}$$

E.Y. Lam et al. offer a rigorous mathematical analysis using a doubly stochastic model of the images and conclude from the high-order statistics analysis that AC coefficients have a fat-tail distribution [9]. For a general kurtosis value, a generalized Gaussian distribution is usually a better fit. It has a probability density function

$$\boldsymbol{p}_{\boldsymbol{v},b}(\boldsymbol{x}) = \frac{\boldsymbol{v}}{2b\Gamma(\frac{1}{\boldsymbol{v}})}\exp\left\{-\left(\frac{|\boldsymbol{x}|}{b}\right)^{\boldsymbol{v}}\right\} \tag{2}$$

For more information on generalized Gaussian model on DCT coefficients and parameter estimation of the model, see [10,11] for detail. It should be noted that both Gaussian distributions and Laplacian distributions are special cases of generalized Gaussian distributions with shape factor $\boldsymbol{v}$ of 2 and 1, respectively.

3. JPEG-JSTEG STEGANOGRAPHIC ALGORITHM

There are many advantages using images in JPEG format as carrier-image in steganographic applications. First, JPEG is a popular and widely-used image file format and has become a de facto standard for network image transmission. If we apply JPEG images to data hiding, the stego-image will draw less attention of suspect than that with most other formats. Second, the wide control available over image quantization makes it very difficult to establish whether or not the inaccuracies which do appear are caused by steganographic data or by lower-quality quantization. Finally, JPEG images can also offer a considerable data hiding capacity for steganographic messages. For example, standard color 512*512 Lena image, after compressed using JPEG algorithm with a quality factor of 80, produces a JPEG image file of size 56.8 kilo-bytes and its data hiding capacity is as much as 7.5 kilo-bytes using JPEG-JSteg algorithm.

JPEG-JSteg algorithm is a typical steganographic algorithm using JPEG file as carrier-image proposed by D. Upham [12]. After quantization of DCT coefficients, JPEG-JSteg replaces the least significant bits (LSB) of the quantized DCT coefficients by the secret message bits. The embedding mechanism skips all coefficients with the values 0 or 1.

Let $\boldsymbol{C}=\{C_0,C_1,\ldots, C_{n-1}\}$ denote the set of all quantized DCT coefficients of a specified carrier image and $\boldsymbol{M}=\{M_0,M_1,\ldots, M_{m-1}\}$ denote the message bits to be embedded. Now we construct a subset $\boldsymbol{S}=\{C_{l(0)},C_{l(1)},\ldots,C_{l(m-1)}\}$ ($m=n$, $l(m-1)<n$) from the set $\boldsymbol{C}$. For all the elements in $\boldsymbol{S}$ except 0 and 1, we substitute M_i for the LSB of $C_{l(i)}$. The subset $\boldsymbol{S}$ can be obtained by sequential or randomly selecting given number of elements from the set $\boldsymbol{C}$.

The JSteg algorithm proposed by D. Upham selects pixels to be replaced sequentially from the set $\boldsymbol{C}$ to obtain the subset $\boldsymbol{S}$, that is, for all $0 \le \boldsymbol{i} \le \boldsymbol{m}-1$, we have $l(i)=i$. We call that sequential JSteg algorithm. Sequential JSteg is easy to implement but has a much serious problem on security. The reason is that there is an obvious difference on statistical properties between the alternated part and the unalternated part of the stego-image. Based on that fact, A. Westfeld proposes a steganographic detection algorithm.

The other method is to select the elements randomly. First, we generate a pseudo-random number sequence $k(0), k(1),\ldots, k(m-1)$, and let $l(0)=k(0)$; $l(i)=l(i-1)+k(i)$, $0 \le \boldsymbol{i} \le \boldsymbol{m}-1$. Adjust the mean value of the pseudo-random number sequence and make $l(m)$ equal approximately to n, it will be assured that the elements in $\boldsymbol{S}$ can scatter randomly in the set $\boldsymbol{C}$. The receiver can reconstruct the set $\boldsymbol{S}$ from $\boldsymbol{C}$ by the same pseudo-random number seed and generator. We called it random JSteg algorithm. Because the secret message bits are scattered randomly in the carrier-image, random JSteg algorithm is more secure compared with sequential JSteg.

For a set of integer $\boldsymbol{G}=\{G_0,G_1,\ldots,G_{n-1}\}$, define $\|\boldsymbol{G}\|$ as the number of elements in the set, $\boldsymbol{h}_i(\boldsymbol{G})$ as the number of elements that equal to i in the set. We know that the maximum data hiding capacity are $\|\boldsymbol{C}\| - \boldsymbol{h}_0(\boldsymbol{C}) - \boldsymbol{h}_1(\boldsymbol{C})$ bits for both sequential and random JSteg algorithms.

4. IMAGE STEGANALYSIS AGAINST JSTEG-LIKE ALGORITHM

Westfeld pointed out that JSteg introduces a dependency between the value's frequencies of occurrence [4,5]. JSteg influences pairs of the coefficient's frequency of occurrence, as Figure 2 shows. Let $\boldsymbol{h_i}$ be the histogram of quantized JPEG coefficients. The assumption for a stego-image produced by JSteg algorithm is that adjacent frequencies $\boldsymbol{h_{2i}}$ and $\boldsymbol{h_{2i+1}}$ are similar except for i=0.

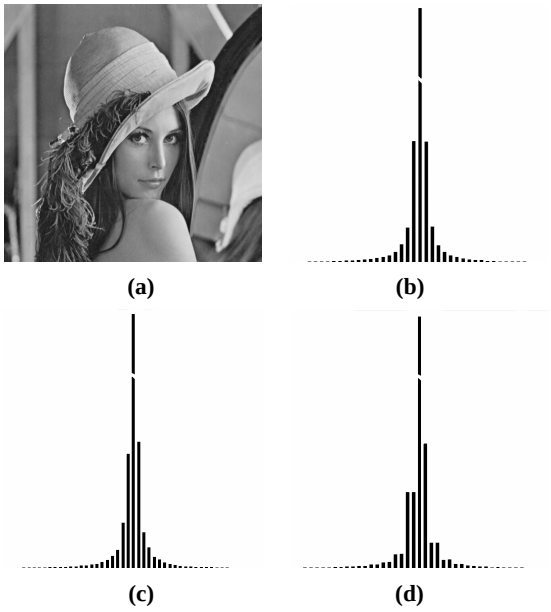


Figure 2. (a)Standard image "Lena" after JPEG compression; (b) Quantized DCT coefficients distribution; (c) DCT coefficients distribution with 1.5 KB data embedded; (d) DCT coefficients distribution with 7.5 KB data embedded.

Figure 2 shows the change on the statistical distributions of quantized DCT coefficients owing to message embedding. Figure 2 (a) shows standard image Lena after JPEG compression; (b) shows the quantized DCT coefficients distribution of (a); (c) shows the DCT coefficients distribution of (a) with 1.5 kilo-bytes data embedded; (d) shows the DCT coefficients distribution of (a) with 7.5 kilo-bytes (approximately maximum hiding capacity of this image) data embedded.

Based on the statistical difference between carrier-images and stego-images Westfeld designed a chi-square test to determine the existence of hidden messages in stego-images. However, if the subset $\boldsymbol{S}$ is selected randomly rather than sequentially, this chi-square test becomes less effective unless majority of DCT coefficients have been used for message embedding. Therefore, his approach is only effective for sequential JSteg algorithm and ineffective for random JSteg algorithm. Furthermore, Westfeld's approach is computation-consuming because chi-square test must be performed many times for accurate message length estimation.

In this paper, we propose an approach to determine the existence of hidden messages embedded using JSteg algorithm based on the statistical model of DCT coefficients. Our approach is applicable for both traditional sequential JSteg algorithm and random JSteg algorithm. By simple computation, we can also estimate the length of hidden message bits.

For a set of integer $\boldsymbol{G}=\{G_0, G_1, \ldots, G_{n-1}\}$, define $\boldsymbol{f_0}(\boldsymbol{G})$ as the total number of positive even numbers and negative odd numbers in $\boldsymbol{G}$, and $\boldsymbol{f_1}(\boldsymbol{G})$ as the total number of positive odd numbers and negative even numbers in $\boldsymbol{G}$. Note that zero elements in $\boldsymbol{G}$ are not included when calculating $\boldsymbol{f_0}(\boldsymbol{G})$ and $\boldsymbol{f_1}(\boldsymbol{G})$. Formally, they can be expressed as

$$\boldsymbol{f}_0(\boldsymbol{G}) = \sum_{\boldsymbol{i}>0, \boldsymbol{i} \bmod 2=0} \boldsymbol{h}_i(\boldsymbol{G}) + \sum_{\boldsymbol{i}<0, \boldsymbol{i} \bmod 2=1} \boldsymbol{h}_i(\boldsymbol{G}) \tag{3}$$

$$\boldsymbol{f}_1(\boldsymbol{G}) = \sum_{\boldsymbol{i}>0, \boldsymbol{i} \bmod 2=1} \boldsymbol{h}_i(\boldsymbol{G}) + \sum_{\boldsymbol{i}<0, \boldsymbol{i} \bmod 2=0} \boldsymbol{h}_i(\boldsymbol{G}) \tag{4}$$

Based on the study on the statistical distribution of DCT coefficients in Section 2 we can assume that the quantized DCT coefficients of JPEG image distribute symmetrically around zero. Experiments show that it is a rational assumption because for most images the DCT coefficients distribution satisfies symmetry around zero. Therefore, for the set of quantized DCT coefficients of JPEG images, we have

$$\boldsymbol{f}_0(\boldsymbol{C}) \approx \boldsymbol{f}_1(\boldsymbol{C}) \tag{5}$$

Furthermore, we can assume that the chi-square statistic

$$c^2(\boldsymbol{C}) = \frac{(\boldsymbol{f}_1(\boldsymbol{C}) - \boldsymbol{f}_0(\boldsymbol{C}))^2}{\boldsymbol{f}_1(\boldsymbol{C}) + \boldsymbol{f}_0(\boldsymbol{C})} \tag{6}$$

follows a chi-square distribution with 1 degree of freedom.

Moreover, we note that the message bits are usually compressed and/or encrypted prior to being embedded, so the zeros and ones are uniformly distributed in the messages bits set $\boldsymbol{M}$. Therefore, we know that the set $\boldsymbol{S}$ has a distribution similar to the one shown in Figure 2(d) and we get

$$\boldsymbol{f}_1(\boldsymbol{S}) - \boldsymbol{f}_0(\boldsymbol{S}) \approx \boldsymbol{h}_1(\boldsymbol{S}) \tag{7}$$

Let $\overline{\boldsymbol{S}} = \boldsymbol{C} - \boldsymbol{S}$ denote the complement of $\boldsymbol{S}$ in $\boldsymbol{C}$. We can assume that the statistical distribution of the set $\overline{\boldsymbol{S}}$ satisfies symmetry around zero, whether the elements of set $\boldsymbol{S}$ is randomly selected or sequentially selected from $\boldsymbol{C}$. Therefore, we have

$$\boldsymbol{f}_1(\overline{\boldsymbol{S}}) \approx \boldsymbol{f}_0(\overline{\boldsymbol{S}}) \tag{8}$$

Note that $\boldsymbol{C} = \boldsymbol{S} \cup \overline{\boldsymbol{S}}$ and Combine the equation (7) and (8), we get

$$\boldsymbol{f}_1(\boldsymbol{C}) - \boldsymbol{f}_0(\boldsymbol{C}) = (\boldsymbol{f}_1(\boldsymbol{S}) + \boldsymbol{f}_1(\overline{\boldsymbol{S}})) - (\boldsymbol{f}_0(\boldsymbol{S}) + \boldsymbol{f}_0(\overline{\boldsymbol{S}})) \approx \boldsymbol{h}_1(\boldsymbol{S}) \tag{9}$$

Let c_a^2 denote the upper tabulated value of chi-square distribution at significance level a and 1 degree of freedom. For a stego-image that contains hidden messages embedded using JSteg algorithm, we have

$$c^2(\boldsymbol{C}) > c_a^2 \text{ and } \boldsymbol{f}_1(\boldsymbol{C}) - \boldsymbol{f}_0(\boldsymbol{C}) > 0 \tag{10}$$

Equation (10) can be used as a judgment criterion for the existence of hidden messages embedded using the JSteg Algorithm in a given images. Furthermore, we can reliably estimate the length of the message bits according to equation (9).

Define the embedding ratio β as the proportion of the message bits length to the maximum embedding capacity ($\|\boldsymbol{C}\|$- $\boldsymbol{h}_0(\boldsymbol{C})$- $\boldsymbol{h}_1(\boldsymbol{C})$ for JSteg algorithm). Now, we discuss how to estimate the embedding ratio β for sequential JSteg algorithm and random JSteg algorithm, respectively.

(1)For sequential JSteg algorithm

Construct a subset $\boldsymbol{S}'$ of $\boldsymbol{C}$ in the same way as sequential JSteg algorithm, and ensure that $\boldsymbol{h}(\boldsymbol{S}')$ is equal to $\boldsymbol{f}_1(\boldsymbol{C})$- $\boldsymbol{f}_0(\boldsymbol{C})$. Therefore, the embedding ratio

$$\beta \approx \frac{\|\boldsymbol{S}'\| - \boldsymbol{h}_1(\boldsymbol{S}') - \boldsymbol{h}_0(\boldsymbol{S}')}{\|\boldsymbol{C}\| - \boldsymbol{h}_1(\boldsymbol{C}) - \boldsymbol{h}_0(\boldsymbol{C})} \tag{11}$$

If $\boldsymbol{f}_1(\boldsymbol{C}) - \boldsymbol{f}_0(\boldsymbol{C}) \geq \boldsymbol{h}_1(\boldsymbol{C})$, it can be considered that the embedding ratio β is equal to 1.

(2)For random JSteg algorithm

For random JSteg algorithm, we assume that before the message embedding the set $\boldsymbol{S}$ has a statistical distribution similar to that of the set $\boldsymbol{C}$ for the randomly selection of the elements of $\boldsymbol{S}$. Because the message embedding has not made any change on the elements equal to 1 in the set $\boldsymbol{S}$ and $\boldsymbol{C}$, the embedding ratio β can be approximated by the ratio of $\boldsymbol{h}_1(\boldsymbol{S})$ to $\boldsymbol{h}_1(\boldsymbol{C})$, that is

$$\beta = \frac{\|\boldsymbol{S}\| - \boldsymbol{h}_1(\boldsymbol{S}) - \boldsymbol{h}_0(\boldsymbol{S})}{\|\boldsymbol{C}\| - \boldsymbol{h}_1(\boldsymbol{C}) - \boldsymbol{h}_0(\boldsymbol{C})} \approx \frac{\boldsymbol{h}_1(\boldsymbol{S})}{\boldsymbol{h}_1(\boldsymbol{C})} \approx \frac{\boldsymbol{f}_1(\boldsymbol{C}) - \boldsymbol{f}_0(\boldsymbol{C})}{\boldsymbol{h}_1(\boldsymbol{C})} \tag{12}$$

If $\boldsymbol{f}_1(\boldsymbol{C}) - \boldsymbol{f}_0(\boldsymbol{C}) \geq \boldsymbol{h}_1(\boldsymbol{C})$, it can be considered that the embedding ratio β is equal to 1.

Now we get the procedures of the algorithm for detection of hidden messages and message length estimation as follows:

(1)Construct the set $\boldsymbol{C}$ of quantized DCT coefficients from the given test JPEG image.

(2)Calculate $\boldsymbol{f}_0(\boldsymbol{C})$ and $\boldsymbol{f}_1(\boldsymbol{C})$ on the set $\boldsymbol{C}$ first, and then calculate the chi-square statistic $\chi^2(\boldsymbol{C})$ according to equation (6). Based on the judgment criterion given in equation (10), decide whether the given test image contains hidden messages at a predefined significance level. If not, we get the embedding ratio β =0 and the algorithm ends up.

(3)Calculate the embedding ratio β according to equation (11) or (12) for sequential JSteg and random JSteg, respectively. If β >1.0, let β equal to 1 and the algorithm ends up.

5. EXPERIMENTAL RESULTS

We do secret messages embedding, detection and length estimation experiments on the CBIR image database from Washington University (854 JPEG images totally) [13]. Secret messages to be embedded into JPEG images are randomly cut from a piece of cipher-text. We do the following two experiments on the CBIR image database:

Experiment 1: (1)Calculate the JSteg hiding capacity of the given JPEG image; (2)Embed variable length of messages into all the images in the database using sequential JSteg algorithm and the length of those messages are 0, 10, 20, 50, 80, and 100 percent of hiding capacity of corresponding carrier image; (3)Estimate the length of hidden messages in stego-images produced in step (2) using the algorithms introduced in section 4.

Experiment 2: (1)Calculate the JSteg hiding capacity of the given JPEG image; (2) Embed variable length of messages into all the images in the database using random JSteg algorithm and the length of those messages are 0, 10, 20, 50, 80, and 100 percent of hiding capacity of corresponding carrier image; (3) Estimate the length of hidden messages in stego-images produced in step (2) using the algorithms introduced in section 4.

The significant level for chi-square test in equation (10) is set to 0.05. Figure 3 (a) and (b) show the statistical distribution of estimated messages length (represented in percent of maximum data hiding capacity). For the convenience of display, all of the experimental data are shown in one figure. In the figure the vertical axis stands for the frequency that the corresponding value occurs. The Gaussian-like peak at 0 and 1 correspond to the statistical distribution of estimated message length of original carrier-images and stego-images with messages bits of maximum data hiding capacity embedded, respectively. The middle four Gaussian-like peaks correspond to message length estimation results when the length embedded messages are 10, 20, 50, and 80 percent of hiding capacity, respectively.

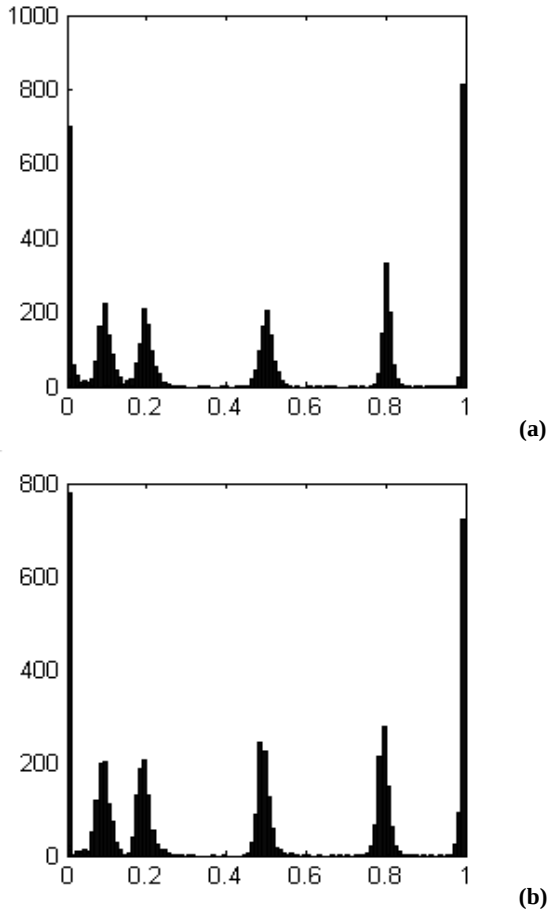


Figure 3. Histogram of estimated size of secret messages: (a)for sequential JSteg Algorithm; (b)for random JSteg Algorithm;

Table 1 gives some statistical data on experiment 1 and 2. The "± 0.05" column shows the proportion of the images for which the message length estimation results are between ± 0.05 of actual values; The "± 0.025" column shows the proportion of the images for which the message length estimation results are between ± 0.025 of actual values; The "<0.05" column shows the proportion of the images for which the message length estimation results are less than 0.05; the last two column list the mean and variance of estimated value, respectively.

Table I Statistical Results of Our Steganalytic Technique

(a) Experiment No.1

	±0.05	±0.025	<0.05	Mean	Variance
0%	96.8%	92.4%	96.8%	0.0085	2.379E-03
10%	93.0%	79.3%	4.1%	0.1031	7.064E-03
20%	92.3%	77.0%	2.5%	0.1995	5.957E-03
50%	94.0%	80.6%	2.7%	0.4896	8.490E-03
80%	95.4%	91.6%	2.6%	0.8007	1.748E-02
100%	98.7%	98.5%	0.5%	0.9930	4.702E-03

(b) Experiment No.2

	±0.05	±0.025	<0.05	Mean	Variance
0%	95.2%	92.0%	95.2%	0.0190	1.474E-02
10%	94.6%	77.5%	3.3%	0.1049	1.347E-02
20%	94.4%	80.1%	2.3%	0.2021	1.061E-02
50%	95.1%	85.4%	1.8%	0.4923	8.049E-03
80%	96.5%	89.3%	1.4%	0.7842	9.454E-03
100%	98.2%	97.4%	0.4%	0.9908	3.922E-03

From Table 1 we know that our approach can assure that the message length estimation error is less than ±5% of hiding capacity in the probability of 92% and less than ±2.5% of hiding capacity in the probability of 80%. If you just want to know the existence of hidden messages in a given image and don't care about the length of hidden messages, you can simply perform a threshold (such as 5%) operation on the estimated message length. From Table 1, we know that for existence judgment of hidden messages the false positive rate and false negative rate are all less than 5% when we select 5% as the threshold.

Our approach is easier to implement and with less computation amount than Westfeld's approach. In our experiments, total 208 seconds are used for messages length estimation on all the images in the database(854 images, 211565 Kilo- Bytes totally) and the processing speed is 1017 Kilo-Bytes per second.

6. CONCLUSIONS

We propose a fast and effective steganalytic technique based on statistical distributions of DCT coefficients which is aimed at two kinds of popular JSteg-like steganographic systems, sequential JSteg and random JSteg for JPEG images. Our approach can not only determine the existence of hidden messages in JPEG images reliably, but also estimate the amount of hidden messages exactly. Its advantages also include simplicity, computational efficiency and easy implementation of real-time detection.

7. REFERENCES

[1] F.A.P.Petitcolas, R.J.Anderson, and M.G.Kuhn, Information Hiding -- A Survey. Proceeding of IEEE, vol. 87, no. 7, pp. 1062-1078, June1999.

[2] Neil F. Johnson, Sushil Jajodia. Steganalysis of Images Created Using Current Steganography Software. LNCS Vol.1525, pp273-289, Springer- Verlag, 1998.

[3] Jiri Fridrich and M. Goljan. Practical Steganalysis: State of the Art. SPIE Vol. 4675, EI2002, January, 2002, pp.1-13.

[4] A.Westfeld and A. Pfitzmann. Attacks on Steganographic Systems. IHW 99, Dresden, Germany, 1999.

[5] A.Westfeld. F5-A Steganographic Algorithm: High Capacity Despite Better Steganalysis. IHW'01, LNCS Vol2137, Springer-Verlag, Berlin, Heidelberg, 2001.

[6] N. Provos and P. Honeyman. Detecting Steganographic Content on the Internet. ISOC NDSS'02, San Diego, CA, February 2002.

[7] N.D.Memon, I.Avcibas, B.Sankur. Steganalysis Based on Image Quality Metrics. SPIE Vol. 4314, 2001, San Jose, California, USA.

[8] H. Farid. Detecting Steganographic Message in Digital Images. Report TR2001-412, Dartmouth College, Hanover, NH, 2001

[9] E.Y. Lam, and J.W. Goodman. A Mathematical Analysis of the DCT Coefficient Distributions for Images. IEEE Trans. on Image Processing, vol 9 , No. 10, pp. 1661-1666, 2000.

[10] R. L. Joshi and T. R. Fischer. Comparison of generalized Gaussian and Laplacian modeling in DCT image coding. IEEE Signal Processing Letters, vol.2, pp. 81-82, May 1995.

[11] F. Müller. Distribution shape of two-dimensional DCT coefficients of natural images. Electronics Letters, 29(22):1935-1936, October 1993.

[12] JPEG-JSteg-V4, http://www.funet.fi/pub/crypt/ steganography/jpeg-jsteg-v4.diff.gz.

[13] CBIR Image Database, University of Washington, http:// ww.cs.washington.edu/research/magedatabase/ roundtruth/.

[Biography] Tao Zhang was born in China in 1977. He received the M.S.degree in signal and information processing from University of Information Engineering, Zhengzhou, China, in 2000. He is now a Ph.D student and his research interests include information hiding, image processing and computer vision.

Xijian Ping was born in China in 1953. He received the M.S. degree in communication engineering from Beijing University of Aeronautics and Astronautics. He is now a doctor's advisor in University of Information Engineering, Zhengzhou, China. His research interests include information hiding, image processing and computer vision.