Defining Cost Functions for Adaptive Steganography at the Microscale

Kejiang Chen, Weiming Zhang, Hang Zhou, Nenghai Yu
CAS Key Laboratory of Electromagnetic Space Information
University of Science and Technology of China, Hefei, China
Email: chenkj@mail.ustc.edu.cn, zhangwm@ustc.edu.cn
zh2991@mail.ustc.edu.cn, ynh@ustc.edu.cn

Guorui Feng
School of Communication and Information Engineering
Shanghai University, Shanghai, China
Email: grfeng@shu.edu.cn

Abstract—In the framework of minimizing embedding distortion steganography, the definition of cost function almost determines the security of the method. Generally speaking, texture areas would be assigned low cost, while smooth areas with high cost. However, the prior methods are still not precise enough to capture image details. In this paper, we present a novel framework of defining cost function for adaptive steganography at the microscale. The proposed framework is designed by using a “microscope” to highlight fine details in an image so that distortion definition can be more refined. Experiments show that by adopting our framework, the current steganographic methods (WOW, UNIWARD, HILL) will achieve better performances on resisting the state-of-the-art steganalysis.

I. INTRODUCTION

Steganography is the art of hiding messages in objects without drawing suspicion from steganalysis [1], [2]. Currently, the vast majority of work on steganography has focused on digital images. With the purpose of minimizing statistical detectability, modern steganography can be formulated as a source coding problem that minimizes embedding distortion [3]. Syndrome-trellis codes (STCs) provide a general methodology for embedding while minimizing an arbitrary additive distortion function with a performance near the theoretical bound [4].

As for content-adaptive steganography, how to define the cost function becomes one of the most important research issues. Taking into account of Adversary’s attack method, the cost function of HUGO [5] is defined as the weighted sum of difference between feature vectors extracted from a cover image and its stego version in SPAM [6] feature space. In this way, pixels after modification which make the feature vectors vary widely will be assigned a higher cost. The embedding changes of HUGO will be made within texture regions and along edges. WOW [7] assigns low costs to pixels in regions that are easily modelable, while pixels in textural regions that are difficult to predict by directional filters have smaller costs. S-UNIWARD [8] has a slightly modified filter bank from WOW. S-UNIWARD and WOW have similar performance and they are more secure than HUGO. HILL [9] is realized by using a high-pass filter and two low-pass filters, making more embedding changes concentrated in textural areas. It outperforms S-UNIWARD under the detection by the powerful steganalysis which employs [10]. MiPOD [11] under a model-driven framework also has an empirical security. The generalized Gaussian function is utilized to model noise residuals of pixels. All these adaptive algorithms consider pixels independently and cluster the modifications in the texture areas.

The state-of-the-art methods have exploited pixels in texture areas for hiding information. By comparing the cover image and the corresponding distortion, we are able to find some pixels with high cost values inside texture areas. However, these areas are probably suitable for concealing data, and should be assigned with low costs. To some extent, the methods mentioned above are still not precise enough to capture image details. In this case, the cost function can be further developed.

We proposed a novel framework of steganography which aims to create a fine distortion. With the help of a “Microscope”, the texture regions can be highlighted so that we can capture fine details of images precisely. The processed image is called auxiliary image. Then we utilize existing steganography methods to define distortion on the auxiliary image. The defined distortion will be smoothed by a low-pass filter and assigned to the cover image. Finally, the information hiding would be well implemented by STCs. The algorithm, which is based on the above framework, is called MS (Microscope) algorithm. We find that techniques in image enhancement such as unsharp masking (UM) [12] can act as the microscope. Experimental results show that the steganographic methods using the proposed framework perform better than not using, in resisting steganalysis with both SRM and selection-channel-aware maxSRMd2 [13].

The rest of this paper is organized as follows. After introducing notations, we review the preliminaries and the framework of minimizing additive distortion. In Section III we present a new steganographic framework which defines the distortion with the assistance of a microscope. Results of comparative experiments are elaborated in Section IV to demonstrate the effectiveness of the proposed framework. Conclusion and future work are given in Section V.
II. PRELIMINARIES AND PRIOR WORK

A. Notations

Throughout the paper, matrices, vectors and sets are written in bold face. The cover image (of size $n_1 \times n_2$) is denoted by $X = (x_{i,j})_{n_1\times n_2}$, where the signal $x_{i,j}$ is an integer, such as the gray value of a pixel. $Y = (y_{i,j})_{n_1\times n_2}$ denotes the stego image. The embedding operation on $x_{i,j}$ is formulated by the range $I_i$. An embedding operation is called binary if $|I_i| = 2$ and ternary if $|I_i| = 3$ for all $i,j$. For example, the ±1 embedding operation is ternary embedding with $I_i = \{\min(x_{i,j} - 1, 0), x_{i,j}, \max(x_{i,j} + 1, 255)\}$, where “0” denotes no modification.

B. Minimal Distortion Steganography

In the model established in [4], the distortion of modifying a pixel $x_{i,j}$ to $y_{i,j}$ can be simply denoted by $d_{i,j}(X, y_{i,j})$. It’s supposed that $d_{i,j}(X, x_{i,j}) = 0$ and $d_{i,j}(X, x_{i,j} - 1) = d_{i,j}(X, x_{i,j} + 1) = d_{i,j} \in \{0, \infty\}$. The overall distortion of the image can be calculated as follows:

$$D(X, Y) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} d_{i,j}|x_{i,j} - y_{i,j}|.$$  \hspace{1cm} (1)

Denote $\pi(y_{i,j})$ as the probability of changing $x_{i,j}$ to $y_{i,j}$. For a given message length $m$, the sender wants to minimize the average distortion (1), one can simulate optimal embedding by assigning

$$\pi(y_{i,j}) = \frac{\exp(-\lambda d_{i,j}(X, y_{i,j}))}{\sum_{y_{i,j}} \exp(-\lambda d_{i,j}(X, y_{i,j}))},$$  \hspace{1cm} (2)

where the scalar parameter $\lambda > 0$ determined by the payload constraint

$$m = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \pi(y_{i,j}) \log \frac{1}{\pi(y_{i,j})}.$$  \hspace{1cm} (3)

For additive distortion, there exist practical coding methods to embed messages, such as STCs [4], which can approach the performance of optimal embedding.

III. PROPOSED METHOD

A. Motivation

Generally speaking, content-adaptive steganography assigns low costs in texture regions while high costs in smooth areas. From this point of view, grasping the distribution of the texture areas in an image counts for a lot. By comparing the cover image and the corresponding distortion, we are able to find some pixels with high cost values inside texture areas. However, these areas are probably suitable for concealing data, and should be assigned with low costs. Furthermore, Fig. 2 (a) is the cover image and Fig. 2 (b) is the corresponding embedding changes. It is easy to find that there are some pixels modified in smooth areas, while the texture areas should have carried these message bits. These all indicate that the current steganographic methods do not capture fine detail of image well.

Fortunately, many methods in image enhancement do favor to search texture areas. On the basis of the techniques, we proposed a novel framework of steganography, and it will be studied in detail in next section.

B. A novel framework for steganography

Since the current steganography cannot seize texture areas exactly, there is still a long way for us to improve the security of steganography. Just as mentioned in Section II, the additive distortion is the foundation of content-adaptive steganography. However, former distortion is not proper enough, which becomes severe in highly texture regions. So we attempt to design a new framework to help steganographic algorithms locate secure area more precisely. The proposed framework can be implemented in five steps as shown in Fig. 3.

1) Magnify the cover image $X$ so as to highlight fine details or to enhance details. The operation of enlarging images is not resizing but filtering. Many techniques of image process such as image sharpening can act as a “microscope”. The processed image is called “auxiliary image”, denoted by $X'$.

2) Utilize distortion definition methods in existing steganography algorithm (WOW, UNIWARD, HILL, and etc.) to calculate the distortion $D'$ on the auxiliary image $X'$.

3) In order to spread the low costs of textual pixels to their neighbourhood, we employ a low-pass filter $L$ to smooth the distortion $D'$. For easy implementation, average filter
We denote the smoothed distortion by $D'_s$.  

4) Assign the smooth distortion $D'_s$ to the cover image $X$. When it comes to the saturated pixels in the cover image $X$, the distortion should be adjusted. There are essentially three options for pixel values at the boundary of the dynamic range as mentioned in [15]. Restricting the polarity of changes is adopted. If $x_{i,j}=0$, then $d_{i,j}(X, x_{i,j} - 1) = \infty$; if $x_{i,j}=255$, then $d_{i,j}(X, x_{i,j} + 1) = \infty$. The final distortion is denoted by $D$.

5) With the help of STCs and distortion $D$, steganography can be well implemented on the cover image $X$.

The former embedding method adopting proposed framework would be prefixed with “MS”, such as MS-WOW, MS-UNIWARD and MS-HILL. In the framework, performances of different microscopes vary greatly, so the selection of microscope is of great importance. We will discuss it in the next subsection.

C. The selection of “Microscope”

The principal objective of “Microscope” is to highlight fine details in an image or to enhance details that has been blurred, either in error or as a natural effect of a particular method of image acquisition. Image enhancement is to the choice, such as edge enhancement, histogram equalization, unsharp masking and etc. Edge enhancement is an extremely common technique used to make images appear sharper. Histogram equalization performs its operation by remapping the gray levels of the image based on the probability distribution of the input gray levels. Unsharp masking (UM) is a widely used technique for improving the perceptual quality of an image by emphasizing its high-frequency components [16]. We conduct some contrast experiments, and find that unsharp masking is most suitable. To some degree, unsharp masking highlights fine details as well as maintains the original characteristics of the image.

In the linear UM algorithm [12], the enhanced image $y(n, m)$ is obtained from the input image $x(n, m)$ as

$$y(n, m) = x(n, m) + \alpha \ast z(n, m),$$

where $z(n, m)$ is the correction signal as the output of a high-pass filter and $\alpha$ is the positive scaling factor that controls the level of contrast enhancement achieved at the output. Base on the high-pass filter as shown in Fig. 5, $z(n, m)$ can be obtained by

$$z(n, m) = 8x(n, m) - x(n - 1, m - 1) - x(n - 1, m) - x(n - 1, m + 1) - x(n + 1, m) - x(n + 1, m + 1) - x(n, m - 1) - x(n, m + 1).$$

The enhanced image are shown in Fig. 6. The enhanced image (a) is sharpened by UM algorithms. It’s apparent that the enhanced image owns more details than the original image as shown in Fig. 2(a). (b) is the corresponding embedding changes of the original image with a fixed payload 0.5 using MS-UNIWARD. Note that embedding changes: +1 = white, -1 = black. We can find that there are very few pixels changed in the smooth area.
D. Distortion Smoothness

With a microscope, we are able to seize the texture areas more accurately. We make a further enhancement of the algorithm by incorporating the Spreading Rule [17], so that the modifications will be clustered more accurately in the texture areas. Spreading rule indicates that the costs of modifying neighbouring elements should be similar, which has been successfully utilized in HILL by smoothing the distortion functions. Moderating costs (reducing high costs and increasing low costs) also helps defend against knowing attackers such as content-aware steganalysis [18]. In our framework, we smooth the distortion \( D' \) of the auxiliary image by adopting a low-pass filter \( L \).

E. Pseudo-code Procedure

To further clarify the framework of steganography at the microscale, in Algorithm 1 we provide a pseudo-code that describes the implementation of information hiding and extraction.

Algorithm 1 Microscope steganography

Input: A cover image \( X \) with \( N \) pixels; \( L \) bits of message \( m \) which determines the relative payload of target \( \gamma = L/N \).

Output: The stego image \( Y \).

1: Sharpen the cover image \( X \) into auxiliary image \( X' \) using linear unsharp masking with scaling factor \( \alpha \).
2: Utilize the distortion definition in existing steganography methods (WOW, UNIWARD, HILL, etc.) to calculate the distortion \( D' \) on the auxiliary image \( X' \).
3: Acquire the smoothed distortion \( D'_s \) by smoothing the distortion \( D' \) with average filter.
4: Assign the distortion \( D'_s \) to the cover image \( X \). The distortion should be adjusted when it comes to saturated pixels in the cover image \( X \). If \( x_{i,j} = 0 \), then \( d_{i,j}(X, x_{i,j} - 1) = \infty \); if \( x_{i,j} = 255 \), then \( d_{i,j}(X, x_{i,j} + 1) = \infty \). The adjusted distortion will be the final distortion \( D \).
5: Embed \( L \) bits of message \( m \) into cover image \( X \) with STCs according to the final distortion \( D \), and finally output the stego image \( Y \).

IV. Experiment

A. Setups

All experiments in this paper are carried out on BOSSbase 1.01 [14] containing 10,000 grayscale \( 512 \times 512 \) images. The detectors are trained as binary classifiers implemented using the FLD ensemble with default settings. A separate classifier is trained for each embedding algorithm and payloads. The ensemble by default minimizes the total classification error probability under equal priors \( P_{E} = \min_{P_{FA}, P_{MD}} \frac{1}{2} (P_{FA} + P_{MD}) \), where \( P_{FA} \) and \( P_{MD} \) are the false-alarm probability and the missed-detection probability respectively. The ultimate security is qualified by average error rate \( P_{E} \) averaged over ten 5000/5000 database splits, and larger \( P_{E} \) means stronger security. The two feature sets used are SRM [10] and its selection-channel-aware version maxSRMd2 [13]. As for maxSRMd2, the method we estimate embedding change probability is exactly the same as the proposed steganography method. All tested embedding algorithms are simulated at their corresponding payload-distortion bound for payloads \( R \in \{0.05, 0.1, 0.2, 0.3, 0.4, 0.5\} \) bpp (bits per pixel).

B. Determining the parameters of MS steganography

In the experiment, linear unsharp masking is adopted as the microscope, and average filter is implemented for distortion smoothness. As for MS-UNIWARD, first we explore the
Fig. 10: Detection error for WOW and MS-WOW when steganalyzing with SRM (a) and maxSRMd2 (b).

Fig. 11: Detection error for UNIWARD and MS-UNIWARD when steganalyzing with SRM (a) and maxSRMd2 (b).

Fig. 12: Detection error for HILL and MS-HILL when steganalyzing with SRM (a) and maxSRMd2 (b).

TABLE I
Detectability in terms of $P_E$ versus embedded payload size in bits per pixel (bpp) for Prior art and applied to our framework on BOSSbase 1.01 using the FLD ensemble classifier with two feature sets.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Embedding Method</th>
<th>0.05</th>
<th>0.1</th>
<th>0.2</th>
<th>0.3</th>
<th>0.4</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRM</td>
<td>WOW</td>
<td>.4569 ± .0024</td>
<td>.4035 ± .0021</td>
<td>.3203 ± .0025</td>
<td>.2556 ± .0030</td>
<td>.2061 ± .0026</td>
<td>.1680 ± .0027</td>
</tr>
<tr>
<td></td>
<td>MS-WOW</td>
<td>.4650 ± .0022</td>
<td>.4186 ± .0032</td>
<td>.3418 ± .0023</td>
<td>.2797 ± .0024</td>
<td>.2312 ± .0017</td>
<td>.1854 ± .0018</td>
</tr>
<tr>
<td></td>
<td>UNIWARD</td>
<td>.4542 ± .0024</td>
<td>.4021 ± .0024</td>
<td>.3199 ± .0026</td>
<td>.2574 ± .0017</td>
<td>.2031 ± .0026</td>
<td>.1640 ± .0025</td>
</tr>
<tr>
<td></td>
<td>MS-UNIWARD</td>
<td>.4692 ± .0015</td>
<td>.4274 ± .0035</td>
<td>.3487 ± .0031</td>
<td>.2869 ± .0018</td>
<td>.2377 ± .0031</td>
<td>.1927 ± .0019</td>
</tr>
<tr>
<td></td>
<td>HILL</td>
<td>.4688 ± .0023</td>
<td>.4340 ± .0034</td>
<td>.3632 ± .0025</td>
<td>.2996 ± .0023</td>
<td>.2482 ± .0022</td>
<td>.2038 ± .0017</td>
</tr>
<tr>
<td></td>
<td>MS-HILL</td>
<td>.4702 ± .0018</td>
<td>.4347 ± .0032</td>
<td>.3618 ± .0027</td>
<td>.3009 ± .0029</td>
<td>.2478 ± .0031</td>
<td>.2023 ± .0021</td>
</tr>
<tr>
<td>maxSRMd2</td>
<td>WOW</td>
<td>.3546 ± .0022</td>
<td>.3010 ± .0031</td>
<td>.2341 ± .0025</td>
<td>.1896 ± .0027</td>
<td>.1553 ± .0031</td>
<td>.1306 ± .0025</td>
</tr>
<tr>
<td></td>
<td>MS-WOW</td>
<td>.4336 ± .0013</td>
<td>.3851 ± .0018</td>
<td>.3187 ± .0021</td>
<td>.2625 ± .0019</td>
<td>.2211 ± .0017</td>
<td>.1826 ± .0023</td>
</tr>
<tr>
<td></td>
<td>UNIWARD</td>
<td>.4189 ± .0024</td>
<td>.3651 ± .0038</td>
<td>.2896 ± .0028</td>
<td>.2350 ± .0021</td>
<td>.1913 ± .0026</td>
<td>.1556 ± .0021</td>
</tr>
<tr>
<td></td>
<td>MS-UNIWARD</td>
<td>.4516 ± .0037</td>
<td>.4033 ± .0024</td>
<td>.3376 ± .0031</td>
<td>.2766 ± .0024</td>
<td>.2340 ± .0023</td>
<td>.1924 ± .0037</td>
</tr>
<tr>
<td></td>
<td>HILL</td>
<td>.4244 ± .0023</td>
<td>.3765 ± .0031</td>
<td>.3120 ± .0029</td>
<td>.2628 ± .0022</td>
<td>.2180 ± .0025</td>
<td>.1853 ± .0024</td>
</tr>
<tr>
<td></td>
<td>MS-HILL</td>
<td>.4504 ± .0025</td>
<td>.4068 ± .0023</td>
<td>.3393 ± .0032</td>
<td>.2840 ± .0018</td>
<td>.2388 ± .0028</td>
<td>.2030 ± .0023</td>
</tr>
</tbody>
</table>
optimal value of scaling factor $\alpha$ of unsharped masking. We set the size of average filter $L = 3 \times 3$ and keep it invariant. Fig. 7 shows the effect of different scaling factor $\alpha$ with a fixed payload of 0.2 bpp on empirical security. The result indicates that $\alpha = 2$ performs best. Since the scaling factor $\alpha$ has been set, an experiment is carried out to obtain the best filter size for $L$. The results are shown in Fig. 8. It is suggested that $L = 7 \times 7$ outperforms other sizes. Under the same experiment condition conducted in MS-WOW and MS-HILL, the optimal parameters are the same. In other words, the parameters ($\alpha = 2, L = 7 \times 7$) can serve as empirical parameters.

C. The effectiveness of the microscope

We conduct some comparative experiments to investigate the effectiveness of microscope in the MS algorithm. “AVG” represents the steganographic method with a single operation of average filter. Three steganography experiments (UNIWARD, AVG-UNIWARD, MS-UNIWARD) are carried out under the steganalytic feature set SRM. According to the results shown in Fig. 9, it is obvious that MS-UNIWARD performs far better than AVG-UNIWARD. It demonstrates that microscope in the algorithm contributes largely to the promotion of the performance.

D. Application to prior methods

The proposed framework is applied to the prior steganography methods. In this paper, WOW, UNIWARD and HILL are chosen as the steganography methods, for they are representative. The parameters of them have been discussed in the previous subsection. As shown in Fig. 10 and Fig. 11, MS-WOW performs better than WOW by about 1.0-2.5% steganalyzing with SRM and 5.0-8.5% steganalyzing with maxSRMd2. MS-UNIWARD performs better than UNIWARD by about 1.5-3.5% against SRM and 3.5-5.0% against maxSRMd2. It can be observed from Fig. 12 that MS-HILL and HILL have similar performances against SRM, while MS-HILL has an apparent improvement over HILL against maxSRMd2.

Table I shows the average total probability of error $P_E$ and its standard deviation for a range of payloads. Note that, the one using MS algorithm offers better security than not using. The experiment results mentioned above support the effectiveness of the proposed framework.

V. Conclusions

In this paper, we propose a new framework which defines cost functions for adaptive steganography at the microscale. Before distortion definition, the cover image would be preprocessed with a microscope. Therefore, prior steganographic methods can seize the texture areas more precisely, so that the distortion can be defined more accurately. In our study, unsharp masking plays the role of the microscope. We also make a further enhancement of the MS algorithm incorporating the spreading rule by smoothing the embedding distortion on the auxiliary image. The experiment results verified that the proposed framework does work.

Since image enhancement techniques vary greatly, we will try to explore the effectiveness of other methods. In addition, our work is discussed in additive distortion steganography, so we intend to generalize an extension of this work in non-additive distortion steganography in our future study.

ACKNOWLEDGMENT

This work was supported in part by the Natural Science Foundation of China under Grant 61572452, Grant 61502007 and Grant 61373151, in part by the China Postdoctoral Science Foundation under Grant 2015M582015, and in part by the Strategic Priority Research Program of the Chinese Academy of Sciences under Grant XDA06030601.

REFERENCES