SCIENCE CHINA
Information Sciences

• **LETTER** •

May 2016, Vol. 59 059103:1–059103:3
doi: 10.1007/s11432-016-5530-z

Novel cover selection criterion for spatial steganography using linear pixel prediction error

Wei HUANG[1]* & Xianfeng ZHAO[2]

[1]*Software School, Xiamen University, Xiamen* 361005*, China;*
[2]*State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing* 100093*, China*

Received October 15, 2015; accepted November 24, 2015; published online March 4, 2016

Citation Huang W, Zhao X F. Novel cover selection criterion for spatial steganography using linear pixel prediction error. Sci China Inf Sci, 2016, 59(5): 059103, doi: 10.1007/s11432-016-5530-z

Dear editor,

In recent years, steganography and steganalysis [1] of empirical covers have been developing rapidly. Empirical cover, such as image and other multimedia, is one of the most commonly used content in daily life, however, the model of images is still unknown. Therefore, the combat against image steganography and steganalysis would be challenging yet beneficial.

Cover selection refers to the technology for steganographers to distribute payloads among a set of images. Therefore, it is significant to design a metric (i.e. distortion) to indicate the detection error when cover objects are used in steganography. It is widely accepted that the detection accuracy of steganalysis can be influenced by many factors. Böhme [1] proposed that various moderating factors are particularly relevant to the detection accuracy of steganalytic methods, however, the precise relations remained opaque. Kouider et al. [2] designed an adaptive steganographic method by building the detectability map of a classifier as an oracle (also called ASO) to select covers. Nonetheless, this closed-loop design is time-consuming. Researchers [3] listed a few important questions when moving steganography and steganalysis from the laboratory into the real world; one of these is listed under Open Problems 9 as investigating approaches to "perform cover selection, if at all".

In this letter, we propose a novel cover selection criterion for spatial steganography by using linear pixel prediction error. We update Böhme's work [1] by raising a criterion that is more accurate to indicate the detection rate and is compatible with highly undetectable modern steganography [4–6]. The linear prediction error (LPE) model can be utilized for modeling the relationship among image pixels in the spatial domain. Its prediction error, i.e., the linear pixel prediction error, is used as a "sum-of-cost" measure to predict the detection accuracy of a detector. Experimental results show that, images with high LPE always turn out to have a low detection accuracy. Moreover, the Spearman's correlation coefficient of LPE and detection accuracy is, on average, greater than main criteria (entropy, and local variances), by a value of 0.05. Therefore, our proposed method is superior to the main criteria in most cases. The contribution of this study is as follows: (a) a novel cover selection criterion LPE by modeling the linear relationship among image pixels is proposed; (b) our procedure is open-loop because no feature extraction or classifiers is required; (c) it is effec-

* Corresponding author (email: whuang@xmu.edu.cn)
The authors declare that they have no conflict of interest.

tive in most of spatial steganography.

Linear pixel prediction error. The statistical distribution of pixel values in natural images is not statistically independent. However, thus far the description of the exact model of natural images is unavailable in literature. Such uncertainty pertaining pixel values stem from the variety in nature of images and the noise generated in image acquisition and encoding.

Pixels of natural images always have dependencies on adjacent pixels and have limited relevance with the pixels in distant regions. Therefore, the strict definition of neighbouring pixels refers to the pixels directly next to the pixel in eight directions, i.e. two vertical, two horizontal, and four diagonal neighbours, given by

$$\mathcal{N}\left(I_{u,v}\right) = \left\{I_{u+\Delta u, v+\Delta v}, |\Delta u| \leqslant 1, |\Delta v| \leqslant 1\right\},$$

where $I_{u,v} \in \{0, 1, \ldots, 255\}$ denotes the grey scale at the position (u, v). We can extend the definition to the k-layer neighbouring pixels, and is given by

$$\mathcal{N}\left(I_{u,v}, k\right) = \left\{I_{u+\Delta u, v+\Delta v}, |\Delta u| \leqslant k, |\Delta v| \leqslant k\right\},$$

where $k = 1, 2, \ldots, \min\left(M, N\right)$ and $k+1 \leqslant u \leqslant M-k, k+1 \leqslant v \leqslant N-k$.

A pixel value $I_{i,j}$ is somewhat related to its neighbouring pixels $\mathcal{N}\left(I_{i,j}, k\right)$. Keeping this in mind, we construct a predictor to estimate the pixel values. For the sake of convenience in computation, the pixel matrix is reshaped to a vector $\boldsymbol{y} = \{y_1, \ldots, y_n\}$. Each element in $\boldsymbol{y}$ is viewed as the sum of two components, the predictable and the unpredictable, i.e., $y_i = \hat{y}_i + \varepsilon_i = f\left(\boldsymbol{x}_i, \boldsymbol{\beta}\right) + \varepsilon_i$, where the dependent variable y_i is the true value of predicted values $\hat{y}_i$, the independent variable $\boldsymbol{x}_i$ is the vector of the neighbouring pixel values of y_i, f is prediction model with $\boldsymbol{\beta}$ as its parameters, and ε_i is the random error. For computational convenience, we select the multiple linear regression model $y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \cdots + \beta_n x_{i,m} + \varepsilon_i$, which can be expressed in the matrix form

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{1}$$

with the dependent value $\boldsymbol{y} = \{y_1, y_2, \ldots, y_n\}^{\mathrm{T}}$, the regression coefficient $\boldsymbol{\beta} = \{\beta_0, \beta_1, \beta_2, \ldots, \beta_m\}^{\mathrm{T}}$, the random error $\boldsymbol{\varepsilon} = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n\}^{\mathrm{T}}$, and the independent value

$$\boldsymbol{X} = \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \\ \vdots \\ \boldsymbol{x}_n \end{pmatrix} = \begin{pmatrix} 1 & x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ 1 & x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ & & \vdots & & \\ 1 & x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{pmatrix}.$$

The first column of the matrix $\boldsymbol{X}$ is $\boldsymbol{1}_n$ which is comprised of 1's. In each row $\boldsymbol{x}_i$, except the first element is one, the remaining elements $(x_{1,1}, x_{1,2}, \ldots, x_{1,m})$ are the k-layer neighbouring pixels of y_i with $m = (2k+1)^2 - 1$. Then, the regression coefficients $\boldsymbol{\beta}$ can be solved using the least squares principle:

$$\boldsymbol{b} = \arg\min_{\boldsymbol{\beta}} \left[\left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}\right)\right],$$

and $\boldsymbol{b}$ is the least square estimation of $\boldsymbol{\beta}$. Proved by Theorems 3.12 and 3.13 in [7], the unbiased least squares estimation of $\boldsymbol{\beta}$ for the multiple linear regression model (1) is given by

$$\boldsymbol{b} = \left(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}\right)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{y}.$$

Here, we assume $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ to be a non-singular matrix, which means that column vectors of $\boldsymbol{X}$ are independent.

While random error $\boldsymbol{\varepsilon}$ is not observable, the residual $e_i = y_i - \hat{y}_i$ can be directly calculated as the distance from the calculated response $\hat{y}_i = \boldsymbol{x}_i\boldsymbol{\beta}$ to the observed value y_i. Now, we define the linear pixel prediction error (LPE), as the mean squared error (MSE) $E = \frac{1}{n}\sum_{i=1}^{n} e_i^2$ of the linear pixel predictor we have described. Figure 1(a) shows the flow chart of the process for calculating LPE.

In essence, the prediction error of the linear pixel predictor indicates the extent to which the image pixels follow a certain pattern. Pixel values in complicated texture area gain more diversity than those in simple texture area. Therefore, the possibility of detecting a modification or disturbance in simple texture area is relatively higher than that in a complicated texture area.

The proposed scheme can also be viewed from the aspect of image processing. The matrix $\boldsymbol{b}$ is the filter in which b_i is the ith filter coefficient. These filter coefficients are obtained using regression on the source image and this process can be viewed as a filter of the image with least squared error. The prediction error is the differential image between a filtered image and its source.

In principle, the criteria for two approaches—entropy and local variances are limited to adjacent pixels. The entropy metric calculates the entropy using the first-ordered TPM of adjacent pixel values. LPE uses a more powerful model than LV as well as entropy. LV can be viewed as a special case of LPE when $y = x + \varepsilon$ is adopted as the model. Unlike LV, LPE uses a linear model which has more parameters to choose from, and in this model each pixel relies on more neighboring pixels (e.g. 8 pixels if $k = 1$). Hence, LPE has a better modeling ability to select noisy images for spatial steganography.

Experiments. We conduct our experiment using BOSS v0.92 [6] and BOWS2 image databases,

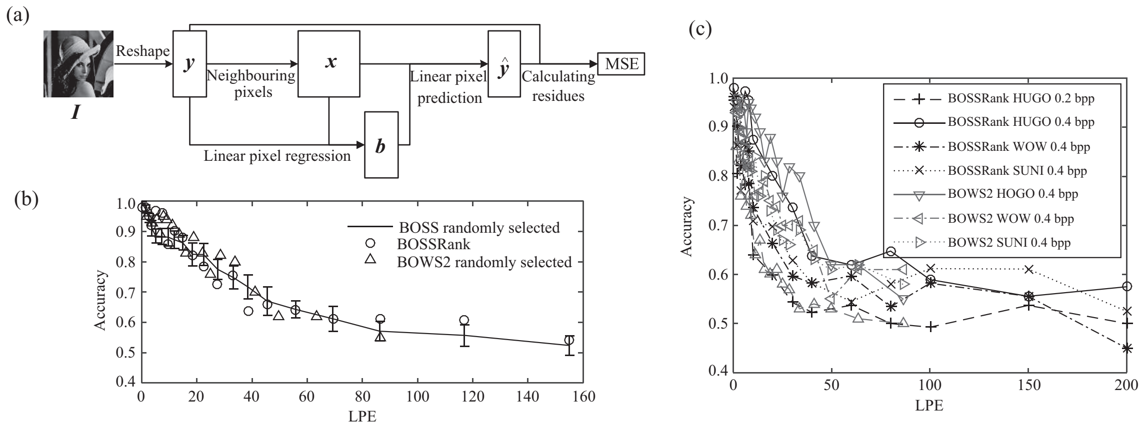

Figure 1 (a) Flow chart for implementation of the proposed scheme; (b) detection accuracy vs. LPE with HUGO 0.4 bpp; (c) detection accuracy vs. LPE with HUGO, WOW, and S-UNIWARD.

which contains 512×512 grey-scale images. Stego sets are generated by embedding randomly generated bytes using different steganographic schemes (HUGO [6], WOW [4], S-UNIWARD [5], etc.) with different embedding rates (0.05–0.60 bpp). They are classified by the ensemble Fisher linear discriminant (FLD) classifiers [8] with HOLMES features.

To demonstrate the stability of the correlations between LPE and accuracy, we randomly select images and divide them into small subsets by LPE interval through a set of thresholds. For the sake of convenience, we used error bars to indicate the deviation along the LPE-Accuracy curve to verify the consistencies in the tendencies under different conditions. The results are plotted graphically as shown in Figure 1(b). The graph indicates that the accuracy declines steadily as LPE rises, with small standard deviation ranges.

Spearman's rank correlation coefficient, which can be calculated using both the criterion and accuracy, is used to evaluate the monotonicity. The result with different steganography and embedding rates (Figure 1(c) and Table S1) shows that the detection accuracy declines as LPE is raised and the prediction curve decreases monotonically. In most of the cases, the LPE index is more relative than local variances and entropy, and using LPE in cover selection can achieve better precision in terms of choosing images that are difficult to analyze accurately. Similar effectiveness is also evident in the test of traditional steganography (LSB, QIM, etc.) and different image sizes (Table S2).

Conclusions. Inspired from content-adaptive steganographic design, this study proposed linear prediction error (LPE) as an evaluation criterion of image texture, and used it to select the most undetectable cover images. Experiments show that, in most of the cases, the proposed measure achieves a higher correlation than the local variance and entropy metrics. In our future work, we will focus our attention on quantised DCT domain.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 61402390, 61170281), Strategic Priority Research Program of Chinese Academy of Sciences (Grant No. XDA06030600).

Supporting information Tables S1 and S2. The supporting information is available online at info. scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Böhme R. Principles of modern steganography and steganalysis. In: Advanced Statistical Steganalysis. Berlin: Springer, 2010. 11–77
2 Kouider S, Chaumont M, Puech W. Technical points about adaptive steganography by oracle (ASO). In: Proceedings of IEEE 20th European Signal Processing Conference (EUSIPCO), Bucharest, 2012. 1703–1707
3 Ker A D, Bas P, Böhme R, et al. Moving steganography and steganalysis from the laboratory into the real world. In: Proceedings of 1st ACM Workshop on Information Hiding and Multimedia Security, New York, 2013. 45–58
4 Holub V, Fridrich J. Designing steganographic distortion using directional filters. In: IEEE International Workshop on Information Forensics and Security, Tenerife, 2012. 234–239
5 Holub V, Fridrich J, Denemark T. Universal distortion function for steganography in an arbitrary domain. EURASIP J Inf Secur, 2014, 1: 1–13
6 Bas P, Filler T, Pevný T. Break our steganographic system: the ins and outs of organizing BOSS. In: Proceedings of 13th International Conference on Information Hiding, Prague, 2011. 59–70
7 Yan X, Su X G. Linear Regression Analysis: Theory and Computing. Singapore: World Scientific, 2009
8 Kodovsky J, Fridrich J, Holub V. Ensemble classifiers for steganalysis of digital media. IEEE Trans Inf Foren Secur, 2012, 7: 432–444