

基于无方向性滤波器的空域自适应隐写算法

王龙飞, 郭继昌, 田煜衡

天津大学电子信息工程学院, 天津 300072

摘要 为了确定自适应隐写中载体图像复杂区域, 提高像素间相关性, 提出了一种利用无方向性滤波器设计的空域自适应隐写算法。利用无方向性滤波器对载体图像进行滤波计算, 获得图像中难以被建模检测分析的复杂区域, 再利用高斯低通滤波器对复杂区域进行平滑处理, 增加嵌入信息区域相邻像素间的相关性, 进而得到损失函数, 最后按照损失函数通过校验格编码完成信息嵌入。实验结果表明, 当信息嵌入率较小时, 该算法的抗检测性能与 S-UNIWARD 算法相近; 当信息嵌入率较大时, 该算法优于 S-UNIWARD 算法。

关键词 图像处理; 自适应隐写; 无方向性滤波器; 损失函数; 空域

中图分类号 TP391 文献标识码 A

doi: 10.3788/LOP54.021003

Spatial Adaptive Steganography Based on Non-directional Filter

Wang Longfei, Guo Jichang, Tian Yuheng

School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China

Abstract In order to obtain the complex regions of cover image in adaptive steganography and improve the coherence between pixels, a novel adaptive steganography algorithm based on non-directional filter is proposed. The complex regions, which are difficult to be modeled and detected, are obtained by filter calculation on images through using a non-directional filter. The Gaussian low-pass filter is used for smoothing the complex regions, and enhance the coherence of different pixels in those regions and the cost function is established. Finally, according to the cost function, the messages are embedded by syndrome trellis codes. Experimental results demonstrate that the performance of anti-detection by the proposed algorithm is similar as that of the S-UNIWARD algorithm when the information embedding rate is small, and the proposed algorithm is better than the S-UNIWARD algorithm when the information embedding rate is large.

Key words image processing; adaptive steganography; non-directional filter; cost function; spatial domain

OCIS codes 100.3008; 110.1085; 120.2440

1 引言

隐写技术通过修改一部分载体元素嵌入想要传达的信息, 从而实现消息发送者和接收者之间的安全通信。隐写算法可以分为自适应隐写算法和非自适应隐写算法。非自适应隐写算法不需要考虑载体图像特性, 具有操作简单, 易于实现的特点, 如经典算法最低有效位 (LSB) 替换隐写^[1]及其改进算法 LSB 匹配隐写^[2], 但算法缺点同样明显, 隐藏后的秘密信息容易被隐写检测算法^[3-5]捕捉。统计检测的隐写分析方法的不断发展促使学者们提出了大量的基于载体图像内容特征的自适应隐写算法^[6-13]。与非自适应隐写算法随机地在载体图像中嵌入信息的方式相比, 自适应隐写算法可以根据载体图像自身特性将信息嵌入到载体图像中的纹理复杂、不易被隐写分析算法检测到的区域, 安全性更高。自适应隐写算法有多种设计方式, 其中应用较广且比较成功的为基于损失函数的隐写算法, 如文献[6]和文献[9-11], 其设计过程可分为两步: 1)

收稿日期: 2016-08-30; 收到修改稿日期: 2016-10-18

基金项目: 国家 973 计划(2014CB340400)、天津市自然科学基金(15JCYBJC15500)

作者简介: 王龙飞(1989—), 男, 硕士研究生, 主要从事信息隐藏方面的研究。E-mail: 15202218085@163.com

导师简介: 郭继昌(1966—), 男, 博士, 教授, 主要从事智能视频图像分析、识别及处理, 滤波器理论及设计等方面的研究。

E-mail: jcguo@tju.edu.cn(通信联系人)

设计损失函数;2) 根据损失函数在载体图像中通过隐写编码嵌入信息。对于隐写编码,Filler 等^[14]提出了一种校验格编码(STC)算法,既可以应用在空域又能应用在变换域,其编码性能接近理论上的最优,这使得自适应隐写算法设计问题简化成损失函数的设计问题。

当今一些自适应隐写算法的损失函数大多是启发式的。HUGO 算法^[6]利用载体图像和载密图像在减法像素邻接矩阵 SPAM^[15]特征空间中的差异指导损失函数设计,其缺点是算法过于复杂,且容易在单一方向的边缘嵌入秘密信息。WOW 算法^[9]利用方向性滤波器设计损失函数,从水平、垂直和对角线三个方向计算载体图像的残差,通过残差大小分配损失值。Holub 等^[10]将 WOW 算法从空间域推广到任意域,提出了 UNIWARD 算法,空间域的 UNIWARD 算法(S-UNIWARD)与 WOW 算法相比,两者在损失函数上的变化非常小,两种算法的抗隐写分析性能也非常接近,WOW 算法和 S-UNIWARD 算法改善了单一方向边缘嵌入信息的情况,但其未考虑像素间相关性,且容易将秘密信息嵌入到载体图像中的平滑区域。尽管以往大部分的自适应隐写算法能够利用到载体图像中的复杂区域用于隐藏信息。然而,有些可以被用于嵌入信息的载体区域并不能被充分利用,另外,为了嵌入足够量的信息,一些不适合作为嵌入信息的平滑区域不得不被嵌入信息,降低了算法的安全性。

在 WOW 和 S-UNIWARD 基础上,本文提出了一种利用无方向性滤波器设计的空域自适应隐写算法。首先利用无方向性滤波器^[16-19]确定载体图像中比较适于嵌入信息的纹理丰富区域,然后用高斯低通滤波器对此区域进行滤波平滑,增强像素间的相关性,进而得到损失分布函数,最后根据损失分布在载体图像中用 STC 嵌入秘密信息。安全性能分析实验结果表明,信息嵌入率较小时,该算法抗检测性能接近 S-UNIWARD 算法;信息嵌入率较大时,该算法优于 S-UNIWARD 算法。

2 相关基础

2.1 最小加性失真

在设计自适应隐写算法时,通常利用不同像素修改后造成的损失和来定义损失函数。本文通过最小化加性失真^[20]设计损失函数。记载体图像为 \mathbf{X} , 大小为 $m \times n$, 载体图像中的像素点为 $x_{i,j} \in \{0, \dots, 255\}^{m \times n}$ 。记载密图像为 \mathbf{Y} , 大小为 $m \times n$, 载密图像中的像素点为 $y_{i,j} \in \{0, \dots, 255\}^{m \times n}$, $\rho_{i,j}(\mathbf{X}, y_{i,j})$ 表示将载体像素 $x_{i,j}$ 改变为在载密像素 $y_{i,j}$ 造成的失真, 为了简化设计, 认为 $\rho_{i,j}(\mathbf{X}, x_{i,j} - 1) = \rho_{i,j}(\mathbf{X}, x_{i,j} + 1) = \rho_{i,j} \in [0, +\infty)$, $\rho_{i,j}(\mathbf{X}, x_{i,j}) = 0$ 。

则在载体图像中嵌入信息造成的失真 $D(\mathbf{X}, \mathbf{Y})$ 可定义为

$$D(\mathbf{X}, \mathbf{Y}) = \sum_{i=1}^m \sum_{j=1}^n \rho_{i,j} |x_{i,j} - y_{i,j}|. \quad (1)$$

设 k 为秘密信息的嵌入长度, $\pi(y_{i,j})$ 为载体像素 $x_{i,j}$ 被修改成载密像素 $y_{i,j}$ 的概率。为了最小化(1)式中的嵌入失真, 再进行优化

$$\pi(y_{i,j}) = \frac{\exp[-\lambda \rho_{i,j}(\mathbf{X}, y_{i,j})]}{\sum_{y \in \tau_{i,j}} \exp[-\lambda \rho_{i,j}(\mathbf{X}, y_{i,j})]}, \quad (2)$$

参数 λ 满足

$$k = \sum_{i=1}^m \sum_{j=1}^n \sum_{y \in \tau_{i,j}} \pi(y_{i,j}) \log \frac{1}{\pi(y_{i,j})}. \quad (3)$$

2.2 校验格编码

STC^[14]是一种特殊的矩阵编码,也是目前性能最好的一种隐写编码方法。

假设在载体 \mathbf{x} 中嵌入秘密信息 \mathbf{m} 可得到载密图像 \mathbf{y} , 则使用 STC 编码嵌入信息时应满足 $\mathbf{H}\mathbf{y}^T = \mathbf{m}$, \mathbf{H} 为由若干个大小为 $h \times w$ 的 $\hat{\mathbf{H}}$ 以行为单位从上到下依次向下平移构成的稀疏奇偶校验矩阵。 $\hat{\mathbf{H}}$ 根据共享密钥随机生成, 其参数 h 将会影响到编码的速度和效率, 计算复杂度随 h 增加指数增长; w 影响编码的嵌入效率, 它与信息嵌入率 α 有如下关系: $w = 1/\alpha$ 。为了直观表现出 \mathbf{H} 的结构, 下面举例说明。

若 $\hat{\mathbf{H}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, 则

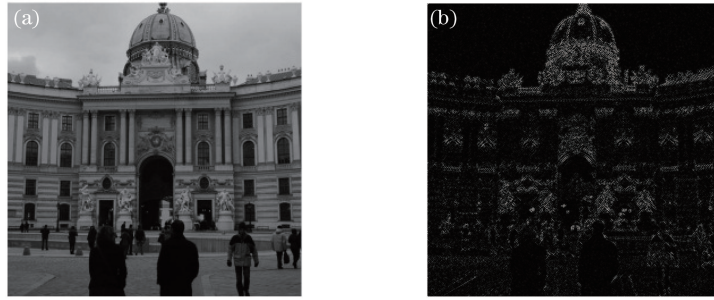


图1 载体图像和利用 F 确定的复杂区域。(a)载体图像;(b)残差图像

Fig. 1 Cover image and complex zones determined by F . (a) Cover image; (b) residual image

在设计损失函数时,通过引入像素间相关性来解决这一问题,利用高斯低通滤波器对载体图像的残差进行平滑处理,可以将像素点的损失权重扩散到相邻像素,增加相邻像素之间相关性,在一定程度上减小加性损失函数的近似失真,提升算法抗隐写检测性能。

为验证方案可行性,进行了一组实验。首先从 BOSSBase 1.01^[21] 图像库中随机选出 5000 张图片用于隐写加密,信息嵌入率为 0.4 bit/pixel,然后提取载体和载密图像的 SRMQ1 特征^[16],最后利用集成分类器^[22]进行测试。算法安全性能用检测错误率 P_E 衡量, $P_E = \min\{[P_{FA} + P_{MD}(P_{FA})]/2\}$, 其中, P_{FA} 表示虚警率, P_{MD} 表示漏检率。检测错误率越大,算法安全性能越高。实验中高斯低通滤波器阶数分别为 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 滤波器参数 σ 与其阶数一致。滤波器阶数为 1 时,表示不考虑像素间的相关性。实验结果如图 2 所示。

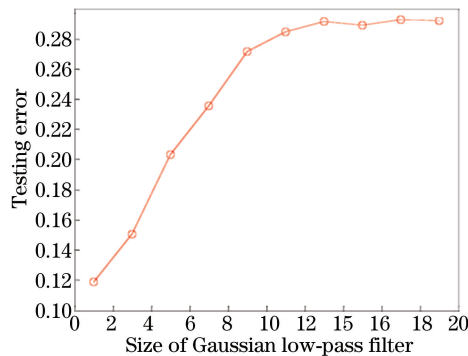


图2 高斯低通滤波器对算法安全性能的影响

Fig. 2 Effect of Gaussian low-pass filter to proposed algorithm security

实验结果表明,使用高斯低通滤波器可以有效提高算法的抗检测能力,且随着滤波器阶数增大,被平滑处理的像素区域增大,像素间的相关性随之增强,算法安全性能也会相应提升,当滤波器阶数达到 13 阶时,检测错误率达到最高,并且错误率不会随着滤波器阶数的继续增大而发生明显变化,这说明此时像素间的相关性已经趋于饱和。实验证明此方案可行,本文算法中 L 选用 13 阶的高斯低通滤波器,滤波器参数 σ 与其阶数一致。

4 实验分析

实验分为三个部分:信息嵌入位置对比示例、时间复杂度对比分析以及抗隐写检测性能对比分析。实验计算机及软件配置如下: Intel Core i3 2.40 GHz CPU, 4.00 GB RAM, 软件平台为 MATLAB R2014a。

选用 BOSSBase 1.01 作为实验图像库,分别利用 12753 维 SRMQ1 和 34671 维 SRM 特征^[16]进行检测分析算法安全性能。实验采用集成分类器作为分类器,对于载体图像和对应生成的载密图像,随机选取 50% 图像用于训练, 50% 图像用于测试。安全性能用最小平均分类错误率表示,错误率越大,则算法抗隐写分析能力越强,安全性越高,反之亦然。当嵌入信息足够少时,分类器会随机判定所检测的图像是否存在隐藏信息,即错误率接近 0.5。

用于对比实验的自适应隐写算法有 4 种:本文算法、HUGO 的改进算法 HUGO-BD 算法^[7]、S-UNIWARD 算法^[10]和 HILL 算法^[11],对比算法的参数设置均为原文献中的默认值。所有实验共生成加密图片 49 万张,提取到用于实验分析的 SRMQ1 特征 17.3 GB,SRM 特征 38.8 GB。

4.1 信息嵌入位置对比示例

从图像库中选择一张图片作为载体图片,此图片含有丰富的纹理信息以及边缘信息,实验嵌入信息量为 0.4 bit/pixel,图 3 给出不同算法对载体图像隐写后秘密信息的分布,图 3(a)为载体图像,从图 3(b)~(e)依次为 4 种不同算法在载体图像中嵌入信息的情况,图像越暗,表示嵌入信息的可能性越小,反之则表示嵌入信息的可能性越大。由图 3 定性分析可知,4 种隐写算法都较好地利用了载体图像中纹理丰富的复杂区域,实现了秘密信息的自适应嵌入。不同的是,HUGO-BD 算法将部分信息嵌入到了图像中单一方向的边缘区域(如柱子边缘、柱子上方的水平线边缘和人影轮廓等),S-UNIWARD 算法与 HUGO-BD 算法相比,单一方向的边缘区域嵌入信息减少。本文算法与这两种算法相比,既减少了在易于建模进行隐写检测的单一方向边缘(柱子和水平线等)嵌入信息,又通过增强像素间相关性的方法集中了秘密信息分布。HILL 算法嵌入的秘密信息集中在了载体图像的纹理丰富的复杂区域,轮廓较为清晰。几种算法安全性能上的具体差异将通过抗检测性能实验给出。

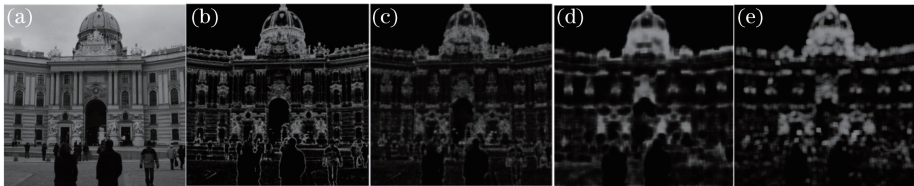


图 3 信息嵌入分布。(a)载体图像;(b) HUGO-BD;(c) S-UNIWARD;(d) HILL;(e)本文算法

Fig. 3 Distribution of information insertion. (a) Cover image; (b) HUGO-BD;

(c) S-UNIWARD; (d) HILL; (e) proposed method

4.2 时间复杂度对比分析

从 BOSSBase 1.01 图像库中任取 1000 张图片,在信息嵌入率为 0.4 bit/pixel 的条件下,使用 4 种自适应隐写算法对所选图片进行信息嵌入,记录下每种算法完成信息嵌入所需要的平均时间,如表 1 所示。

表 1 4 种隐写算法平均耗时

Table 1 Average time cost of four steganographic algorithms

Algorithm	HUGO-BD	S-UNIWARD	HILL	Proposed
Time /s	283.69	1.61	1.26	0.90

由表 2 中可知,使用吉布斯构造的 HUGO-BD 算法的时间复杂度最高,且远大于另外三种算法;所提算法与 S-UNIWARD 和 HILL 相比,均在时间复杂度上存在一定优势,这是因为 S-UNIWARD 算法需要分别从水平、垂直和对角线三个方向计算损失函数,而本文算法中 F 为无方向性滤波器,在计算损失函数时,不需要考虑方向性;HILL 算法在设计损失函数时,需要使用三个滤波器经三次卷积计算,而本文算法只需两个滤波器两次卷积即可。

4.3 算法的安全性对比分析

为全面检测衡量所提算法抵抗隐写分析的能力,做了两组实验,对比 4 种空域自适应隐写算法在 11 种不同信息嵌入率条件下抵抗低维富模型 SRMQ1 特征和高维富模型 SRM 特征隐写分析的能力。选用 BOSSBase 1.01 中 10000 张图片用于实验,在进行分类检测时,对于载体图像和对应生成的载密图像,一半用于训练,另一半用于测试。实验信息嵌入率分别为 0.05,0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9,1.0 bit/pixel 11 种情况。

实验选取了 11 种不同的信息嵌入率进行抗检测性能分析,图 4 和 5 分别为抗 SRMQ1 和 SRM 隐写检测结果,其中,图 4(a)、5(a)是嵌入率为 0~0.5 bit/pixel 时的抗检测性能分析图,图 4(b)、5(b)是嵌入率为 0.5~1.0 bit/pixel 时的抗检测性能分析图。

由图 4 和 5 可以发现,4 种算法在低维富模型特征 SRMQ1 和高维富模型特征 SRM 下的抗检测性能基本一致,且 SRM 特征的检测性能要好于 SRMQ1。4 种算法中,HUGO-BD 算法抗富模型隐写分析能力最

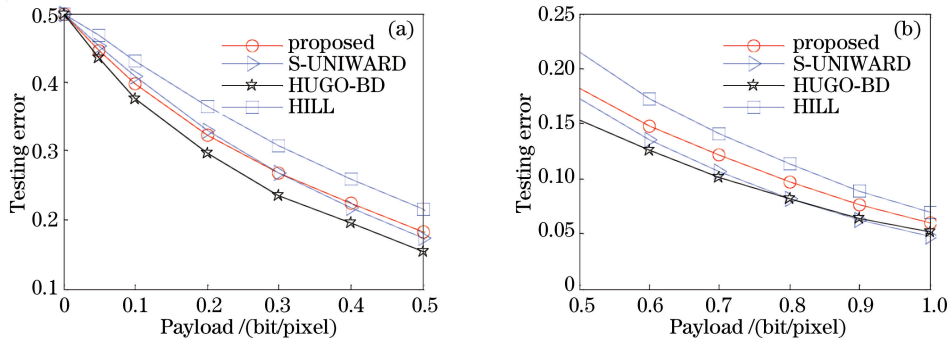


图 4 抗 SRMQ1 隐写检测结果

Fig. 4 Detection results of SRMQ1 steganalysis

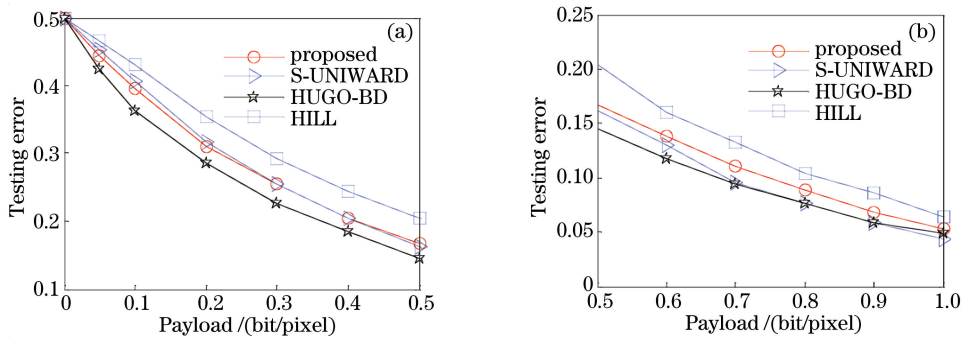


图 5 抗 SRM 隐写检测结果

Fig. 5 Detection results of SRM steganalysis

差, HILL 算法最强。当信息嵌入率较小时(有效载荷小于 0.3 bit/pixel), 所提算法抗富模型隐写分析能力与 S-UNIWARD 算法相近, 当信息嵌入率较大时(有效载荷大于 0.3 bit/pixel) 优于 S-UNIWARD 算法, 且随嵌入率增加, 算法的安全性能优势增大。这是因为在隐写分析中性能良好的无方向性滤波器 F 同样适合在隐写算法中被用于提取载体图像的复杂区域, 且随嵌入信息量增加, 相邻像素间的相关性越为重要, 所提算法充分考虑到此点。从隐写分析的角度看, 当今主流的隐写检测方法检测的是因嵌入秘密信息在图像中造成的“波动”, 本文算法通过高斯低通滤波器的“平滑”作用可以有效地将这种波动变得平缓, 从而有效降低了加密信息被检测到的可能性。

5 结 论

在 WOW 和 S-UNIWARD 算法基础上提出了一种利用无方向性滤波器设计的空域自适应隐写算法, 首先利用无方向性滤波器计算载体图像残差, 再通过高斯低通滤波器对残差进行平滑滤波得到损失函数, 最后利用 STC 按照损失分布完成信息嵌入。实验结果表明, 本文算法在较小的信息嵌入率条件下与 S-UNIWARD 算法相近, 在较大的信息嵌入率条件下优于 S-UNIWARD 算法, 从而验证了本文方案的可行性。同时, 与 HILL 算法相比, 本文算法在时间复杂度上存在一定优势, 但两者在抗检测性能上依然存在差距, 在未来的工作中考虑使用更精确的残差提取方法来改善算法安全性能。

参 考 文 献

- [1] Petitcolas F A P, Anderson R J, Kuhn M G. Information hiding-a survey[J]. Proceedings of the IEEE, 1999, 87(7): 1062-1078.
- [2] Ker A D. Steganalysis of LSB matching in grayscale images[J]. IEEE Signal Processing Letters, 2005, 12(6): 441-444.
- [3] Fridrich J, Goljan M, Du R. Detecting LSB steganography in color, and gray-scale images[J]. IEEE Multimedia, 2001, 8(4): 22-28.

- [4] Liu Xueqian, Ping Xijian, Zhang Tao, *et al.* Steganalysis of LSB matching based on wavelet feature of filtering restoration[J]. *Journal of Data Acquisition & Processing*, 2010, 25(4): 505-511.
刘学谦, 平西建, 张涛, 等. 基于滤波复原的小波特征 LSB 匹配隐写分析方法[J]. *数据采集与处理*, 2010, 25(4): 505-511.
- [5] Xu M, Li T, Ping X. Steganalysis of LSB matching based on wavelet denoising estimation in grayscale image[C]. *International Conference on Future Generation Communication and Networking*, 2008, 1: 106-109.
- [6] Pevný T, Filler T, Bas P. Using high-dimensional image models to perform highly undetectable steganography[C]. *International Conference on Information Hiding*, 2010: 161-177.
- [7] Filler T, Fridrich J. Gibbs construction in steganography [J]. *IEEE Transactions on Information Forensics and Security*, 2010, 5(4): 705-720.
- [8] Luo W, Huang F, Huang J. Edge adaptive image steganography based on LSB matching revisited [J]. *IEEE Transactions on Information Forensics and Security*, 2012, 5(2): 201-214.
- [9] Holub V, Fridrich J. Designing steganographic distortion using directional filters[C]. *IEEE International Workshop on Information Forensics and Security*, 2012: 234-239.
- [10] Holub V, Fridrich J, Denemark T. Universal distortion function for steganography in an arbitrary domain[J]. *Eurasip Journal on Information Security*, 2014, 2014(1): 1-13.
- [11] Li B, Wang M, Huang J, *et al.* A new cost function for spatial image steganography [C]. *IEEE International Conference on Image Processing (ICIP)*, 2015: 4206-4210.
- [12] Al-Shatanawi O M, Emam N N E. A new image steganography algorithm based on MLSB method with random pixels selection[J]. *International Journal of Network Security & Its Applications*, 2015, 7(2): 37-53.
- [13] Sedighi V, Cogramme R, Fridrich J. Content-adaptive steganography by minimizing statistical detectability[J]. *IEEE Transactions on Information Forensics and Security*, 2016, 11(2): 221-234.
- [14] Filler T, Judas J, Fridrich J. Minimizing additive distortion in steganography using syndrome-trellis codes[J]. *IEEE Transactions on Information Forensics and Security*, 2011, 6(3): 920-935.
- [15] Pevný T, Bas P, Fridrich J. Steganalysis by subtractive pixel adjacency matrix[J]. *IEEE Transactions on Information Forensics and Security*, 2010, 5(2): 215-224.
- [16] Fridrich J, Kodovsky J. Rich models for steganalysis of digital images[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(3): 868-882.
- [17] Qian Y, Dong J, Wang W, *et al.* Deep learning for steganalysis via convolutional neural networks[C]. *Proceedings of SPIE*, 2015, 9409: 94090J.
- [18] Xu G, Wu H Z, Shi Y Q. Ensemble of CNNs for steganalysis: an empirical study[C]. *The 4th ACM Workshop on Information Hiding and Multimedia Security*, 2016: 103-107.
- [19] Xu G, Wu H Z, Shi Y Q. Structural design of convolutional neural networks for steganalysis [J]. *IEEE Signal Processing Letters*, 2016, 23(5): 708-712.
- [20] Fridrich J, Filler T. Practical methods for minimizing embedding impact in steganography[C]. *SPIE*, 2007, 6505: 650502.
- [21] Bas P, Filler T, Pevný T. Break our steganographic system: the ins and outs of organizing BOSS[C]. *International Conference on Information Hiding*, 2011: 59-70.
- [22] Kodovský J, Fridrich J, Holub V. Ensemble classifiers for steganalysis of digital media [J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(2): 432-444.