

## Review of distortion function design in spatial domain

Jun Li

Key Laboratory of Network & Information Security under the Chinese Armed Police Force,  
Electronic Department, Engineering College of the Armed Police Force, Xi'an 710086, China

E-mail: lijun9250lj@163.com

**Keywords:** Steganography; distortion function; cost; spatial domain;

**Abstract.** In modern steganography, designing a distortion function to get the cost for every cover elements is an open problem. Many researchers have proposed a lot of different valuable distortion functions from different point of view. In this paper, the distortion function are classified to three categories, the first is from the view of steganalysis to design distortion function, the second is the distortion function are designed by filters bank, the third is the distortion function is designed begin with a well model of images. It is hoped that the summarized result in this paper is valuable for designing other better distortion function.

### Introduction

In steganography, the sender communicates with the receiver by hiding their messages in trusted media, Such as digital images, so that it is hard to distinguish between the cover image and the stego image. Normally, there are two ways to minimize the statistical detectability in empirical covers. The first one relies on a defined cover image model, which is preserved by embedding progress. Such steganography will be perfectly secure with respect to this model, however, since no perfect model of digital images exists, this approach can be usually well detected by steganalysis which working outside of the model[1-4]. The second one for steganography is usually by minimizing the embedding distortions function[5-7]. Since there exist a general framework for this embedding paradigm with established payload distortion bounds as well as near-optimal practical coding schemes, building an embedding scheme has been essentially reduced to the distortion design. In recent years, the second method become more and more popular, because this approach is more flexible and enables development of steganographic methods driven by the detection performance of steganalysis. We focus on the second method in this paper.

Fig 1 is the framework of minimizing distortion steganography. The data embedding process consists of a distortion function and an efficient encoding method. The sender embed the given payload into the cover images and get the stego image. The distortion is obtained by first assigning a cost of changing each cover pixel and then computing the total distortion as a sum of costs of all modified pixels. The larger costs are assigned to pixels where the detection is expected to be easier, and on the contrary, the smaller costs assigned to pixels where the detection is expected to be harder. So that the embedding changes usually concentrate in textured or noisy regions. An embedding algorithm designed by Syndrome Trellis Code (STC) [5,6] coding can provide near-optimal practical coding schemes to embed messages. Within this framework, the only task left to the sender is essentially the design of the distortion function.

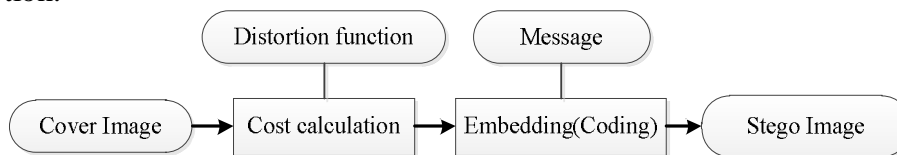


Fig 1 Framework of Minimizing Distortion Steganography

In the next section, we introduce the details of the framework for minimizing distortion steganography. In Section 3, we summarized several method of design distortion function, and we divide them to three types. The problem of distortion design and the future development direction are given in section 4.

### Minimizing distortion steganography

In spatial domain, the cover and stego image can be expressed as  $X = (x_1, \dots, x_n)$ ,  $Y = (y_1, \dots, y_n) \in \{0, 1\}^n$  respectively, the cost by the process of embedding is defined by follow:

$$D(\mathbf{X}, \mathbf{Y}) = \|\mathbf{X} - \mathbf{Y}\|_r = \sum_{i=1}^n r_i |x_i - y_i| \quad (1)$$

Where  $0 \leq r_i \leq \infty$  is the cost of changing pixel  $x_i$  to  $y_i$ . When  $r_i = \infty$  means that this pixel is the so called wet pixel which is not allowed to change [8]. For the additive distortion function  $r$  in (1), paper [Error! Bookmark not defined.] has the following separation principle:

Let  $r = (r_i)_{i=1}^n$ ,  $0 \leq r_i \leq \infty$  is the set of constants defining the additive distortion measure (1) for  $i \in \{1, \dots, n\}$ . Let  $0 \leq m \leq n$  be the number of bits we want to communicate by using a binary embedding operation. The minimal expected distortion has the following form:

$$D_{min} = (m, n, \mathbf{r}) = \sum_{i=1}^n p_i r_i, \quad p_i = \frac{e^{-r_i}}{1 + e^{-r_i}} \quad (2)$$

Where  $p_i$  is the probability of changing the  $i$ th pixel. The parameter  $\lambda$  is obtained by solving following embedding capacity equation:

$$-\sum_{i=1}^n (p_i \log_2 p_i + (1 - p_i) \log_2 (1 - p_i)) = m \quad (3)$$

The importance of above principle is in the separation of the image model (needed for calculating  $r_i$ ) and the coding algorithm used in a practical implementation. By virtue of this separation, better steganographic algorithms can be derived by using better coding or by using a better image model. The optimal coding can be simulated by flipping each pixel with probability  $p_i$  as defined in (2), which is the so-called simulated optimal embedding method. In 2010, Filler, Judas, and Fridrich proposed the practical near-optimal embedding method syndrome-trellis codes [Error! Bookmark not defined., Error! Bookmark not defined.]. So the designer of steganography can propose practical algorithms which are minimizing embedding impact with the use of STCs.

### Some distortion functions for spatial steganography in recent years

With the help of STC, the remaining works leaving for researches is only the design of distortion function. Paper [9] formulates the process of cost assignment into two steps: (1) determining a priority profile, (2) specifying a cost-value distribution. However, in this paper, the distortion function are classified to three categories, the first is from the view of steganalysis to design distortion function, the second is the distortion function are designed by filters bank, the third is the distortion function is designed begin with a well model of images.

### Steganalysis based method

HUGO (Highly Undetectable steGO) [10] is the first steganography algorithm using the framework of minimizing distortion. The methodology for designing distortion function of HUGO is that the features for steganalysis can serve as a good precursor of the image model to determine the embedding costs. HUGO defines the distortion as a weighted norm between higher-order statistics of pixel differences in cover and stego images, with high weights assigned to well-populated bins and low weights to sparsely populated bins that correspond to more complex content. The steganalysis features used in HUGO is SPAM (Subtractive Pixel Adjacency Matrix) [11], which model dependencies

between neighboring pixels by means of higher-order markov chains. The embedding changes of HUGO will be made concentrate in texture regions and along edges. The success of HUGO has been approved in BOSS competition [12].

### **Filters based method**

Many other distortion functions are often assigning the pixel costs by quantifying the impact of making an embedding change on outputs of one or more filters (noise residuals).

WOW (Wavelet Obtained Weights)[13] employ a bank of directional high-pass filters to obtain the so-called directional residuals, which are related to the predictability of the pixel in a certain direction. By measuring the impact of embedding on every directional residual and by suitably aggregating these impacts, WOW force the embedding cost to be high where the content is predictable in at least one direction (smooth areas and along edges) and low where the content is unpredictable in every direction (e.g., in textured or noisy areas). As a result, WOW has resolved the problem that the HUGO has existed, which is the distortion function of HUGO concentrates the embedding changes not only in textures but also in edges, however, the content along an edge can usually be well modeled using locally polynomial models [14-16].

The cost proposed in paper [17] named S-UNIWARD (Spatial - UNIversal WAvelet Relative Distortion), is in the form of a sum of relative changes between the stego and cover images represented in the wavelet domain like the WOW. Therefore, S-UNIWARD is similarity to WOW and has the almost same statistic security against steganalysis, but is simpler and suitable for embedding in other domain (JPEG and Side-JPEG).

Although WOW and S-UNIWORD have exploited more pixels in texture areas for hiding data, authors of algorithm HILL (High-pass, Low-pass, and Low-pass)[18] argue that some pixels in texture areas, which may be suitable for carrying data, should assigned with high costs. Intuitively, embedding changes made in such pixels should be more secure than in pixels located in smooth areas. In this sense, the cost function of HILL ensuring all pixels within textural regions have relatively low costs, and is realized by using a high-pass filter and two low-pass filters, making more embedding changes concentrated in textural areas, and thus achieves a better performance on resisting the steganalysis over prior works, such as HUGO, WOW, and S-UNIWARD.

### **Model driven method**

How to construct a well statistic model for modeling cover image is a difficult problem for a long time. Recently, many approaches design distortion function begin with a good statistic image model.

Different to WOW, S-UNIWARD, or HILL, MG (Multivariate Gaussian)[19] adopt a different strategy in which the cover is modeled as a sequence of independent but not necessarily identically distributed quantized Gaussians. Then, the costs (the probabilities of modifying each pixel or change rates) are computed analytically from the estimated model in order to minimize the Kullback-Leibler divergence between the cover and stego distributions or, equivalently, the power of an optimal statistical test. It becomes tractable to compute the optimal embedding change probabilities (costs) using the method of Lagrange multipliers.

MVGG (Multivariate Generalized Gaussian)[20] use the similar framework with MG, and has two improvement. Firstly, MVGG explore the possibility to further improve the empirical security by using a more general cover model. Rather than use the multivariate Gaussian distribution to model pixels, MVGG employ the Multivariate Generalized Gaussian model. Secondly, as the generalized Gaussian can have thicker tails, it makes sense to allow embedding changes with amplitude larger than 1 to embed a larger payload in pixels from textured areas, MVGG use pentary embedding rather than ternary embedding which is used in MG, As a result, MVGG algorithm can improve security especially for large payloads.

In the base of MV and MVGG, paper [21] propose the MiPOD (Minimizing the Power of Optimal Detector) algorithm, which is based on a locally-estimated multivariate Gaussian cover image model

that is sufficiently simple to derive a closed-form expression for the power of the most powerful detector of content-adaptive LSB matching. At the same time, the proposed model is complex enough to capture the non-stationary character of natural images. The closed-form expression for detectability within the chosen model is used to obtain new fundamental insight regarding the performance limits of empirical steganalysis detectors built as classifiers.

## Conclusion

Minimizing the statistical detectability with respect to the additive distortion function is an open problem. The most difficult issue is the image model can never be known precisely. In this paper, the steganographic scheme based on the principle of minimizing embedding distortion is described. It can be seen that the current design of the distortion function, although a bit, but the general lack of rigorous theoretical basis, many algorithms based on a variety of assumptions. It is hoped that the content of this paper will be helpful to the design of distortion function in the future.

## References

- [1] R. Böhme. Improved Statistical Steganalysis Using Models of Heterogeneous Cover Signals. PhD thesis, Faculty of Computer Science, Technische Universität Dresden, Germany, 2008.
- [2] R. Böhme. Advanced Statistical Steganalysis. Springer-Verlag, Berlin Heidelberg, 2010.
- [3] A. Westfeld, “F5-A steganographic algorithm,” in Information Hiding (Lecture Notes in Computer Science), vol. 2137. Berlin, Germany: Springer-Verlag, 2001, pp. 289–302.
- [4] P. Sallee, “Model-based steganography,” in Digital Watermarking (Lecture Notes in Computer Science), vol. 2939. Berlin, Germany: Springer-Verlag, 2003, pp. 154–167.
- [5] Filler T, Judas J, Fridrich J, (2011) Minimizing additive distortion in steganography using Syndrome-Trellis Codes. IEEE T INF FOREN SEC, 6(3):920-935. doi: 10.1109/TIFS.2011.2134094.
- [6] Filler T, Fridrich J, (2010) Minimizing additive distortion functions with non-binary embedding operation in steganography. in: Proceedings of 2nd IEEE Workshop on Information Forensics and Security(WIFS'2010), Vol. 41, 2010, pp.1-6. doi: 10.1109/WIFS.2010.5711444.
- [7] Fridrich J, Filler T, (2007) Practical methods for minimizing embedding impact in steganography. in: Proceedings of Elec-tronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, vol. 6505, 2007, pp. 02-03. doi:10.1117/12.697471.
- [8] Fridrich J, Golian M, Soukal D, (2006) Wet paper codes with improved embedding efficiency. IEEE T INF FOREN SEC, 1(1):102-110. doi: 10.1109/TIFS.2005.863487.
- [9] B. Li, S. Tan, M. Wang, and J. Huang, “Investigation on cost assignment in spatial image steganography,” IEEE TIFS, vol. 9, pp. 1264–1277, August 2014.
- [10] Pevný T, Filler T, Bas P, (2010) Using high-dimensional image models to perform highly undetectable steganography. in: Proceedings of the 12th International Workshop on Information Hiding(IH'2010), Springer Verlag, Calgary, Vol. 6387, 2010, pp. 161-177. doi: 10.1007/978-3-642-16435-4\_13.
- [11] T. Pevný, P. Bas, and J. Fridrich. Steganalysis by subtractive pixel adjacency matrix. IEEE Transactions on Information Forensics and Security, 5(2):215–224, June 2010.
- [12] T. Filler, T. Pevný, and P. Bas, “BOSS (Break Our Steganography System),” <http://www.agents.cz/boss>, July 2010.

- [13] Holub V, Fridrich J, (2012) Designing steganographic distortion using directional filters. in: Proceedings of IEEE Workshop on Information Forensic and Security(WIFS'2012), 2012, pp.234-239. doi: 10.1109/WIFS.2012.6412655.
- [14] Fridrich J, Kodovsky J, (2012) Rich Models for Steganalysis of Digital Images. IEEE T INF FOREN SEC, 7(3):868-882. doi: 10.1109/TIFS.2012.2190402.
- [15] J. Fridrich, J. Kodovský, M. Goljan, and V. Holub. Steganalysis of content-adaptive steganography in spatial domain. In Information Hiding, 13th Int. Conf., volume 6958 of Springer LNCS, pages 102–117, 2011
- [16] G. Gül and F. Kurugollu. A new methodology in steganalysis : Breaking highly undetectable steganography (HUGO). In Information Hiding, 13th Int. Conf., volume 6958 of Springer LNCS, pages 71–84, 2011.
- [17] Holub V, Fridrich J, Denmark T, (2014) Universal distortion function for steganography in an arbitrary domain. Eurasip Journal on Information Security, 1:1-13. doi: 10.1186/1687-417X-2014-1.
- [18] Li B, Wang M, Huang J, Li X, (2014) A new cost function for spatial image steganography. in: Proceedings of IEEE Inter-national Conference on Image Process(ICIP'2014), pp.4206-4210. doi: 10.1109/ICIP.2014.7025854.
- [19] J. Fridrich and J. Kodovský. Multivariate Gaussian model for designing additive distortion for steganography. In Proc. IEEE ICASSP, Vancouver, BC, May 26–31, 2013.
- [20] Sedighi V, Fridrich J, Cogranne R, (2015) Content-adaptive pentary steganography using the multivariate generalized Gauss-ian cover model. in: Proceedings of SPIE Media Watermarking, Security and Forensics 2015, SPIE, Vol. 9409, pp. 01-13. doi: 10.1117/12.2080272.
- [21] Vahid Sedighi, Remi Cogranne, Jessica Fridrich. Content-Adaptive Steganography by Minimizing Statistical Detectability.IEEE TIFS, vol 11, no.2, pp.221-234,2016.