A CONTENT-ADAPTIVE ± 1-BASED STEGANOGRAPHY BY MINIMIZING THE DISTORTION OF FIRST ORDER STATISTICS

Xinlu Gui, Xiaolong Li and Bin Yang

Institute of Computer Science and Technology, Peking University, Beijing 100871, China

ABSTRACT

Least significant bit (LSB) matching is a well-known steganographic method with advantages of high payload, good visual/statistical imperceptibility and extreme ease of implementation. However, by utilizing the distortion of one-dimensional histogram or the generated additive embedding noise, some steganalyzers can perceive the existence of covert communication to some extent. Due to this, we extend the LSB matching steganography by minimizing the distortion of first order statistics (i.e., one-dimensional histogram) and adaptively embedding data into noise regions. With these extensions, our method significantly improves the stego-security. The experimental results also prove its superiority over some state-of-the-art steganographic methods against various steganalyzers.

Index Terms— Steganography, LSB matching, histogram analysis, adaptive embedding, stego-security

1. INTRODUCTION

Steganography is a technique of covert communication, whose goal is to embed secret message into cover data (e.g., digital images) in such a way that the stego data cannot be discerned except for the intended recipients. As the contrary technique of steganography, steganalysis aims to detect the existence of secret message.

There are two widely used steganographic algorithms for digital images: least significant bit (LSB) replacement and LSB matching. They both have advantages of high payload, good visual imperceptibility and extreme ease of implementation. The embedding procedure of LSB replacement is simple: just modify the LSB of pixels in the payload as the corresponding bit of secret message. The asymmetry property of one-dimensional histogram of this embedding method leaves a clue to attackers and some recent work has shown that LSB replacement can be easily detected even when the embedding rate (secret message bits embedded per pixel) is low [1,2]. The LSB matching embedding is similar to LSB replacement: if the secret message bit does not match the LSB of the corresponding cover pixel value, the pixel value is randomly increased or decreased by 1. This symmetrical embedding procedure makes LSB matching more secure in terms of resisting steganalysis.

Though it is a difficult task, remarkable progress has been made on steganalysis of LSB matching. In [3], Harmsen and Pearlman proposed a steganalysis method based on the center of the mass of the histogram characteristic function (HCF-COM) to detect the additive-noise-based steganography. Afterwards, Ker [4] proposed an effective approach on the basis of HCF-COM and the calibration (downsample) technique. Then in [5], Li *et al.* suggested calculating the calibration-based steganalyzers on the difference image, which is defined as the difference of adjacent pixels in the original image. Besides, Zhang *et al.* [6] proposed a new method to detect LSB matching based on the statistics of amplitude of local extreme (ALE) of image's histogram. The authors observed that ALE would decrease after LSB matching embedding. Cancelli *et al.* [7] extended Zhang *et al.*'s work to the two-dimensional histogram. They experimentally demonstrated that the novel steganalyzer was much more reliable than the original one described in [6]. The ALE-based steganalyzers [6, 7] were further investigated by Gao *et al.* in [8]. Moveover, besides the targeted steganalyzers designed specifically for LSB matching, there also exist blind steganalyzers that are intended to detect a wide range of steganographic algorithms [9–11]. For instance, the wavelet absolute moment (WAM) steganalyzer proposed in [9] is reported to outperform Ker's methods [4]. In summary, although LSB matching is more secure than LSB replacement, it still can be perceived to some extent.

To resist the statistic-based steganalysis of LSB matching, an effective approach is to minimize the distortion of first order statistics (i.e., one-dimensional histogram) [12]. As is mentioned in [12], for a cover image I_c, one hope to get a stego image I_s which satisfies

$$I_s = \arg\min \|h_c - h_s\|_{l^2} \tag{1}$$

where h_c and h_s are one-dimensional histograms of the cover and stego image, respectively. By this approach, instead of randomly increasing or decreasing the pixel value, the embedding procedure is carried out under a pre-determined vector (which is dependent on the cover image) whose components are the probabilities of increasing/decreasing the pixel values. This method can effectively resist histogram analysis compared to LSB matching. However, it fails to resist noise-analysis-based steganalysis.

To remedy this drawback of the method in [12], we propose a new embedding scheme which can minimize the distortion of first order statistics, and adaptively select the embedding locations based on image content in the meantime. The main idea of content-adaptive steganography is that it is more secure to modify pixels in noise regions than smooth regions by the same amount [13]. In this way, the novel embedding method can effectively resist both histogram analysis and noise analysis, thus more difficult to detect.

The rest of this paper is organized as follows. We describe our algorithm in detail in Section 2. Then in Section 3, we analyze the security of our method and show the experimental results. The final conclusion is drawn in Section 4.

2. THE PROPOSED METHOD

Typically, data embedding is a procedure that introduces a certain type of noise into the cover image. The noise will distort the statistical distribution of cover image. According to Cachin's theory [14], one should minimize the distortion of statistical distribution as much as possible to achieve higher level of security. Besides, as previously

Corresponding author: Bin Yang, e-mail: yang_bin@pku.edu.cn

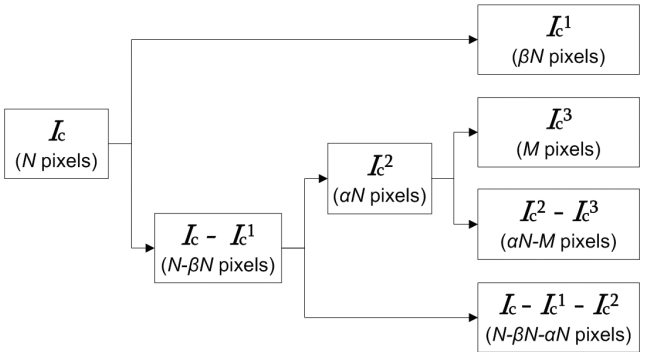


Fig. 1. Pixels classification. The pixels of I_c^1 remain unchanged in data embedding and they are used to compute the noisy levels of pixels in $I_c - I_c^1$. The pixels of I_c^3 will be modified due to data embedding, and the modification direction (increasing or decreasing the pixel value) will be specifically selected, to minimize the distortion of first order statistics.

mentioned in Section 1, one should embed data into relatively noisy locations to resist noise analysis. Considering these perspectives, we propose a steganographic algorithm that can both minimize the distortion of first order statistics and keep the smooth regions of cover image unchanged. As a result, this method can resist both histogram-analysis-based and noise-analysis-based detection. The embedding procedure of our method includes three steps: noisy level computation, pixels classification and histogram-adapted data embedding.

Above all, we compute the noisy level. Consider a gray-scale cover image I_c. Assume the embedding rate is $\alpha < 1$. We will embed αN bits into N pixels, where N is the total number of image pixels. To begin with, we choose a parameter $\beta \in (0, 1-\alpha)$, and select βN pixels in I_c according to a secret key. The chosen βN pixels are denoted as a set I_c^1. Then, for each pixel $x \notin I_c^1$, we compute its noisy level using the pixels in I_c^1. More specifically, take the smallest integer t such that the $(2t+1) \times (2t+1)$ neighborhood of x contains at least 4 pixels of I_c^1 (assume these pixels are $\{y_1, ..., y_m\} \subset I_c^1$), then the noisy level of x is defined as

$$NL(x) = \max_{1 \le i \le m} y_i - \min_{1 \le i \le m} y_i. \tag{2}$$

The noisy level function will be used to select suitable embedding pixels to carry hidden data.

With noisy level computed, we move to the step of image pixels classification. We first select the most noisy αN pixels from the set $I_c - I_c^1$ according to the noisy levels computed in (2). The selected pixels are denoted as a set I_c^2. Then for each pixel $x \in I_c^2$, assume the corresponding data bit to be embedded is w. We can get the stego pixel:

$$x_w = \begin{cases} x & \text{if } x \bmod 2 = w, \\ x+1 \text{ or } x-1 & \text{if } x \bmod 2 \neq w. \end{cases} \tag{3}$$

Particularly, when $x \bmod 2 \neq w$, the choice between $x+1$ or $x-1$ is not random in our method. We will introduce how to determine it in the next step. Denote the pixels which need to be changed in data embedding as a set I_c^3, i.e., $x \in I_c^3 \Leftrightarrow x \bmod 2 \neq w$. To sum up, we show the classification of image pixels in Fig. 1.

Now, we present the histogram-adapted data embedding procedure. Notice that, only the M pixels in I_c^3 need modification, while the other $N - M$ pixels will keep unchanged. Assume here the histograms of I_c, $I_c - I_c^3$ and I_c^3 are h_c, h_c^1 and h_c^2, respectively:

$$h_c(k) = \#\{(i,j) \in I_c : I_c(i,j) = k\} \tag{4}$$

$$h_c^1(k) = \#\{(i,j) \in I_c - I_c^3 : I_c(i,j) = k\} \tag{5}$$

$$h_c^2(k) = \#\{(i,j) \in I_c^3 : I_c(i,j) = k\} \tag{6}$$

where $0 \le k \le 255$. It is obvious that $h_c(k) = h_c^1(k) + h_c^2(k)$, and only the histogram h_c^2 will change after data embedding. Instead of randomly modifying the pixel value by 1 in the conventional LSB matching, we carry out the embedding procedure under a pre-determined vector (which is dependent on the cover image) whose components are the probabilities of increasing the pixel values. The pre-determined vector is achieved by solving an optimization problem of minimizing the distance between h_c and h_s. We assume that for each pixel value k, it changes to $k+1$ with the probability of a_k or $k-1$ with the probability of $1-a_k$. It is clear that $a_0 = 1$ and $a_{255} = 0$.

We now discuss how to determine the vector $\mathbf{a} = (a_1, ..., a_{254})^t$.

Consider the histograms of stego image: h_s, h_s^1 and h_s^2, whose definitions are exactly the same as the histograms of cover image defined in (4)-(6). For each k, we have $h_s(k) = h_s^1(k) + h_s^2(k)$, $h_s^1(k) = h_c^1(k)$, and

$$h_s^2(k) = a_{k-1} h_c^2(k-1) + (1 - a_{k+1}) h_c^2(k+1). \tag{7}$$

Then,

$$\|h_c - h_s\|_{l^2}^2 = \|h_c^2 - h_s^2\|_{l^2}^2 = \sum |h_c^2(k) - h_s^2(k)|^2$$

$$= \sum |h_c^2(k) - a_{k-1} h_c^2(k-1) - (1 - a_{k+1}) h_c^2(k+1)|^2. \tag{8}$$

The problem of minimizing $\|h_c - h_s\|_{l^2}^2$ can be reformulated as a quadratic programming problem and the solution is independent of embedding rate α. More concretely, the quadratic programming problem is:

$$\begin{cases} \text{minimize} & \frac{1}{2}\mathbf{a}^t \mathbf{Q} \mathbf{a} + \mathbf{c}^t \mathbf{a} \\ \text{subject} & 0 \le a_k \le 1 \end{cases} \tag{9}$$

where $\mathbf{a} = (a_1, ..., a_{254})^t$ is the vector to be determined, $\mathbf{Q} = (q_{i,j})_{254 \times 254}$ is a positive-definite matrix with $q_{i,j} = 0$ if $|i-j| \neq 0, 2$, and

$$q_{i,i} = 4(h_c^2(i))^2 \tag{10}$$

$$q_{i-1,i+1} = q_{i+1,i-1} = -2h_c^2(i-1)h_c^2(i+1) \tag{11}$$

$\mathbf{c} = (c_1, ..., c_{254})^t$ is a vector with, if $i \neq 2, 254$,

$$c_i = 2h_c^2(i)(h_c^2(i+2) - h_c^2(i+1) - h_c^2(i) + h_c^2(i-1)) \tag{12}$$

and

$$c_2 = 2h_c^2(2)(h_c^2(4) - h_c^2(3) - h_c^2(2) + h_c^2(1) - h_c^2(0)) \tag{13}$$

$$c_{254} = 2h_c^2(254)(-h_c^2(255) - h_c^2(254) + h_c^2(253)). \tag{14}$$

This quadratic programming problem can be solved efficiently by the ellipsoid method [15].

Finally, we summarize our data embedding procedure:

- Select βN pixels (denoted as I_c^1) from the cover image I_c using a secret key shared by the encoder and decoder. Calculate the noisy level for each pixel in $I_c - I_c^1$ according to (2).
- Choose the most noisy αN pixels from $I_c - I_c^1$ as the data embedding locations (denoted as I_c^2). Compare the pixel values of I_c^2 with the secret message and define the pixels to be changed as a set I_c^3.
- Calculate the histogram of I_c^3. Solve the quadratic programming problem in (9) to determine the vector $\mathbf{a}$.

- For each pixel in I_c^3, increase the pixel value k by 1 with possibility a_k or decrease it by 1 with possibility $1 - a_k$.

In particular, it should be mentioned that, through the whole data embedding process, pixels in I_c^1 remain unchanged, and the noisy level computation step is fully based on the set I_c^1, which is the key point of our method.

The data extraction procedure of our method is as follows. The decoder first locates the set I_c^1 according to the secret key. Then it computes the noisy levels of pixels in $I_c - I_c^1$ just as the encoder. Finally, the decoder selects the most noisy pixels as embedding locations (i.e., to determine the set I_c^2) and rebuilds the secret message by linking up the LSBs of the chosen pixel values.

3. EXPERIMENTAL RESULTS

The security of steganography is commonly measured by its performance against steganalysis. Existing steganalysis methods can be classified into two categories: targeted and blind. Targeted steganalyzers normally pick out some special statistical magnitudes for consideration which are related to the data embedding procedure, while blind steganalyzers generally extract some exquisite features to measure the noise introduced by data embedding. It has been proved that LSB matching can be perceived, to some extent, by both targeted and blind steganalyzers.

Our work improves LSB matching by minimizing the distortion of first order statistics and adaptively selecting embedding locations, which helps LSB matching achieve higher security level. Targeted steganalyzers on LSB matching, such as calibrated HCF-COM [4] and ALE [6], utilize various characters to measure the distortion of cover's histogram after LSB matching embedding. As we minimize this distortion, the characters picked out for classifying are not sensitive, which makes the steganalyzers less reliable. In addition, our method implements content-adaptive embedding by calculating noisy levels, which can effectively resist noise analysis. Since we embed message into noisy regions, the noise introduced to the image is much more unperceived. Thus the blind steganalyzers relying on noise measurement will lose efficacy to some extent.

Our experiments are conducted as below.

- **Image set:** We download 3000 images from the USDA NRCS Photo Gallery[1]. After the download, these images are changed into gray-scale format. Then, in each image, a square area with maximum size is cropped. Finally, each cropped image is downsampled to 512×512 pixels.
- **Compared embedding methods:** To evaluate the proposed method, the conventional LSB matching, the optimized embedding method [12] which minimizes the distortion of first order statistics and the recently proposed content-adaptive embedding method [13] are adopted.
- **Steganalyzers:** Our experiments introduce three detection methods: the targeted steganalyzer calibrated HCF-COM computed on the difference image [5], the classical blind steganalyzer WAM [9] and the blind steganalyzer proposed in [16].
- **Parameters:** Two parameters are taken into consideration. One is the embedding rate α, which is chosen as 0.5 and 0.3 in our experiments. In each experiment, the three compared methods and our method are conducted under the same embedding rate. The other is the parameter β. We try several values and will discuss its influence on the result later.

[1] http://photogallery.nrcs.usda.gov

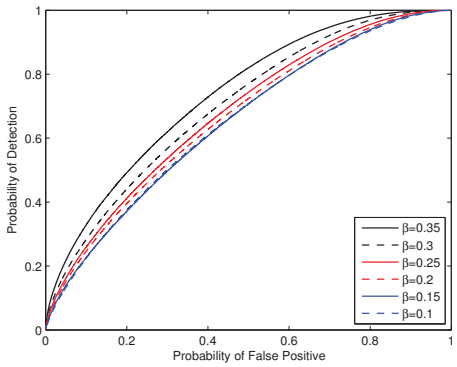


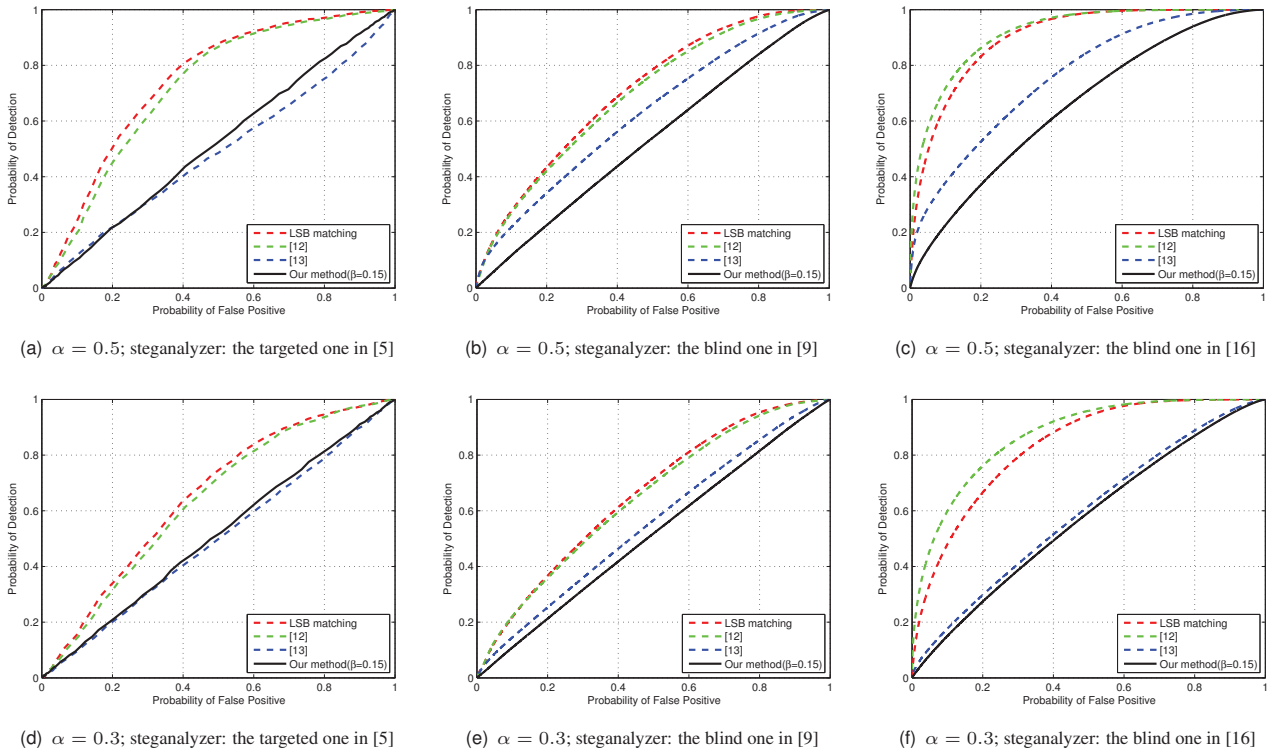
Fig. 2. ROC curves of different parameter β. Here, the embedding rate is 0.5.

Above all, we try several values of β at an embedding rate of 0.5 to make a decision on this parameter. The receiver operating characteristic (ROC) curves for steganalyzer [16] is shown in Fig. 2. The area under ROC curve (AUC) measures the general probability of correct classification between cover and stego image. Notice that a high AUC value indicates excellent discrimination and low embedding security. The figure demonstrates that the value of β plays an important role here. It can be observed that the result of $\beta = 0.35$ is much inferior to that of $\beta = 0.15$. We try the same values for two more steganalyzers and another embedding rate, and get similar results that $\beta = 0.15$ outperforms other values. Based on above experiments, we can conclude that $\beta = 0.15$ is a better choice, so the following experiments are conducted with $\beta = 0.15$.

Further experimental results are shown in Fig. 3, which reveals that our scheme (black solid line) is superior to the others. Specifically, for targeted steganalyzer in Fig. 3(a)(d), the ROC curve of the method minimizing the first order distortion in [12] is slightly below that of LSB matching, while curves of the content-adaptive method in [13] and our method are close to straight lines, which implies that the steganalyzer can hardly perceive the covert message. When it comes to blind steganalyzer in Fig. 3(b)(c)(e)(f), our method distinctly excels the LSB matching and method in [12], and besides, is visibly better than the method in [13]. We furthermore explore the reasons. Commonly blind steganalyzers detect the covert message by measuring variation of noise after embedding. Our improvement remarkably limit the variation by content-adaptive embedding in the noisy regions; in contrast, the method in [12] does not change the noise introduced. Our method also minimizes the distortion of first order statistics compared with the method in [13]. In conclusion, the experimental results prove that our method can effectively resist both targeted and blind steganalyzers.

4. CONCLUSION

In this paper, we proposed an improved ± 1-based embedding method by both minimizing the distortion of first order statistics and adaptively embedding into noisy regions. According to Cachin's theory [14], the less distortion of the statistical distributions between image and stego image, the higher level of security can be achieved. At the same time, we guaranteed that the message is embedded into noisy regions, making noise-analysis-based steganalysis invalid. Experiments showed that our method significantly improved the level of security. The proposed method is more secure than the conventional LSB matching or some state-of-the-art steganographic methods against both the targeted and blind steganalyzers. The fu-



(a) $\alpha = 0.5$; steganalyzer: the targeted one in [5]

(b) $\alpha = 0.5$; steganalyzer: the blind one in [9]

(c) $\alpha = 0.5$; steganalyzer: the blind one in [16]

(d) $\alpha = 0.3$; steganalyzer: the targeted one in [5]

(e) $\alpha = 0.3$; steganalyzer: the blind one in [9]

(f) $\alpha = 0.3$; steganalyzer: the blind one in [16]

Fig. 3. Comparison of ROC curves of different embedding methods for different steganalyzers.

ture work will focus on how to efficiently minimize the distortion of higher order statistics.

5. REFERENCES

[1] J. Fridrich, M. Goljan, and R. Du, "Detecting LSB steganography in color and gray-scale images," *IEEE Multimedia*, vol. 8, no. 4, pp. 22–28, October–December 2001.

[2] A. D. Ker, "A general framework for structural steganalysis of LSB replacement," in *Proc. of the 7th International Workshop on Information Hiding*, 2005, vol. 3727 of *LNCS*, pp. 296–311.

[3] J. J. Harmsen and W. A. Pearlman, "Steganalysis of additive noise modelable information hiding," in *Security and Watermarking of Multimedia Contents V*, 2003, vol. 5020 of *SPIE*, pp. 131–142.

[4] A. D. Ker, "Steganalysis of LSB matching in grayscale images," *IEEE Signal Process. Lett.*, vol. 12, no. 6, pp. 441–444, June 2005.

[5] X. Li, T. Zeng, and B. Yang, "Detecting LSB matching by applying calibration technique for difference image," in *Proc. of the 10th Workshop on Multimedia & Security*, 2008, pp. 133–138.

[6] J. Zhang, I. J. Cox, and G. Doërr, "Steganalysis for LSB matching in images with high-frequency noise," in *Proc. IEEE MMSP*, 2007, pp. 385–388.

[7] G. Cancelli, G. Doërr, I. J. Cox, and M. Barni, "Detection of +-1 LSB steganography based on the amplitude of histogram local extrema," in *Proc. IEEE ICIP*, 2008, pp. 1288–1291.

[8] Y. Gao, X. Li, B. Yang, and Y. Lu, "Detecting LSB matching by characterizing the amplitude of histogram," in *Proc. IEEE ICASSP*, 2009, pp. 1505–1508.

[9] M. Goljan, J. Fridrich, and T. Holotyak, "New blind steganalysis and its implications," in *Security, Steganography, and Watermarking of Multimedia Contents VIII*, 2006, vol. 6072 of *Proc. SPIE*, pp. 1–13.

[10] S. Lyu and H. Farid, "Steganalysis using higher-order image statistics," *IEEE Trans. Inf. Forens. Security*, vol. 1, no. 1, pp. 111–119, March 2006.

[11] H. Gou, A. Swaminathan, and M. Wu, "Noise features for image tampering detection and steganalysis," in *Proc. IEEE ICIP*, 2007, vol. 6, pp. 97–100.

[12] Y. Lu, X. Li, and B. Yang, "A $\pm$1-based steganography by minimizing the distortion of first order statistics," in *Proc. IIH-MSP*, 2009, pp. 754–758.

[13] W. Luo, F. Huang, and J. Huang, "Edge adaptive image steganography based on LSB matching revisited," *IEEE Trans. Inf. Forens. Security*, vol. 5, no. 2, pp. 201–214, June 2010.

[14] C. Cachin, "An information-theoretic model for steganography," *Information and Computation*, vol. 192, no. 1, pp. 41–56, July 2004.

[15] D. G. Luenberger and Y. Ye, ," in *Linear and nonlinear programming.*, 2008, New York, USA: Springer.

[16] B. Li, J. Huang, and Y. Q. Shi, "Textural features based universal steganalysis," in *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, 2008, vol. 6819 of *Proc. SPIE*, pp. 681912–681912–12.